PHYSIQUE STATISTIQUE

B. DIU, C. GUTHMANN, D. LEDERER, B. ROULET



Ludwig Bolzmann. 1844-19

HERMANN ÉDITEURS DES SCIENCES ET DES ARTS

Éléments de physique statistique

COLLECTION ENSEIGNEMENT DES SCIENCES, 37

COLLECTION ENSEIGNEMENT DES SCIENCES

1.	Henri Cartan	Théorie élémentaire des fonctions analytiques
3.	Laurent Schwartz	Méthodes mathématiques pour les sciences physiques
5.	Roger Godement	Cours d'algèbre
7.	Gustave Choquet	L'enseignement de la géométrie
14.	Paul Malliavin	Géométrie différentielle intrinsèque
16.	Claude Cohen-Tannoudji et *al.*	Mécanique quantique (2 volumes)
17.	François Chapeville et *al.*	Biochimie
20.	Jean-Pierre Lafon	Les formalismes fondamentaux de l'algèbre commutative
24.	Jean-Pierre Lafon	Algèbre commutative
25.	Jean Bussac, Paul Reuss	Traité de neutronique
27.	Luc Valentin	Le monde subatomique
28.	Luc Valentin	Physique subatomique II
29.	Luc Valentin	L'univers mécanique
30.	Pierre Laszlo	Cours de chimie organique
31.	Michel Sakarovitch	Optimisation. I. Graphes et programmation
32.	Michel Sakarovitch	Optimisation. II. Programmation discrète
33.	Pierre Bergé et *al.*	L'ordre dans le chaos
35.	M. Blanchard-Desce et *al.*	Chimie organique expérimentale
36.	Ch. Vidal, H. Lemarchand	La réaction créatrice
37.	Bernard Diu *et al.*	Éléments de physique statistique
38.	Jacques Baranger et *al.*	Analyse numérique
39.	Luc Valentin	Noyaux et particules. Modèles et symétries
40.	F. Diener, G. Reeb	Analyse non standard
41.	P. Papon, J. Leblond	Thermodynamique des états de la matière
42.	Laurent Schwartz	Analyse I. Théorie des ensembles et topologie
43.	Laurent Schwartz	Analyse II. Calcul différentiel et équations différentielles
44.	Laurent Schwartz	Analyse III. Calcul intégral
45.	Laurent Schwartz	Analyse IV. Applications de la théorie de la mesure
46.	Christian Vidal *et al.*	Au-delà de l'équilibre
47.	Laurent Jolivet	La déformation des continents
48.	Sextant	Optique expérimentale
49.	Éric Suraud	Physique des collisions nucléaires
50.	Pierre Bergé et *al.*	L'espace chaotique
51.	Éric Suraud	La matière nucléaire. Des étoiles aux noyaux
52.	Florence Daumarie et *al.*	Florilège de chimie pratique
53.	Hervé Lemarchand et *al.*	Thermodynamique de la chimie
54.	Patrick Iglesias	Symétries et moment
55.	Florence Daumarie et *al.*	Cap*ès* chimie
56.	Michel Joffre	Électrophysiologie moléculaire I. Le *patch clamp*
57.	Michel Joffre	Électrophysiologie moléculaire II. Les canaux ioniques

Bernard Diu, Claudine Guthmann,
Danielle Lederer, Bernard Roulet

Éléments de physique statistique

HERMANN ÉDITEURS DES SCIENCES ET DES ARTS

Bernard Diu	*Professeur à l'université Paris VII; il est co-auteur, avec Claude Cohen-Tannoudji et Franck Laloë, de* Mécanique quantique, *publié chez Hermann.*
Claudine Guthmann	*Professeur à l'université Paris VII.*
Danielle Lederer	*Maître de conférences à l'université Paris VII.*
Bernard Roulet	*Professeur à l'université Paris VII.*

Publié avec le concours du ministère de l'Éducation nationale, de l'Enseignement supérieur et de la Recherche (D.I.S.T.N.B.)

Nouveau tirage 2001
ISBN 2 7056 6065 8

© 1989, HERMANN, ÉDITEURS DES SCIENCES ET DES ARTS, 293 RUE LECOURBE, 75015 PARIS

Toute reproduction ou représentation de cet ouvrage, intégrale ou partielle, serait illicite sans l'autorisation de l'éditeur et constituerait une contrefaçon. Les cas strictement limités à usage privé ou de citation sont régis par la loi du 11 mars 1957.

Structure de l'ouvrage : « Mode d'emploi »

Ce livre se compose de *deux parties* imbriquées : les chapitres et les compléments. Les six *chapitres* constituent un *cours*, exposant de façon progressive les notions, méthodes et résultats fondamentaux; leur compréhension ne nécessite la connaissance d'aucun des compléments. Les *compléments* sont de courtes *monographies* illustrant ou prolongeant le cours, *indépendantes* les unes des autres (sauf exception); chacun d'eux peut être compris à partir des chapitres qui le précèdent; les compléments qui suivent un chapitre donné peuvent être abordés dans un ordre quelconque (un court texte de présentation oriente le choix), et il n'est en aucun cas conseillé de les lire tous avant d'entreprendre le chapitre suivant.

Chaque chapitre ou complément se présente sous la forme d'un *texte principal* accompagné de *notes de bas de page.* Le but est de suggérer et de faciliter une *« lecture à deux niveaux » :* le texte principal, lu cursivement en ignorant les notes, donne une première idée globale du sujet; si l'on désire le comprendre plus à fond, on en reprendra l'étude en s'aidant cette fois des notes. On ne cherchera toutefois pas à les assimiler toutes du premier coup : la plupart sont simples (les calculs et démonstrations techniques sont par exemple systématiquement renvoyés en note), mais quelques-unes s'adressent aux lecteurs déjà avertis.

Les *exercices* proposés après chaque ensemble chapitre-compléments sont groupés par thèmes et classés, dans chaque groupe, par ordre de difficulté croissante.

Le symbole $\widehat{=}$ signifie « égale par définition ».
Le sigle C.D.L. renvoie à C. Cohen-Tannoudji, B. Diu et F. Laloë, *Mécanique quantique,* Hermann (1985).

Table

Avant-propos XI

CHAPITRE I
Notions et outils de base de la mécanique statistique 1

I. *Introduction : le macroscopique et le microscopique* 3
II. *Probabilités, information et entropie statistique* 8
III. *États d'un système de taille macroscopique* 12
IV. *Principe ergodique et ensembles statistiques* 26

Compléments du chapitre I 34

A. *Ordres de grandeur* 35
B. *Distribution statistique d'une fonction de variables aléatoires* 53
C. *La distribution binomiale dans la limite des grands nombres* 62
D. *Marche au hasard et mouvement brownien* 72
E. *Densité d'états d'un gaz parfait* 82
F. *États électroniques dans les solides cristallins* 85
G. *Information et entropie* 98
H. *L'opérateur-densité* 114
Exercices I 130

CHAPITRE II
Système isolé à l'équilibre - Distribution microcanonique 143

I. *La distribution microcanonique* 145
II. *Propriétés et comportement d'un système isolé* 153
III. *Équilibre entre sous-systèmes d'un système isolé* 158

Compléments du chapitre II 172

A. *Couplage entre systèmes de taille macroscopique* 173
B. *Le cristal paramagnétique parfait. Températures négatives* 180
C. *Le gaz parfait classique : entropie de mélange et paradoxe de Gibbs* 190
D. *Transformations thermodynamiques et variations d'entropie* 201
E. *Mouvement macroscopique d'un système totalement isolé en équilibre statistique* 221
F. *Distribution et entropie microcanoniques en mécanique classique* 236
Exercices II 242

CHAPITRE III
Système en équilibre avec un thermostat. Description canonique 253

I. *Situation et distribution canoniques* 255
II. *Fonction de partition et énergie libre* 267
III. *Le formalisme canonique à la limite thermodynamique* 280
IV. *Description canonique des systèmes de particules indépendantes* 286

Compléments du chapitre III 308

A. *Paramagnétisme et diamagnétisme des systèmes de particules indépendantes* 309
B. *Les gaz parfaits polyatomiques* 329
C. *La théorie cinétique des gaz* 350
D. *Ensemble T-p* 365
E. *Chaleur spécifique des solides* 378
F. *Les solutions diluées* 399
G. *L'équation de van der Waals et la transition gaz-liquide* 415
H. *Systèmes de particules chargées* 434
J. *Le ferromagnétisme* 444
K. *Transition ordre-désordre dans les alliages binaires* 478
L. *Chaleur et travail. Le «second principe» de la thermodynamique* 497
Exercices III 506

CHAPITRE IV
Évolution vers l'équilibre et irréversibilité 535

I. *L'équation maîtresse* 537
II. *Principales implications de l'équation maîtresse* 545

Compléments du chapitre IV 561

A. *Notions succinctes de théorie ergodique* 562
B. *Perturbations aléatoires d'un système «isolé» : discussion physique et ordres de grandeur* 572
C. *Propriétés de l'équation maîtresse* 582
D. *«Postulat d'évolution» et équation de Schrödinger* 590
E. *Évolution d'un gaz classique dilué : l'équation de Boltzmann* 603
F. *Équations générales d'évolution d'un fluide classique* 621
Exercices IV 630

CHAPITRE V
Système en équilibre avec un réservoir de particules. Description grand-canonique 645

I. *Situation et distribution grand-canoniques* 647

II. *Grande fonction de partition et grand potentiel* 654
III. *Le formalisme grand-canonique à la limite thermodynamique* 663
IV. *Description grand-canonique d'un système de particules indépendantes identiques et indiscernables* 667

Compléments du chapitre V 675

A. *Adsorption* 676
B. *Équilibre d'un système macroscopique dans un champ de forces extérieur* 688
C. *Les gaz réels et le développement du viriel* 704
D. *Le postulat fondamental sous sa forme « généralisée »* 717
E. *Utilisation du formalisme des opérateurs-densités* 729
F. *Fluctuations dans un système à l'équilibre statistique* 745
Exercices V 757

CHAPITRE VI
Gaz parfaits quantiques 771

I. *Gaz de fermions indépendants* 774
II. *Gaz de bosons indépendants* 794
III. *Gaz de photons. Thermodynamique du rayonnement* 818

Compléments du chapitre VI 833

A. *Propriétés électroniques des métaux et des semi-conducteurs* 834
B. *Les naines blanches* 860
C. *Potentiel chimique et chaleur spécifique d'un gaz de bosons libres au voisinage de la température de Bose* 874
D. *L'hélium à basse température* 878
E. *Émission et absorption de rayonnement par un atome. Coefficients d'Einstein* 906
F. *Absorption et émission de rayonnement par un corps macroscopique. Loi de Kirchhoff* 911
G. *Le rayonnement cosmique à trois kelvins* 919
Exercices VI 925

Appendices 945

Appendice I Notions élémentaires sur les probabilités 947
Appendice II Description quantique d'un système physique 965
Appendice III Description classique d'un système physique 981
Appendice IV La méthode des multiplicateurs de Lagrange 990

Index 993

Avant-propos

Ce livre est issu d'un enseignement de mécanique statistique et thermodynamique que nous avons donné pendant plusieurs années au niveau de la maîtrise de physique. Nous avons été au départ très fortement influencés par le cours polycopié que Philippe Nozières a rédigé lorsque, après la réforme des programmes universitaires de 1967, il a inauguré cet enseignement de second cycle à la Faculté des Sciences de Paris.

L'idée directrice de notre démarche est de présenter d'abord la mécanique statistique élémentaire pour en déduire ensuite la thermodynamique. La mécanique statistique constitue en effet, avec la mécanique quantique et la relativité, l'un des « piliers » de la physique moderne; il est important que les étudiants la perçoivent d'emblée comme une théorie fondamentale, et non pas comme une simple tentative pour justifier *a posteriori* la thermodynamique. La thermodynamique elle-même y gagne en retour une compréhension plus juste et plus profonde de ses « principes » et de ses méthodes. Le présent livre expose la mécanique statistique et explicite son articulation avec la thermodynamique, mais il n'aborde pas le cours de thermodynamique proprement dite; celui-ci devra évidemment, pour tirer le meilleur profit des connaissances acquises par les étudiants, être fondé, non pas sur l'énoncé des traditionnels « principes », mais sur les propriétés fondamentales de l'entropie, démontrées en mécanique statistique.

La rédaction de cet ouvrage est guidée par deux préoccupations principales. En premier lieu, nous nous sommes efforcés de présenter tous les sujets abordés de façon suffisamment explicite et claire pour les rendre accessibles à des étudiants du niveau de la maîtrise de physique. Ensuite, nous avons essayé d'illustrer au fur et à mesure les notions théoriques par des applications physiques. Évidemment, ces deux exigences ne sont que rarement compatibles au premier abord, mais notre travail de mise au point a constamment eu pour but d'expliquer clairement et sans tricherie, au niveau élémentaire que nous nous sommes fixé, les phénomènes physiques que nous considérons comme importants et intéressants.

Du point de vue technique, le livre réunit deux parties distinctes mais imbriquées : les chapitres et les compléments. En outre, nous avons fait dans la rédaction un usage systématique des notes de bas de page, pour alléger le texte principal. La raison d'être de ces procédés et le « mode d'emploi » qui en découle sont présentés sur une page spéciale. La lecture attentive de cette page est vivement conseillée avant toute utilisation de l'ouvrage.

En ce qui concerne le fond, nous avons opté pour une présentation de la mécanique statistique qui distingue très nettement les différentes distributions statistiques (microcanonique, canonique, grand-canonique, ...) des états microscopiques d'un système à l'équilibre suivant la situation physique dans laquelle est placé ce système (isolé, en contact avec un thermostat, avec un réservoir de particules, ...). Évidemment, ces diverses descriptions donnent les mêmes résultats si la taille du système est suffisamment grande (limite thermodynamique), ce qui est le cas dans la plupart des applications. Mais les hypothèses fondamentales et les méthodes spécifiques de la mécanique statistique se dégagent mieux dans cette présentation; elles sont en tout cas mieux perçues par les étudiants car elles apparaissent de façon plus progressive et plus explicite, et sont donc plus faciles à cerner et à assimiler.

La mécanique statistique a pour but d'expliquer le comportement des systèmes macroscopiques à partir de leurs caractéristiques microscopiques. C'est de façon générale la mécanique quantique qui décrit les propriétés et l'évolution des systèmes physiques à l'échelle microscopique. La mécanique statistique est donc construite sur cette description quantique. Pour des raisons pédagogiques, cependant, nous avons voulu éviter que les idées physiques et les raisonnements propres à la mécanique statistique ne soient obscurcis par des difficultés de compréhension liées à la mécanique

quantique. C'est pourquoi nous exposons la mécanique statistique dans un formalisme simplifié (probabilités des états propres de l'hamiltonien) qui ne nécessite que des connaissances très élémentaires de mécanique quantique. Seuls quelques compléments, qui introduisent puis développent le formalisme des opérateurs-densités, et certaines notes font appel à des notions plus avancées. En fait, le formalisme simple que nous utilisons partout ailleurs est essentiellement équivalent à celui qui s'appuie sur l'opérateur-densité; ce dernier ne s'avère vraiment nécessaire que dans certains problèmes techniquement compliqués comme la mise en œuvre de méthodes d'approximation un peu sophistiquées ou la démonstration rigoureuse de résultats théoriques généraux. En outre, le cadre que nous avons choisi se prête mieux à la construction progressive, chez l'étudiant, d'une intuition physique en mécanique statistique.

Ce livre s'adresse en premier lieu aux étudiants de second cycle de physique des Universités : ce sont eux qui définissent le niveau scientifique et technique du lecteur potentiel. Certains compléments, toutefois, se situent plutôt au niveau du troisième cycle, bien qu'ils soient totalement compréhensibles à partir des chapitres qui les précèdent. Mais la juxtaposition et l'imbrication des chapitres et des compléments, du texte principal et des notes, des notions théoriques et des applications physiques, ainsi que la diversification des structures (très hiérarchisées ou au contraire « éclatées »), visent à faciliter des utilisations variées : nous espérons que l'étudiant pourra apprendre seul le sujet en suivant le déroulement des chapitres et en étudiant quelques compléments, que l'enseignant pourra bâtir un cours de maîtrise ou de D.E.A. en puisant dans les chapitres et les compléments, et que le chercheur débutant pourra consulter le livre avec profit sur tel point particulier.

Nous n'avons pas cherché à établir une bibliographie complète : les ouvrages de physique statistique sont très nombreux. Contentons-nous ici de citer quelques livres que nous avons trouvés particulièrement utiles au niveau de la maîtrise de physique.

En français :

R. Balian, *« Du microscopique au macroscopique »,* Cours de Physique Statistique de l'École Polytechnique, tomes 1 et 2. Ellipses (1982).

L. Landau et E. Lifchitz, *Physique Statistique,* Édition M. I. R., Moscou (1967).

En anglais :

F. Reif, *Fundamentals of Statistical and Thermal Physics,* McGraw-Hill, New York (1965).

R. Kubo, *Statistical Mechanics,* North-Holland, Amsterdam (1971).

D. McQuarrie, *Statistical Mechanics,* Harper and Row, New York (1976).

S. K. Ma, *Statistical Mechanics,* World Scientific, Singapore (1985).

G. Wannier, *Statistical Physics,* Wiley, New York (1966).

K. Huang, *Statistical Mechanics,* Wiley, New York (1967).

C. Kittel and H. Kroemer, *Thermal Physics,* Freeman, San Francisco (1980).

Nous tenons à remercier chaleureusement les collègues avec qui nous avons enseigné ces dernières années, et tout particulièrement Françoise Balibar, Odile Betbeder-Matibet et Laurence Viennot, ainsi que Christiane Caroli : leurs remarques, leurs suggestions et leurs encouragements nous ont été précieux. Nous sommes également très reconnaissants à Patricia Méralli, Stéphane Dupin, Laure-Anne Némirouvsky et Marie-Claire Bégot, qui ont dactylographié avec beaucoup de patience notre manuscrit, et à Romain Lécuiller qui a bien voulu dessiner les figures.

I

I. INTRODUCTION : LE MACROSCOPIQUE ET LE MICROSCOPIQUE

A. « Définitions » et bref historique
   1. La physique macroscopique
   2. L'« hypothèse atomique » et ses développements
   3. Ordres de grandeur et « définitions »

B. La mécanique statistique
   1. Le problème à résoudre : du microscopique au macroscopique
   2. La méthode statistique
   3. Mécanique statistique et mécanique quantique

II. PROBABILITÉS, INFORMATION ET ENTROPIE STATISTIQUE

A. Définition de l'entropie statistique

B. Principales propriétés
   1. Positivité et symétrie
   2. Minimum
   3. Maximum (nombre d'événements possibles fixé)
   4. Variation avec le nombre d'événements possibles
   5. Additivité

III. ÉTATS D'UN SYSTÈME DE TAILLE MACROSCOPIQUE

A. Description quantique du système
   1. L'état microscopique : un « état pur »
   2. L'état macroscopique : un « mélange statistique »
      a. Définition d'un mélange statistique d'états purs
      b. Caractérisation de l'état macroscopique

B. La densité d'états (microscopiques)
   1. Définition et utilité de la densité d'états
      a. Sommation sur les états et sommation sur les énergies
      b. Passage à une intégrale sur l'énergie
   2. Exemple : particules libres enfermées dans une boîte macroscopique
      a. Particule unique
      b. Système de N particules libres et indépendantes
   3. Généralisation à un système de N éléments faiblement couplés

C. Description d'un système classique
   1. État microscopique classique
   2. État macroscopique

IV. PRINCIPE ERGODIQUE ET ENSEMBLES STATISTIQUES

A. Évolution temporelle d'un système macroscopique
   1. Analyse d'un exemple simple
      a. Évolution au niveau microscopique : « agitation thermique » et fluctuations
      b. Évolution au niveau macroscopique
   2. Généralisation à un système macroscopique quelconque

B. Moyennes sur les fluctuations microscopiques
   1. Moyennes temporelles
      a. Définition
      b. Description de l'évolution macroscopique
   2. Moyennes sur un ensemble
      a. Notion d'ensemble statistique
      b. Caractérisation d'un ensemble statistique
      c. Évolution de l'ensemble dans le temps
   3. Discussion : le principe ergodique

CHAPITRE I

Notions et outils de base de la mécanique statistique

I. INTRODUCTION : LE MACROSCOPIQUE ET LE MICROSCOPIQUE

La mécanique statistique, dont l'initiateur fut Ludwig Boltzmann (1844-1906), permet de comprendre comment les propriétés des corps à l'échelle courante, que nous qualifierons de macroscopique, s'expliquent à partir de celles de leurs constituants microscopiques (atomes, molécules, électrons, ions). Avant d'introduire les notions et outils qui serviront de base à l'étude de la mécanique statistique, nous allons consacrer ce premier paragraphe à en préciser le but général et à indiquer l'idée fondamentale qui permet d'atteindre ce but, au moins en principe.

A. « Définitions » et bref historique

1. LA PHYSIQUE MACROSCOPIQUE

Jusqu'à la seconde moitié du dix-neuvième siècle, la science avait pour objet l'étude de phénomènes directement perceptibles par les sens de l'homme, même si la logique de cette étude amenait le plus souvent les scientifiques à remplacer la perception sensorielle directe par des mesures objectives effectuées à l'aide d'appareils de plus en plus sophistiqués. Dans le champ de la physique, on étudiait ainsi la mécanique, l'électricité, le magnétisme, l'optique, l'acoustique, mais aussi la thermodynamique et les changements d'état de la matière, dont nous reparlerons en détail dans la suite. Il est important de noter que les différents domaines de recherche que nous venons d'énumérer restaient, au dix-neuvième siècle, pratiquement indépendants les uns des autres, car les divers aspects de la réalité auxquels on s'intéressait alors semblaient n'avoir aucun lien entre eux; en particulier, les thermodynamiciens étaient fort loin d'imaginer (la résistance farouche opposée par la plupart d'entre eux aux idées de Boltzmann le montre à l'évidence) que la théorie qu'ils construisaient pût être rattachée à l'un quelconque des autres domaines, à la mécanique en particulier.

Bien entendu, la physique macroscopique, c'est-à-dire la physique à l'échelle humaine où à une échelle plus vaste (astronomie, mécanique céleste), a survécu au dix-neuvième siècle. Elle continue à poser — et à résoudre — des problèmes extrêmement intéressants, et son développement n'est sans doute pas près de s'arrêter.

2. L'« HYPOTHÈSE ATOMIQUE » ET SES DÉVELOPPEMENTS

Mais le dix-neuvième siècle vit aussi l'accession au niveau scientifique d'une idée philosophique fort ancienne, puisqu'elle remonte à l'Antiquité. L'étude quantitative des réactions chimiques révéla des lois (loi des « proportions définies », loi des « proportions multiples ») qui s'interprétaient de façon très convaincante dans le cadre de l'*« hypothèse atomique »*, selon laquelle les divers réactifs sont des amas ou des agrégats de constituants microscopiques (atomes et molécules). Cette hypothèse, considérée d'abord comme une façon de présenter les choses (« tout se passe comme

si... »), puisqu'il semblait impensable de pouvoir mettre en évidence directement les atomes, gagna progressivement du terrain au cours du dix-neuvième siècle et finit par s'imposer définitivement au début du vingtième sous la forme d'une « théorie atomique », maintenant universellement admise.

Malgré ses succès, la physique macroscopique était ainsi condamnée à perdre tôt ou tard son caractère fondamental au profit d'une *physique microscopique :* les phénomènes observés à notre échelle sont des conséquences, plus ou moins directes et identifiables comme telles, de phénomènes et de réalités sous-jacents mettant en jeu les constituants microscopiques des objets considérés; ce sont par conséquent ces constituants microscopiques et leurs propriétés qu'il convient d'étudier si l'on recherche une compréhension fondamentale de l'ensemble du monde physique.

Née au début de ce siècle, la physique microscopique s'est révélée extrêmement riche et a connu des développements aussi spectaculaires qu'imprévus : mise en évidence de la structure de l'atome, puis de celle du noyau, découverte de la mécanique quantique, étude des propriétés microscopiques des solides et des liquides, puis plus récemment recherche des particules élémentaires...

3. ORDRES DE GRANDEUR ET « DÉFINITIONS »

Essayons maintenant de préciser, au moins semi-quantitativement, les contours des deux domaines, macroscopique et microscopique.

Le passage du microscopique au macroscopique fait intervenir le *nombre d'Avogadro*

$$N_A = 6{,}02 \times 10^{23} \,, \tag{I.1}$$

introduit de la façon suivante. L'étude quantitative des réactions chimiques conduit à attribuer à chaque corps pur une masse caractéristique, celle de ce que nous appelons aujourd'hui une *mole*[1] du corps pur considéré : ainsi, une mole d'eau a pour masse 18 g, une mole d'hydrogène (moléculaire) 2 g, etc. Une mole est donc une quantité macroscopique. Mais, comme nous l'indiquions au paragraphe 2, la possibilité même d'associer de telles masses caractéristiques aux divers corps purs est une manifestation de leur caractère « granulaire », c'est-à-dire de l'existence des molécules : le rapport 18 g/2 g entre les masses d'une mole d'eau et d'une mole d'hydrogène est tout simplement celui qui existe entre les masses des molécules H_2O et H_2 correspondantes. Le nombre d'Avogadro (I.1) est alors le rapport, pour un corps pur quelconque, entre la masse d'une mole et celle d'une molécule de ce corps, c'est-à-dire le *nombre de molécules*[2] *constituant une mole.*

Le nombre d'Avogadro est *fantastiquement grand.* C'est ce qui explique que l'existence des atomes et molécules ait été difficile à mettre en évidence, et qu'elle ait pu faire l'objet de polémiques passionnées. L'énormité du nombre d'Avogadro implique une *séparation nette des ordres de grandeur* entre le domaine macroscopique et le domaine microscopique : on pourra par exemple définir sans difficulté une fraction de mole suffisamment petite pour être inappréciable dans une mesure de masse au niveau macroscopique (disons 10^{-8} ou 10^{-9} mole), mais suffisamment grande pour contenir encore un très grand nombre de molécules (10^{14} à 10^{15} dans cet exemple).

L'étude de chaque domaine particulier de la physique s'accompagne d'un *choix d'unités adaptées au domaine :* disons qu'une unité est considérée comme adaptée si les nombres qui mesurent la grandeur physique correspondante sont la plupart du temps compris entre 10^{-4} et 10^{+4} (avec éventuellement des « pointes » occasionnelles

1. La mole est définie à partir de l'isotope 12 du carbone : par définition, la masse d'une mole de ^{12}C est exactement 12 g.

2. La notion de mole et la signification du nombre d'Avogadro se généralisent aux cas où la structure du corps considéré n'est pas moléculaire, mais ionique.

jusqu'à 10^{-6} ou 10^{+6}); si ces nombres étaient systématiquement très grands ou très petits, on changerait d'unité pour les ramener à des valeur plus raisonnables. Cette notion d'unités adaptées va nous permettre de « définir » de façon semi-quantitative les termes « macroscopique » et « microscopique ».

Dans la suite, seront dits *macroscopiques* les phénomènes et grandeurs pour lesquels les unités adaptées sont essentiellement celles du système légal M.K.S.A., ou éventuellement leurs premiers multiples ou sous-multiples : ainsi, les longueurs macroscopiques seront exprimées en *mètres* (ou en centimètres, ou même en millimètres), les énergies macroscopiques en *joules* (ou millijoules). Bien entendu, on est encore, *a fortiori,* dans le domaine macroscopique s'il faut utiliser des unités plus grandes que les précédentes. Au contraire, nous qualifierons de *microscopiques* les phénomènes et grandeurs se situant au niveau des atomes ou molécules, ou éventuellement à un niveau inférieur : ici, les unités adaptées seront l'*électron-volt* ($1\ \text{eV} = 1{,}6 \times 10^{-19}$ J) pour les énergies[3], et l'*angström* ($1\ \text{Å} = 10^{-10}$ m) pour les longueurs.

Ces questions d'ordres de grandeur sont reprises en détail dans le complément I.A.

B. La mécanique statistique

1. LE PROBLÈME A RÉSOUDRE : DU MICROSCOPIQUE AU MACROSCOPIQUE

Il ne fait pas de doute — sans que ceci constitue en aucune manière un jugement de valeur — que *la physique microscopique est plus fondamentale que la physique macroscopique :* la nature profonde des phénomènes doit être recherchée au niveau microscopique; essentiellement, le niveau macroscopique perçoit seulement certaines conséquences du comportement des constituants microscopiques. Une loi macroscopique (telle que la loi d'Ohm ou la loi de Boyle-Mariotte) ne peut être qu'un « reflet » des lois microscopiques, et on doit chercher « d'où elle vient ».

Dans certains cas, les propriétés microscopiques sont directement reconnaissables dans leurs manifestations macroscopiques. Par exemple, le spectre d'énergie de l'ion Cu^{++} explique de manière convaincante la couleur bleue des sels de cuivre. Ou bien, la force s'exerçant entre l'électron et le proton de l'atome d'hydrogène est la même que celle qui a été découverte par Coulomb au niveau macroscopique[4]. Ceci ne signifie pas que l'on puisse alors se dispenser d'examiner comment se fait le passage du microscopique au macroscopique : il faudra expliquer pourquoi ce changement radical d'ordres de grandeur ne modifie pas fondamentalement certaines lois ou propriétés. Mais on s'attend à une explication simple. Pourtant, même dans des cas de ce genre, l'origine profonde, microscopique, des phénomènes se manifeste par des

3. On reste dans le microscopique si les énergies atteignent le MeV (10^6 eV) ou même le GeV (10^9 eV) : il y a en effet un facteur 10^{19} entre l'électron-volt et le joule. Cependant, on entre avec les MeV et les GeV dans le domaine relativiste, que nous aborderons peu dans la suite.

4. Lorsqu'on électrise une boule de dimensions macroscopiques, on met en jeu de l'ordre de 10^{10} électrons. La force entre deux boules de ce type est donc environ 10^{20} fois plus intense qu'entre un électron et un proton placés à la même distance, ce qui la rend mesurable macroscopiquement mais ne change pas sa nature et ses caractéristiques. Bien entendu, dans cette situation, un électron de plus ou de moins sur 10^{10} ne fait pas une différence appréciable : au niveau macroscopique, la quantification de la charge (c'est-à-dire son caractère « granulaire ») passe totalement inaperçue, et on peut développer une théorie macroscopique de l'électrostatique et de l'électrocinétique dans laquelle les quantités d'électricité sont traitées comme continues (on parle de quantités d'électricité « infinitésimales » dq, qui sont pourtant très grandes devant la charge élémentaire).

effets caractéristiques, qui n'ont pas leur place dans une théorie purement macroscopique : par exemple, l'agitation thermique, au niveau microscopique, des électrons dans les conducteurs donne lieu à un « bruit » dans les circuits électriques (c'est-à-dire à un courant aléatoire, changeant sans arrêt de sens, d'intensité faible mais parfaitement décelable).

On pourrait conclure naïvement de ce qui précède que seule importe la physique microscopique, puisque sa compréhension entraînerait aussitôt celle de ses manifestations macroscopiques. Tout au plus — mais ce serait déjà une restriction importante — admettrait-on qu'une explication quantitative détaillée de tel phénomène macroscopique puisse nécessiter des calculs trop longs pour pouvoir être vraiment menés à bien, même à l'aide d'ordinateurs. Des situations de ce genre sont courantes en physique : par exemple, sans sortir du cadre microscopique, on considère que la structure de l'atome de plomb (82 électrons) est comprise en principe lorsqu'on a développé la théorie de l'atome d'hydrogène.

Mais *une telle conclusion serait fondamentalement incorrecte.* Il existe en effet de nombreux phénomènes macroscopiques qui échappent à une explication directe, même seulement qualitative, à partir des lois microscopiques. C'est le cas par exemple pour la liquéfaction d'un gaz, ou pour le fonctionnement d'une machine à vapeur. Ici, les manifestations macroscopiques de la structure et des propriétés du système au niveau microscopique sont loin d'être aisément identifiables comme telles, et apparaissent même comme totalement inattendues : *le passage du microscopique au macroscopique se présente comme hautement non trivial,* et comme devant nécessiter une étude approfondie et particulière.

2. LA MÉTHODE STATISTIQUE

Ce problème étant posé, comment peut-on espérer le résoudre? La méthode qui va être utilisée tire parti du fait que ce passage du microscopique au macroscopique met en jeu des nombres extrêmement grands.

Considérons, pour fixer les idées, 1 mm^3 d'un gaz comme l'air, à la pression atmosphérique et à la température ordinaire. Il contient environ[5]

$$N \simeq 3 \times 10^{16} \tag{I.2}$$

molécules. Supposons, pour que l'argument soit plus aisément formulable, que les mouvements (microscopiques) de ces molécules soient décrits par la mécanique newtonienne (voir complément I.A, § III.2.b). Pour caractériser l'état d'une molécule particulière à un instant donné, il faut alors au moins 6 nombres (mesurant les 3 coordonnées de sa position et les 3 composantes de sa vitesse). Donc, pour fixer l'état, à un instant donné, du système que constitue ce mm^3 de gaz, il faudrait connaître environ 2×10^{17} nombres. Ceci est *totalement exclu :* songez par exemple qu'un ordinateur qui écrirait 40 nombres par ligne et une ligne par cm devrait utiliser 50 milliards de km de papier pour répertorier cet état instantané! Même si c'était envisageable, la masse d'informations correspondantes serait strictement *inexploitable*[6].

---

5. Dans les « conditions normales » de température et de pression (0 °C et 1 atmosphère), une mole d'un gaz occupe un volume de 22,4 l approximativement. Donc

$$N = \frac{N_A}{22{,}4 \times 10^6 \text{ mm}^3} \simeq 3 \times 10^{16}/\text{mm}^3 .$$

6. Les ordinateurs les plus puissants sont actuellement capables de « suivre » le mouvement de systèmes comportant tout au plus quelques milliers de particules. On reste ainsi très au-dessous du nombre N de la formule (I.2).

Ces nombres faramineux sont à mettre en opposition avec le nombre restreint de caractéristiques macroscopiques que l'on peut envisager de mesurer sur cet échantillon de gaz : volume, pression, température, chaleur spécifique, viscosité, indice de réfraction... Si l'on tente une énumération, même en faisant preuve d'imagination, on reste loin du 10^{17} précédent.

On comprend alors que le passage du microscopique au macroscopique ne puisse pas toujours se faire de façon directe, comme cela avait été envisagé au paragraphe 1, c'est-à-dire en calculant « par la force brute » le comportement de systèmes de « plusieurs » objets microscopiques. *Les caractéristiques individuelles* de ces objets *sont à la fois inaccessibles* (puisqu'il est même inconcevable de pouvoir les répertorier) *et sans intérêt* pour expliquer les propriétés macroscopiques des systèmes. On conçoit aussi que ces propriétés puissent être *qualitativement différentes* de celles d'un ensemble de 2, 3, 10 ou 100 objets microscopiques, et donc que l'explication de certains phénomènes macroscopiques à partir des lois de la physique microscopique s'avère effectivement non triviale.

Mais le fait même d'avoir affaire à des grands nombres suggère l'*utilisation de méthodes probabilistes.* En effet, si les caractéristiques microscopiques de si grands systèmes sont inaccessibles, c'est que les mesures macroscopiques sont forcément « grossières » : elles ne sont pas sensibles au comportement individuel des objets microscopiques, mais seulement à des *moyennes* (d'un type qui reste à préciser) sur ces grands nombres d'objets[7]. Les méthodes probabilistes destinées à expliciter le lien entre le microscopique et le macroscopique constituent ce qu'on appelle la *mécanique statistique.* C'est à l'étude de ces méthodes et à leur application dans des situations concrètes qu'est consacré ce livre (les notions et résultats essentiels de la théorie des probabilités sont rappelés dans l'appendice I; les compléments I.B, I.C et I.D développent quelques points un peu plus techniques).

On pourrait être déçu d'apprendre que la mise en œuvre au niveau macroscopique des lois microscopiques ne puisse se faire « que » de façon probabiliste, ce qui donne une impression d'approximatif et d'incomplet. Pourtant, l'*énormité des nombres* qui interviennent est rassurante à ce sujet : la « loi des grands nombres », qui régit les applications des méthodes probabilistes, aboutit à des *prédictions pratiquement certaines,* c'est-à-dire à une *description pratiquement exacte* du niveau macroscopique.

Pour préciser un peu, reprenons l'exemple de l'échantillon de gaz évoqué ci-dessus. Supposons que nous prélevions ainsi 1 mm^3 d'air dans une pièce. On s'attend à ce qu'il contienne $N = 3 \times 10^{16}$ molécules (formule (I.2)). On sait cependant que l'agitation incessante des molécules d'air dans la pièce conduit à des *fluctuations* du nombre de celles qui sont contenues à un instant donné dans le volume considéré : à certains moments, ce nombre est inférieur au nombre N attendu, à d'autres il lui est supérieur. Mais il a les plus grandes chances d'être compris entre $N - \sqrt{N}$ et $N + \sqrt{N}$ (les fluctuations sont d'ordre $\sqrt{N}$. Voir appendice I, § II.1.c.γ). Il faut alors noter

7. Pour illustrer ce point fondamental, prenons encore un exemple. Considérons un fil conducteur faisant partie d'un circuit électrique. Les électrons de conduction qu'il contient s'agitent dans tous les sens, de façon désordonnée, à des vitesses de plusieurs centaines de km/s. En l'absence de générateur dans le circuit, c'est-à-dire en l'absence de champ électrique dans le fil, le courant est nul parce que, « en moyenne », il y a à chaque instant autant d'électrons qui se dirigent dans un sens que dans le sens opposé. Lorsqu'un générateur maintient dans le fil un champ électrique, cet équilibre est rompu à cause de la force électrique s'exerçant dans le même sens sur tous les électrons; il y a maintenant, « en moyenne », plus d'électrons se dirigeant dans le sens de la force qu'en sens inverse, et c'est cette différence qui se manifeste sous la forme du courant électrique : l'ensemble du « nuage » d'électrons se déplace en bloc dans le fil. Bien entendu, l'agitation thermique ne cesse pas pour autant. Il faut d'ailleurs souligner que la vitesse d'ensemble du « nuage » (vitesse moyenne des électrons) n'est que de quelques mm/s, et donc considérablement plus petite que les vitesses individuelles; c'est pourtant elle qui régit le courant macroscopique.

que, bien que $\sqrt{N} \simeq 1{,}7 \times 10^8$ soit lui-même grand, la *fluctuation relative* $\sqrt{N}/N$ est de l'ordre de $0{,}6 \times 10^{-8}$, c'est-à-dire impossible à mettre en évidence par des mesures macroscopiques de densité (les variations de densité que l'on peut constater d'un point à l'autre d'une pièce ont des causes totalement étrangères à ces fluctuations statistiques : inhomogénéités de température et de pression, influence de la pesanteur, ...).

3. MÉCANIQUE STATISTIQUE ET MÉCANIQUE QUANTIQUE

Un point reste à clarifier, dans cette discussion des liens entre le microscopique et le macroscopique. Quand il a été question, au paragraphe 2, du mouvement des molécules d'un gaz, nous avons supposé qu'il était décrit correctement par la mécanique newtonienne. C'est effectivement le cas, comme le montre le complément I.A. Mais cette situation est très particulière : de façon générale, on sait que, *au niveau microscopique,* c'est la *mécanique quantique* qu'il convient d'utiliser. C'est ce que nous ferons dans la suite, sauf cas particuliers analogues au précédent. Les connaissances de mécanique quantique dont nous aurons vraiment besoin restent à un niveau très élémentaire. Elles sont rappelées dans l'appendice II.

Une des caractéristiques essentielles de la mécanique quantique est qu'elle ne donne elle-même que des prédictions de nature probabiliste. Il convient cependant de bien comprendre que cette nouvelle intervention des probabilités est fondamentalement différente, dans son origine, de celle que nous avons discutée plus haut. En effet, *la description quantique d'un système,* même microscopique, *est probabiliste dans son principe :* même si l'on possède sur un système toute l'information qu'il est théoriquement possible d'obtenir (par la mesure simultanée d'un ensemble complet d'observables qui commutent), les prédictions que fournit la mécanique quantique quant aux résultats de mesures ultérieures pouvant être effectuées sur ce système s'expriment en termes de probabilités. La mécanique statistique vient ensuite y superposer ses propres méthodes probabilistes, puisqu'*il n'est pas possible de connaître parfaitement l'état quantique d'un système macroscopique.* Les probabilités interviennent donc à deux niveaux : au niveau microscopique, parce que l'état véritable du système doit en général être décrit par la mécanique quantique; au niveau macroscopique, parce que cet état véritable n'est pas connu et qu'on n'a que des renseignements fragmentaires sur ce qu'il peut être.

II. PROBABILITÉS, INFORMATION ET ENTROPIE STATISTIQUE

La mécanique statistique est fondée sur l'utilisation des *probabilités,* utilisation justifiée par l'énormité du nombre d'Avogadro (§ I.B.2).

Or, de façon générale, lorsqu'on traite par la technique des probabilités les résultats possibles d'une expérience, c'est qu'on n'en connaît pas l'issue de façon certaine à l'avance : on ne possède pas toute l'information qui serait nécessaire pour prédire exactement le résultat de cette expérience[8]. La *théorie de l'information,* qui s'est développée au cours des trente dernières années, enseigne comment *mesurer quantitativement l'information,* à l'aide de l'*entropie statistique,* dont nous donnerons ici la définition; celle-ci, déjà découverte par Boltzmann, est à la base de l'*interprétation microscopique de l'entropie thermodynamique.*

La théorie de l'information démontre l'expression de l'entropie statistique à partir d'hypothèses (tout à fait raisonnables) sur les propriétés de la fonction qui doit servir à mesurer l'information. Nous suivrons ici la démarche inverse : nous poserons

8. Ce manque d'information peut résulter de circonstances contingentes (si ces circonstances changeaient, on pourrait posséder toute l'information nécessaire à la prédiction exacte du résultat), ou de principes fondamentaux comme c'est le cas en mécanique quantique (§ I.B.3 et appendice II).

cette expression et expliciterons ensuite ses principales propriétés, montrant ainsi que l'entropie statistique peut légitimement prétendre mesurer l'information. Le complément I.G donne quelques précisions supplémentaires sur l'information et l'entropie statistique.

A. Définition de l'entropie statistique

Considérons un ensemble discret fini[9] d'événements possibles e_m ($m = 1, 2, \ldots, M$), de probabilités respectives P_m ($m = 1, 2, \ldots, M$) vérifiant bien sûr les conditions

$$0 \leqslant P_m \leqslant 1 \quad \forall m \,, \tag{II.1.a}$$

$$\sum_{m=1}^{M} P_m = 1 \,. \tag{II.1.b}$$

Si l'une de ces probabilités, P_{m_0}, est égale à 1, toutes les autres sont nulles : l'événement e_{m_0} est certain, tous les autres impossibles. Dans ce cas, notre information sur le résultat de l'expérience est complète : il va être à coup sûr e_{m_0}. Par rapport à ce cas très particulier, toute autre distribution de probabilités $\{P_m\}$ correspond à un certain *manque d'information,* que nous voulons chiffrer. C'est ce que fait l'entropie statistique $S(P_1, P_2, \ldots, P_M)$ définie par[10]

$$\boxed{S(P_1, P_2, \ldots, P_M) \triangleq -k \sum_{m=1}^{M} P_m \ln P_m} \,, \tag{II.2}$$

avec[11]

$$P_m \ln P_m = 0 \quad \text{si} \quad P_m = 0 \,. \tag{II.3}$$

Cette définition comporte un facteur k *positif* que l'on peut choisir à sa convenance. En mécanique statistique, on prend pour k la *constante de Boltzmann*

$$k = 1{,}38 \times 10^{-23} \text{ J/K} \,. \tag{II.4}$$

J désigne ici le *joule*, unité d'énergie, et K le *degré kelvin*, unité de température. L'entropie statistique obtenue avec le choix (II.4) possède donc des dimensions (celles de k). L'avantage de ce choix apparaîtra plus tard : l'entropie statistique ainsi définie s'identifiera exactement, dans les situations d'équilibre, avec l'entropie thermodynamique.

---

9. Nous étendrons la définition au cas d'un ensemble d'événements possibles infini dénombrable, à condition que la somme de la formule (II.2) converge. Le passage au continu pose ici des problèmes spécifiques (cf. complément I.G, § III), mais nous n'utiliserons par la suite la formule que dans le cas discret.

10. L'entropie statistique permet également de mesurer la quantité d'information gagnée ou perdue dans une opération, ce qui explique qu'elle soit à la base de la théorie de l'information et de la communication.

11. $\ln P_m$ n'est pas défini pour $P_m = 0$, mais

$$\lim_{P_m \to 0} P_m \ln P_m = 0 \,,$$

ce qui justifie le choix (II.3). Ainsi, les événements impossibles ne contribuent pas à l'entropie statistique. Cette propriété est indispensable pour des raisons de cohérence : on peut toujours rajouter à l'ensemble d'événements possibles un nombre quelconque d'événements de probabilité nulle; l'information manquante ne doit pas changer de valeur dans une telle manipulation.

B. Principales propriétés

Il est facile de se convaincre, à partir de la définition (II.2), que l'entropie statistique possède toutes les propriétés que l'on est en droit d'attendre d'une *mesure du manque d'information* que l'on a sur un système.

1. POSITIVITÉ ET SYMÉTRIE

Remarquons tout d'abord que la condition (II.1.a) vérifiée par les probabilités P_m entraîne que S est *définie positive :*

$$S \geqslant 0 \quad \forall \{P_m\} . \tag{II.5}$$

D'autre part, S est une fonction *totalement symétrique* de ses M arguments : par exemple,

$$S(P_2, P_1, P_3, \ldots, P_M) = S(P_1, P_2, P_3, \ldots, P_M) . \tag{II.6}$$

2. MINIMUM

L'entropie statistique atteint son minimum, qui est *nul* (cf. inégalité (II.5)), lorsque l'une des probabilités vaut 1, les autres étant par conséquent toutes nulles :

$$S(0, 0, \ldots, 1, 0, \ldots, 0) = 0 . \tag{II.7}$$

Effectivement, le cas envisagé ici est celui que nous avons présenté au début du paragraphe A comme comportant une information complète, et il est donc naturel que le manque d'information y prenne la valeur nulle.

3. MAXIMUM (NOMBRE D'ÉVÉNEMENTS POSSIBLES FIXÉ)

On démontre aisément[12] que, pour M *fixé,* la fonction $S(P_1, P_2, \ldots, P_M)$ est *maximum lorsque les M événements sont équiprobables,* c'est-à-dire lorsque

$$P_1 = P_2 = \ldots = P_M = \frac{1}{M} \tag{II.8}$$

(la dernière égalité provenant de la condition de normalisation (II.1.b.)). Ce maximum vaut

$$S_{\max} = k \ln M , \tag{II.9}$$

comme on le voit immédiatement en portant les valeurs (II.8) dans la définition (II.2).

C'est effectivement dans le cas où toutes les probabilités sont égales que le manque d'information est maximum : rien ne permet alors d'indiquer que tel ou tel événement a plus de chances de se produire que les autres.

---

12. Pour chercher le maximum de la fonction $S(P_1, P_2, \ldots, P_M)$ pour M fixé, il faut tenir compte du fait que *les M variables P_m ne sont pas indépendantes,* puisqu'elles doivent vérifier la condition de normalisation (II.1.b). Étant donné la simplicité de cette relation, on peut éliminer l'une des M variables en l'exprimant en fonction des $(M-1)$ autres, qui sont alors indépendantes (exercice I.8). Mais il existe une méthode générale pour traiter ce genre de problèmes, dite *méthode des multiplicateurs de Lagrange* (appendice IV), et c'est elle que nous allons appliquer ici.

Nous cherchons les extrema de la fonction $S(P_1, P_2, \ldots, P_M)$ sachant que les M variables sont liées par la relation

$$\sum_{m=1}^{M} P_m - 1 = 0 . \tag{1}$$

La méthode de Lagrange consiste à chercher les extrema de la fonction

$$\widetilde{S}_a(P_1, P_2, \ldots, P_M) \widehat{=} S(P_1, P_2, \ldots, P_M) - a \left[\sum_{m=1}^{M} P_m - 1 \right] \tag{2}$$

4. VARIATION AVEC LE NOMBRE D'ÉVÉNEMENTS POSSIBLES

Pour caractériser la variation de l'entropie statistique en fonction du nombre M d'événements possibles, il faut choisir une distribution particulière de probabilités qui ne fausse pas la comparaison entre valeurs différentes de M. La distribution uniforme (II.8) obtenue au paragraphe 3 précédent répond à cette exigence, et c'est pourquoi il est intéressant de constater que *la valeur maximale de l'entropie statistique croît avec le nombre d'événements possibles,* comme le montre la formule (II.9).

Ici encore, cette propriété est conforme à l'idée intuitive que l'on se fait d'une mesure du manque d'information : si l'on est confronté à un ensemble d'événements équiprobables, l'information que l'on possède sur le système est évidemment d'autant moindre que le nombre d'événements possibles est plus grand.

5. ADDITIVITÉ

Considérons deux ensembles *indépendants* d'événements possibles,

$$\{e_m;\ m=1,\ 2,\ \ldots,\ M\}\ ;\quad \{e'_{m'};\ m'=1,\ 2,\ \ldots,\ M'\}\ , \tag{II.10}$$

et les probabilités correspondantes

$$\{P_m;\ m=1,\ 2,\ \ldots,\ M\}\ ;\quad \{P'_{m'};\ m'=1,\ 2,\ \ldots,\ M'\}\ , \tag{II.11}$$

convenablement normalisées :

$$\sum_{m=1}^{M} P_m = 1\ ;\quad \sum_{m'=1}^{M'} P'_{m'} = 1\ . \tag{II.12}$$

Les résultats possibles d'une expérience « double », dans laquelle se produisent l'un des événements e_m et l'un des $e'_{m'}$, ont pour probabilités (d'après la loi (15) de l'appendice I)

$$P_{m,m'} \mathrel{\widehat{=}} P(e_m \text{ et } e'_{m'}) = P_m \,.\, P'_{m'}\ . \tag{II.13}$$

D'après la définition (II.2), le manque d'information concernant l'expérience double est mesuré par l'entropie statistique

$$S(P_{m,m'};\ m=1,\ 2,\ \ldots,\ M;\ m'=1,\ 2,\ \ldots,\ M') = -k \sum_{m=1}^{M} \sum_{m'=1}^{M'} P_{m,m'} \ln P_{m,m'}\ . \tag{II.14}$$

---

en faisant comme si les M variables étaient indépendantes; a est un paramètre (« multiplicateur ») qui sera ensuite déterminé en imposant la condition (1) à la solution trouvée. On peut montrer que cette solution est bien alors celle du problème initial (voir appendice IV).

Dérivons donc la fonction (2) par rapport à chacune des variables; compte tenu de l'expression (II.2) de S, nous obtenons

$$\frac{\partial}{\partial P_n} \widetilde{S}_a = -k \left[\ln P_n + 1 + \frac{a}{k} \right]\ ;\quad n=1,\ 2,\ \ldots,\ M\ . \tag{3}$$

Annuler ces M dérivées est immédiat :

$$\ln P_n = -\left(1 + \frac{a}{k}\right)\ . \tag{4}$$

On en déduit aussitôt que les M probabilités sont égales entre elles. En imposant finalement la relation (1), on trouve bien le résultat (II.8).

Ce raisonnement montre seulement que la distribution de probabilités (II.8) conduit à un extremum de S. Mais, comme cet extremum est unique et que nous connaissons des distributions qui rendent S inférieure à la valeur (II.9) correspondante (voir paragraphe 2), il ne peut s'agir que d'un maximum.

Si l'on porte dans cette expression les valeurs (II.13) des probabilités composées $P_{m,m'}$, on trouve

$$\begin{aligned} S(P_{m,m'}) &= -k \sum_{m=1}^{M} \sum_{m'=1}^{M'} P_m P'_{m'} \ln (P_m P'_{m'}) \\ &= -k \sum_{m=1}^{M} P_m \sum_{m'=1}^{M'} P'_{m'} [\ln P_m + \ln P'_{m'}] \\ &= -k \left\{ \sum_{m'=1}^{M'} P'_{m'} \sum_{m=1}^{M} P_m \ln P_m + \sum_{m=1}^{M} P_m \sum_{m'=1}^{M'} P'_{m'} \ln P'_{m'} \right\} \\ &= S(P_m) + S(P'_{m'}) \end{aligned} \tag{II.15}$$

(nous avons introduit la notation simplifiée $S(P_m)$ pour l'entropie statistique, et avons utilisé, dans la dernière étape, les conditions de normalisation (II.12)).

Le manque d'information sur l'expérience double est ainsi *la somme des manques d'information* sur chacun des deux ensembles indépendants d'événements.

III. ÉTATS D'UN SYSTÈME DE TAILLE MACROSCOPIQUE

Les probabilités, que nous étudions succinctement dans l'appendice I, vont être utilisées en mécanique statistique pour décrire l'état macroscopique des systèmes. Nous considérons d'abord le cas général où la description du système est fondée sur la mécanique quantique (§ A), et introduisons la notion de « densité d'états » (§ B); le paragraphe C examine ensuite le cas où les propriétés (même microscopiques) du système sont bien décrites par la mécanique newtonienne classique.

A. Description quantique du système

1. L'ÉTAT MICROSCOPIQUE : UN « ÉTAT PUR »

On apprend en mécanique quantique élémentaire (voir appendice II) que l'état d'un système quelconque est caractérisé par un *vecteur-ket,* appartenant à l'espace des états du système. Ce ket $|\psi(t)\rangle$ dépend du temps t, et son évolution au cours du temps est donnée par l'équation de Schrödinger

$$i\hbar \frac{\mathrm{d}}{\mathrm{d}t} |\psi(t)\rangle = H |\psi(t)\rangle , \tag{III.1}$$

où H est l'opérateur hamiltonien du système considéré.

Si l'hamiltonien H est indépendant du temps (ce qui est très souvent le cas), on dit qu'on a affaire à un *système conservatif.* L'évolution dans le temps d'un état quelconque du système s'obtient alors très simplement si l'on a auparavant résolu *l'équation aux valeurs propres de l'hamiltonien,*

$$H |\varphi_\ell\rangle = E_\ell |\varphi_\ell\rangle . \tag{III.2}$$

Dans cette équation, les valeurs propres E_ℓ sont les *valeurs possibles de l'énergie* du système. Les kets propres $|\varphi_\ell\rangle$ associés sont appelés *« états stationnaires »* du système : on montre facilement (appendice II, § II.2) que, si le système se trouve à un instant t_0 donné dans un tel état $|\varphi_\ell\rangle$, il y demeure indéfiniment et ses propriétés sont indépendantes du temps.

Ces considérations sont valables pour n'importe quel système, grand ou petit. Les états qui sont ainsi caractérisés par un ket de l'espace des états sont appelés *« états purs »* (par opposition aux mélanges statistiques que nous envisagerons au

paragraphe 2). Rappelons que, pour préparer un système dans un état pur, il faut en général effectuer sur ce système une mesure simultanée d'un « ensemble complet d'observables qui commutent » (C.D.L., p. 235) : l'ensemble des résultats obtenus dans une telle mesure suffit à déterminer *le* ket (à un facteur près, sans importance physique) qui décrit l'état du système après la mesure.

Soulignons un point important. Une valeur possible E_ℓ de l'énergie est en général *« dégénérée »*, c'est-à-dire qu'il existe en général plusieurs vecteurs propres linéairement indépendants qui correspondent à la même valeur propre E_ℓ, dans l'équation (III.2). Le nombre $g(E_\ell)$ de ces kets propres indépendants associés à la même énergie E_ℓ est appelé *« degré de dégénérescence du niveau E_ℓ »* (si $g(E_\ell)=1$, le niveau E_ℓ est « non dégénéré »). Dans les cours de mécanique quantique, on affecte le plus souvent un indice supplémentaire aux kets propres pour tenir compte de ce fait (appendice II, § II.2) :

$$H\left|\varphi_\ell^i\right\rangle = E_\ell\left|\varphi_\ell^i\right\rangle \; ; \quad i=1,\ 2,\ \ldots,\ g(E_\ell)\ . \tag{III.3}$$

Nous adopterons ici une *convention différente,* pour éviter la lourdeur due aux indices multiples : *l'indice ℓ repérera les états* stationnaires; si plusieurs états $|\varphi_\ell\rangle$ correspondent à la même énergie, cela se traduira par le fait que plusieurs valeurs propres E_ℓ seront égales; par exemple, si $|\varphi_1\rangle$ et $|\varphi_2\rangle$ sont kets propres de H avec la même valeur propre E, on écrira quand même

$$\begin{aligned} H\left|\varphi_1\right\rangle &= E_1\left|\varphi_1\right\rangle\ , \\ H\left|\varphi_2\right\rangle &= E_2\left|\varphi_2\right\rangle\ , \end{aligned} \tag{III.4}$$

mais on saura que

$$E_1 = E_2 = E\ . \tag{III.5}$$

Pour un système de taille macroscopique, on appellera *« état microscopique »* (on dit quelquefois *« micro-état »*) l'état quantique pur dans lequel il se trouve à un instant donné. Bien entendu, comme au paragraphe I.B.2 où nous raisonnions dans le cadre de la mécanique newtonienne, il est impossible et d'ailleurs sans intérêt de connaître exactement l'état microscopique d'un système de taille macroscopique : le nombre d'observables qu'il faudrait mesurer simultanément est à peine moins grand que celui des paramètres définissant l'état du système en mécanique classique (cf. formule (I.2)). On se contentera donc d'une information (très) incomplète sur l'état du système, ce qui se traduira par l'intervention des probabilités dans la description de cet état imparfaitement connu (voir paragraphe I.B.3 ci-dessus).

2. L'ÉTAT MACROSCOPIQUE : UN « MÉLANGE STATISTIQUE »

a. Définition d'un mélange statistique d'états purs

Pour traiter des situations dans lesquelles l'état (pur) d'un système n'est pas connu, mais où l'on a sur lui une information de type probabiliste, on est amené à généraliser la notion d'état quantique : on dira que l'état du système est un *mélange statistique* des états purs $|\psi_1\rangle$, $|\psi_2\rangle$, ..., $|\psi_m\rangle$, ... avec les probabilités P_1, P_2, ..., P_m, ... si l'on sait que le système a la probabilité P_1 de se trouver dans l'état $|\psi_1\rangle$, la probabilité P_2 d'être dans l'état $|\psi_2\rangle$, ... Bien entendu, ces probabilités doivent vérifier les conditions habituelles :

$$0 \leqslant P_m \leqslant 1 \quad \forall m\ , \tag{III.6.a}$$

$$\sum_m P_m = 1\ . \tag{III.6.b}$$

Précisons comment l'on calcule concrètement les prédictions physiques concernant les propriétés du système lorsque son état est un tel mélange statistique. Supposons par exemple que l'on s'apprête à mesurer une grandeur physique A, et que l'on veuille connaître la probabilité pour obtenir dans cette mesure la valeur a de cette grandeur; nous prendrons pour simplifier une valeur propre a non dégénérée de A, et nous noterons $|u\rangle$ le vecteur propre correspondant. Le raisonnement sera alors le suivant : si le système était dans l'état pur $|\psi_m\rangle$, la probabilité cherchée serait (si nous avons pris soin de normer à 1 $|\psi_m\rangle$ et $|u\rangle$)

$$\mathcal{P}_m(a) = |\langle u|\psi_m\rangle|^2 \; ; \tag{III.7}$$

comme cette éventualité (le système est dans l'état pur $|\psi_m\rangle$) a pour probabilité P_m, il faut multiplier (III.7) par P_m pour trouver la probabilité pour que le système soit dans l'état $|\psi_m\rangle$ *et* qu'on trouve a dans la mesure; enfin, la probabilité totale pour qu'on trouve a (le système étant dans n'importe quel état pur) s'obtient en sommant sur tous les états $|\psi_m\rangle$ possibles :

$$\mathcal{P}(a) = \sum_m P_m \mathcal{P}_m(a) = \sum_m P_m |\langle u|\psi_m\rangle|^2 \; . \tag{III.8}$$

On obtient de façon analogue la valeur moyenne $\overline{A}$ d'une grandeur physique A dans la situation où l'état du système est un mélange statistique : dans l'état $|\psi_m\rangle$, la valeur moyenne de A (que nous noterons pour simplifier A_m) serait

$$A_m = \langle \psi_m | A | \psi_m \rangle \; ; \tag{III.9}$$

la valeur moyenne globale $\overline{A}$ s'obtient comme ci-dessus en pondérant ce résultat par la probabilité P_m de l'éventualité correspondante, et en sommant sur toutes ces éventualités :

$$\overline{A} = \sum_m P_m A_m \; . \tag{III.10}$$

Lorsque l'état du système est un mélange statistique, on a donc tout simplement à effectuer des *moyennes sur la distribution statistique* $\{P_m\}$ *des états purs,* selon les techniques explicitées dans l'appendice I, des diverses propriétés, grandeurs ou quantités que l'on sait calculer dans chaque état pur.

b. Caractérisation de l'état macroscopique

Les mélanges statistiques d'états purs, que nous venons d'introduire au paragraphe a, sont très utiles même pour des systèmes de taille microscopique. Mais, par suite des raisons indiquées à la fin du paragraphe 1 ci-dessus, ils sont indispensables pour la description quantique des systèmes de taille macroscopique : l'état quantique d'un tel système est nécessairement un mélange statistique d'états purs (c'est-à-dire d'états microscopiques, voir paragraphe 1), que l'on appelle son *« état macroscopique »,* ou *« macro-état ».*

Le formalisme le mieux adapté à la caractérisation et à la manipulation des mélanges statistiques d'états purs est celui des opérateurs-densités, qui est décrit succinctement dans le complément I.H. C'est donc dans le cadre de ce formalisme que se développe le plus naturellement la mécanique statistique. Nous ne l'utiliserons pourtant pas ici, par souci pédagogique : nous choisirons une méthode moins générale, mais plus simple et plus intuitive, réservant pour certains compléments les considérations plus abstraites basées sur les opérateurs-densités.

Dans la suite donc, nous caractériserons l'état macroscopique d'un système comme un *mélange statistique d'états purs stationnaires.* De façon plus précise, nous supposerons connus les états propres $|\varphi_\ell\rangle$ de l'hamiltonien H du système et les énergies E_ℓ correspondantes (équation (III.2)). Pour simplifier, *nous noterons simplement* (ℓ) *les états stationnaires* $|\varphi_\ell\rangle$[13]. Nous n'envisagerons que ces états (ℓ) comme états microscopiques possibles du système. Dès lors, un état macroscopique sera caractérisé

par *la donnée de l'ensemble des probabilités* P_ℓ pour que le système se trouve dans chacun des états stationnaires (ℓ) possibles[14]. La raison pour laquelle cette description est en fait suffisante, au moins au niveau élémentaire où nous nous situons ici, découle du rôle primordial que joue l'énergie comme grandeur physique, tant en mécanique statistique qu'en physique microscopique[15].

B. La densité d'états (microscopiques)

Pour un système de taille macroscopique, il est assez souvent suffisant de connaître le nombre d'états microscopiques (ℓ) d'énergie donnée, et ce nombre est fourni par la « densité d'états » du système, que nous allons maintenant définir, puis étudier sur des exemples simples (cette étude est poursuivie dans les compléments I.E et I.F).

1. DÉFINITION ET UTILITÉ DE LA DENSITÉ D'ÉTATS

a. Sommation sur les états et sommation sur les énergies

Nous avons conclu le paragraphe A en indiquant que les propriétés d'un système de taille macroscopique s'obtiennent en effectuant des moyennes sur les états microscopiques possibles; concrètement, on aura à calculer des *sommes sur les états stationnaires* (ℓ), de la forme

$$\overline{A} = \sum_{(\ell)} P_\ell A_\ell \,, \tag{III.11}$$

où P_ℓ est la probabilité pour que le système soit dans l'état microscopique (ℓ) et A_ℓ est la valeur prise dans cet état par la grandeur qui nous intéresse. Insistons à nouveau sur la nécessité de *ne pas confondre niveaux d'énergie et états stationnaires :* il existe en général plusieurs états microscopiques (ℓ) distincts (décrits par des kets linéairement indépendants) associés à une même énergie. Au départ, la sommation dans une expression de la forme (III.11) porte sur les états (ℓ).

Assez souvent cependant, la quantité à sommer ne dépend en fait que de l'énergie E_ℓ de l'état (ℓ). Supposons que ce soit le cas, et posons

$$P_\ell A_\ell = f(E_\ell) \,. \tag{III.12}$$

---

13. Le symbole (ℓ) représente en général un ensemble de plusieurs « nombres quantiques », ceux qui sont nécessaires pour spécifier complètement un état propre de H. Pour un système macroscopique, leur nombre est fantastiquement grand (cf. § B.2.b.α).

14. Nous serons amenés plus loin (cf. § IV.A.2 ci-dessous et chapitre IV, § I.A) à constater que les états (ℓ) que nous allons utiliser, états propres de l'hamiltonien H, ne sont pas rigoureusement stationnaires. Ceci est dû au fait que, pour un système de taille macroscopique, il est impossible de connaître exactement son hamiltonien; l'opérateur H qui nous sert pour définir les états (ℓ) est donc forcément un hamiltonien approché, et ses états propres (ℓ) ne sont qu'approximativement stationnaires.

15. Particulariser comme nous le faisons les états propres (ℓ) de l'hamiltonien H revient à les choisir comme base dans l'espace des états du système. L'importance du rôle de l'énergie en physique se traduit par le fait que les opérateurs-densités de la mécanique statistique à l'équilibre sont diagnonaux dans une telle base, c'est-à-dire que les états stationnaires (ℓ) sont de toute façon privilégiés.

Considérons maintenant une valeur déterminée E_ℓ de l'énergie. Si $g(E_\ell)$ est le nombre d'états (ℓ) distincts qui lui sont associés (degré de dégénérescence), la somme (III.11) comportera $g(E_\ell)$ termes distincts qui seront égaux à $f(E_\ell)$. On pourra alors *remplacer la somme sur les états* (ℓ) *par une somme sur les niveaux d'énergie,* à condition de tenir compte du facteur $g(E_\ell)$ que nous venons d'expliciter; nous écrirons

$$\overline{A} = \sum_{(\ell)} P_\ell A_\ell = \sum_{E_\ell} g(E_\ell) f(E_\ell) \,. \tag{III.13}$$

Comme la notation simplifiée que nous utilisons (un seul indice, ℓ, pour les états stationnaires) introduit ici une légère ambiguïté, précisons que la sommation sur E_ℓ consiste à écrire un terme et un seul pour chaque niveau, c'est-à-dire pour chaque valeur possible de l'énergie : si, comme dans (III.5), deux énergies E_ℓ sont égales, il leur correspond deux termes dans la somme sur les états (ℓ), mais un seul dans la somme sur les niveaux E_ℓ.

b. Passage à une intégrale sur l'énergie

Bien que grands, les systèmes étudiés par la mécanique statistique sont *finis,* c'est-à-dire qu'il existe toujours une « boîte », de volume macroscopique mais fini, qui englobe totalement le système. Ceci a pour conséquence, comme on le montre facilement en mécanique quantique, que *les niveaux d'énergie* de tels systèmes *forment un spectre discret*[16]. C'est pour cela que, dans la formule (III.13), la somme sur les niveaux est une somme discrète Σ, comme celle sur les états.

Cependant, pour un système de dimensions macroscopiques, *l'écart entre deux niveaux consécutifs est très faible :* en effet, la quantification de l'énergie n'est pas perceptible directement à l'échelle macroscopique.

Considérons alors un intervalle d'énergie δE possédant les deux propriétés suivantes :

(i) il est suffisamment petit pour être *inappréciable au niveau macroscopique,* c'est-à-dire que

$$\delta E \ll \delta_0 E \,, \tag{III.14}$$

où $\delta_0 E$ est l'incertitude qui affecte les mesures de l'énergie macroscopique;

(ii) il est suffisamment grand pour contenir *un grand nombre de niveaux microscopiques* E_ℓ, c'est-à-dire qu'il est très grand devant l'écart entre deux niveaux consécutifs :

$$\delta E \gg \left| E_{\ell+1} - E_\ell \right| \,. \tag{III.15}$$

Ces deux exigences ne sont pas contradictoires, et peuvent être simultanément vérifiées pour un système de dimensions macroscopiques. Soit $\delta n(E)$ le *nombre d'états microscopiques* (ℓ) *dont l'énergie* E_ℓ *est comprise entre* E *et* $E + \delta E$. Par définition, la *densité d'états* $\rho(E)$ du système considéré est telle que

$$\delta n(E) \mathrel{\widehat{=}} \rho(E)\, \delta E \,. \tag{III.16.a}$$

Bien entendu, nous utiliserons ici, comme toujours en pareil cas, la notation différentielle :

$$\mathrm{d}n(E) \mathrel{\widehat{=}} \rho(E)\, \mathrm{d}E \tag{III.16.b}$$

est le nombre d'états (ℓ) dont l'énergie est comprise entre E et $E + \mathrm{d}E$.

La connaissance de la densité d'états $\rho(E)$ permet d'évaluer les sommes de la forme (III.13). On s'attend en effet à ce que, pour un système de taille macroscopique,

16. Cette propriété découle directement de la condition que la probabilité de présence des constituants du système (molécules d'un gaz, par exemple) doit s'annuler à l'extérieur de la boîte qui englobe le système : comme dans le cas simple du puits de potentiel carré à une dimension, l'équation aux valeurs propres de l'hamiltonien n'a de solution vérifiant ces conditions aux limites que pour certaines valeurs discrètes de l'énergie (cf. C.D.L., p. 350).

la fonction $f(E_\ell)$ à sommer (formule (III.12)) varie de façon inappréciable entre deux niveaux consécutifs, c'est-à-dire qu'elle prenne pratiquement la même valeur $f(E)$ pour tous les niveaux dont l'énergie est comprise entre E et $E+\mathrm{d}E$. La contribution à la somme (III.13) de l'ensemble de ces niveaux est donc $f(E)$ multipliée par le nombre d'états correspondants $\rho(E)\,\mathrm{d}E$. Pour sommer sur tous les niveaux possibles, il faut ensuite sommer les contributions de tous les intervalles $(E, E+\mathrm{d}E)$ pour E variant de sa borne inférieure E_0 (« niveau fondamental ») à $+\infty$, ce qui revient par définition à calculer une *intégrale sur l'énergie :*

$$\overline{A} \widehat{=} \sum_{(\ell)} P_\ell A_\ell = \int_{E_0}^{+\infty} f(E)\rho(E)\,\mathrm{d}E\,. \tag{III.17}$$

Remarques :

(i) Soulignons encore que $\rho(E)\,\mathrm{d}E$ est le nombre d'*états* dont l'énergie est comprise entre E et $E+\mathrm{d}E$, pas le nombre de niveaux.

(ii) Il est parfois commode de calculer d'abord le *nombre $\Phi(E)$ d'états dont l'énergie est inférieure à E.* Si l'on connaît la fonction $\Phi(E)$, la densité d'états $\rho(E)$ s'en déduit aussitôt, puisque le nombre d'états $\mathrm{d}n\,(E)$ dont l'énergie est comprise entre E et $E+\mathrm{d}E$ est

$$\mathrm{d}n\,(E) = \Phi(E+\mathrm{d}E) - \Phi(E)\,. \tag{III.18}$$

La densité d'états est donc la dérivée de la fonction $\Phi(E)$:

$$\rho(E) = \frac{\mathrm{d}\Phi(E)}{\mathrm{d}E}\,. \tag{III.19}$$

(iii) Il doit être clair d'après ce qui précède que la formule (III.17) donne une *valeur approchée* de la somme initiale (III.11) (complétée par la condition (III.12)). Pour un système de dimensions macroscopiques, cette approximation est *excellente dans la plupart des cas.* On conçoit cependant qu'elle pourra devenir mauvaise si la fonction $f(E_\ell)$ ne varie pas lentement dans certains domaines d'énergie, par exemple pour les tout premiers niveaux; dans un tel cas, la somme discrète (III.13) n'est pas forcément bien reproduite par l'intégrale (III.17). S'il y a doute quant à la validité d'une formule approchée comme (III.17), il faut revenir à l'expression initiale (III.13).

2. EXEMPLE : PARTICULES LIBRES ENFERMÉES DANS UNE BOÎTE MACROSCOPIQUE

Voici un exemple très simple. Il permettra de préciser certains points qui apparaissent plus clairement sur une situation concrète, et de donner une idée de la manière dont la densité d'états $\rho(E)$ varie avec l'énergie E.

a. Particule unique

Nous commençons par un système tellement simple qu'il n'est pas vraiment macroscopique, puisqu'il est constitué d'une seule particule libre (sans spin). Mais nous supposons que cette particule est enfermée dans une boîte de dimensions macroscopiques, ce qui permet de définir valablement une densité d'états. Nous prendrons une boîte parallélépipédique, d'arêtes L_x, L_y, L_z et de volume $V = L_x \,.\, L_y \,.\, L_z$.

L'étude quantique de ce système est résumée dans l'appendice II (§ II.4). Si l'on impose des conditions aux limites périodiques sur les parois de la boîte, on peut prendre pour fonctions d'ondes stationnaires de la particule considérée les ondes planes

$$\varphi_{\vec{K}}(\vec{r}) = \frac{1}{\sqrt{V}}\,\mathrm{e}^{i\vec{K}\cdot\vec{r}}\,; \tag{III.20}$$

$\hbar\vec{K}$ étant l'*impulsion* de la particule dans l'état stationnaire $\varphi_{\vec{K}}$, l'énergie associée à cet état vaut

$$E_{\vec{K}} = \frac{\hbar^2 \vec{K}^2}{2m}\,. \tag{III.21}$$

Les impulsions possibles sont quantifiées : les seuls vecteurs $\vec{K}$ permis par les conditions périodiques sont de la forme

$$\vec{K} = \ell_x \frac{2\pi}{L_x} \vec{e}_x + \ell_y \frac{2\pi}{L_y} \vec{e}_y + \ell_z \frac{2\pi}{L_z} \vec{e}_z \,, \tag{III.22}$$

où les paramètres ℓ_x, ℓ_y, ℓ_z prennent *toutes les valeurs entières relatives.* L'indice que nous appelons (ℓ) dans le cas général est donc ici un *ensemble de trois indices entiers :*

$$(\ell) \equiv (\ell_x,\ \ell_y,\ \ell_z) \,. \tag{III.23}$$

Les énergies (III.21) forment un *spectre discret*[17]. Mais on constate aussi facilement que les écarts entre niveaux consécutifs sont d'autant plus faibles que les dimensions de la boîte L_x, L_y, L_z sont plus grandes; pour des dimensions macroscopiques, *ces écarts sont totalement inappréciables*[18].

Ceci permet de définir et de calculer la densité d'états. D'après (III.21), fixer l'énergie revient à fixer le module du vecteur $\vec{K}$. Compter les états (ℓ_x, ℓ_y, ℓ_z) dont l'énergie est comprise entre E et $E + \mathrm{d}E$ revient donc à compter les vecteurs $\vec{K}$ permis dont le module est compris entre K et $K + \mathrm{d}K$, avec

$$K = \sqrt{\frac{2mE}{\hbar^2}} \,, \tag{III.24.a}$$

$$\mathrm{d}K = \frac{1}{2} \sqrt{\frac{2m}{\hbar^2}}\, E^{-1/2}\, \mathrm{d}E \,. \tag{III.24.b}$$

Ce problème se résout facilement[19] par des considérations géométriques dans l'espace à trois dimensions des vecteurs $\vec{K}$. On trouve que le nombre cherché, qui est par

---

17. Ceci est dû aux restrictions sur les valeurs possibles des paramètres ℓ_x, ℓ_y, ℓ_z, et ces restrictions proviennent à leur tour des conditions aux limites : dans le paragraphe II.4 de l'appendice II, les équations (26) imposent seulement que u, v, w soient des exponentielles imaginaires, et ce sont les conditions (29) ou (35) qui fixent les valeurs permises aux vecteurs d'onde. On vérifie donc concrètement sur cet exemple l'affirmation générale énoncée au début du paragraphe 1.b.

18. Pour un électron ($m \simeq 10^{-30}$ kg) enfermé dans une boîte cubique de côté $L = 1$ cm, l'écart entre deux niveaux consécutifs

$$E_{\ell_x+1,\ \ell_y,\ \ell_z} - E_{\ell_x,\ \ell_y,\ \ell_z} = (2\ell_x + 1) \frac{2\pi^2\hbar^2}{mL^2} \tag{1}$$

vaut environ 10^{-33} J pour ℓ_x de l'ordre de 1. Nous verrons (III, § V.C) que, à haute température T, l'ordre de grandeur des entiers ℓ (ℓ_x, ℓ_y ou ℓ_z) importants est donné par

$$\frac{2\pi^2\hbar^2}{mL^2}\, \ell^2 \leqslant \frac{1}{2}\, kT \,. \tag{2}$$

Pour de telles valeurs de ℓ_x, on voit que

$$E_{\ell_x+1,\ \ell_y,\ \ell_z} - E_{\ell_x,\ \ell_y,\ \ell_z} \leqslant \left[kT \frac{4\pi^2\hbar^2}{mL^2} \right]^{1/2} \,. \tag{3}$$

Même à des températures de plusieurs milliers de kelvins, cette différence d'énergie ne dépasse pas 10^{-6} eV.

Noter que les valeurs obtenues seraient encore plus faibles pour des particules de masse plus élevée (atomes, molécules, ...).

19. Les vecteurs $\vec{K}$ qui nous intéressent ont leur extrémité située en l'un des nœuds d'un réseau tridimensionnel (fig. 1) constitué de mailles élémentaires dont les arêtes sont parallèles aux axes

.../...

définition égal à $\rho(E)\,\mathrm{d}E$, vaut

$$\rho(E)\,\mathrm{d}E = \frac{V}{2\pi^2}\,K^2\,\mathrm{d}K\,, \tag{III.25}$$

ce qui donne, compte tenu des relations (III.24),

$$\rho(E) = \frac{V\sqrt{2}}{2\pi^2\hbar^3}\,m^{3/2}E^{1/2}\,. \tag{III.26}$$

Nous expliquons dans l'appendice II (§ II.4.c.α) pourquoi il est plus commode d'utiliser, comme nous l'avons fait ici, des conditions aux limites périodiques sur les

---

et mesurent respectivement $2\pi/L_x$, $2\pi/L_y$ et $2\pi/L_z$. Nous voulons évaluer le nombre de vecteurs pour lesquels cette extrémité tombe dans l'intervalle entre les deux sphères centrées à l'origine et de rayons K et $K+\mathrm{d}K$, donnés par (III.24). Pour des valeurs macroscopiques de L_x, L_y, L_z, les dimensions de la maille élémentaire sont très petites devant l'accroissement $\mathrm{d}K$ considéré, de sorte que le nombre cherché est très grand. On peut alors l'obtenir (à quelques unités près, ce qui est sans importance vu la grandeur du nombre lui-même) de la façon suivante.

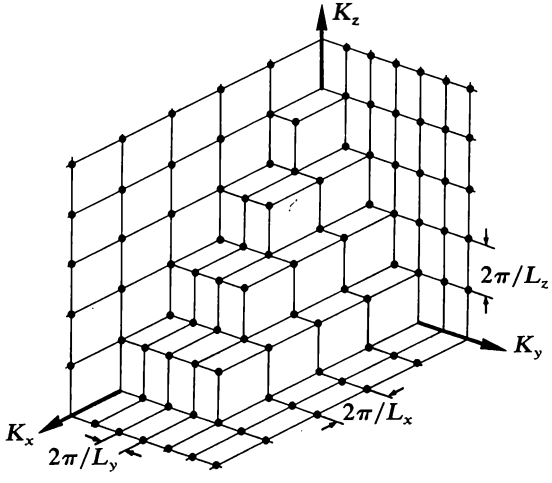


FIGURE 1
Construction géométrique des extrémités « permises » pour les vecteurs $\vec{K}$ repérant les états stationnaires.

La coquille sphérique comprise entre les sphères de rayons K et $K+\mathrm{d}K$ a pour volume, dans l'espace des $\vec{K}$,

$$4\pi K^2\,\mathrm{d}K\,. \tag{1}$$

Le nombre de mailles élémentaires incluses dans cette région de l'espace des $\vec{K}$ est, à peu de chose près, égal au nombre de fois que son volume contient celui de la maille élémentaire, qui vaut

$$\frac{2\pi}{L_x}\cdot\frac{2\pi}{L_y}\cdot\frac{2\pi}{L_z} = \frac{8\pi^3}{V}\,. \tag{2}$$

On obtient ainsi

$$\frac{4\pi K^2\,\mathrm{d}K}{8\pi^3/V}\,. \tag{3}$$

Ce sont les nœuds du réseau, c'est-à-dire les sommets des mailles élémentaires, que nous voulons compter. Or chaque maille possède 8 sommets, mais chaque sommet appartient à 8 mailles distinctes. Il y a donc en moyenne, sur un grand nombre de mailles, un sommet par maille. Le nombre (3) est par conséquent, à une très bonne approximation, celui que nous cherchions.

parois de la boîte plutôt que des conditions « strictes ». On montre sans difficulté[20] que le résultat concernant la densité d'états reste inchangé si l'on raisonne à partir des conditions aux limites « strictes ». C'est pourquoi, lorsque nous avons affaire à des particules libres enfermées dans une boîte, *nous utilisons systématiquement les conditions aux limites périodiques.*

b. Système de *N* particules libres et indépendantes

Nous considérons maintenant le cas, plus réaliste que celui du paragraphe a, de N particules libres (sans spin) enfermées dans une boîte de taille macroscopique. Les N particules sont identiques entre elles, et nous supposons la boîte parallélépipédique (d'arêtes L_x, L_y et L_z, de volume $V = L_x . L_y . L_z$).

α. États du système et états de chaque particule

Les N particules sont *indépendantes* si l'hamiltonien H du système est simplement la somme de N termes dont chacun se rapporte à une seule particule :

$$H = \sum_{i=1}^{N} H_i \,, \tag{III.27}$$

où H_i ne dépend que des observables attachées à la particule numérotée i; *les termes d'interaction* entre les diverses particules, qui dépendraient des observables associées à plusieurs d'entre elles, *sont ici nuls.* Dans ce cas, on montre en mécanique quantique (appendice II, § III.2.a) que l'*on peut construire les états propres de H, par produit (tensoriel), à partir des états propres des N hamiltoniens individuels* H_i (il faut bien sûr prendre N états individuels, un pour chaque particule), *les énergies correspondantes étant sommes des N énergies individuelles.*

Ici, nous supposons en plus que chaque particule est *libre,* ce qui signifie que l'hamiltonien H_i s'écrit simplement

$$H_i = \frac{\vec{P}_i^{\,2}}{2m} \,, \tag{III.28}$$

où $\vec{P}_i$ est l'opérateur impulsion de la particule (i). Nous pouvons donc utiliser pour chaque particule les résultats du paragraphe a ci-dessus.

Mais il apparaît alors une *difficulté importante :* les particules étant identiques, il faut appliquer aux états du système le *postulat de symétrisation* (appendice II, § IV.2); ceci complique considérablement le problème, jusqu'à rendre sa solution exacte impossible, au moins sous la forme où nous l'avons énoncé. Nous allons raisonner ici *comme si les N particules étaient toutes de nature différente,* nous réservant de revenir ultérieurement sur les implications de l'identité des particules[21].

---

20. Dans le cas des conditions aux limites strictes (appendice II, § II.4.b), on peut encore remplacer la donnée des trois nombres (ℓ_x, ℓ_y, ℓ_z) par celle d'un vecteur $\vec{K}$, défini ici par

$$\vec{K} \widehat{=} \ell_x \frac{\pi}{L_x} \vec{e}_x + \ell_y \frac{\pi}{L_y} \vec{e}_y + \ell_z \frac{\pi}{L_z} \vec{e}_z \,; \tag{1}$$

l'énergie a alors pour expression

$$E_{\vec{K}} = \frac{\hbar^2 \vec{K}^2}{2m} \tag{2}$$

comme en (III.21), mais les valeurs prises par chacun des indices ℓ_x, ℓ_y et ℓ_z sont seulement les *nombres entiers positifs* (zéro exclu).

Le calcul de la densité d'états est alors analogue à celui de la note 19, avec les deux différences suivantes. D'une part, la maille élémentaire dans l'espace des $\vec{K}$ a ici pour volume $(\pi/L_x)\cdot(\pi/L_y)\cdot(\pi/L_z)$, c'est-à-dire 8 fois moins que dans le cas précédent. D'autre part, les vecteurs $\vec{K}$ permis doivent avoir leurs trois composantes positives; le volume de la coquille sphérique comprise entre les sphères de rayons K et $K + \mathrm{d}K$ doit donc être divisé par 8. Ces deux facteurs se compensent pour laisser le résultat inchangé.

21. Le cas des systèmes de particules identiques et indiscernables sera traité d'abord de façon

.../...

Un état stationnaire du système est alors caractérisé par la donnée de N vecteurs ordinaires

$$\{\vec{K}_1, \vec{K}_2, \ldots, \vec{K}_N\}, \tag{III.29}$$

où $\hbar\vec{K}_i$ est l'impulsion de la particule (i) si nous choisissons des conditions aux limites périodiques. L'énergie de cet état vaut

$$E = \varepsilon_1 + \varepsilon_2 + \ldots + \varepsilon_N, \tag{III.30}$$

où

$$\varepsilon_i = \frac{\hbar^2 \vec{K}_i^2}{2m} \tag{III.31}$$

est l'énergie de la particule (i). Les composantes de chacun des vecteurs $\vec{K}_i$ doivent être des multiples entiers de $2\pi/L_x$, $2\pi/L_y$ et $2\pi/L_z$ respectivement. Ainsi, le symbole (ℓ) que nous utilisons pour caractériser les états microscopiques du système représente ici $3N$ indices, puisqu'il faut 3 entiers ℓ_x, ℓ_y, ℓ_z pour repérer l'état individuel de chacune des N particules.

β. Densité d'états du système

Nous voulons compter les états stationnaires du système dont l'énergie est comprise entre E et $E + \mathrm{d}E$. La méthode géométrique qui nous a servi au paragraphe a peut être généralisée à l'espace à $3N$ dimensions des vecteurs (III.29); ce calcul est effectué dans le complément I.E. Nous allons raisonner ici de manière différente, et déduire la densité d'états du système, que nous noterons $\rho_N(E)$, de la densité d'états d'une seule particule, $\rho(\varepsilon)$.

Commençons par calculer le nombre $\Phi_N(E)$ d'états du système dont l'énergie est inférieure à E. Un raisonnement simple[22] fournit l'expression

$$\Phi_N(E) = \int_0^{+\infty} \mathrm{d}\varepsilon_1\, \mathrm{d}\varepsilon_2 \ldots \mathrm{d}\varepsilon_N\, \rho(\varepsilon_1)\rho(\varepsilon_2) \ldots \rho(\varepsilon_N)\theta(E - \varepsilon_1 - \varepsilon_2 - \ldots - \varepsilon_N), \tag{III.32}$$

---

approchée dans le complément II.C et au paragraphe IV.A.3 du chapitre III (approximation de Maxwell-Boltzmann), puis plus complètement au chapitre V, paragraphe IV (statistiques quantiques).

22. Pour un système de $N = 2$ particules indépendantes, *si l'on ignore les effets du postulat de symétrisation,* le nombre d'états dans lesquels l'énergie de la particule (1) est comprise entre ε_1 et $\varepsilon_1 + \mathrm{d}\varepsilon_1$ et celle de la particule (2) entre ε_2 et $\varepsilon_2 + \mathrm{d}\varepsilon_2$ est le produit

$$\rho(\varepsilon_1)\, \mathrm{d}\varepsilon_1 \cdot \rho(\varepsilon_2)\, \mathrm{d}\varepsilon_2 \tag{1}$$

des nombres d'états de chacune des particules qui vérifient la condition correspondante : en effet, à cause de l'indépendance des deux particules, à chaque état de la particule (1) dont l'énergie appartient à l'intervalle $(\varepsilon_1, \varepsilon_1 + \mathrm{d}\varepsilon_1)$ on peut associer, pour construire un état du système, n'importe quel état de la particule (2) dont l'énergie appartient à $(\varepsilon_2, \varepsilon_2 + \mathrm{d}\varepsilon_2)$. Dans la formule (1), c'est la même densité d'états $\rho(\varepsilon)$ qui apparaît pour les deux particules, puisque celles-ci sont identiques. Dans la formule (III.32), le produit

$$\rho(\varepsilon_1)\, \mathrm{d}\varepsilon_1 \cdot \rho(\varepsilon_2)\, \mathrm{d}\varepsilon_2 \cdot \ldots \cdot \rho(\varepsilon_N)\, \mathrm{d}\varepsilon_N \tag{2}$$

qui figure dans l'intégrale est de même le nombre d'états du système des N particules indépendantes (traitées comme si elles étaient non identiques) pour lesquels l'énergie de la particule (j) est comprise entre ε_j et $\varepsilon_j + \mathrm{d}\varepsilon_j$, avec $j = 1, 2, \ldots, N$.

Pour obtenir $\Phi_N(E)$, il faut sommer tous les nombres (2) correspondant à toutes les valeurs de $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_N$ telles que

$$\varepsilon_1 + \varepsilon_2 + \ldots + \varepsilon_N < E, \tag{3}$$

puisque l'énergie d'un état du système est la somme des énergies individuelles des particules (quand ces dernières sont indépendantes). La formule (III.32) fait cela en sommant sur toutes les valeurs possibles pour chaque ε_j (qui varie de 0, pour $\vec{K}_j = \vec{0}$, à l'infini), après avoir multiplié par 1 les contributions qui vérifient la condition (3), et par 0 celles qui la violent.

où $\theta(x)$ est la « fonction de Heaviside » :

$$\theta(x) \widehat{=} \begin{cases} 0 & \text{si } x<0 ; \\ 1 & \text{si } x>0 . \end{cases} \tag{III.33}$$

On sait (formule (III.19)) que la densité d'états $\rho_N(E)$ cherchée est la dérivée de $\Phi_N(E)$:

$$\rho_N(E) = \frac{\mathrm{d}}{\mathrm{d}E} \Phi_N(E) . \tag{III.34}$$

Or, E figure au second membre de (III.32) uniquement dans la fonction θ, dont la dérivée est connue[23] :

$$\frac{\mathrm{d}}{\mathrm{d}x} \theta(x) = \delta(x) , \tag{III.35}$$

où δ est la « fonction de Dirac ». On obtient donc

$$\rho_N(E) = \int_0^{+\infty} \mathrm{d}\varepsilon_1 \, \mathrm{d}\varepsilon_2 \ldots \mathrm{d}\varepsilon_N \, \rho(\varepsilon_1)\rho(\varepsilon_2) \ldots \rho(\varepsilon_N)\delta(E - \varepsilon_1 - \varepsilon_2 - \ldots - \varepsilon_N) . \tag{III.36}$$

Ce qui nous intéresse principalement, c'est le comportement de $\rho_N(E)$ comme fonction de l'énergie E du système. Ce comportement peut être obtenu par le raisonnement suivant.

Dans la formule (III.36), la fonction δ assure que la somme des ε_i est égale à E. *Grosso modo,* on s'attend donc à ce que chaque ε_i soit de l'ordre de E/N. Comme ε_i a les dimensions d'une énergie, on est amené à poser

$$\varepsilon_i = \frac{E}{N} \eta_i , \tag{III.37}$$

où *les nouvelles variables* η_i *sont sans dimension.* Dans l'intégrale (III.32), E est bien sûr fixée, de sorte que

$$\mathrm{d}\varepsilon_i = \frac{E}{N} \mathrm{d}\eta_i . \tag{III.38}$$

On sait d'autre part (formule (III.26)) que

$$\rho(\varepsilon) = AV\varepsilon^{1/2} , \tag{III.39.a}$$

avec $$A = \frac{\sqrt{2}}{2\pi^2\hbar^3} m^{3/2} . \tag{III.39.b}$$

On obtient donc

$$\Phi_N(E) = \left(\frac{E}{N}\right)^N A^N V^N \left(\frac{E}{N}\right)^{N/2} \int_0^{+\infty} \mathrm{d}\eta_1 \mathrm{d}\eta_2 \ldots \mathrm{d}\eta_N [\eta_1\eta_2 \ldots \eta_N]^{1/2} \times \theta\left[\frac{E}{N}(N - \eta_1 - \eta_2 - \ldots - \eta_N)\right]$$

$$= A^N V^N \left(\frac{E}{N}\right)^{3N/2} \int_0^{+\infty} \left[\prod_{i=1}^{N} \mathrm{d}\eta_i \eta_i^{1/2}\right] \quad \theta\left[N - \sum_{i=1}^{N} \eta_i\right] \tag{III.40}$$

car $$\theta(ax) = \theta(x) \quad \text{pour} \quad a>0 . \tag{III.41}$$

23. Voir par exemple C.D.L., appendice II, paragraphe 3.a.

L'intégrale qui reste à calculer est purement numérique. Il suffit alors de dériver par rapport à E pour trouver la densité d'états sous la forme

$$\rho_N(E) = C(N)\, V^N \left(\frac{E}{N}\right)^{\frac{3N}{2}-1}, \tag{III.42}$$

où $C(N)$ est un coefficient numérique, indépendant du volume V et de l'énergie E. On peut calculer ce coefficient (cf. complément I.E), et on trouve

$$C(N) = \frac{3}{2}\left(\frac{mN}{2\pi\hbar^2}\right)^{3N/2} \frac{1}{\Gamma\left(\frac{3N}{2}+1\right)} \simeq \frac{1}{2}\sqrt{\frac{3}{N\pi}}\left(\frac{m\,\mathrm{e}}{3\pi\hbar^2}\right)^{3N/2} \tag{III.43}$$

L'aspect le plus important du raisonnement précédent est l'*utilisation de l'analyse dimensionnelle :* il faut seulement comprendre que chaque ε_i, ayant les dimensions d'une énergie, doit être proportionnel à E; on trouve alors immédiatement, à partir de (III.32) et (III.39), le comportement de $\Phi_N(E)$, simplement en comptant les différentielles et les $\rho(\varepsilon)$.

La formule (III.42) montre que, pour N grand, $\rho_N(E)$ est une *fonction très rapidement croissante* de E. Nous serons amenés à en prendre le logarithme, qui est grand (linéaire en N) mais varie lentement avec E (comme ln E).

Attention : la formule (III-43) suppose les N particules discernables. Pour des particules identiques et donc indiscernables, on peut obtenir un résultat approché en divisant $C(N)$ par N ! (cf. note 21).

3. GÉNÉRALISATION A UN SYSTÈME DE *N* ÉLÉMENTS FAIBLEMENT COUPLÉS

Le raisonnement du paragraphe précédent s'appuie principalement sur l'indépendance des N particules du système : c'est elle qui permet d'écrire la formule de départ (III.32), dans laquelle l'énergie du système est la somme des énergies individuelles des N particules et le nombre d'états du système le produit des nombres d'états de chaque particule. Le fait qu'il s'agisse de particules libres est en réalité secondaire, si l'on s'en tient aux ordres de grandeur. C'est ce que nous allons montrer dans ce paragraphe.

Considérons un système constitué de N éléments faiblement couplés. Ceci signifie que l'hamiltonien H du système est par exemple de la forme

$$H = \sum_{i=1}^{N} H_i + \sum_{\substack{i,j=1 \\ i\neq j}}^{N} H_{ij}, \tag{III.44}$$

où H_{ij} fait intervenir à la fois les constituants (i) et (j), mais que *les termes d'interaction* H_{ij} *sont petits* devant les termes individuels H_i[24]. On peut alors, en première approximation, négliger les termes d'interaction, et l'on est ainsi ramené à un système d'éléments indépendants[25].

Commençons par étudier l'un de ces éléments, supposé isolé et décrit par un hamiltonien H. Si les niveaux d'énergie correspondants étaient équidistants (avec un espacement δ) et de même degré de dégénérescence, comme c'est le cas pour l'oscillateur harmonique à une dimension, le nombre $\Phi(\varepsilon)$ d'états distincts dont l'énergie est inférieure à ε serait de la forme

$$\Phi(\varepsilon) \propto \varepsilon/\delta, \tag{III.45}$$

24. Dans de nombreux systèmes, les interactions ne sont pas négligeables. Pourtant, même pour de tels systèmes, on arrive parfois à se ramener au cas considéré ici : il existe des méthodes permettant de remplacer l'hamiltonien H par un hamiltonien approché où cette fois les termes individuels sont grands devant les termes de couplage entre constituants; nous en verrons des exemples dans les compléments III.E à III.K.

25. On peut dans certains cas améliorer l'approximation en traitant ensuite les termes d'interaction par la théorie des perturbations (voir par exemple le complément V.E).

c'est-à-dire linéaire en ε. Dans le cas général, les valeurs propres de H ne sont pas équidistantes, et la formule (III.45) fait intervenir l'espacement moyen δ des valeurs propres inférieures à ε; d'ailleurs, cet espacement moyen doit tenir compte du fait que le degré de dégénérescence n'est pas non plus le même pour tous les niveaux. En bref, δ dépend en général de ε. Cependant, cette dépendance n'est pas si rapide qu'elle puisse changer radicalement l'évaluation (III.45). *Grosso modo,* on aura donc toujours

$$\Phi(\varepsilon) \propto \varepsilon^{\alpha} , \tag{III.46}$$

avec $\alpha \sim 1$. (III.47)

La densité d'états $\rho(\varepsilon)$ s'en déduit par dérivation :

$$\rho(\varepsilon) \propto \varepsilon^{\alpha-1} . \tag{III.48}$$

Passons maintenant au système de N éléments. Dans la mesure où leurs interactions sont négligeables, on peut reprendre la formule (III.32) et l'expression (III.36) qui s'en déduit pour la densité d'états $\rho_N(E)$ du système. Comme nous l'avons souligné à la fin du paragraphe 2.b.β, chaque différentielle $d\varepsilon_i$ amène dans (III.32) un facteur E/N, et chaque $\rho(\varepsilon_i)$ un facteur $(E/N)^{\alpha-1}$ (d'après (III.48)). On obtient ainsi

$$\Phi_N(E) \propto \left(\frac{E}{N}\right)^{N\alpha} . \tag{III.49}$$

La dépendance en énergie de $\rho_N(E)$ est donc semblable à celle que nous avons trouvée plus haut pour des particules libres et indépendantes. C'est l'indépendance (le plus souvent approchée) des divers éléments qui est la propriété essentielle pour aboutir à un résultat de ce type; le fait qu'il ne s'agisse plus de particules libres remplace simplement, dans l'exposant, $3N/2$ par $N\alpha$, où α ne diffère jamais beaucoup de 1.

C. Description d'un système classique

De façon générale, c'est la mécanique quantique qui permet d'étudier les propriétés microscopiques d'un système quelconque, grand ou petit. Il existe cependant des cas où l'on peut valablement utiliser la mécanique newtonienne classique, car, dans certaines conditions, elle décrit avec une bonne approximation le comportement du système (voir paragraphe I-4 du complément I.A). Nous allons indiquer ici les bases de la mécanique statistique classique[26].

1. ÉTAT MICROSCOPIQUE CLASSIQUE

En mécanique classique (voir appendice III), l'état d'une particule ponctuelle à un instant donné est caractérisé par sa position $\vec{r}$ et sa vitesse $\vec{v}$ à cet instant, ou plutôt, dans le formalisme hamiltonien, par la position $\vec{r}$ et l'impulsion $\vec{p}$, moment conjugué de la position. Si donc on a affaire à un système constitué de N particules, son état classique instantané est défini par la donnée de $6N$ paramètres :

$$\{\vec{r}_1, \vec{r}_2, \ldots, \vec{r}_N ; \vec{p}_1, \vec{p}_2, \ldots, \vec{p}_N\} , \tag{III.50}$$

$\vec{r}_j$ et $\vec{p}_j$ étant la position et l'impulsion de la particule numérotée j.

Ceci se généralise à un système quelconque. Pour connaître, à un instant donné, l'état du système (c'est-à-dire la position et la vitesse de chacun de ses points), il faut se donner un certain nombre n de paramètres q_i $(i = 1, 2, \ldots, n)$ que l'on appelle *coordonnées généralisées,* et (dans le formalisme hamiltonien) les n *moments conjugués*

26. Historiquement, la mécanique statistique est apparue avant la mécanique quantique, et elle a donc été fondée sur la mécanique classique.

p_i des paramètres q_i. L'état instantané du système est donc caractérisé par les $2n$ grandeurs

$$\{q_1, q_2, \ldots, q_n; p_1, p_2, \ldots, p_n\} . \tag{III.51}$$

Le mouvement du système est décrit par l'évolution temporelle de son état, c'est-à-dire par les $2n$ fonctions du temps[27]

$$\{q_i(t), p_i(t); i = 1, 2, \ldots, n\} . \tag{III.52}$$

A chaque *grandeur physique* que l'on peut mesurer sur le système est associée une fonction des variables d'état q_i et p_i (cette fonction pouvant en outre dépendre explicitement du temps), qui donne la valeur de cette grandeur lorsqu'on connaît l'état (III.51). C'est ainsi qu'à l'*énergie totale* du système est associée la *fonction de Hamilton* $\mathcal{H}(q_i, p_i; t)$. Par exemple, pour un système de N particules ponctuelles dans lequel les forces dérivent d'une énergie potentielle $U(\vec{r}_1, \vec{r}_2, \ldots, \vec{r}_N)$, la fonction de Hamilton s'écrit simplement

$$\mathcal{H}(\vec{r}_1, \vec{r}_2, \ldots, \vec{r}_N; \vec{p}_1, \vec{p}_2, \ldots, \vec{p}_N) = \sum_{j=1}^{N} \frac{\vec{p}_j^{\,2}}{2m_j} + U(\vec{r}_1, \vec{r}_2, \ldots, \vec{r}_N) \tag{III.53}$$

(m_j étant la masse de la particule (j)) : c'est la somme des énergies cinétiques des N particules et de l'énergie potentielle totale.

Il est commode d'imaginer un espace à $2n$ dimensions ($6N$ dans le cas d'un système de N particules ponctuelles), dit *« espace des phases »*, dans lequel chaque état possible du système est représenté par un point, de coordonnées $q_1, q_2, \ldots, q_n, p_1, p_2, \ldots, p_n$. Le mouvement du système se traduit par un déplacement du point représentatif[28] dans l'espace des phases, où il décrit une trajectoire caractérisée par la donnée des fonctions (III.52).

Ainsi donc, dans les cas où la mécanique classique décrit correctement les propriétés du système étudié, son état microscopique est caractérisé par la donnée de $2n$ variables, que l'on peut considérer comme les coordonnées d'un point dans l'espace des phases associé au système considéré.

2. ÉTAT MACROSCOPIQUE

Si le nombre $2n$ de paramètres nécessaires à la description complète du système est grand (c'est-à-dire par exemple si le nombre N de particules qui le constituent est grand), il est impossible de connaître exactement son état microscopique. On fait alors appel aux probabilités pour définir un état macroscopique, dont la donnée sera suffisante pour calculer les propriétés macroscopiques du système, les seules qui soient vraiment accessibles.

Dans le cas de la mécanique classique, les états microscopiques possibles forment un continuum, commodément représenté par l'espace des phases (§ 1) : l'indice (ℓ) qui repérait les états quantiques est ici remplacé par l'ensemble des $2n$ variables continues (III.51). L'état macroscopique est donc caractérisé par une *densité de probabilité dans l'espace des phases,* c'est-à-dire par une fonction des coordonnées généralisées q_i et de leurs impulsions conjuguées p_i, fonction qui dépend en outre explicitement du temps dans le cas général :

$$w(q_1, q_2, \ldots, q_n; p_1, p_2, \ldots, p_n; t) . \tag{III.54}$$

---

27. Ces fonctions ne sont bien sûr pas quelconques : elles doivent vérifier les *équations du mouvement,* équations différentielles qui expriment la loi fondamentale de la mécanique newtonienne (appendice III, § I.2 et § II.2).

28. Attention : les p_i, comme les q_i, sont ici des coordonnées de *position* du point représentatif dans l'espace des phases; la vitesse de ce point sur sa trajectoire est le vecteur à $2n$ dimensions de composantes $\{\dot{q}_i, \dot{p}_i; i = 1, 2, \ldots, n\}$.

Cette fonction donne la probabilité

$$dP(q_1, q_2, \ldots, q_n; p_1, p_2, \ldots, p_n; t) = w(q_1, q_2, \ldots, q_n; p_1, p_2, \ldots, p_n; t)$$
$$dq_1\, dq_2 \ldots dq_n\, dp_1\, dp_2 \ldots dp_n \qquad \text{(III.55)}$$

pour que, à l'instant t, l'état microscopique du système ait son point représentatif dans le volume infinitésimal $dq_1\, dq_2 \ldots dp_1\, dp_2 \ldots dp_n$ situé au point $\{q_1, \ldots, q_n; p_1, \ldots, p_n\}$ de l'espace des phases.

Bien entendu, cette probabilité est normalisée : à chaque instant t, on a[29]

$$\int w(q_i, p_i; t) \prod_{i=1}^{n} dq_i\, dp_i = 1 , \qquad \text{(III.56)}$$

l'intégrale portant sur tout l'espace des phases. La valeur moyenne d'une grandeur physique quelconque $f(q_i, p_i; t)$ se calcule par la formule habituelle :

$$\bar{f} = \int w(q_i, p_i; t) f(q_i, p_i; t) \prod_{i=1}^{n} dq_i\, dp_i . \qquad \text{(III.57)}$$

Cette valeur moyenne ne dépend plus alors que du temps t.

L'appendice III (§ II.3.b) énonce et démontre une propriété importante de la densité de probabilité dans l'espace des phases (équation de Liouville).

IV. PRINCIPE ERGODIQUE ET ENSEMBLES STATISTIQUES

Nous allons préciser dans ce paragraphe IV l'hypothèse sur laquelle est fondée la description probabiliste des systèmes en mécanique statistique. Nous utiliserons d'abord un exemple simple pour présenter cette hypothèse, puis nous indiquerons comment elle se généralise et comment elle se traduit par l'introduction des probabilités. Nous terminerons par une discussion succincte de cette hypothèse fondamentale.

A. Évolution temporelle d'un système macroscopique

1. ANALYSE D'UN EXEMPLE SIMPLE

Considérons tout d'abord un gaz, constitué d'un grand nombre N, fixé, de molécules identiques. Le mouvement de ces molécules, au niveau microscopique, peut être traité par la mécanique classique (voir complément I.A, § III.2), ce qui nous permettra de raisonner de façon plus intuitive.

Lorsqu'on examine l'évolution d'un gaz au cours du temps, on est amené à distinguer deux types de mouvements.

a. Évolution au niveau microscopique : « agitation thermique » et fluctuations

Supposons en premier lieu que le gaz considéré soit à l'équilibre du point de vue macroscopique. Ce sera le cas si les conditions extérieures qui lui sont imposées (volume du récipient, température, ...) ne varient pas, et si l'on a attendu suffisamment longtemps après le remplissage du récipient.

Nous savons que, en réalité, cet équilibre macroscopique cache une agitation constante au niveau microscopique : les molécules sont en mouvement incessant, à

29. L'intégrande de la formule (III.56) n'est rien d'autre que l'expression (III.55) écrite dans une notation simplifiée.

des vitesses assez considérables (~ 500 m/s); elles se choquent entre elles et heurtent les parois du récipient, changeant ainsi très souvent de vitesse, en grandeur et en direction. L'apparent équilibre est dû au fait que les molécules sont extrêmement nombreuses, et qu'elles subissent de nombreuses collisions : il y a ainsi, à chaque instant, pratiquement autant de molécules qui se trouvent dans la moitié droite du récipient que dans la moitié gauche, et pratiquement autant qui sont animées d'une vitesse $\vec{v}$ que de la vitesse $-\vec{v}$. Ce mouvement au niveau microscopique est souvent appelé *«agitation thermique»*, car les vitesses qu'il fait intervenir sont d'autant plus importantes que la température est plus élevée.

L'agitation thermique s'accompagne de *fluctuations*. Par exemple, le nombre de molécules qui se trouvent dans la moitié droite du récipient n'est pas exactement et constamment égal à $N/2$: il fluctue au cours du temps autour de cette valeur moyenne, l'amplitude de ces fluctuations étant petite en valeur relative $(\sim \sqrt{N}/N)$, mais non nulle. De même, le nombre de molécules animées de la vitesse $\vec{v}$ est tantôt légèrement supérieur, tantôt légèrement inférieur à celui des molécules qui ont à cet instant la vitesse $-\vec{v}$, car ces nombres fluctuent eux aussi au cours du temps. Rappelons que ces fluctuations dues à l'agitation au niveau microscopique peuvent être mises en évidence expérimentalement, ce qui démontre la réalité des molécules et de leur mouvement.

Pour concrétiser les arguments qui vont suivre, imaginons que nous puissions compter, à chaque instant t, les molécules qui se trouvent dans une région déterminée et fixe du récipient contenant le gaz (fig. 2.a)[30]. Les nombres $n(t)$ obtenus, portés en fonction du temps t, ont l'allure représentée sur la figure 2.b : fluctuations de faible amplitude relative autour d'une valeur centrale n_0.

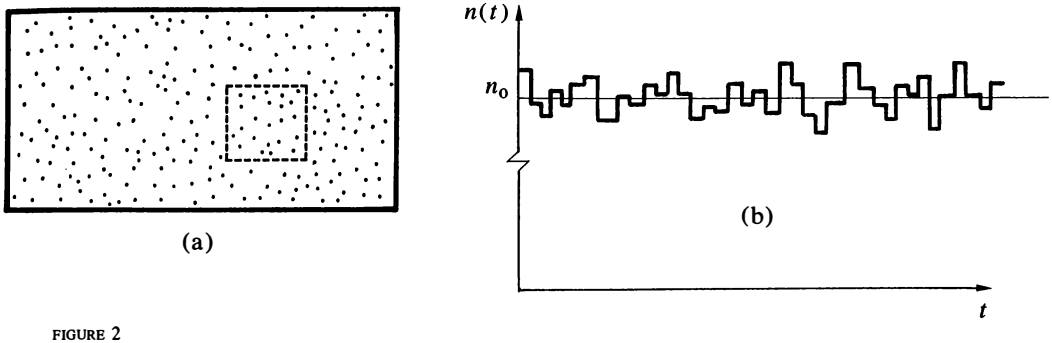


FIGURE 2

Dans un gaz à l'équilibre macroscopique, on mesure le nombre de molécules $n(t)$ qui se trouvent à l'instant t dans une région fixe du récipient, délimitée par un trait tireté sur la figure a. L'allure des variations de $n(t)$ est indiquée sur la figure b.

b. Évolution au niveau macroscopique

Le gaz considéré au paragraphe a était à l'équilibre du point de vue macroscopique, mais ce n'est pas toujours le cas. Imaginons par exemple que le récipient contenant le gaz soit séparé en deux compartiments par une cloison amovible

30. Pour «compter» ainsi les molécules, une expérience réelle devrait mesurer la *densité* du gaz dans la région considérée, avec une précision suffisante pour être sensible aux fluctuations de cette grandeur. Celle-ci a bien entendu été choisie pour la commodité de l'exposition; toute autre grandeur macroscopique aurait essentiellement le même comportement.

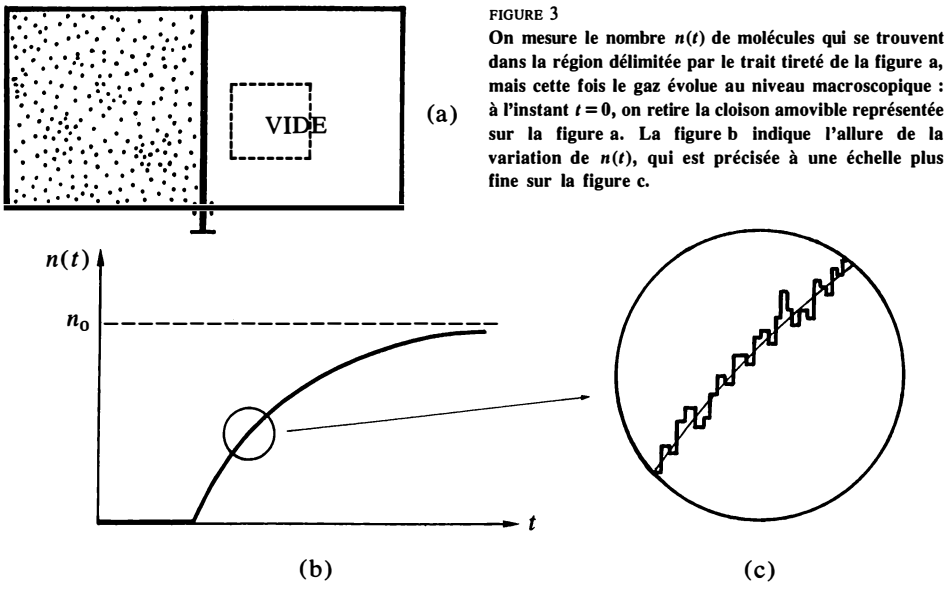


FIGURE 3
On mesure le nombre $n(t)$ de molécules qui se trouvent dans la région délimitée par le trait tireté de la figure a, mais cette fois le gaz évolue au niveau macroscopique : à l'instant $t = 0$, on retire la cloison amovible représentée sur la figure a. La figure b indique l'allure de la variation de $n(t)$, qui est précisée à une échelle plus fine sur la figure c.

(fig. 3.a). On prépare le système de telle sorte que l'un de ces compartiments soit vide, le gaz se trouvant dans l'autre. Si, à un instant donné, on retire la cloison, le gaz va se répandre dans le compartiment initialement vide, et occuper finalement tout le volume disponible. La variation au cours du temps du nombre $n(t)$ de molécules contenues dans la même région du récipient va évidemment être ici qualitativement différente de celle qui était observée à l'équilibre macroscopique : l'expansion du gaz se traduit par une variation macroscopique de n, qui passe ainsi de 0 à n_0 (fig. 3.b). Mais si l'on regarde ce qui se passe à une échelle plus fine (fig. 3.c), on constate que cette variation macroscopique se fait elle aussi de façon irrégulière, car l'agitation au niveau microscopique est bien sûr toujours présente.

2. GÉNÉRALISATION À UN SYSTÈME MACROSCOPIQUE QUELCONQUE

Dans l'exemple que nous venons d'analyser, le mouvement microscopique du gaz peut être décrit par la mécanique classique, ce qui facilite la compréhension intuitive des phénomènes. Mais, en général, la description d'un système à l'échelle microscopique fait intervenir la mécanique quantique.

Nous considérons dans ce cas (voir paragraphe III.A.2.b) les états propres (ℓ) de l'hamiltonien H associé au système, d'énergies E_ℓ. Si l'hamiltonien d'un système est indépendant du temps, ses états propres sont stationnaires : lorsque le système se trouve, à un instant donné, dans un état propre de l'hamiltonien, il y demeure indéfiniment, et ses propriétés ne varient pas au cours du temps. Il faut cependant tenir compte ici du fait fondamental que, *pour un système de taille macroscopique, l'hamiltonien n'est qu'imparfaitement défini.* Il y a à cela plusieurs raisons, tant théoriques que pratiques. La principale tient à ce que le système que l'on veut isoler, pour l'étudier, du reste de l'univers, ne peut l'être de façon rigoureuse : des influences extérieures, négligeables à l'échelle macroscopique, devraient être prises en compte dans l'hamiltonien microscopique, et ceci est bien sûr impossible. Par exemple, nous discutons dans l'appendice II (§ II.4.c.α) les conditions aux limites imposées par le fait que le système est enfermé dans un récipient de volume macroscopique : si les dimensions de ce récipient sont bien définies à l'échelle macroscopique, elles sont loin de l'être avec une précision suffisante pour pouvoir être considérées comme

parfaitement fixées dans un calcul microscopique. De même, le mouvement microscopique des constituants du système est affecté par des forces exercées sur eux par des corps extérieurs au système, forces qu'il n'est pas possible de recenser complètement dans l'hamiltonien H, et dont l'intensité est d'ailleurs négligeable à l'échelle macroscopique. Nous reviendrons sur ces effets au chapitre IV, et préciserons leur ordre de grandeur dans le complément IV.B.

Ils ont pour conséquence que l'hamiltonien H est forcément approché, et ses états propres (ℓ) ne sont pas véritablement stationnaires : *le système passe constamment, au cours du temps, d'un état* (ℓ) *à l'autre*[31], et ceci de façon incontrôlable. Il en résulte, comme dans l'exemple du gaz (§ 1), des fluctuations temporelles des diverses grandeurs physiques. Comme au paragraphe 1, ces fluctuations sont présentes lorsque le système évolue macroscopiquement, comme lorsqu'il est à l'équilibre à cette échelle.

B. Moyennes sur les fluctuations microscopiques

La mécanique statistique se fixe pour but de décrire de façon aussi précise que possible l'évolution macroscopique des systèmes tels que ceux que nous avons considérés au paragraphe A, mais nous savons qu'il est impossible, et d'ailleurs inutile, de « suivre » en détail leur mouvement microscopique. Ce dernier sera donc traité par des techniques probabilistes : on effectuera des moyennes sur les fluctuations engendrées par l'agitation microscopique.

1. MOYENNES TEMPORELLES

a. Définition

Examinons d'abord le cas où le système étudié (par exemple le gaz considéré au paragraphe A.1.a) est à l'équilibre du point de vue macroscopique. La variation dans le temps d'une grandeur quelconque $f(t)$ a alors l'allure indiquée pour $n(t)$ dans la figure 2.b.

Au niveau macroscopique, nous nous intéressons principalement à la valeur centrale f_0 autour de laquelle $f(t)$ fluctue, car l'amplitude des fluctuations est faible en valeur relative. Pour déterminer cette valeur centrale, il suffit d'étudier le système pendant un intervalle de temps τ suffisamment long pour qu'il s'y produise un grand nombre de fluctuations. On calculera alors la moyenne, au sens ordinaire, des valeurs prises par la grandeur au cours de cet intervalle de temps τ : si, entre t et $t+\mathrm{d}t$, on mesure la valeur $f(t)$, celle-ci doit être affectée du « poids » $\mathrm{d}t/\tau$, et l'on doit ensuite sommer sur toutes les valeurs de t entre le début t_0 et la fin $t_0+\tau$ du temps d'observation. On aboutit ainsi à la définition de la *valeur moyenne dans le temps* $\widehat{f}$ de la grandeur f[32] :

$$\widehat{f} \widehat{=} \lim_{\tau \to \infty} \frac{1}{\tau} \int_{t_0}^{t_0+\tau} f(t)\,\mathrm{d}t \,. \qquad \text{(IV.1)}$$

---

31. En mécanique quantique, l'état (pur) d'un système est en général une combinaison linéaire des états de base (ici les états (ℓ)). C'est simplement par commodité que nous raisonnons comme si les états (ℓ) étaient les seuls états microscopiques possibles (cf. § III.A.2).

32. La valeur moyenne dans le temps $\widehat{f}$ est *a priori* différente de la valeur moyenne probabiliste définie dans l'appendice I (§ I.4.b). En effet, on répète ici la mesure de f sur le même système, au lieu de prendre à chaque fois un nouveau système (toujours préparé de la même façon) pour que chaque expérience soit indépendante des précédentes. On peut cependant utiliser la définition (IV.1) pour calculer également, par une formule analogue à la définition (29) de l'appendice I, l'écart quadratique moyen d'une grandeur, qui caractérise l'amplitude de ses fluctuations temporelles.

Si le système est à l'équilibre au niveau macroscopique, la valeur trouvée $\hat{f}$ est indépendante de l'instant initial t_0 choisi.

b. Description de l'évolution macroscopique

Peut-on généraliser la définition précédente au cas où le système considéré évolue à l'échelle macroscopique, comme dans l'exemple du paragraphe A.1.b?

Cette généralisation se heurte immédiatement à une difficulté fondamentale, qui se comprend facilement à partir de la figure 3.b. L'évolution macroscopique dure un temps fini, au bout duquel le système se stabilise à un nouvel état d'équilibre (sur la figure 3.b, $n(t)$ atteint la valeur n_0, autour de laquelle il fluctue ensuite comme dans la figure 2.b). La contribution de cette période transitoire à la somme (IV.1) décroît lorsque τ croît, pour devenir finalement négligeable. La valeur moyenne trouvée à partir de (IV.1) est celle qui correspond au nouvel état d'équilibre; elle ne garde plus trace de l'évolution macroscopique du système, que nous cherchions pourtant à décrire.

Cependant si, comme le suggère la figure 3.c, les fluctuations microscopiques sont beaucoup plus rapides que l'évolution macroscopique, on peut modifier la formule (IV.1) de façon à la rendre utilisable. De façon plus précise, soit τ_m un temps caractéristique de l'agitation microscopique : τ_m, défini en ordre de grandeur, est un intervalle de temps au cours duquel le système fluctue un petit nombre de fois. Soit de même τ_M un temps caractéristique de l'évolution macroscopique. Supposons qu'il soit possible de choisir des intervalles de temps τ vérifiant simultanément les deux conditions

$$\tau_m \ll \tau \ll \tau_M \,. \tag{IV.2}$$

Alors, les valeurs moyennes calculées sur de tels intervalles τ selon (IV.1) (sans la limite $\tau \to \infty$) dépendront du temps t_0, cette dépendance correspondant à l'évolution macroscopique du système. Pour que les conditions (IV.2) puissent être vérifiées, il faut évidemment que τ_M soit de plusieurs ordres de grandeur supérieur à τ_m.

2. MOYENNES SUR UN ENSEMBLE

Depuis Gibbs (1839-1903), on remplace en mécanique statistique les moyennes temporelles sur les fluctuations d'un système déterminé par des moyennes d'ensemble, faisant intervenir un grand nombre de systèmes identiques.

a. Notion d'ensemble statistique

Supposons que, au lieu d'étudier un système macroscopique unique, nous en construisions des répliques : nous prenons par exemple plusieurs récipients identiques, nous les remplissons avec le même nombre de molécules d'un même gaz, et nous les plaçons dans les mêmes conditions. Dans la limite où le nombre de ces répliques devient très grand, elles constituent un *ensemble statistique.*

Considérons alors cet ensemble de systèmes *à un instant donné.* Dans le cas où ils sont le siège d'une évolution macroscopique (comme au paragraphe A.1.b), on s'attend à ce que, à l'instant considéré, tous ces systèmes soient parvenus au *même stade de cette évolution macroscopique :* si, dans un grand nombre de systèmes identiques à celui de la figure 3.a, on retire au même moment et de la même façon la cloison intermédiaire, l'expansion du gaz se traduira dans chacun d'eux par la même variation macroscopique de $n(t)$, selon la figure 3.b. Au contraire, on s'attend à ce que les divers systèmes de l'ensemble se trouvent, au même instant, à des *stades différents de l'évolution microscopique* fluctuante. En effet, les systèmes constituant l'ensemble statistique ne sont identiques qu'à l'échelle macroscopique, car c'est à cette échelle qu'on contrôle leur préparation : il est par exemple impossible de s'assurer que les divers récipients contiennent exactement le même nombre de molécules, à l'unité près; de même, leur volume n'est identique qu'à la précision des mesures

macroscopiques. Par conséquent, les systèmes de l'ensemble se trouvent dès leur préparation dans des états microscopiques différents, et leur évolution fluctuante au niveau microscopique suit des courbes différentes, bien qu'analogues.

b. Caractérisation d'un ensemble statistique

Reprenant la description quantique des états microscopiques, nous supposons que chacun des systèmes identiques (au niveau macroscopique) de l'ensemble statistique se trouve dans un état propre (ℓ) de l'hamiltonien H. D'après ce qui vient d'être dit, les divers systèmes se trouvent dans des états microscopiques (ℓ) différents[33]. Soient $\mathcal{N}$ le nombre total de systèmes de l'ensemble, et $\mathcal{N}_\ell$ le nombre de ceux qui sont, à un instant donné, dans l'état microscopique (ℓ).

Supposons que l'on choisisse une grandeur déterminée f, et que l'on mesure sa valeur dans chacun des systèmes de l'ensemble, pour en faire ensuite la moyenne (somme des $\mathcal{N}$ valeurs obtenues divisée par $\mathcal{N}$). On obtient ainsi la *moyenne d'ensemble* de la grandeur f à l'instant considéré. Il est clair que, si f prend la valeur f_ℓ dans l'état (ℓ), cette moyenne d'ensemble vaut

$$\bar{f}_{\mathcal{N}} = \frac{1}{\mathcal{N}} \sum_{(\ell)} \mathcal{N}_\ell f_\ell \,, \tag{IV.3}$$

où la somme porte sur les divers états (ℓ) représentés dans l'ensemble. Dans la limite où $\mathcal{N}$ tend vers l'infini, le rapport $\mathcal{N}_\ell/\mathcal{N}$ devient une probabilité :

$$P_\ell \mathrel{\widehat{=}} \lim_{\mathcal{N}\to\infty} \frac{\mathcal{N}_\ell}{\mathcal{N}} \,, \tag{IV.4}$$

et la moyenne d'ensemble $\bar{f}_{\mathcal{N}}$ tend vers la valeur moyenne calculée avec ces probabilités :

$$\lim_{\mathcal{N}\to\infty} \bar{f}_{\mathcal{N}} = \bar{f} = \sum_{(\ell)} \mathrm{P}_\ell f_\ell \,. \tag{IV.5}$$

A strictement parler, P_ℓ est la probabilité pour que, si l'on choisit un système au hasard dans l'ensemble, ce système se trouve dans l'état microscopique (ℓ). En pratique cependant, on n'a jamais accès à un nombre infini (ni même grand) de systèmes identiques, mais à un seul d'entre eux. On interprète alors la probabilité P_ℓ définie par (IV.4) comme la *probabilité pour que le système étudié se trouve dans l'état microscopique* (ℓ) à l'instant considéré.

Ainsi, c'est la notion d'ensemble statistique qui permet de définir de façon précise les probabilités P_ℓ que nous avons introduites dans les paragraphes précédents. Inversement, l'ensemble statistique associé à un système placé dans des conditions déterminées est entièrement caractérisé, à un instant donné, par la distribution statistique des états microscopiques, c'est-à-dire par l'ensemble des probabilités P_ℓ pour tous les états (ℓ) possibles[34]. P_ℓ est souvent appelée *«probabilité d'occupation»* de l'état (ℓ)[35].

---

33. Les conditions extérieures imposées au système (à l'échelle macroscopique) restreignent les états microscopiques possibles. Mais le nombre des états (ℓ) compatibles avec ces conditions est toujours énorme.

34. Les états (ℓ) incompatibles avec les conditions extérieures ont une probabilité nulle. Suivant la commodité, on peut inclure ou non ces états parmi les «états possibles».

35. Lorsque le mouvement microscopique du système considéré peut être décrit par la mécanique classique, l'ensemble statistique associé est caractérisé par une densité de probabilité $w(q_1, q_2, \ldots, q_n; p_1, p_2, \ldots, p_n; t)$ dans l'espace des phases (voir paragraphe III.C).

c. Évolution de l'ensemble dans le temps

Dans la définition (IV.4), $\mathcal{N}_\ell$ est le nombre de systèmes de l'ensemble qui se trouvent, *à un instant donné,* dans l'état (ℓ); les probabilités P_ℓ dépendent donc *a priori* du temps : $P_\ell(t)$. Noter cependant que, d'après (IV.4), elles vérifient constamment la condition de normalisation

$$\sum_{(\ell)} P_\ell(t) = 1 \,. \tag{IV.6}$$

Nous étudierons au chapitre IV l'évolution dans le temps des probabilités d'occupation $P_\ell(t)$; nous nous contenterons ici de considérations générales.

Supposons d'abord que les probabilités P_ℓ soient constantes dans le temps :

$$\frac{\mathrm{d}P_\ell}{\mathrm{d}t} = 0 \quad \forall(\ell) \,. \tag{IV.7}$$

La moyenne d'ensemble $\bar{f}$ d'une grandeur f quelconque (elle-même indépendante de t) est alors constante dans le temps. Si les moyennes d'ensemble décrivent correctement les propriétés macroscopiques du système (voir paragraphe 3 ci-dessous), ces dernières seront, dans le cas présent, indépendantes du temps, c'est-à-dire que le système sera *à l'équilibre macroscopique.* De façon plus précise, nous prendrons la condition (IV.7) comme *définition* de l'équilibre macroscopique, que nous avions jusqu'ici traité de façon essentiellement intuitive[36].

Noter cependant que chacun des systèmes de l'ensemble évolue dans le temps, changeant sans arrêt d'état microscopique (ℓ). La condition (IV.7) implique seulement que la proportion $\mathcal{N}_\ell/\mathcal{N}$ de systèmes de l'ensemble qui se trouvent dans l'état (ℓ) ne varie pas au cours du temps (dans la limite où $\mathcal{N} \to \infty$), c'est-à-dire qu'il y a en moyenne autant de systèmes de l'ensemble qui quittent l'état (ℓ) durant un intervalle de temps δt que de systèmes qui arrivent dans ce même état (ℓ) au cours du même intervalle de temps. Il s'agit donc d'un *équilibre statistique :* ce sont les propriétés de l'ensemble, pas celles des systèmes individuels (au niveau microscopique), qui sont indépendantes du temps.

Si au contraire les probabilités $P_\ell(t)$ varient dans le temps, il en sera de même des valeurs moyennes d'ensemble :

$$\bar{f}(t) = \sum_{(\ell)} f_\ell P_\ell(t) \,. \tag{IV.8}$$

D'après les considérations du paragraphe a ci-dessus, on s'attend à ce que ceci traduise une *évolution macroscopique.*

3. DISCUSSION : LE PRINCIPE ERGODIQUE

Les moyennes d'ensemble introduites au paragraphe 2 présentent un avantage considérable par rapport aux moyennes temporelles du paragraphe 1 : en prenant *tous les systèmes de l'ensemble au même instant t,* on élimine la difficulté fondamentale des moyennes temporelles (§ 1.b); l'évolution macroscopique, variation dans le temps des valeurs moyennes $\bar{f}(t)$, ne gêne plus le calcul des moyennes sur les fluctuations microscopiques.

Mais la notion de moyenne d'ensemble pose problème sur un autre plan fondamental. En effet, on a toujours à étudier en pratique *un* système unique, et c'est

36. Les propriétés d'un système peuvent être indépendantes du temps sans qu'il soit à l'équilibre : c'est le cas par exemple pour un conducteur parcouru par un courant, si celui-ci est stationnaire. La différence entre un état d'équilibre et un régime stationnaire mais hors d'équilibre réside dans les flux (de charge, d'énergie, ou de particules...) qui ne sont pas tous nuls dans le second cas.

effectivement à cause des fluctuations *temporelles* qu'elle produit que l'agitation microscopique peut être traitée par des moyennes. Se pose donc la question de l'*équivalence des moyennes d'ensemble et des moyennes temporelles,* ces dernières étant les seules que l'on puisse vraiment mesurer.

Au paragraphe 2.a, nous avons donné des arguments de plausibilité en faveur de cette équivalence. Effectivement, « on s'attend à ce que » les divers systèmes d'un ensemble statistique évoluent de façon identique au niveau macroscopique, de sorte qu'une moyenne d'ensemble conserve cette évolution macroscopique. Mais il n'est pas évident que les divers états microscopiques (ℓ) représentés à un instant donné dans l'ensemble statistique le soient avec des fréquences $\mathcal{N}_\ell/\mathcal{N}$ reproduisant exactement celles qui apparaissent au cours des fluctuations temporelles d'un système unique. Nous discuterons cette question plus en détail dans le complément IV.A, où nous verrons qu'elle est loin d'être parfaitement résolue à l'heure actuelle.

La mécanique statistique est fondée sur le *principe ergodique,* qui consiste à postuler l'équivalence entre moyennes d'ensemble et moyennes temporelles. Nous accepterons ce principe, admettant que sa validité découle de la vérification expérimentale des conséquences qu'on en tire : dorénavant, ce sont les notions introduites au paragraphe 2 qui nous serviront de point de départ.

COMPLÉMENTS DU CHAPITRE I

I.A. Ordres de grandeur	Très physique. Facile à lire, plus difficile à bien comprendre. Pourra être repris après l'étude des chapitres II et III.
I.B. Distribution statistique d'une fonction de variables aléatoires	Compléments sur les probabilités. Nécessitent la connaissance des notions rappelées dans l'appendice I.
I.C. La distribution binomiale dans la limite des grands nombres	I.B et I.C : un peu techniques; à étudier lorsque le problème se présentera; le raisonnement développé dans I.C servira de modèle dans plusieurs autres situations.
I.D. Marche au hasard et mouvement brownien	I.D : facile, intéressant à cause des multiples applications de la «marche au hasard».
I.E. Densité d'états d'un gaz parfait	Facile. Démontre un résultat annoncé dans le chapitre.
I.F. États électroniques dans les solides cristallins	Important par le problème physique qu'il aborde. Quelques difficultés techniques, mais pas insurmontables.
I.G. Information et entropie	Développe la notion importante d'entropie statistique introduite dans le chapitre, et discute en particulier son application aux distributions continues de probabilité. Assez facile.
I.H. L'opérateur-densité	Introduit l'outil adapté à une formulation de la mécanique statistique quantique plus générale et plus puissante que celle qui est proposée dans le chapitre. Nécessite une bonne connaissance des techniques de base de la mécanique quantique.

COMPLÉMENT I.A

Ordres de grandeur

I. Ordres de grandeur microscopiques
   1. Les constantes physiques fondamentales
   2. L'atome isolé
      a. Unités naturelles de longueur et d'énergie
      b. Énergie d'ionisation et rayon des atomes
   3. Molécules et solides
      a. Énergies de liaison et distances interatomiques
      b. Énergies de vibration et de rotation
      c. Moments électriques et magnétiques
   4. Mécanique classique ou mécanique quantique?

II. Le passage au macroscopique
   1. Ordres de grandeur macroscopiques
   2. Échelle des températures
   3. Champs électriques et magnétiques

III. Un exemple : le gaz parfait classique
   1. Définitions
   2. Validité de la description classique
      a. Condition de validité
      b. Cas d'un gaz de molécules
      c. Cas du gaz d'électrons dans un métal
   3. Comparaison entre énergie cinétique et énergie potentielle

Ce complément est destiné à familiariser le lecteur avec l'ordre de grandeur des quantités physiques qui interviendront dans toute la suite de ce cours. Plutôt que de dresser un catalogue, nous montrerons comment retrouver rapidement la plupart de ces valeurs par des arguments fondés principalement sur l'analyse dimensionnelle. Nous essaierons ainsi de souligner l'intérêt des méthodes qualitatives pour obtenir une estimation de l'ordre de grandeur du résultat attendu dans la résolution d'un problème. Ces méthodes, en effet, non seulement permettent de justifier les approximations nécessaires, mais encore donnent une bien meilleure compréhension de la physique sous-jacente; elles aident par conséquent à la mise en œuvre de la solution quantitative et à l'analyse du résultat obtenu[1].

I. ORDRES DE GRANDEUR MICROSCOPIQUES

1. LES CONSTANTES PHYSIQUES FONDAMENTALES

On trouve sur les pages de garde de ce livre les valeurs des constantes physiques fondamentales. Si on néglige les corrections relativistes, les propriétés de la matière au niveau atomique (que nous qualifions de *microscopiques*) font intervenir quatre constantes seulement[2] :

— la charge de l'électron, de valeur absolue

$$e = 1{,}602 . 10^{-19}\ \text{C} \simeq 1{,}6 . 10^{-19}\ \text{C}\ , \qquad \text{(A.1)}$$

— la masse de l'électron

$$m_e = 9{,}109 . 10^{-31}\ \text{kg} \simeq 10^{-30}\ \text{kg}\ , \qquad \text{(A.2)}$$

1. Au cours de la discussion, il nous arrivera d'anticiper sur la suite de ce cours. Nous demandons au lecteur d'admettre certains résultats qui seront justifiés plus tard, et nous lui conseillons de revenir à ce complément après lecture des chapitres correspondants.

2. Comme nous nous limitons aux ordres de grandeur, nous négligeons l'énergie de liaison des noyaux : la masse d'un noyau contenant A nucléons est alors $A \times M_N$.

— la masse du nucléon (proton ou neutron)

$$M_N = 1{,}67 . 10^{-27} \text{ kg} \simeq 1{,}7 . 10^{-27} \text{ kg} , \tag{A.3}$$

— la constante de Planck

$$\hbar = \frac{h}{2\pi} = 1{,}055 . 10^{-34} \text{ J.s} \simeq 10^{-34} \text{ J.s} . \tag{A.4}$$

A ces constantes il faut ajouter, dans le système d'unités M.K.S.A., la permittivité du vide

$$\varepsilon_0 = \frac{10^{-9}}{36\pi} . \tag{A.5}$$

En fait, les charges électriques ne peuvent intervenir que par paires *via* l'interaction coulombienne; plutôt que la charge e, la grandeur physique à considérer est donc souvent la combinaison

$$q^2 \widehat{=} \frac{e^2}{4\pi\varepsilon_0} = 2{,}310 \times 10^{-28} \text{ J.m} \simeq 2{,}3 \times 10^{-28} \text{ J.m} . \tag{A.6}$$

2. L'ATOME ISOLÉ

a. Unités naturelles de longueur et d'énergie

La masse du nucléon étant considérablement plus grande que celle de l'électron ($M_N/m_e = 1\,836$), il est raisonnable de penser que le mouvement du noyau ne joue pas un rôle essentiel dans les propriétés atomiques : on peut traiter le noyau comme infiniment lourd et fixe dans l'espace. Le « rayon » et l'énergie de l'atome ne doivent donc faire intervenir que les trois constantes[3] m_e, $\hbar$ et q^2. Il est alors facile de se convaincre[4] que les seules combinaisons de ces constantes qui ont les dimensions d'une longueur et d'une énergie sont respectivement

$$a_0 = \frac{\hbar^2}{m_e q^2} = 0{,}53 . 10^{-10} \text{ m} = 0{,}53 \text{ Å} , \tag{A.7}$$

et

$$E_0 = \frac{m_e q^4}{\hbar^2} = 4{,}37 . 10^{-18} \text{ J} = 27{,}3 \text{ eV} . \tag{A.8}$$

a_0 et E_0 apparaissent ainsi comme les *unités naturelles*[5] de longueur et d'énergie de la physique atomique : on s'attend à ce que toute longueur et toute énergie du domaine atomique soient proportionnelles respectivement à a_0 et E_0, le coefficient de proportionnalité restant dans des limites raisonnables autour de 1 (disons entre 10^{-3} et 10^3).

---

3. De ce point de vue, $\hbar$ apparaît bien comme la « constante manquante » en mécanique classique. Il est en effet impossible d'obtenir, par exemple, une *longueur caractéristique* avec les seules constantes m_e et q^2. Évidemment, la mécanique classique fait intervenir des longueurs, mais ces longueurs n'ont pas d'« échelle naturelle » : par exemple l'orbite d'un satellite qui tourne autour de la Terre peut avoir, suivant sa vitesse, un rayon quelconque, compris entre celui de la Terre et l'infini.

4. Il suffit d'exprimer les trois constantes m_e, $\hbar$ et q^2 à partir des grandeurs fondamentales : longueur, temps et masse et de chercher les valeurs de α, β et γ telles que la combinaison $[m_e^\alpha \hbar^\beta q^{2\gamma}]$ soit homogène à une longueur ou à une énergie. Dans les deux cas, la solution est unique.

5. Ces deux unités naturelles peuvent d'ailleurs être reliées « de façon naturelle » : E_0 est l'énergie d'interaction q^2/a_0 de deux charges élémentaires e situées à la distance a_0.

En pratique on utilise, plutôt que le « hartree » E_0, l'unité

$$Ry \widehat{=} \frac{E_0}{2} = \frac{m_e q^4}{2\hbar^2} \simeq 13{,}6 \text{ eV} , \tag{A.9}$$

appelée le rydberg. En effet Ry et égal à l'énergie d'ionisation de l'atome d'hydrogène depuis son état fondamental ; le « rayon de Bohr » a_0 est le « rayon » correspondant à cet état. Il est donc commode de choisir a_0 et Ry comme références.

b. Énergie d'ionisation et rayon des atomes

L'atome d'hydrogène, nous venons de le voir, est caractérisé dans son état fondamental par un rayon a_0 et une énergie d'ionisation Ry. Si on considère des atomes plus lourds, de nombre atomique Z, des arguments simples permettent de se faire une idée de la dépendance en Z des dimensions et des énergies.

Supposons en effet que l'on construise l'atome en apportant un à un les électrons autour du noyau. Pour le premier électron, il est clair que le résultat est le même que pour l'atome d'hydrogène, à condition de tenir compte du fait que la charge du noyau est maintenant Ze; il suffit donc de remplacer q^2 par Zq^2 dans les expressions (A.9) et (A.7) pour obtenir l'énergie de liaison du premier électron :

$$E_1 \simeq Z^2 Ry , \tag{A.10}$$

et sa distance moyenne au noyau :

$$r_1 \simeq \frac{a_0}{Z} . \tag{A.11}$$

Pour Z grand, l'énergie de liaison est donc considérablement plus grande et le rayon plus petit que pour l'atome d'hydrogène. Si on ajoute l'électron suivant, il sera aussi très lié au noyau; en effet l'énergie de répulsion entre les deux électrons (qui se trouvent, en moyenne, à une distance d'ordre r_1 l'un de l'autre) est Z fois plus faible que l'énergie de liaison de chacun d'eux au noyau. Lorsqu'on apporte ainsi le $(n+1)$-ième électron, l'ion formé par le noyau et les n premiers électrons se comporte comme un « noyau » de charge $(Z-n)e$; les électrons sont donc de moins en moins fortement liés au noyau, et se situent à des distances de plus en plus grandes. Lorsqu'on ajoute le dernier électron, l'ion a une charge e; on s'attend donc à ce que *le « rayon » de l'atome soit de l'ordre de* a_0, *et l'énergie d'ionisation*[6] *de l'ordre de* Ry.

Il est important de souligner que les arguments qualitatifs précédents ne peuvent donner que l'échelle des phénomènes; ils ne dispensent en aucune manière d'une théorie quantitative correcte. C'est seulement en résolvant l'équation de Schrödinger que l'on montrera que a_0 est bien le rayon de l'atome d'hydrogène et Ry son énergie d'ionisation. On est sûr que, si l'équation de Schrödinger ne fait intervenir que les quantités m_e, q^2 et $\hbar$, le rayon des atomes s'exprimera en unité de a_0; mais le coefficient de proportionnalité pourrait être assez différent de 1 : si, par exemple, nous avions défini a_0 avec h plutôt que $\hbar$, il n'aurait donné le rayon de l'atome d'hydrogène qu'au facteur $4\pi^2$ ($\simeq 40$) près. On ne peut donc jamais être sûr de ne pas se « tromper » d'un facteur numérique important. Cependant un bon choix (par exemple $\hbar$ plutôt que h), guidé par la pratique[7], donne généralement le bon ordre de grandeur, c'est-à-dire le résultat correct à un facteur près compris entre 10 et 1/10.

---

6. L'énergie d'ionisation est l'énergie minimum qu'il faut fournir pour arracher un électron à l'atome, et c'est évidemment l'électron le moins lié qui se détache le premier.

7. Ce choix est évidemment d'autant plus facile qu'on connaît déjà le résultat!

C'est ce qu'on vérifie sur le tableau I, qui donne le rayon et l'énergie d'ionisation de quelques atomes : ils restent comparables à a_0 et Ry dans tous les cas. On note cependant des différences significatives entre les atomes : les gaz rares (He et Ne) sont particulièrement stables, les éléments qualifiés en chimie d'« électropositifs » perdent facilement un électron (Li, Na), les éléments « électronégatifs » moins facilement (O, Cl).

TABLEAU I
Énergie d'ionisation et rayon pour quelques atomes.

Atome	Énergie d'ionisation (eV)	Rayon atomique (Å)
H	13,6	0,53
He	24,5	0,30
Li	5,4	1,67
C	11,2	0,65
O	13,6	0,46
Ne	21,5	0,37
Na	5,1	1,80
Cl	13,0	0,75

3. MOLÉCULES ET SOLIDES

a. Énergies de liaison et distances interatomiques

Étant donné une molécule diatomique, quelles sont les grandeurs qui fixent son énergie de liaison et la distance entre les deux noyaux?

En première approximation, on peut considérer les noyaux comme immobiles à une distance fixe d l'un de l'autre. Ces noyaux sont entourés par le « nuage » des électrons périphériques des deux atomes; l'énergie de ces électrons est une fonction de d. L'état stable de la molécule est obtenu en cherchant la valeur de d qui rend l'énergie totale minimum. Donc, les seules constantes fondamentales qui vont intervenir sont, là encore, m_e, $\hbar$ et q^2. Il en résulte que les « unités naturelles » de longueur et d'énergie sont encore a_0 et Ry, de sorte que les longueurs et énergies caractéristiques de la molécule sont, en ordre de grandeur, a_0 et Ry. C'est ce que confirme le tableau II.

TABLEAU II
Énergie de dissociation et distance internucléaire pour quelques molécules diatomiques.

Molécule	Énergie de dissociation (eV)	Distance entre noyaux (Å)
H_2	4,5	0,75
N_2	9,8	1,09
O_2	5,1	1,20
Cl_2	2,5	1,99
HCl	4,4	1,27
CO	11,2	1,13
NaCl	3,6	2,51

On note cependant que les énergies de dissociation des molécules sont sensiblement plus faibles que les énergies d'ionisation des atomes : il est généralement plus facile de dissocier une molécule que d'ioniser un atome; d'autre part les distances entre noyaux sont systématiquement supérieures à a_0[8].

Les mêmes arguments tiennent évidemment pour des molécules plus compliquées ou pour des solides : l'ordre de grandeur des distances internucléaires est l'angström, celui des énergies de cohésion l'électron-volt (pour un solide, l'énergie de cohésion est l'énergie de liaison par molécule; voir tableau III).

TABLEAU III
Énergie de cohésion et distance entre proches voisins pour quelques éléments à l'état solide.

Solide	Énergie de cohésion (eV)	Distance entre proches voisins (Å)
C (diamant)	7,36	1,5
Na	1,13	3,7
Al	3,34	2,9
Si	4,64	2,3
Ar	0,08	3,8
Fe	4,29	2,5
Te	2,0	2,9

b. Énergies de vibration et de rotation

En dehors des énergies d'excitation électroniques dont nous venons de parler, une molécule présente des excitations de vibration et de rotation. Celles-ci mettent en jeu le mouvement de la molécule par rapport à son centre de masse et font donc directement intervenir la masse des noyaux atomiques.

Pour une molécule diatomique comportant deux noyaux de masses M_1 et M_2, la séparation du mouvement du centre de masse laisse un seul paramètre pour le mouvement relatif, la *masse réduite*

$$M_R = \frac{M_1 M_2}{M_1 + M_2} . \tag{A.12}$$

Nous l'exprimerons à l'aide de son unité naturelle, la masse M_N du nucléon :

$$M_R = \alpha M_N , \tag{A.13}$$

où α est un paramètre qui dépend explicitement des nombres de nucléons A_1 et A_2 des deux noyaux :

$$\alpha = \frac{A_1 A_2}{A_1 + A_2} . \tag{A.14}$$

8. Cet effet peut être attribué à la répulsion électrostatique entre les deux noyaux. Lorsque leur distance d est supérieure à quelques Å, ils n'interagissent pratiquement pas, car la charge de chacun d'eux est « écrantée » par les électrons qui l'entourent. Mais, lorsque d est suffisamment faible pour que les deux « nuages » électroniques s'interpénètrent, les électrons périphériques adoptent des « orbites » englobant les deux noyaux, et ceux-ci « se voient » alors l'un l'autre. Pour les électrons, certaines « orbites » moléculaires autour des deux noyaux sont énergétiquement plus favorables que les « orbites » atomiques autour d'un seul noyau. C'est ce qui permet finalement la formation d'une molécule stable. Mais cette tendance est contrecarrée par la répulsion électrostatique entre les noyaux, d'autant plus forte que d est plus petit. D'où les « grandes » valeurs de d (en unités de a_0) et les « faibles » valeurs de l'énergie de dissociation (en unités de Ry).

α. Énergies de vibration

Les vibrations de faible amplitude de la molécule sont contrôlées par la dérivée seconde du potentiel d'interaction entre les deux atomes au voisinage de la distance d'équilibre[9]. Or nous avons vu que l'échelle de ce potentiel est le rydberg Ry, et qu'il varie sur des longueurs caractéristiques de l'ordre de a_0. On peut donc considérer que les mouvements de vibration de la molécule se font, en première approximation, dans un potentiel harmonique dont la constante de rappel K a pour ordre de grandeur[10]

$$K \simeq \frac{Ry}{a_0^2} . \tag{A.15}$$

La fréquence de vibration correspondante est

$$\omega_v = \sqrt{\frac{K}{M_R}} . \tag{A.16}$$

Il est commode d'exprimer ce résultat en unité d'énergie en multipliant ω_v par $\hbar$[11]. On obtient finalement, en utilisant les expressions (A.7) et (A.9),

$$\hbar\omega_v \simeq \hbar \sqrt{\frac{Ry}{a_0^2 M_R}} = Ry \sqrt{\frac{m_e}{M_N}} \times \frac{1}{\sqrt{\alpha}} \tag{A.17}$$

(nous négligeons un facteur 2, sans importance dans une estimation d'ordre de grandeur), avec

$$Ry \sqrt{\frac{m_e}{M_N}} \simeq 0{,}32 \text{ eV} . \tag{A.18}$$

---

9. Soit $V(d)$ le potentiel d'interaction entre les deux atomes lorsque leur distance est égale à d. En développant $V(d)$ au second ordre, au voisinage de la valeur d'équilibre d_0 (qui annule la dérivée première de $V(d)$), on obtient

$$V(d) = V(d_0) + \frac{1}{2}(d - d_0)^2 V''(d_0) + \dots . \tag{1}$$

Pour les petites oscillations autour de d_0, $V(d)$ se comporte donc comme un potentiel harmonique de constante de rappel

$$K = V''(d_0) . \tag{2}$$

10. Les seules grandeurs caractéristiques du potentiel $V(d)$ sont Ry et a_0. Sa dérivée première $V'(d)$, dont la dimension est le quotient d'une énergie par une longueur, est mesurée en unité de Ry/a_0; on obtient de même Ry/a_0^2 pour sa dérivée seconde $V''(d)$. Comme toujours, nous supposons le coefficient de proportionnalité d'ordre 1 (remarquons que dans certains cas ce coefficient est nul; c'est le cas de $V'(d)$ pour $d = d_0$).

11. Rappelons que, en mécanique quantique, les niveaux d'énergie de cet oscillateur harmonique sont

$$\varepsilon_n^v = \left(n + \frac{1}{2}\right) \hbar\omega_v ,$$

où n prend toutes les valeurs entières positives ou nulle.

Le tableau IV donne les valeurs expérimentales de $\hbar\omega_v$ pour quelques molécules diatomiques. Dans tous les cas ces valeurs restent proches de l'estimation donnée par (A.17)[12].

TABLEAU IV

Quantums de vibration et de rotation pour quelques molécules diatomiques. α est la masse réduite (en unité de masse du nucléon). On a indiqué entre parenthèses les valeurs « prédites » par les estimations (A.17) et (A.20).

Molécule	Masse réduite $\alpha = M_R/M_N$	Quantum de vibration $\hbar\omega_v$ (eV)	Quantum de rotation $\hbar\omega_r$ (eV)
H_2	0,5	0,53 (0,45)	$7{,}5\times10^{-3}$ (15×10^{-3})
D_2	1	0,39 (0,32)	$3{,}8\times10^{-3}$ (7×10^{-3})
N_2	7	0,29 (0,12)	$2{,}5\times10^{-4}$ (10^{-3})
O_2	8	0,19 (0,11)	$1{,}8\times10^{-4}$ (9×10^{-4})
Cl_2	17,8	0,07 (0,08)	3×10^{-5} (4×10^{-4})
HCl	0,97	0,36 (0,32)	$1{,}3\times10^{-3}$ (7×10^{-3})
NO	7,5	0,23 (0,12)	2×10^{-4} (10^{-3})

Pour des molécules comprenant plus de deux atomes, les résultats sont plus compliqués; il existe alors plusieurs modes propres de vibration, mais les ordres de grandeur restent les mêmes.

Dans un solide, le couplage entre les atomes donne lieu à un très grand nombre de modes propres (appelés phonons) dont les fréquences sont très voisines les unes des autres. Elles s'étendent sur des « bandes » pratiquement continues de valeurs, dont l'échelle est encore donnée par (A.17). Le tableau V donne, pour quelques solides, une des fréquences caractéristiques des vibrations dans les cristaux : la fréquence de Debye ω_D (π/ω_D est l'ordre de grandeur du temps mis par une onde sonore pour parcourir une distance interatomique).

TABLEAU V

Fréquence de Debye ω_D pour quelques solides. Entre parenthèses l'estimation fournie par la formule (A.17), dans laquelle α est remplacée par A.

Solide	Masse atomique A	Fréquence de Debye $\hbar\omega_D$ (eV)
Be	9	$8{,}6\times10^{-2}$ (11×10^{-2})
Na	23	$1{,}3\times10^{-2}$ (7×10^{-2})
Al	27	$3{,}4\times10^{-2}$ (6×10^{-2})
Si	28	$5{,}4\times10^{-2}$ (6×10^{-2})
Ar	40	$0{,}7\times10^{-2}$ (5×10^{-2})
Cu	63,5	$2{,}7\times10^{-2}$ (4×10^{-2})
NaCl	58,5	$2{,}8\times10^{-2}$ (4×10^{-2})

12. On pourrait, pour chaque molécule, raffiner l'estimation en tenant compte dans (A.15) des valeurs expérimentales données par le tableau II.

β. Énergies de rotation

La rotation d'une molécule diatomique autour de son centre de masse fait intervenir la masse réduite M_R et la distance internucléaire d, d'ordre a_0. Si on suppose la molécule *rigide*, M_R et d sont les seules grandeurs qui fixent les propriétés de rotation de la molécule (elles n'interviennent d'ailleurs que par la combinaison $I = M_R d^2$, moment d'inertie par rapport au centre de masse). On ne dispose donc, pour déterminer une fréquence caractéristique de cette rotation, que des constantes M_R, d et $\hbar$. La seule combinaison de ces constantes homogène à une fréquence est[13]

$$\omega_r = \frac{\hbar}{M_R d^2} . \tag{A.19}$$

Là encore, c'est plutôt le quantum d'énergie $\hbar\omega_r$ que l'on considérera; il est alors commode d'exprimer $\hbar\omega_r$ en terme de Ry en remplaçant d par a_0 dans (A.19). Il vient immédiatement (à un facteur 2 près)

$$\hbar\omega_r \simeq Ry \frac{m_e}{M_N} \times \frac{1}{\alpha} , \tag{A.20}$$

avec $$Ry \frac{m_e}{M_N} \simeq 7{,}4 . 10^{-3} \text{ eV} . \tag{A.21}$$

Les énergies de rotation[14] sont donc considérablement plus faibles que les énergies de vibration. C'est ce qu'on vérifie sur le tableau IV[15].

c. Moments électriques et magnétiques

α. Moments électriques

Dans certaines molécules, le barycentre des charges positives ne coïncide pas avec celui des charges négatives; la molécule présente alors un moment dipolaire électrique $\vec{p}$. Les charges caractéristiques sont d'ordre e, les distances d'ordre a_0; l'unité naturelle de moment dipolaire microscopique est donc

$$p_0 = e a_0 \simeq 10^{-29} \text{ C.m} . \tag{A.22}$$

Évidemment, pour certaines molécules, le moment dipolaire est nul en l'absence de champ électrique extérieur : c'est le cas par exemple des molécules constituées de deux atomes identiques (N_2, O_2, H_2, ...). Mais, si une molécule présente un moment dipolaire permanent, celui-ci sera égal à quelques 10^{-29} C.m, le facteur de proportionnalité étant généralement compris entre 10 et 1/10. Le tableau VI donne les valeurs de p pour quelques molécules[16].

---

13. On obtient la même estimation en écrivant que l'énergie de rotation $I\omega_r^2/2$ est approximativement égale à $\hbar\omega_r$.

14. Les énergies de rotation d'une molécule sont données, en mécanique quantique, par

$$\varepsilon_j^r = j(j+1) \frac{\hbar^2}{2I} ,$$

où j prend toutes les valeurs entières positives ou nulle.

15. (A.20) donne bien l'ordre de grandeur de la valeur expérimentale de ω_r. L'estimation est toutefois systématiquement trop grande. Ceci est dû au fait que la distance d entre les noyaux est toujours supérieure à a_0 (voir tableau II et note 8). Là encore on pourrait raffiner l'estimation en utilisant les données du tableau II.

16. On remarque que le moment dipolaire est grand (par rapport à l'unité naturelle) lorsque l'un des atomes est très « électronégatif » (Cl) et l'autre très « électropositif » (Na ou Cs). Le moment dipolaire de la molécule CO_2 est nul par symétrie : les trois atomes sont alignés, les O encadrant le C. Dans H_2O, au contraire, les atomes forment un triangle isocèle, et l'atome « électronégatif » O attire vers lui le « nuage » électronique, qui laisse une charge positive sur les H.

TABLEAU VI
Moment dipolaire électrique de quelques molécules.

Molécule	Moment dipolaire (10^{-29} C.m)
H_2O	0,62
HCl	0,36
NaCl	3,00
CsCl	3,48
NH_3	0,49
CO	0,04
CO_2	0,0

β. Moments magnétiques

De la même façon, si un atome ou un ion possède un moment cinétique total non nul, le mouvement correspondant de particules chargées est équivalent à un courant électrique, et l'atome (ou l'ion) présente un moment magnétique permanent $\vec{\mu}$. L'unité naturelle de moment magnétique atomique est le *magnéton de Bohr*[17]

$$\mu_B = \frac{e\hbar}{2m_e} = 0{,}927 \cdot 10^{-23}\ \text{J/T} \simeq 10^{-23}\ \text{J/T}\ . \tag{A.23}$$

Pour un atome (ou un ion) libre, le moment magnétique s'écrit

$$\vec{\mu} = -\mu_B (\vec{L} + 2\vec{S})\ , \tag{A.24}$$

où $\vec{L}$ et $\vec{S}$ sont respectivement le moment cinétique orbital total et le spin total de l'atome, en unités de $\hbar$. La projection de $\vec{\mu}$ sur un axe de quantification vaut donc quelques μ_B. La situation est plus complexe pour un atome dans un cristal, mais l'ordre de grandeur est le même.

---

17. On retrouve ce résultat par la méthode habituelle : un moment magnétique est le produit d'une intensité de courant (homogène à une charge électrique divisée par un temps) par une surface; c'est donc ici e, elle-même, qui intervient et non plus q^2; il faut alors chercher une combinaison de e, m_e et $\hbar$ homogène à un moment magnétique. Là encore la solution est unique.

On peut aussi raisonner de façon plus « physique » en évaluant l'intensité i équivalente au mouvement d'une charge e décrivant une orbite de rayon a_0 (c'est-à-dire la charge e multipliée par le nombre de révolutions par seconde) :

$$i \simeq e\frac{v}{a_0}\ , \tag{1}$$

où v est la « vitesse » de l'électron sur son orbite. En ordre de grandeur,

$$\frac{1}{2} m_e v^2 \simeq Ry\ ; \tag{2}$$

d'où
$$\mu \simeq i a_0^2 \simeq a_0 e \sqrt{\frac{Ry}{m_e}}\ . \tag{3}$$

En remplaçant Ry et a_0 par leurs valeurs (A.7) et (A.9), on trouve

$$\mu \simeq \frac{e\hbar}{m_e} \tag{4}$$

qui, à un facteur 2 près, est bien le magnéton de Bohr μ_B. Insistons encore une fois : ce qui précède *n'est pas un calcul* de la valeur du magnéton de Bohr, ce n'est qu'une méthode pour obtenir l'échelle naturelle des moments magnétiques atomiques.

Les noyaux possèdent aussi, en général, un moment cinétique, et donc un moment magnétique. L'ordre de grandeur de ce dernier est obtenu en remplaçant m_e par M_N dans (A.23). Les moments magnétiques nucléaires sont donc environ 1 000 fois plus petits que les moments magnétiques atomiques (ils valent quelques 10^{-26} J/T).

4. MÉCANIQUE CLASSIQUE OU MÉCANIQUE QUANTIQUE?

Les propriétés de la matière au niveau microscopique sont décrites par la mécanique quantique. Il est cependant des cas où la mécanique classique est suffisante pour donner une description correcte des phénomènes observés. Cette description classique est généralement plus facile à mettre en œuvre que la description quantique; il est donc important de connaître un critère simple qui permettra de l'utiliser chaque fois que cela sera justifié.

La constante caractéristique de la mécanique quantique est la *constante de Planck* $\hbar$. Cette constante a les dimensions d'une *action,* c'est-à-dire du produit d'un temps par une énergie, ou bien du produit d'une longueur par une quantité de mouvement, ou encore d'un moment cinétique. Le critère est alors le suivant :

Si toutes les variables naturelles d'un système physique qui ont les dimensions d'une action sont très supérieures à la constante de Planck $\hbar$*, le système sera décrit correctement dans le formalisme de la mécanique classique.*

Ce critère est presque toujours vérifié par les objets de taille macroscopique; mais il existe aussi des cas où il l'est au niveau microscopique[18]. Il faut d'ailleurs remarquer que la mécanique classique (améliorée par quelques notions simples de mécanique quantique) peut souvent être utilisée avec profit pour obtenir des estimations d'ordres de grandeur, même dans des cas où il ne saurait être question de l'employer dans des calculs quantitatifs; c'est ce que nous avons déjà fait plusieurs fois dans ce complément.

II. LE PASSAGE AU MACROSCOPIQUE

1. ORDRES DE GRANDEUR MACROSCOPIQUES

L'ordre de grandeur de la plupart des quantités physiques macroscopiques s'obtient simplement en multipliant les valeurs des quantités microscopiques correspondantes par le nombre d'Avogadro

$$N_A = 6{,}02 \times 10^{23} \ . \tag{A.25}$$

Donnons quelques exemples :

(i) Dans une réaction chimique, des molécules se dissocient pour former de nouvelles molécules. Les énergies associées à ce réarrangement des atomes doivent donc être de l'ordre de l'énergie de liaison des molécules, c'est-à-dire de 1 à 10 eV par processus élémentaire. Les *chaleurs de réaction chimique* sont ainsi de l'ordre de $(1 \text{ à } 10) \times N_A$ eV par mole, c'est-à-dire 10^5 à 10^6 joules par mole. On sait par exemple que

$$H_2 + \frac{1}{2} O_2 \longrightarrow H_2O + 288 \text{ kJ} \ ,$$

$$C + O_2 \longrightarrow CO_2 + 405 \text{ kJ} \ ,$$

$$H_2 + Cl_2 \longrightarrow 2\,HCl + 183 \text{ kJ} \ ,$$

etc.

18. Certains systèmes macroscopiques exceptionnels, tels que l'hélium superfluide, ne vérifient pas le critère, et présentent des propriétés typiquement quantiques sur une échelle macroscopique. Inversement, nous montrerons au paragraphe III.2 que la plupart des gaz peuvent être décrits par la mécanique classique, même au niveau microscopique.

(ii) De même, l'énergie nécessaire pour vaporiser une mole d'un cristal à température et pression constantes, appelée *chaleur latente de sublimation,* est du même ordre de grandeur que l'énergie de cohésion du cristal multipliée par le nombre d'Avogadro. D'après le tableau III, on s'attend à trouver $(0{,}1 \text{ à } 10) \times N_A$ eV par mole, soit 10^4 à 10^6 J par mole[19]. C'est bien ce que l'on constate sur le tableau VII.a.

Les *chaleurs latentes de vaporisation* (passage de l'état liquide à l'état vapeur) sont légèrement plus faibles, mais l'ordre de grandeur est encore donné par l'argument ci-dessus (tableau VII.b).

TABLEAU VII

Chaleurs latentes. VII.a : Chaleurs latentes de sublimation pour quelques métaux (la température n'est pas précisée, car entre 0 K et la température de fusion, L_s ne change que de quelques pour-cent). VII.b : Chaleurs latentes de fusion (L_f) et de vaporisation (L_v) pour quelques corps, mesurées respectivement à la température de fusion (T_f) et de vaporisation (T_v), pour une pression d'une atmosphère.

Métal	L_s (kJ/mole)
K	82
Fe	436
Cu	338
Ag	284
W	878
Au	367
Pb	197

VII.a

Solide ou liquide	T_f (K)	L_f (kJ/mole)	T_v (K)	L_v (kJ/mole)
N_2	63	0,7	77	5,6
O_2	55	0,5	90	6,8
Cl_2	172	6,4	238	18,4
Ar	84	1,2	87,3	6,3
Hg	234	2,3	630	54,5
H_2O	273	6,0	373	40,6
NH_3	198	7,7	240	23,2

VII.b

Les *chaleurs latentes de fusion* (passage du solide au liquide) sont, elles, systématiquement plus faibles. Elles valent entre 0,5 et 50 kJ par mole (tableau VII.b). Ceci s'explique par le fait que, dans le liquide, l'énergie de liaison entre molécules n'est pas négligeable comme dans le gaz : elle est certes plus faible que l'énergie de cohésion du solide, mais pas très différente.

(iii) Certains corps, dits *ferroélectriques,* présentent une polarisation électrique (moment dipolaire électrique par unité de volume) permanente à l'échelle macroscopique. D'autres, les *ferromagnétiques,* ont de même une aimantation (moment magnétique par unité de volume) non nulle. Ces propriétés proviennent du fait que les moments microscopiques s'orientent parallèlement les uns aux autres sur des distances macroscopiques au lieu de se distribuer au hasard. Ainsi, le moment par unité de volume est égal au produit d'un moment microscopique par le nombre de ces moments par unité de volume. Pour une distance interatomique d_0 de l'ordre de 2 Å, l'ordre de grandeur de la polarisation d'un ferroélectrique sera

$$\mathcal{P}_0 = p_0 \cdot \frac{1}{d_0^3} \simeq 1 \text{ C/m}^2 , \tag{A.26}$$

et celui de l'aimantation d'un ferromagnétique

$$\mathcal{M}_0 = \mu_B \times \frac{1}{d_0^3} \simeq 10^6 \text{ A/m} . \tag{A.27}$$

19. Attention, la chaleur latente n'est pas reliée si simplement à l'énergie de cohésion. En effet, les atomes d'un cristal possèdent une énergie de vibration d'origine thermique, et c'est seulement la différence entre l'énergie de cohésion et cette énergie de vibration qui constitue la chaleur latente. Cette dernière dépend donc de la température; néanmoins on obtient un ordre de grandeur correct à partir de la seule énergie de cohésion.

Par exemple, les cristaux de titanate de baryum ($BaTiO_3$) présentent une polarisation de 0,26 C/m^2 à la température ordinaire; l'aimantation du fer, à la même température, est de $1{,}7 \times 10^6$ A/m.

Évidemment, de telles considérations n'expliquent pas pourquoi le fer, par exemple, est ferromagnétique, encore moins pourquoi il devient paramagnétique au-dessus de 1 043 K. Néanmoins, elles indiquent, dans la mesure où les ordres de grandeur sont compatibles, qu'on peut espérer comprendre le phénomène macroscopique qu'est le ferromagnétisme en termes des moments magnétiques des atomes constitutifs.

2. ÉCHELLE DES TEMPÉRATURES

Le passage au macroscopique se fait aussi de façon plus subtile par l'intermédiaire de la notion de température. Cette notion sera définie dans le chapitre II, et étudiée en détail dans les chapitres suivants; disons seulement ici que la *température absolue* T (mesurée en degré Kelvin, noté K) est égale à la température Celsius t augmentée de 273,15 :

$$T\,(\mathrm{K}) = t\,(^\circ\mathrm{C}) + 273{,}15\ . \tag{A.28}$$

A chaque valeur de la température on peut associer une *énergie caractéristique* microscopique, donnée par le produit kT de la température absolue T par la constante de Boltzmann

$$k = 1{,}38 . 10^{-23}\ \mathrm{J/K}\ . \tag{A.29}$$

Nous verrons que, pour un système maintenu à la température T, l'énergie kT fournit un *ordre de grandeur de référence* auquel il est intéressant de comparer les énergies du système[20]. Il est donc important de savoir « traduire » rapidement une température en énergie ou vice versa, et d'avoir en tête les quelques chiffres suivants :

— 1 K est équivalent à $\simeq 10^{-4}$ eV ($0{,}8617 \times 10^{-4}$ eV)

— 1 eV est équivalent à $\simeq$ 12 000 K (11 605 K)

— la température ordinaire (20 °C $\simeq$ 293 K) est équivalente à $\simeq 1/40$ eV.

Enfin, pour une appréciation de l'échelle des températures absolues, nous renvoyons au tableau VIII qui donne quelques températures de référence (en K et en eV).

3. CHAMPS ÉLECTRIQUES ET MAGNÉTIQUES

Nous terminerons ce paragraphe en donnant quelques ordres de grandeur des champs électriques et magnétiques accessibles au laboratoire.

Au *niveau macroscopique,* les plus forts champs électriques réalisables actuellement sont de l'ordre de 10^7 V/m; les plus forts champs magnétiques sont au plus de 30 teslas (c'est-à-dire 300 000 gauss) en continu, mais peuvent atteindre des valeurs proches de 100 T pendant de très courts intervalles de temps.

Ces valeurs restent très inférieures à celles de ces mêmes champs au niveau microscopique. Par exemple, le champ électrique qui retient l'électron dans l'atome d'hydrogène est

$$E_{\mathrm{at}} \simeq \frac{e}{4\pi\varepsilon_0 a_0^2} \simeq 5 . 10^{11}\ \mathrm{V/m}\ . \tag{A.30}$$

---

20. Le fait que l'on puisse établir une équivalence entre température et énergie ne signifie *pas* que la température est une énergie, ni que l'énergie d'un système macroscopique à la température T est le produit de kT par le nombre d'atomes qu'il contient (nous verrons que la réalité est plus complexe). Cependant, il est parfois vrai (mais pas toujours!) que l'énergie moyenne par atome d'un système maintenu à la température T est *de l'ordre de kT,* ce qui peut être utilisé pour des estimations d'ordres de grandeur. Cette propriété, et sa validité, seront étudiées en détail au chapitre III (§ V.C).

TABLEAU VIII
Quelques températures en K et en eV.

	T (K)	kT (eV)
Centre du Soleil	$>5\times10^{6}$	>400
Surface du Soleil	$\sim5\,500$	0,47
Fusion du tungstène	3 650	0,31
Fusion du fer	1 808	0,16
Fusion du plomb	600	$5{,}2\times10^{-2}$
Ébullition de l'eau	373	$3{,}2\times10^{-2}$
Fusion de la glace	273	$2{,}4\times10^{-2}$
Ébullition de l'azote	77	$6{,}6\times10^{-3}$
Fusion de l'azote	63,3	$5{,}5\times10^{-3}$
Ébullition de l'hydrogène	20,3	$1{,}7\times10^{-3}$
Fusion de l'hydrogène	14	$1{,}2\times10^{-3}$
Ébullition de l'hélium 4	4,2	$3{,}6\times10^{-4}$
Transition superfluide de l'hélium 4	2,17	$1{,}9\times10^{-4}$
Transition superfluide de l'hélium 3	$2{,}6\times10^{-3}$	$2{,}2\times10^{-7}$

Cette valeur nous permet de déterminer une limite supérieure pour les champs macroscopiques que l'on peut obtenir au laboratoire : en effet, si le champ électrique créé par un conducteur atteignait la valeur (A.30), la force de pression électrostatique sur le conducteur serait du même ordre que la force de cohésion qui maintient ensemble les atomes, et le conducteur se désintégrerait. De même, le champ magnétique maximal B_{max} que l'on peut espérer obtenir est tel que la force magnétique correspondante soit du même ordre que la force électrique qui maintient l'électron :

$$evB_{max} \simeq eE_{at} \; ; \tag{A.31}$$

puisque, en ordre de grandeur,

$$\frac{1}{2}\, m_e v^2 \simeq Ry \, , \tag{A.32}$$

on trouve

$$B_{max} \simeq E_{at} \sqrt{\frac{m_e}{Ry}} \simeq \frac{\hbar}{e a_0^2} \simeq 3.10^5 \text{ T.} \tag{A.33}$$

La matière qui transporte les courants électriques nécessaires à la réalisation de tels champs magnétiques n'est plus stable.

En fait, aussi bien pour les champs électriques que pour les champs magnétiques, des problèmes se posent bien avant d'atteindre ces limites théoriques; c'est pourquoi les valeurs maximales obtenues au laboratoire sont nettement plus faibles.

III. UN EXEMPLE : LE GAZ PARFAIT CLASSIQUE

1. DÉFINITIONS

Un ensemble de N particules identiques, de masse m, confinées dans un récipient de volume V constitue *un gaz* si, en moyenne, l'énergie cinétique des particules est grande devant leur énergie potentielle d'interaction[21].

21. Le cas inverse, où l'énergie potentielle l'emporte largement sur l'énergie cinétique, est celui du solide. Dans un liquide, les deux énergies sont du même ordre de grandeur (ce qui complique considérablement l'étude théorique).

De plus, ce gaz est dit

(i) *classique,* si on peut décrire le mouvement des particules constitutives dans le formalisme de la mécanique classique;

(ii) *parfait,* si l'énergie potentielle est négligeable.

Si la deuxième condition est vérifiée sans que la première le soit, on a affaire à un *gaz parfait quantique.*

Il est clair que les deux conditions précédentes sont d'autant mieux vérifiées que le gaz est *plus dilué;* mais il faut évidemment préciser cette affirmation pour chacune des deux conditions.

2. VALIDITÉ DE LA DESCRIPTION CLASSIQUE

a. Condition de validité

Comme nous l'avons vu au paragraphe I.4, il nous faut comparer $\hbar$ aux grandeurs physiques du système ayant les dimensions d'une action. Dans le cas du mouvement des molécules d'un gaz, il existe une seule grandeur de ce genre, obtenue en effectuant le produit de la quantité de mouvement mv d'une particule par la distance d entre deux de ces particules (le produit de l'énergie cinétique d'une particule par le temps s'écoulant entre deux collisions est systématiquement plus grand). La condition (i) se traduit alors par la relation

$$mvd \gg \hbar \,. \tag{A.34}$$

Évidemment, le domaine de variation de d et v est *a priori* très vaste : d peut varier entre le diamètre d'une particule et la plus grande dimension du récipient, v entre zéro et l'infini. Mais les vitesses et les distances interparticulaires sont pratiquement toujours proches des valeurs moyennes $\bar{v}$ et $\bar{d}$ de v et d. Il suffit donc que la condition

$$m\bar{v}\bar{d} \gg \hbar \tag{A.35}$$

soit vérifiée. En introduisant la longueur d'onde de de Broglie associée à la vitesse moyenne :

$$\lambda\!\!\!^{-} \widehat{=} \frac{\hbar}{m\bar{v}}, \tag{A.36}$$

on réécrit cette condition sous la forme

$$\lambda\!\!\!^{-} \ll \bar{d} \,. \tag{A.37}$$

La distance moyenne $\bar{d}$ s'obtient facilement :

$$\bar{d} = \left(\frac{V}{N}\right)^{1/3} . \tag{A.38}$$

Quant à la vitesse moyenne $\bar{v}$, nous montrerons au chapitre III (§ V.C) qu'elle dépend de la température du gaz; en *ordre de grandeur,* on a[22]

$$\frac{1}{2} m\bar{v}^2 \simeq kT \,. \tag{A.39}$$

La condition (A.37) est donc équivalente à

$$kT \gg \frac{\hbar^2}{m} \left(\frac{N}{V}\right)^{2/3} . \tag{A.40}$$

22. En fait, la relation (A.39) n'est valable que si le gaz se comporte comme un gaz classique, c'est-à-dire si la condition (A.37) est vérifiée. Ce mode de raisonnement est courant en physique et nous l'utiliserons souvent par la suite. Il consiste à analyser les propriétés d'un système dans le cadre d'une approximation que l'on justifie (ou infirme) ensuite à l'aide des résultats obtenus.

b. Cas d'un gaz de molécules

Considérons par exemple l'hydrogène à température et pression ordinaires. La masse d'une molécule vaut deux fois celle du nucléon, et la longueur d'onde de de Broglie associée est environ, à la température ordinaire (300 K),

$$\lambda\!\!\!^{-} \simeq \frac{\hbar}{\sqrt{4M_N kT}} \simeq 0{,}20 \text{ Å} . \tag{A.41}$$

D'autre part, l'équation d'état du gaz, supposé parfait et classique (cf. note 22),

$$pV = NkT , \tag{A.42}$$

donne[23] pour une pression p de 1 atmosphère (10^5 Pa)

$$\bar{d} = \left(\frac{V}{N}\right)^{1/3} = \left(\frac{kT}{p}\right)^{1/3} \simeq 34 \text{ Å} . \tag{A.43}$$

La condition (A.37) est donc bien vérifiée, et l'approximation classique excellente.

Les molécules des autres gaz ont des masses plus importantes; la longueur d'onde de de Broglie est donc plus petite et l'approximation meilleure. *Tous les gaz ordinaires sont bien décrits par la mécanique classique*[24].

c. Cas du gaz d'électrons dans un métal

Un solide cristallin est constitué d'atomes régulièrement agencés dans l'espace, aux nœuds d'un réseau. Si ce solide est un métal, les électrons périphériques des divers atomes ne restent pas au voisinage d'un site particulier, mais se déplacent de façon quasi-libre dans l'ensemble du cristal; celui-ci est donc constitué d'un réseau d'*ions positifs* baignés par un *«gaz» d'électrons de conduction.*

La masse de l'électron est beaucoup plus petite (d'un facteur 2×1836) que celle de la molécule d'hydrogène considérée au paragraphe b; sa longueur d'onde de de Broglie est donc beaucoup plus grande[25] que (A.41) :

$$\lambda\!\!\!^{-} \simeq 0{,}20\sqrt{2 \times 1836} = 12 \text{ Å} . \tag{A.44}$$

Au contraire, la distance interparticulaire est beaucoup plus petite : s'il y a un électron de conduction par ion (c'est le cas, par exemple, du sodium ou du cuivre), $\bar{d}$ est la distance entre ions proches voisins dans le métal :

$$\bar{d} \simeq 2 \text{ Å} . \tag{A.45}$$

La condition (A.37) n'est donc pas satisfaite. Les électrons d'un métal forment un gaz trop dense pour qu'une description classique soit adéquate. Il est essentiel de les traiter quantiquement en tenant compte explicitement du principe d'exclusion de Pauli. On dit que les électrons de conduction d'un métal forment un *gaz quantique dégénéré.*

---

23. Notons que ce résultat est valable pour n'importe quel gaz parfait classique. Il peut donc être utile de se rappeler que, dans les conditions ordinaires, la distance moyenne entre deux molécules d'un gaz est voisine de 30 Å.

24. L'approximation classique pourrait devenir mauvaise à basse température ou à haute pression (cf. (A.41) et (A.43)). En fait, seul l'hélium reste gazeux à suffisamment basse température (ou haute pression) pour que la question puisse vraiment se poser.

25. Comme dans le cas de l'hydrogène, nous calculons $\lambda\!\!\!^{-}$ en supposant la relation (A.39) satisfaite. En fait, elle ne l'est pas pour les électrons d'un métal; mais elle le serait s'ils se comportaient comme un gaz parfait classique, ce qu'on cherche à déterminer (cf. note 22).

3. COMPARAISON ENTRE ÉNERGIE CINÉTIQUE ET ÉNERGIE POTENTIELLE

Nous limiterons l'analyse de la deuxième condition aux gaz de molécules. Comme nous venons de le voir, cette analyse peut alors s'effectuer dans le cadre classique.

L'énergie cinétique des molécules d'un gaz dépend de la température. Son ordre de grandeur, nous l'avons déjà indiqué (formule (A.39)), est

$$E_{cin} \simeq NkT . \tag{A.46}$$

Reste à évaluer l'énergie potentielle d'interaction, ce qui est plus délicat. L'interaction entre deux molécules d'un gaz dépend de leur distance relative. Lorsque cette distance devient inférieure à leur dimension propre, les deux molécules exercent l'une sur l'autre des forces répulsives extrêmement grandes (c'est essentiellement le principe de Pauli auquel obéissent les électrons qui rend ainsi chaque molécule pratiquement impénétrable aux autres). A longue distance, au contraire, elles s'attirent faiblement (forces de van der Waals). Le potentiel d'interaction $u(r)$ entre deux molécules, tracé en fonction de leur distance relative r, a donc l'allure représentée sur la figure 1.

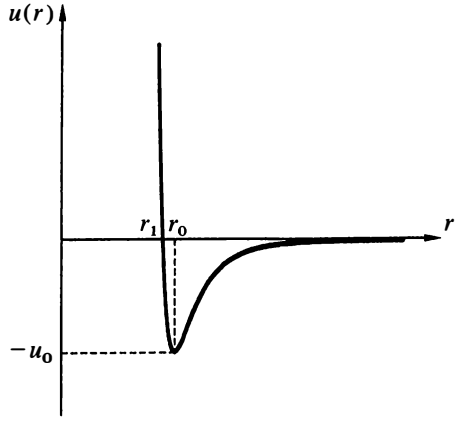


FIGURE 1
Potentiel d'interaction typique entre deux molécules, en fonction de leur distance relative r.

Pour des molécules simples, il est possible de calculer $u(r)$ à l'aide de la mécanique quantique. Mais on utilise, en général, des relations semi-empiriques de la forme

$$u(r) = -u_0 f(r/r_0) ; \tag{A.47}$$

$u(r)$ est donc essentiellement caractérisé par les deux paramètres microscopiques qui définissent son minimum : r_0 et u_0; $f(x)$ est une fonction sans dimension qui rend compte de la forme de la courbe de la figure 1.

Appelons r_1 la distance qui annule $u(r)$. En première approximation[26], nous pouvons dire qu'une molécule donnée à une probabilité nulle d'approcher une autre

26. Cette approximation est d'autant mieux justifiée que l'énergie cinétique des molécules est plus grande devant leur énergie potentielle (c'est justement la condition que nous cherchons à écrire). En effet, à cause de leur attraction à longue distance, deux molécules ont une probabilité plus forte d'être proches qu'éloignées; mais cet effet devient négligeable si l'énergie cinétique est grande.

molécule à une distance inférieure à r_1, et une probabilité uniforme d'en être à une distance supérieure à r_1, en n'importe quel point du récipient. La probabilité $dP(r)$ de trouver deux molécules à une distance r l'une de l'autre (à dr près) est alors

$$dP(r) \simeq \begin{cases} 0 & \text{si } r < r_1 \\ \dfrac{4\pi r^2\, dr}{V} & \text{si } r > r_1 \,. \end{cases} \tag{A.48}$$

On en déduit l'énergie moyenne d'interaction entre ces deux molécules[27] :

$$\bar{u} \simeq -\int_{r_1}^{\infty} u_0 f(r/r_0) \frac{4\pi r^2}{V}\, dr \,, \tag{A.49}$$

que l'on peut écrire, en passant à la variable $x = r/r_0$,

$$\bar{u} \simeq -\frac{4\pi r_0^3}{V} u_0 \int_{x_1}^{\infty} f(x) x^2\, dx \,, \tag{A.50}$$

où $x_1 = r_1/r_0$ est un nombre légèrement inférieur à 1. La fonction $f(x)$ a une portée de quelques unités; elle varie entre 0 et 1 dans l'intervalle d'intégration (voir figure 1), de sorte que l'intégrale qui intervient dans (A.50) est un nombre légèrement inférieur à 1. En ordre de grandeur, on a donc

$$\bar{u} \simeq -\frac{u_0 r_0^3}{V} \,. \tag{A.51}$$

Si N est le nombre de molécules du gaz, il y a

$$\frac{1}{2} N(N-1) \simeq \frac{N^2}{2} \tag{A.52}$$

paires de molécules, et l'énergie d'interaction moyenne de l'ensemble du gaz vaut

$$U \simeq -\frac{N^2}{2} u_0 \frac{r_0^3}{V} \,. \tag{A.53}$$

L'énergie potentielle d'interaction sera donc négligeable devant l'énergie cinétique (A.46), et le gaz considéré comme parfait, si

$$\frac{kT}{u_0} \times \frac{V}{N r_0^3} \gg 1 \,. \tag{A.54}$$

Ainsi, u_0/k apparaît comme l'échelle naturelle de température et r_0^3 comme celle du volume par particule V/N.

A densité N/V fixée, c'est donc à haute température que le gaz se comporte comme un gaz parfait. A température fixée, l'approximation est d'autant meilleure que le gaz est plus dilué.

Il nous reste à préciser les paramètres du potentiel, u_0 et r_0. Puisque ce sont des paramètres du domaine atomique, on s'attend à trouver u_0 de l'ordre de l'eV, r_0 de l'ordre de l'Å. En fait, u_0, qui représente l'énergie d'interaction entre deux molécules *chimiquement stables,* est pour les gaz usuels beaucoup plus faible, et r_0 systématiquement plus grand : u_0 est de l'ordre de quelques centièmes d'eV seulement et r_0 vaut de 3 à 10 Å suivant la taille des molécules. Par exemple, r_0 vaut 3,9 Å pour l'argon, 4,2 Å pour l'azote et 4,8 Å pour le gaz carbonique; les valeurs de u_0 correspondantes sont respectivement 10^{-2} eV (118 K), 8.10^{-3} eV (95 K) et $1{,}7 \times 10^{-2}$ eV (198 K)[28].

---

27. La portée de la fonction $u(r)$ est de quelques r_0 seulement, c'est-à-dire d'ordre microscopique. On peut donc, sans changer la valeur de $\bar{u}$, prolonger l'intégration dans (A.49) jusqu'à l'infini.

28. Les valeurs de u_0 et r_0 dépendent en fait de la fonction $f(x)$ choisie pour représenter $u(r)$. Celles qui sont données ici correspondent au modèle de Lennard-Jones, pour lequel

$$f(x) = \frac{2}{x^6} - \frac{1}{x^{12}} \,.$$

A température ordinaire (300 K) et sous pression atmosphérique, si on prend 10^{-2} eV pour u_0 et 4 Å pour r_0, on obtient (cf. (A.43))

$$\left(\frac{kT}{u_0}\right)\left(\frac{V}{Nr_0^3}\right) \simeq 1\,500 \,. \tag{A.55}$$

L'approximation gaz parfait est donc bien justifiée dans ce cas. Elle peut cependant devenir mauvaise pour des températures plus basses ou des pressions plus élevées.

En résumé, un gaz se comporte comme un gaz parfait classique s'il est suffisamment dilué pour que les deux conditions (A.40) et (A.54) soient simultanément satisfaites. Il faut noter que, pour pratiquement tous les gaz usuels, la condition (A.54) se trouve violée avant la condition (A.40) : sauf circonstances exceptionnelles, un gaz est correctement décrit par la mécanique classique; c'est seulement si la température n'est pas trop basse et si la densité est assez faible qu'on peut négliger les interactions et que l'approximation gaz parfait est justifiée.

COMPLÉMENT I.B

Distribution statistique d'une fonction de variables aléatoires

I. Distribution statistique d'une grandeur fonction d'une seule variable aléatoire
   1. Densité de probabilité d'une fonction strictement monotone
      a. Fonction strictement croissante
      b. Fonction strictement décroissante
   2. Cas général d'une fonction quelconque

II. Distribution statistique d'une grandeur fonction de plusieurs variables aléatoires
   1. Fonction de répartition
   2. Densité de probabilité
   3. Cas particulier : $A(\vec{r})$ est une somme de variables aléatoires indépendantes

III. Exemples
   1. Projection d'un vecteur aléatoire de module constant
   2. Énergie cinétique d'une particule de vitesse aléatoire

Soit une variable aléatoire continue x dont on connaît la densité de probabilité $w(x)$. Considérons une grandeur physique A, fonction de cette variable : $A = A(x)$. Quelle est la densité de probabilité $W(A)$ correspondante? Autrement dit, quelle est la probabilité

$$\mathrm{d}P(A) = W(A)\,\mathrm{d}A \tag{B.1}$$

pour que la grandeur A appartienne à l'intervalle infiniment petit $[A,\ A+\mathrm{d}A]$?

Les formules (24) et (25) de l'appendice I indiquent que cette probabilité s'écrit

$$W(A)\,\mathrm{d}A = \int_{\mathrm{d}D(A)} w(x)\,\mathrm{d}x\ , \tag{B.2}$$

où $\mathrm{d}D(A)$ est le domaine infinitésimal de valeurs de x telles que

$$A \leqslant A(x) \leqslant A + \mathrm{d}A\ . \tag{B.3}$$

Nous montrons dans ce complément comment calculer pratiquement la densité de probabilité $W(A)$ d'une grandeur A, fonction d'abord **d'une seule**, puis de plusieurs variables aléatoires[1]. Nous terminons en traitant **explicitement un exemple** de chaque cas.

I. DISTRIBUTION STATISTIQUE D'UNE GRANDEUR FONCTION D'UNE SEULE VARIABLE ALÉATOIRE

Pour bien mettre en évidence le raisonnement, nous examinerons le cas où A est une fonction monotone de la variable aléatoire x avant de traiter le cas général.

1. Nous supposons que les valeurs de A sont des nombres réels. Si A représente une grandeur appartenant à un espace de dimension supérieure à 1 (A est alors un vecteur de cet espace), on utilise la même méthode en examinant tour à tour chacune de ses composantes.

1. DENSITÉ DE PROBABILITÉ D'UNE FONCTION STRICTEMENT MONOTONE

a. Fonction strictement croissante

Si $A(x)$ est une fonction strictement croissante de x, la relation entre A et x est biunivoque et x est elle-même une fonction bien définie et croissante de A :

$$A = A(x) \text{ croissante} \iff x = x(A) \text{ croissante} . \tag{B.4}$$

La figure 1.a définit alors sans ambiguïté le domaine $\mathrm{d}D(A)$ de la formule (B.2) qui s'écrit

$$W(A)\,\mathrm{d}A = \int_{x(A)}^{x(A+\mathrm{d}A)} w(x')\,\mathrm{d}x' . \tag{B.5}$$

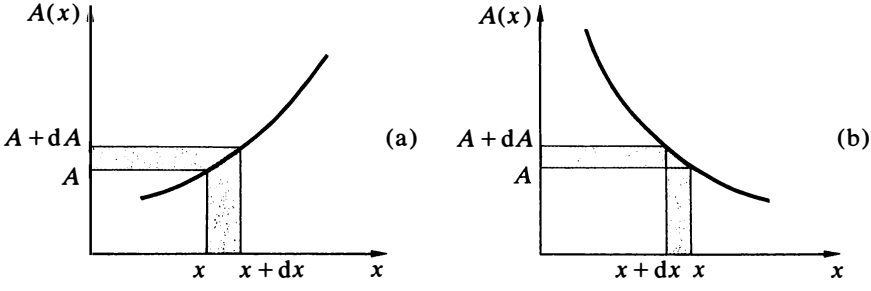


FIGURE 1
Domaine $\mathrm{d}D(A)$ des valeurs de x telles que $A \leqslant A(x) \leqslant A + \mathrm{d}A$ pour une fonction $A(x)$ croissante (fig. a) et pour une fonction $A(x)$ décroissante (fig. b).

Le changement de variable $x' = x(A')$ donne

$$W(A)\,\mathrm{d}A = \int_{A}^{A+\mathrm{d}A} w\left[x(A')\right] \frac{\mathrm{d}x}{\mathrm{d}A'}\,\mathrm{d}A' . \tag{B.6}$$

Mais l'intervalle d'intégration $\mathrm{d}A$ est infinitésimal, de sorte que

$$W(A)\,\mathrm{d}A = w\left[x(A)\right] \frac{\mathrm{d}x}{\mathrm{d}A}\,\mathrm{d}A . \tag{B.7}$$

On peut réécrire cette relation plus succinctement sous la forme

$$W(A)\,\mathrm{d}A = w(x)\,\mathrm{d}x , \tag{B.8}$$

à condition de bien comprendre que, dans le membre de droite, il est sous-entendu que $x = x(A)$.

b. Fonction strictement décroissante

Si $A(x)$ est une fonction strictement décroissante, la relation entre A et x reste biunivoque, mais la fonction $x(A)$ est maintenant une *fonction décroissante de A* :

$$A = A(x) \text{ décroissante} \iff x = x(A) \text{ décroissante} . \tag{B.9}$$

La figure 1.b montre qu'alors le domaine d'intégration $dD(A)$ va de $(x+dx)$ à x[2], de sorte que

$$W(A)dA = \int_{x(A+dA)}^{x(A)} w(x')dx' . \tag{B.10}$$

Le même changement de variable $x' = x(A')$ donne alors

$$W(A)dA = \int_{A+dA}^{A} w[x(A')] \frac{dx}{dA'} dA' , \tag{B.11}$$

c'est-à-dire

$$W(A)dA = -w[x(A)] \frac{dx}{dA} dA , \tag{B.12}$$

ce qu'on peut réécrire de façon analogue à (B.8) sous la forme

$$W(A)dA = -w(x)dx . \tag{B.13}$$

2. CAS GÉNÉRAL D'UNE FONCTION QUELCONQUE

Lorsque A n'est pas une fonction monotone de x, $A = A(x)$ ne définit pas une fonction unique x de A mais autant de fonctions $x_i(A)$ que l'équation en x

$$A - A(x) = 0 \tag{B.14}$$

a de solutions x_i pour A donnée (voir figure 2). Chaque $x_i(A)$ donne une contribution

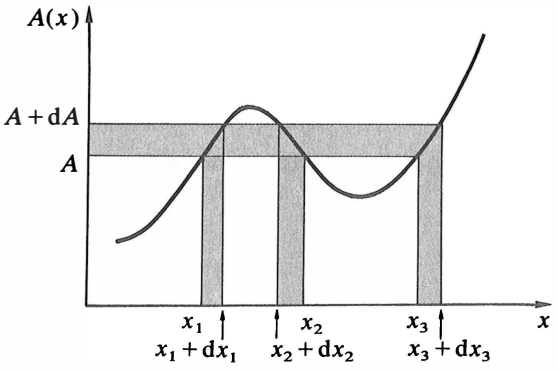


FIGURE 2

Domaine $dD(A)$ des valeurs de x telles que $A \leq A(x) \leq A + dA$ pour une fonction quelconque $A(x)$. Ici le domaine $dD(A)$ se décompose en trois puisque l'équation $A - A(x) = 0$ a trois solutions en x pour la valeur de A considérée : x_1, x_2 et x_3.

2. x étant une fonction décroissante de A, l'accroissement infinitésimal dx associé à dA ($dA > 0$) est ici négatif. La position relative des bornes d'intégration dans (B.10) assure donc que la probabilité $W(A)dA$ est positive.

à $W(A)\mathrm{d}A$ du type (B.7) ou (B.12) suivant qu'il est une fonction croissante ou décroissante de A. De sorte que la probabilité cherchée s'écrit[3],[4]

$$W(A)\mathrm{d}A = \sum_i w\left[x_i(A)\right] \left|\frac{\mathrm{d}x_i}{\mathrm{d}A}\right| \mathrm{d}A \,. \tag{B.15}$$

Montrons comment on retrouve ce résultat à partir des propriétés de la « fonction de Dirac » (c'est cette méthode que nous généraliserons au paragraphe II pour traiter le cas des fonctions de plusieurs variables aléatoires).

Définissons tout d'abord la *fonction de répartition* $P(A)$ de la variable aléatoire A : $P(A)$ est la probabilité pour que la variable aléatoire prenne une valeur inférieure à la valeur donnée A. $P(A)$ se calcule par la « loi d'addition » (appendice I, § I.2.a) :

$$P(A) = \int_{D(A)} w(x)\mathrm{d}x \,, \tag{B.16}$$

où $D(A)$ est le domaine de valeurs de x telles que

$$A(x) \leqslant A \,. \tag{B.17}$$

Utilisons la « fonction de Heaviside » $\theta(x)$:

$$\theta(x) \widehat{=} \begin{cases} 0 & \text{si} \quad x<0 \\ 1 & \text{si} \quad x>0 \end{cases} \tag{B.18}$$

pour réécrire (B.16) sous la forme[5]

$$P(A) = \int_{-\infty}^{+\infty} w(x)\,\theta\left[A - A(x)\right]\mathrm{d}x \,. \tag{B.19}$$

D'autre part, la probabilité $\mathrm{d}P(A)$ pour que A appartienne à l'intervalle infinitésimal $[A,\ A+\mathrm{d}A]$ est évidemment

$$\mathrm{d}P(A) = P(A+\mathrm{d}A) - P(A) \,, \tag{B.20}$$

de sorte que la densité de probabilité $W(A)$ apparaît comme la dérivée de la fonction de répartition :

$$W(A) = \frac{\mathrm{d}P(A)}{\mathrm{d}A} \,. \tag{B.21}$$

Dérivant donc $P(A)$ par rapport à A dans (B.19), on obtient

$$W(A) = \int_{-\infty}^{+\infty} w(x)\,\delta\left[A - A(x)\right]\mathrm{d}x \,. \tag{B.22}$$

---

3. La valeur absolue de la dérivée qui intervient dans (B.15) interdit une écriture simplifiée de cette formule analogue à (B.8) ou (B.13). Il est cependant commode d'utiliser (B.8) et (B.13) comme moyen mnémotechnique pour retrouver (B.15).

4. La formule (B.15) suppose que la fonction $A(x)$ ne présente pas de palier, c'est-à-dire d'intervalle fini $[x_1, x_2]$ tel que, pour tout x appartenant à cet intervalle, $A(x)$ prenne une valeur constante A_0. Si c'était le cas, la probabilité d'obtenir cette valeur A_0 de la variable A serait finie et égale à $\int_{x_1}^{x_2} w(x)\mathrm{d}x$.

5. Si $w(x)$ n'est définie que sur un intervalle fini $[x_1, x_2]$, on convient de la prolonger en la prenant nulle à l'extérieur de cet intervalle.

En effet, la dérivée de la fonction de Heaviside $\theta(x)$ n'est autre que la fonction de Dirac $\delta(x)$. Pour calculer $W(A)$, il suffit alors de savoir[6] que

$$\delta[f(x)]=\sum_i \frac{1}{|f'(x_i)|}\,\delta(x-x_i)\,, \tag{B.23.a}$$

où $f'(x_i)$ est la dérivée de $f(x)$ au point x_i; la somme porte sur l'ensemble des x_i tels que

$$f(x_i)=0\,. \tag{B.23.b}$$

On obtient alors, en utilisant (B.22) et (B.23),

$$W(A)=\sum_i \frac{1}{|A'(x_i)|}\,w(x_i)\,, \tag{B.24.a}$$

où la somme porte sur l'ensemble des x_i tels que

$$A(x_i)-A=0\,. \tag{B.24.b}$$

On vérifie aisément que (B.24) est identique à (B.15).

II. DISTRIBUTION STATISTIQUE D'UNE GRANDEUR FONCTION DE PLUSIEURS VARIABLES ALÉATOIRES

La grandeur physique A est maintenant une fonction de plusieurs variables aléatoires $(x, y, z...)$ dont la distribution est caractérisée par la densité de probabilité $w(x, y, z...)$. Pour alléger l'écriture, nous utiliserons la notation vectorielle (appendice I, § I.1.b) et nous écrirons

$$w(x, y, z, ...)\,dx\,dy\,dz... \mathrel{\widehat{=}} w(\vec{r})\,d\vec{r} \tag{B.25}$$

et

$$A(x, y, z, ...) \mathrel{\widehat{=}} A(\vec{r})\,. \tag{B.26}$$

1. FONCTION DE RÉPARTITION

La fonction de répartition $P(A)$, probabilité pour que la variable aléatoire $A(\vec{r})$ prenne une valeur inférieure à la valeur A, est alors donnée par une relation analogue à (B.16) :

$$P(A)=\int_{D(A)} w(\vec{r})\,d\vec{r}, \tag{B.27}$$

où l'intégrale porte sur le « volume » $D(A)$ défini comme l'ensemble des points $\vec{r}$ tels que

$$A(\vec{r})\leqslant A\,. \tag{B.28}$$

Là encore on peut utiliser la fonction de Heaviside pour réécrire (B.27) sous la forme équivalente

$$P(A)=\int w(\vec{r})\,\theta[A-A(\vec{r})]\,d\vec{r}, \tag{B.29}$$

où l'intégrale porte maintenant sur toutes les valeurs de $\vec{r}$.

6. Voir par exemple C.D.L., appendice II.

2. DENSITÉ DE PROBABILITÉ

La densité de probabilité $W(A)$ se calcule comme dans le paragraphe précédent à partir de la fonction de répartition :

$$W(A)=\frac{\mathrm{d}P(A)}{\mathrm{d}A}\,. \tag{B.30}$$

En explicitant la dérivée par rapport à A dans (B.29), on obtient

$$W(A)=\int w(\vec{r})\,\delta[A-A(\vec{r})]\,\mathrm{d}\vec{r}\,. \tag{B.31}$$

Mais cette fois l'intégrale ne peut se calculer aussi simplement que dans le cas d'une fonction d'une seule variable. En effet, la fonction de Dirac permet d'effectuer l'intégration par rapport à l'une des n composantes du vecteur $\vec{r}$, mais il reste encore à calculer une intégrale de « surface » sur les $(n-1)$ variables restantes. Cette intégrale de surface peut s'écrire[7]

$$W(A)=\int_{s(A)} \mathrm{d}s\,\frac{w(\vec{r})}{|\vec{\nabla}A(\vec{r})|}\,, \tag{B.32}$$

où $|\vec{\nabla}A(\vec{r})|$ est le module du vecteur gradient de $A(\vec{r})$; l'intégrale porte sur la surface $s(A)$ d'équation

$$A(\vec{r})=A\,. \tag{B.33}$$

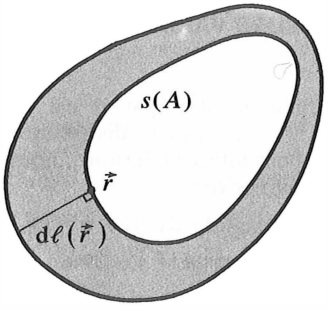


FIGURE 3
Ensemble $\mathrm{d}D(A)$ des points $\vec{r}$ tels que $A\leq A(\vec{r})\leq A+\mathrm{d}A$, délimité par les surfaces $s(A)$ et $s(A+\mathrm{d}A)$; $\mathrm{d}\ell(\vec{r})$ est la distance entre ces deux surfaces au point $\vec{r}$.

7. La probabilité $\mathrm{d}P(A)$ pour que la variable aléatoire $A(\vec{r})$ appartienne à l'intervalle $[A, A+\mathrm{d}A]$ est

$$\mathrm{d}P(A)=\int_{\mathrm{d}D(A)} \mathrm{d}\vec{r}\,w(\vec{r})\,, \tag{1}$$

où $\mathrm{d}D(A)$ est le volume compris entre la surface $s(A)$ d'équation

$$A(\vec{r})=A \tag{2}$$

et la surface $s(A+\mathrm{d}A)$ d'équation

$$A(\vec{r})=A+\mathrm{d}A\,. \tag{3}$$

Soit alors $\mathrm{d}\ell(\vec{r})$ la distance entre ces deux surfaces, mesurée dans la direction orthogonale à $s(A)$ au point $\vec{r}$ (voir figure 3). L'élément de volume $\mathrm{d}\vec{r}$ de la formule (1) peut s'écrire comme le produit de $\mathrm{d}\ell(\vec{r})$ par l'élément de surface $\mathrm{d}s$ de $s(A)$ au point $\vec{r}$, de sorte que

$$\mathrm{d}P(A)=\int_{s(A)} \mathrm{d}s\,\mathrm{d}\ell(\vec{r})\,w(\vec{r})\,. \tag{4}$$

Or $\mathrm{d}\ell(\vec{r})$ est relié à $\mathrm{d}A$ par

$$\mathrm{d}A=|\vec{\nabla}A|\,\mathrm{d}\ell(\vec{r})\,, \tag{5}$$

puisque $\vec{\nabla}A$ est orthogonal à la surface $s(A)$. Il suffit alors de reporter (5) dans (4) pour obtenir (B.32).

Pour calculer la densité de probabilité $W(A)$ on pourra donc, soit calculer l'intégrale de volume (B.27) pour déterminer la fonction de répartition et en déduire $W(A)$ par dérivation, soit calculer directement cette densité de probabilité sous la forme (B.31) ou (B.32). Remarquons que (si le calcul a été fait correctement!) le résultat obtenu pour la densité de probabilité $W(A)$ est automatiquement normé[8] :

$$\int_{-\infty}^{+\infty} W(A)\,\mathrm{d}A = 1 \,. \tag{B.34}$$

3. CAS PARTICULIER : $A(\vec{r})$ EST UNE SOMME DE VARIABLES ALÉATOIRES INDÉPENDANTES

Pour ne pas alourdir inutilement l'exposé, nous supposerons que $A(\vec{r})$ est la somme de *deux* variables aléatoires indépendantes x et y. L'extension à un nombre quelconque de variables est évidente.

Nous cherchons donc à calculer la densité de probabilité $W(A)$ de la fonction

$$A(\vec{r}) \equiv A(x, y) = x + y \tag{B.35}$$

des deux variables aléatoires *indépendantes* x et y. La densité de probabilité $w(\vec{r})$ est alors (appendice I, § I.2.b) le produit des densités de probabilité $u(x)$ et $v(y)$ des deux variables x et y :

$$w(x, y) = u(x).v(y) \,. \tag{B.36}$$

Pour calculer l'intégrale (B.31), il est commode d'utiliser la représentation intégrale suivante de la « fonction de Dirac » (note 6) :

$$\delta(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \mathrm{d}k\ \mathrm{e}^{-ikx} \,. \tag{B.37}$$

On obtient ainsi

$$W(A) = \frac{1}{2\pi} \int \mathrm{d}x \int \mathrm{d}y\ u(x)\,v(y) \int_{-\infty}^{+\infty} \mathrm{d}k\ \mathrm{e}^{ik(x+y-A)} \,, \tag{B.38}$$

c'est-à-dire, en intervertissant l'ordre des intégrations,

$$W(A) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \mathrm{d}k\ \mathrm{e}^{-ikA} \int_{-\infty}^{+\infty} \mathrm{d}x\ u(x)\mathrm{e}^{ikx} \int_{-\infty}^{+\infty} \mathrm{d}y\ v(y)\mathrm{e}^{iky} \,. \tag{B.39}$$

Cette formule s'interprète très simplement si l'on introduit la notion de *fonction caractéristique :* pour une variable aléatoire z quelconque de densité de probabilité $w(z)$, la fonction caractéristique $\widetilde{w}(k)$ est, par définition, la *valeur moyenne de* e^{ikz}, c'est-à-dire (à un facteur près) la *transformée de Fourier* de $w(z)$:

$$\widetilde{w}(k) \mathrel{\widehat{=}} \int \mathrm{d}z\ \mathrm{e}^{ikz} w(z) \,. \tag{B.40}$$

---

8. On vérifie aisément ce résultat à partir de (B.31) :

$$\begin{aligned} \int W(A)\,\mathrm{d}A &= \int \mathrm{d}A \int \mathrm{d}\vec{r}\, w(\vec{r})\,\delta[A - A(\vec{r})] \\ &= \int \mathrm{d}\vec{r}\, w(\vec{r}) \int \mathrm{d}A\ \delta[A - A(\vec{r})] \\ &= \int \mathrm{d}\vec{r}\, w(\vec{r}) \,. \end{aligned}$$

Bien entendu, $w(z)$ peut être retrouvée à partir de la fonction caractéristique, par transformée de Fourier inverse :

$$w(z)=\frac{1}{2\pi}\int \mathrm{d}k\ \mathrm{e}^{-ikz}\widetilde{w}(k)\ . \tag{B.41}$$

La formule (B.39) exprime alors la règle suivante : la fonction caractéristique $\widetilde{W}(k)$ de la grandeur physique A, somme des variables aléatoires indépendantes x et y, est le produit des fonctions caractéristiques de x et de y :

$$\widetilde{W}(k)=\widetilde{u}(k).\widetilde{v}(k)\ . \tag{B.42}$$

III. EXEMPLES

1. PROJECTION D'UN VECTEUR ALÉATOIRE DE MODULE CONSTANT

Soit $\vec{A}$ un vecteur du plan, de module A_0 constant et de direction aléatoire. Nous supposons que toutes les directions sont également probables, c'est-à-dire que si φ $(0\leqslant\varphi<2\pi)$ représente l'angle que fait le vecteur $\vec{A}$ avec une direction Ox donnée, la densité de probabilité correspondante $w(\varphi)$ est constante. La condition de normalisation impose

$$w(\varphi)=\frac{1}{2\pi}\ ,\quad (0\leqslant\varphi<2\pi)\ . \tag{B.43}$$

Quelle est alors la probabilité pour que la projection du vecteur $\vec{A}$ sur l'axe Ox soit comprise entre A_x et $A_x+\mathrm{d}A_x$?

La figure 4 montre qu'il existe deux valeurs de φ : φ_1 et $\varphi_2=2\pi-\varphi_1$ qui correspondent à une valeur donnée de A_x :

$$A_x=A_0\cos\varphi_1=A_0\cos\varphi_2\ . \tag{B.44}$$

L'application de (B.15) donne immédiatement

$$W(A_x)\,\mathrm{d}A_x=\frac{1}{2\pi}\left[\frac{1}{|A_0\sin\varphi_1|}+\frac{1}{|A_0\sin\varphi_2|}\right]\mathrm{d}A_x\ . \tag{B.45}$$

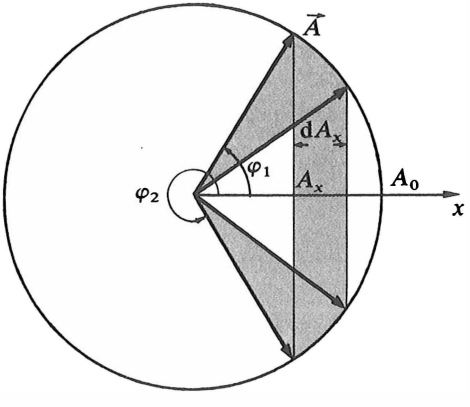


FIGURE 4
Vecteur $\vec{A}$ de module constant A_0 et de direction aléatoire. A la valeur A_x de sa projection sur l'axe Ox correspondent deux valeurs φ_1 et φ_2 de l'angle polaire $(\varphi_2=2\pi-\varphi_1)$.

En remarquant que

$$|\sin \varphi_1| = |\sin \varphi_2| = \left[1 - \left(\frac{A_x}{A_0}\right)^2\right]^{1/2}, \tag{B.46}$$

on obtient finalement

$$W(A_x)\,dA_x = \frac{dA_x}{\pi [A_0^2 - A_x^2]^{1/2}}, \quad (-A_0 \leq A_x \leq A_0). \tag{B.47}$$

La densité de probabilité $W(A_x)$ (représentée sur la figure 5) devient infinie pour $A_x = \pm A_0$. Ce résultat était prévisible à partir de la figure 4 : au voisinage de ces valeurs extrémales, à un petit accroissement dA_x correspond un $d\varphi$ important. Évidemment $W(A_x)$ reste intégrable; en particulier on vérifie aisément la condition de normalisation en intégrant le second membre de (B.47) sur l'intervalle $[-A_0, A_0]$.

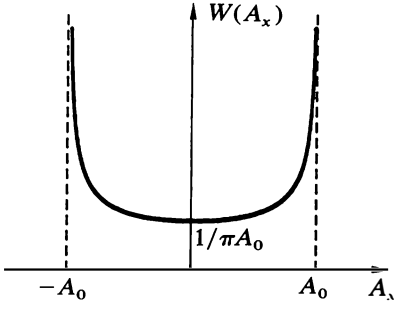


FIGURE 5
Densité de probabilité $W(A_x)$ pour la projection A_x du vecteur aléatoire $\vec{A}$ de la figure 4.

2. ÉNERGIE CINÉTIQUE D'UNE PARTICULE DE VITESSE ALÉATOIRE

On considère une particule libre de masse m dont les composantes v_x, v_y, v_z de la vitesse sur trois axes orthogonaux sont aléatoires, avec une densité de probabilité

$$w(v_x, v_y, v_z) = \left(\frac{\alpha}{\pi}\right)^{3/2} e^{-\alpha(v_x^2 + v_y^2 + v_z^2)}. \tag{B.48}$$

Quelle est la probabilité $dP(E) = W(E)\,dE$ pour que l'énergie cinétique E de cette particule soit comprise entre E et $E + dE$?

Dans ce cas très simple, le lecteur montrera sans difficulté[9] que les relations (B.30), (B.31) ou (B.32) conduisent au résultat suivant :

$$dP(E) = W(E)\,dE = \frac{2}{\sqrt{\pi}} \left(\frac{2\alpha}{m}\right)^{3/2} \sqrt{E}\, e^{-\frac{2\alpha E}{m}}\, dE, \tag{B.49}$$

qui vérifie la condition de normalisation

$$\int_0^\infty W(E)\,dE = 1. \tag{B.50}$$

9. Il suffit de passer en coordonnées sphériques dans l'espace des vitesses, en notant que

$$E = \frac{1}{2} m(v_x^2 + v_y^2 + v_z^2) = \frac{1}{2} mv^2.$$

COMPLÉMENT I.C

La distribution binomiale dans la limite des grands nombres

I. Hypothèses

II. Établissement de l'expression approchée
1. Détermination approchée de la valeur la plus probable de n
2. Développement de ln P autour de la valeur la plus probable
3. Résultat : l'approximation gaussienne de la distribution binomiale

III. Discussion de l'approximation

Dans l'appendice I, nous avons défini deux distributions statistiques particulières : la distribution binomiale et la distribution gaussienne. Nous allons montrer ici que, dans le domaine des grands nombres, la distribution binomiale se comporte pratiquement comme une gaussienne.

I. HYPOTHÈSES

Considérons un problème quelconque de probabilités discrètes. Dans l'ensemble des événements possibles, choisissons-en un particulier, que nous noterons (+), de probabilité P_+; l'événement contraire sera noté (−), sa probabilité étant par définition $P_- = 1 - P_+$. Si l'on effectue N tirages indépendants entre ces événements, la probabilité $P(N, n)$ pour que n de ces expériences fournissent le résultat (+) et les $(N - n)$ autres le résultat (−) est donnée par la loi binomiale (appendice I, § II.1) :

$$P(N, n) = \frac{N!}{n!(N-n)!} P_+^n P_-^{N-n} . \tag{C.1}$$

La valeur moyenne du nombre n est simplement

$$\bar{n} = NP_+ , \tag{C.2}$$

et son écart quadratique moyen

$$\Delta n = \sqrt{NP_+P_-} . \tag{C.3}$$

Nous supposons ici N *très grand devant* 1. Nous supposons en outre que la probabilité P_+ n'est pas trop petite ni trop voisine de 1[1], de sorte que $\bar{n}$ *est lui aussi grand devant* 1, *et éloigné de* N :

$$\bar{n} = NP_+ \gg 1 \; ; \quad N - \bar{n} = N(1 - P_+) \gg 1 . \tag{C.4}$$

1. Si la probabilité P_+ (ou P_-) est très petite devant 1, il existe un domaine intermédiaire où N est grand sans que NP_+ (ou NP_-) le soit. L'expression approchée que nous allons donner ici de la distribution binomiale n'est pas valable dans ce domaine. On peut cependant montrer que la distribution binomiale y est très voisine d'une *« distribution de Poisson »* (voir par exemple W. Feller, *An Introduction to Probability Theory and its Applications,* John Wiley, New York (1950), p. 153). Évidemment, si P_+ est fixée, même très petite ou très voisine de 1, la condition $NP_+P_- \gg 1$ finit toujours par être satisfaite pour des valeurs suffisamment grandes de N; l'approximation gaussienne devient alors valable.

Ces deux conditions peuvent être réunies en une seule :

$$NP_+P_- \gg 1 \,. \tag{C.5}$$

Nous nous intéressons uniquement aux *valeurs de n pas trop éloignées de* $\bar{n}$; plus précisément, nous limiterons l'intervalle de variation de n à quelques écarts quadratiques moyens Δn autour de $\bar{n}$:

$$|n-\bar{n}| \lesssim a\,\Delta n \,, \quad a \text{ de l'ordre de quelques unités} \tag{C.6}$$

(nous nous restreignons à ce domaine car la probabilité $P(N, n)$ devient rapidement très faible lorsque n s'éloigne de $\bar{n}$; voir paragraphe III).

Nous nous proposons enfin de *traiter n comme une variable continue.* Une telle approximation est justifiée si la fonction étudiée varie peu entre $n-1$ et n, c'est-à-dire, dans le cas présent, si

$$|P(N, n)-P(N, n-1)| \ll P(N, n) \,; \tag{C.7}$$

on peut alors valablement considérer la différence entre n et $n-1$ comme infinitésimale. Cette condition est effectivement remplie si les hypothèses précédentes sont vérifiées[2].

II. ÉTABLISSEMENT DE L'EXPRESSION APPROCHÉE

Nous allons expliciter ici, sur l'exemple de la distribution binomiale, une méthode d'analyse que nous appliquerons souvent dans ce livre à des cas analogues. Cette méthode exploite les *propriétés de la fonction logarithme.*

1. DÉTERMINATION APPROCHÉE DE LA VALEUR LA PLUS PROBABLE DE *n*

Considérons le logarithme de la probabilité $P(N, n)$:

$$\ln P(N, n) = \ln N! - \ln n! - \ln(N-n)! + n \ln P_+ + (N-n) \ln P_- \,. \tag{C.8}$$

Le premier avantage de cette opération est d'ordre pratique : la factorielle d'un grand nombre (songeons au nombre d'Avogadro!) est fantastiquement grande; en prendre le

---

2. On trouve facilement, à partir de (C.1), le rapport

$$\frac{P(N, n)}{P(N, n-1)} = \frac{N-n+1}{n} \frac{P_+}{P_-} \,. \tag{1}$$

Les valeurs de n qui nous intéressent sont telles que

$$n = \bar{n} + \nu \,, \tag{2}$$

avec $|\nu| \ll N$

(avec la restriction (C.6), $|\nu|$ est au plus de l'ordre de quelques $\sqrt{N}$). Compte tenu de la valeur de $\bar{n}$, on obtient finalement

$$\frac{P(N, n)}{P(N, n-1)} = \frac{\bar{n}(1-P_+)-(\nu-1)P_+}{(\bar{n}+\nu)P_-} = \frac{1-(\nu-1)/NP_-}{1+\nu/NP_+} = 1 + \mathcal{O}\left(\frac{\nu}{N}\right) \,, \tag{4}$$

de sorte que la condition (C.7) est satisfaite.

logarithme permet de manipuler des nombres d'un ordre de grandeur plus accessible au raisonnement et à l'intuition[3].

Pour N *fixé,* la valeur la plus probable n_m de n est celle qui rend $P(N, n)$ maximum. Mais comme le logarithme est une fonction toujours croissante, on peut trouver n_m en écrivant que $\ln P(N, n)$ est maximum. Pour dériver l'expression (C.8), il suffit de savoir que[4]

$$\frac{\mathrm{d}}{\mathrm{d}n}[\ln n!] \simeq \ln n \,. \tag{C.9}$$

On trouve ainsi

$$\frac{\partial}{\partial n} \ln P(N, n) \simeq \ln \frac{N-n}{n} + \ln \frac{P_+}{P_-} \,. \tag{C.10}$$

La valeur la plus probable n_m est donc donnée approximativement par

$$\ln \frac{n_m}{N-n_m} \simeq \ln \frac{P_+}{1-P_+} \,, \tag{C.11}$$

c'est-à-dire

$$n_m \simeq NP_+ \,. \tag{C.12}$$

Nous verrons plus loin (formule (C.16.b)) que n_m rend bien $\ln P(N, n)$ maximal, et pas seulement extrémal.

---

3. Pour n grand, $n!$ est bien approximé par la *formule de Stirling :*

$$n! \simeq n^n \mathrm{e}^{-n} \sqrt{2\pi n} \tag{1}$$

(en pratique, l'erreur commise lorsqu'on utilise cette formule est inférieure à 2 % dès que n dépasse 5). Donc

$$\ln n! \simeq n \ln n - n + \frac{1}{2} \ln (2\pi n) \tag{2}$$

est en gros de l'ordre de n. D'ailleurs, le dernier terme de la formule (2) devient très rapidement négligeable devant les précédents quand n croît, de sorte que

$$\ln n! \simeq n \ln n - n \quad \text{pour} \quad n \geqslant 50 \tag{3}$$

(pour $n = 50$, l'erreur commise est inférieure à 1,5 %; elle n'est plus que de 6×10^{-4} pour $n = 1\,000$; pour $n = 10^{23}$, il faut comparer $\ln(2\pi n) \simeq 55$ à $\ln n! \simeq 52 \times 10^{23}$). Soulignons que le facteur $\sqrt{2\pi n}$ ne peut pas être omis dans la formule (1), même lorsque son logarithme peut l'être dans (2) (voir note 6).

4. Par définition de la dérivée,

$$\frac{\mathrm{d}}{\mathrm{d}n}[\ln n!] = \lim_{\delta n \to 0} \frac{\ln(n+\delta n)! - \ln n!}{\delta n} \,. \tag{1}$$

Lorsque n peut être traité comme une variable continue, $\delta n = 1$ est suffisamment petit pour donner la valeur de la dérivée, de sorte que

$$\frac{\mathrm{d}}{\mathrm{d}n} \ln(n!) \simeq \ln[(n+1)!] - \ln(n!) \simeq \ln n. \tag{2}$$

Noter d'ailleurs que ce résultat découle également de la formule (3) de la note 3.

Si l'on compare avec la formule (C.2), on voit que

$$n_m \simeq \bar{n} : \tag{C.13}$$

pour les grands nombres, valeur la plus probable et valeur moyenne de n coïncident[5].

2. DÉVELOPPEMENT DE ln *P* AUTOUR DE LA VALEUR LA PLUS PROBABLE

Développons maintenant ln $P(N, n)$ en série de Taylor au voisinage de $n = n_m$. Ici se situe l'avantage crucial du logarithme par rapport à la fonction elle-même : les factorielles de grands nombres comme ceux que nous considérons ici sont non seulement très grandes mais elles varient aussi très vite; *leur logarithme varie beaucoup plus lentement, et son développement permet d'obtenir une bonne approximation dans un domaine considérablement plus large* de valeurs de la variable[6].

---

5. Il est facile de calculer n_m exactement. En effet, le rapport

$$\frac{P(N, n)}{P(N, n-1)} = \frac{(N-n+1)\,P_+}{nP_-} \tag{1}$$

est supérieur ou égal à 1 pour

$$n \leqslant \bar{n} + P_+ \,. \tag{2}$$

La valeur la plus probable n_m est donc égale à *l'entier immédiatement inférieur à* $\bar{n} + P_+$ (si $\bar{n} + P_+$ est un entier, $P(N, n)$ garde la même valeur pour $n_m + 1 = \bar{n} + P_+$ que pour n_m, et commence à décroître seulement pour $n = n_m + 2$). On constate donc que la différence entre la valeur exacte de n_m et la valeur moyenne $\bar{n}$ est toujours inférieure à 1.

6. Les grandeurs de la mécanique statistique ont à ce sujet un *comportement bien particulier*. Notons $G(x)$ une telle grandeur (positive) et $g(x)$ son logarithme :

$$g(x) = \ln G(x) \quad ; \quad G(x) = \exp g(x) \tag{1}$$

On constate alors, de façon pratiquement générale, que les dérivées de $g(x)$ ont des *ordres de grandeur spécifiques* et remarquables :

$$g(x) \sim N \quad ; \quad g'(x) \sim 1 \quad ; \quad g''(x) \sim \frac{1}{N} \quad ; \quad g'''(x) \sim \frac{1}{N^2}, \ldots, \tag{2}$$

où N est un nombre très grand (de l'ordre du nombre d'Avogadro). Supposons d'abord que nous cherchions à développer les fonctions $g(x)$ et $G(x)$ autour d'une valeur x_0 de la variable pour laquelle $g'(x_0) \neq 0$. Il existe alors au moins un nombre réel θ, compris entre 0 et 1, tel que

$$g(x) = g(x_0) + (x - x_0) g'(x_0) + \frac{1}{2}(x - x_0)^2 g''(x_0 + \theta(x - x_0)). \tag{3}$$

D'après les évaluations (2), le troisième terme du membre de droite sera effectivement négligeable devant les deux premiers pour

$$|x - x_0| \ll N. \tag{4}$$

Mais on souhaite obtenir une approximation de $G(x)$, c'est-à-dire de l'exponentielle de $g(x)$. Or, pour que $\exp(a + \varepsilon) \simeq \exp a$, il n'est pas suffisant que ε soit petit devant a, il faut aussi que $\varepsilon \ll 1$. Si on l'applique à la formule (3), cette dernière condition s'avère *plus restrictive que la précédente* (4), puisqu'elle exige que

$$\frac{(x - x_0)^2}{N} \ll 1. \tag{5}$$

En résumé, nous aboutissons à la conclusion :

$$G(x) \simeq G(x_0) \exp[(x - x_0)\, g'(x_0)] \quad \text{pour} \quad |x - x_0| \ll \sqrt{N}. \tag{6}$$

(En poussant plus loin le développement de $g(x)$, on obtiendrait non seulement une meilleure approximation pour $G(x)$, mais aussi un domaine de validité plus vaste).

Que nous donnerait, en comparaison, un développement direct de $G(x)$? Comme

$$G'(x) = g'(x) \exp g(x), \quad G''(x) = \{[g'(x)]^2 + g''(x)\} \exp g(x), \ldots, \tag{7}$$

toutes les dérivées de $G(x)$ *sont du même ordre de grandeur* (toujours selon les estimations de (2)). Il est facile de voir que nous obtiendrions ici

$$G(x) = G(x_0)[1 + (x - x_0)\, g'(x_0)] \quad \text{pour} \quad |x - x_0| \ll 1 \,; \tag{8}$$

Le développement que nous voulons utiliser s'écrit

$$\ln P(N, n) = \ln P(N, n_m) + A_1(n - n_m) + \frac{1}{2} A_2(n - n_m)^2 + \frac{1}{3!} A_3(n - n_m)^3 + \dots , \qquad (C.14)$$

avec $$A_p = \left[\frac{\partial^p}{\partial n^p} \ln P(N, n)\right]_{n = n_m} . \qquad (C.15)$$

Les dérivées successives de $\ln P(N, n)$ se calculent aisément à partir de (C.10); compte tenu de (C.12), on trouve

$$A_1 = 0 , \qquad (C.16.a)$$

$$A_2 = -\frac{1}{NP_+P_-} , \qquad (C.16.b)$$

$$A_3 = \frac{1}{N^2} \frac{P_- - P_+}{P_+^2 P_-^2} , \dots . \qquad (C.16.c)$$

Donc, dans le développement (C.14), *le terme du premier ordre est nul,* car n_m est la position d'un extremum de la fonction. *Le terme du second ordre est négatif* ($A_2 < 0$), de sorte que cet extremum est bien un *maximum.*

Pour les valeurs de n vérifiant la condition (C.6), *les termes d'ordre supérieur à 2 sont négligeables.* En effet, si

$$|n - n_m| \lesssim a\ \Delta n = a\sqrt{NP_+P_-} , \qquad (C.17)$$

---

le domaine de validité de cette approximation est donc considérablement réduit par rapport à celui de la formule (6) (il ne servirait à rien, de ce point de vue, d'inclure des termes d'ordre supérieur dans le développement de $G(x)$).

Souvent (comme dans ce complément I-C), on souhaite approximer $G(x)$ autour d'un maximum, c'est-à-dire au voisinage d'une valeur x_0 telle que

$$g'(x_0) = 0 \ ; \ g''(x_0) < 0. \qquad (9)$$

Le développement de $g(x)$ autour d'un tel point s'écrit

$$g(x) = g(x_0) + \frac{1}{2}(x - x_0)^2 g''(x_0) + \frac{1}{6}(x - x_0)^3 g'''(x_0 + \theta(x - x_0)), \qquad (10)$$

et on en déduit, à partir des comportements (2),

$$g(x) \simeq g(x_0) + \frac{1}{2}(x - x_0)^2 g''(x_0) \quad \text{pour } |x - x_0| << N. \qquad (11)$$

Ici aussi, c'est l'opération d'exponentiation qui régit le domaine de validité de l'approximation qui en résulte pour $G(x)$. Il vient l'*approximation gaussienne*

$$G(x) \simeq G(x_0) \exp[-(x - x_0)^2/2\sigma^2] \quad \text{pour } |x - x_0| << N^{2/3}, \qquad (12)$$

où l'on a posé

$$\sigma^2 = -\frac{1}{g''(x_0)}. \qquad (13)$$

On déduit du comportement (2) de $g''(x_0)$ que

$$\sigma \sim \sqrt{N}. \qquad (14)$$

Comme, pour N grand, $N^{2/3} >> \sqrt{N}$, cette approximation gaussienne est *excellente* : elle reste valable même si $|x - x_0|$ atteint des valeurs plusieurs fois supérieures à σ, c'est-à-dire dans des régions où $G(x)$ a considérablement décrû ; autrement dit, c'est seulement dans « queues » éloignées du maximum que $G(x)$ et son approximation gaussienne peuvent différer appréciablement, mais elles y sont de toute façon toutes deux très petites par rapport à $G(x_0)$.

on peut majorer le terme du troisième ordre par

$$A_3|n-n_m|^3 \leqslant a' \frac{1}{N^2} N^{3/2} = \frac{a'}{\sqrt{N}} \ll 1 , \tag{C.18}$$

a' étant de l'ordre de quelques unités. Ce même argument semi-quantitatif indique que *le terme du second degré est de l'ordre de* 1, et, si l'on calcule les valeurs des coefficients A_p successifs, que les termes d'ordre supérieur p sont d'autant plus petits que p est plus grand.

3. RÉSULTAT : L'APPROXIMATION GAUSSIENNE DE LA DISTRIBUTION BINOMIALE

Le développement (C.14) permet d'approximer $\ln P(N, n)$ par

$$\ln P(N, n) \simeq \ln P(N, n_m) - \frac{1}{2\sigma^2}(n-n_m)^2 , \tag{C.19}$$

avec, d'après (C.12) et (C.13),

$$n_m \simeq \bar{n} = NP_+ , \tag{C.20}$$

et, d'après (C.16.b) et (C.3),

$$\sigma^2 = NP_+P_- = (\Delta n)^2 . \tag{C.21}$$

Ceci fournit (cf. note 6) une expression approchée pour la probabilité $P(N, n)$ elle-même :

$$P(N, n) \simeq C_N \exp\left[-\frac{(n-\bar{n})^2}{2\sigma^2}\right] . \tag{C.22}$$

La constante

$$C_N = P(N, n_m) \tag{C.23}$$

peut être déterminée par normalisation[7], ce qui donne finalement

$$P(N, n) \simeq \frac{1}{\sqrt{2\pi(\Delta n)^2}} \exp\left[-(n-\bar{n})^2/2(\Delta n)^2\right] . \tag{C.24}$$

Dans le domaine des grands nombres, la distribution binomiale se confond pratiquement avec une distribution gaussienne dont les paramètres de position et de largeur coïncident avec la valeur moyenne $\bar{n}$ et l'écart quadratique moyen Δn de la distribution binomiale.

---

7. La variable n, que nous traitons comme continue, est comprise entre 0 et N. Mais la fonction gaussienne à intégrer devient totalement négligeable dès que $|n-\bar{n}|$ dépasse quelques Δn. Son intégrale de 0 à N ne diffère donc pratiquement pas de son intégrale de $-\infty$ à $+\infty$, de sorte que la condition de normalisation s'écrit

$$\int_0^N C_N e^{-(n-\bar{n})^2/2\sigma^2} \, dn \simeq C_N \int_{-\infty}^{+\infty} e^{-(n-\bar{n})^2/2\sigma^2} \, dn = 1 .$$

Elle donne donc

$$C_N = \frac{1}{\sqrt{2\pi\sigma^2}} .$$

III. DISCUSSION DE L'APPROXIMATION

Pour établir l'expression approchée (C.24) de la distribution binomiale, nous avons posé *a priori* la condition (C.6) limitant les valeurs de n acceptables. Examinons de plus près le domaine de validité de l'approximation gaussienne.

Cette approximation, obtenue en tronquant le développement (C.14) de $\ln P(N, n)$ puis en prenant l'exponentielle de ce développement tronqué, sera valable *si les termes négligés sont petits devant* 1 *et devant le terme d'ordre* 2 *qui a été conservé* (cf. note 6).

Considérons le terme d'ordre 3 (cf. (C.16.c)). Il est petit devant 1 (en valeur absolue) si

$$|n - \bar{n}| \ll (NP_+P_-)^{2/3} \; ; \tag{C.25}$$

il est négligeable devant le terme d'ordre 2 si

$$|n - \bar{n}| \ll NP_+P_- \; . \tag{C.26}$$

L'approximation ne présenterait que peu d'intérêt si elle était limitée aux valeurs de n différant à peine de la valeur moyenne $\bar{n}$; c'est pourquoi nous nous intéressons seulement aux cas où

$$NP_+P_- \gg 1 \; . \tag{C.27}$$

Mais alors la relation (C.25) entraîne (C.26); c'est donc la première de ces conditions qui fixe l'intervalle de valeurs de n dans lequel la gaussienne approxime correctement la distribution binomiale[8]. On montre en effet facilement que les termes d'ordre supérieur à 3 sont eux aussi négligeables, *a fortiori,* dans ce domaine.

TABLEAU I

Comparaison numérique entre $(NP_+P_-)^{2/3}$ et Δn, c'est-à-dire entre les conditions (C.25) et (C.6) limitant le domaine de validité en n de l'approximation gaussienne.

NP_+P_-	$(NP_+P_-)^{2/3}$	$(NP_+P_-)^{1/2} = \Delta n$
50	13,6	7,07
100	21,5	10
1 000	100	31,6
10 000	464	100
10^{10}	$4{,}64 \times 10^{6}$	10^{5}
10^{20}	$2{,}15 \times 10^{13}$	10^{10}

8. Dans le cas particulier où

$$P_+ = P_- = \frac{1}{2} \; , \tag{1}$$

le terme d'ordre 3 est identiquement nul. Il faut alors considérer le terme d'ordre 4, ce qui remplace la condition (C.25) par

$$|n - \bar{n}| \ll (NP_+P_-)^{3/4} \; . \tag{2}$$

Il est compréhensible que l'approximation gaussienne soit encore meilleure dans ce cas particulier : la distribution binomiale est alors symétrique par rapport à $\bar{n}$, comme la fonction gaussienne par laquelle on la représente.

La condition (C.25) que nous obtenons ainsi n'est pas celle que nous avions posée initialement en (C.6). Le tableau I permet de comparer numériquement ces deux conditions. On constate qu'en réalité $(NP_+P_-)^{2/3}$ est du même ordre de grandeur que $\Delta n = (NP_+P_-)^{1/2}$ tant que NP_+P_- reste inférieur à 10 000; dans ce domaine, le nombre a figurant dans (C.6) ne peut pas dépasser 2. Ce n'est que pour des valeurs beaucoup plus grandes de NP_+P_- (au moins de l'ordre de 10^{10}) que $(NP_+P_-)^{2/3}$ devient nettement supérieur à Δn de sorte que les limites fixées par la condition (C.25) débordent de façon significative du domaine défini par (C.6). Bien entendu, les bornes du domaine de validité de l'approximation gaussienne dépendent de la précision que l'on exige de cette approximation.

TABLEAU II

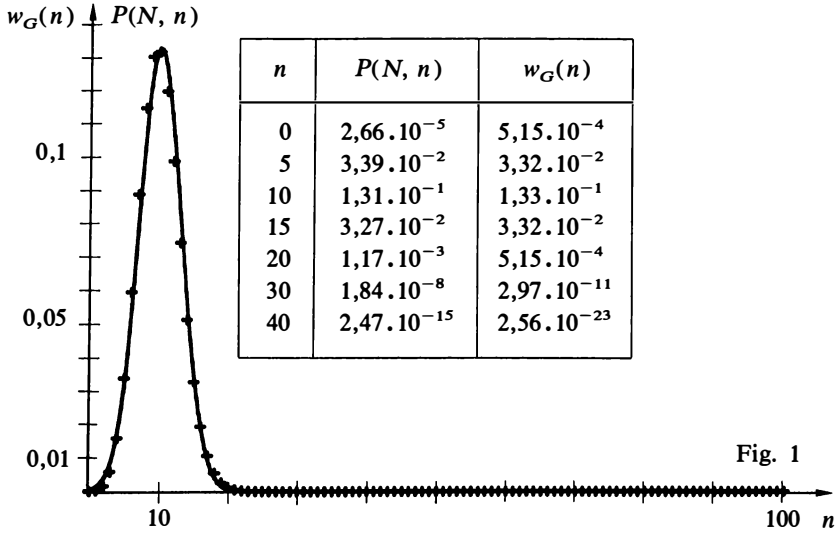
Comparaison numérique entre la distribution binomiale $P(N, n)$ et son approximation gaussienne $w_G(n)$ pour $N = 100\,000$ et $P_+ = 0{,}1$ ($\bar{n} = 10\,000$ et $\Delta n \simeq 100$).

n	$P(N, n)$	$w_G(n)$
9 500 $(\bar{n} - 5\Delta n)$	$3{,}237 \times 10^{-9}$	$3{,}908 \times 10^{-9}$
9 600	$5{,}304 \times 10^{-7}$	$5{,}799 \times 10^{-7}$
9 700	$2{,}745 \times 10^{-5}$	$2{,}833 \times 10^{-5}$
9 800	$4{,}537 \times 10^{-4}$	$4{,}557 \times 10^{-4}$
9 900 $(\bar{n} - \Delta n)$	$2{,}419 \times 10^{-3}$	$2{,}413 \times 10^{-3}$
9 920	$2{,}955 \times 10^{-3}$	$2{,}947 \times 10^{-3}$
9 940	$3{,}451 \times 10^{-3}$	$3{,}443 \times 10^{-3}$
9 960	$3{,}854 \times 10^{-3}$	$3{,}847 \times 10^{-3}$
9 980	$4{,}116 \times 10^{-3}$	$4{,}113 \times 10^{-3}$
10 000 $(\bar{n})$	$4{,}205 \times 10^{-3}$	$4{,}205 \times 10^{-3}$
10 020	$4{,}109 \times 10^{-3}$	$4{,}113 \times 10^{-3}$
10 040	$3{,}841 \times 10^{-3}$	$3{,}847 \times 10^{-3}$
10 060	$3{,}435 \times 10^{-3}$	$3{,}443 \times 10^{-3}$
10 080	$2{,}939 \times 10^{-3}$	$2{,}947 \times 10^{-3}$
10 100 $(\bar{n} + \Delta n)$	$2{,}406 \times 10^{-3}$	$2{,}413 \times 10^{-3}$
10 200	$4{,}576 \times 10^{-4}$	$4{,}557 \times 10^{-4}$
10 300	$2{,}922 \times 10^{-5}$	$2{,}833 \times 10^{-5}$
10 400	$6{,}319 \times 10^{-7}$	$5{,}799 \times 10^{-7}$
10 500 $(\bar{n} + 5\Delta n)$	$4{,}674 \times 10^{-9}$	$3{,}908 \times 10^{-9}$

Il est important également de noter que la probabilité $P(N, n)$ devient très faible (en valeur absolue, et en valeur relative comparée à $P(N, \bar{n})$) dès que n sort du domaine défini par (C.6) ou (C.25) : le tableau II montre par exemple que, pour $N = 100\,000$ et $P_+ = 0{,}1$, la probabilité binomiale tombe à 7 pour mille de sa valeur maximale lorsque $|n - \bar{n}|$ atteint 300, c'est-à-dire environ 3 fois Δn.

On retiendra donc que, *en pratique, l'approximation gaussienne est valable pour NP_+P_- grand devant* 1, *et dans un intervalle de quelques Δn autour de $\bar{n}$, c'est-à-dire pour les valeurs de n ayant une probabilité appréciable*[9].

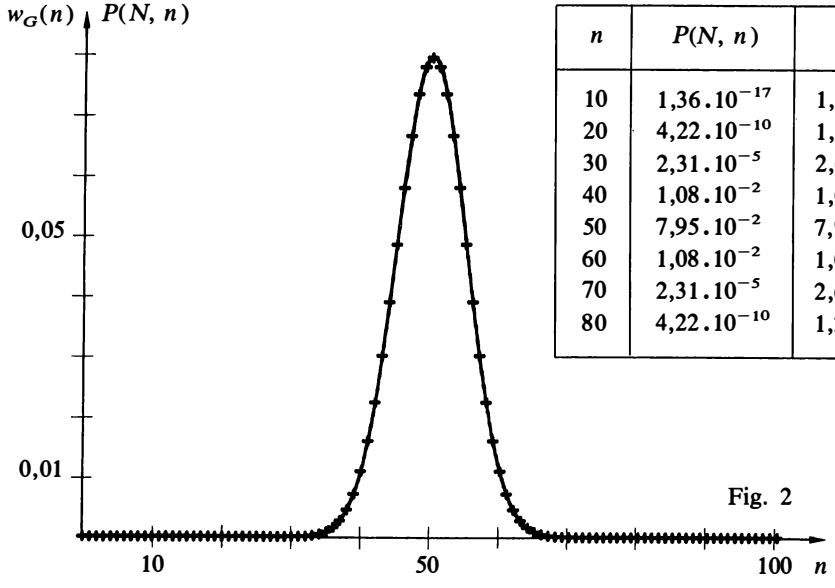
9. On peut être amené à s'intéresser à des valeurs de n très peu probables, c'est-à-dire très différentes de $\bar{n}$ (voisines de 0 ou de N, par exemple.). L'approximation gaussienne n'est alors d'aucun secours; il faut revenir à l'expression initiale de la distribution binomiale.



n	$P(N, n)$	$w_G(n)$
0	$2{,}66.10^{-5}$	$5{,}15.10^{-4}$
5	$3{,}39.10^{-2}$	$3{,}32.10^{-2}$
10	$1{,}31.10^{-1}$	$1{,}33.10^{-1}$
15	$3{,}27.10^{-2}$	$3{,}32.10^{-2}$
20	$1{,}17.10^{-3}$	$5{,}15.10^{-4}$
30	$1{,}84.10^{-8}$	$2{,}97.10^{-11}$
40	$2{,}47.10^{-15}$	$2{,}56.10^{-23}$

Fig. 1

FIGURES 1 et 2
Comparaison entre la distribution binomiale $P(N, n)$ (points étoilés) et son approximation gaussienne (courbe continue) pour $N = 100$. Les encadrés précisent quelques valeurs numériques typiques. La figure 1 correspond à $P_+ = 0{,}1$, soit $\bar{n} = 10$ et $\Delta n = 3$ (NP_+P_- y vaut seulement 9); pour la figure 2, $P_+ = 1/2$, $\bar{n} = 50$ et $\Delta n = 5$, NP_+P_- valant seulement 25.



n	$P(N, n)$	$w_G(n)$
10	$1{,}36.10^{-17}$	$1{,}01.10^{-15}$
20	$4{,}22.10^{-10}$	$1{,}21.10^{-9}$
30	$2{,}31.10^{-5}$	$2{,}68.10^{-5}$
40	$1{,}08.10^{-2}$	$1{,}08.10^{-2}$
50	$7{,}95.10^{-2}$	$7{,}98.10^{-2}$
60	$1{,}08.10^{-2}$	$1{,}08.10^{-2}$
70	$2{,}31.10^{-5}$	$2{,}68.10^{-5}$
80	$4{,}22.10^{-10}$	$1{,}21.10^{-9}$

Fig. 2

Pour préciser concrètement la validité de l'approximation gaussienne, on peut la comparer numériquement, dans quelques cas typiques, aux valeurs exactes de la probabilité binomiale. Dans le tableau II, N est relativement grand mais sans commune mesure, toutefois, avec le nombre d'Avogadro; on constate pourtant que les valeurs exacte et approchée restent très voisines dans un intervalle de plusieurs Δn, ce qui correspond à une variation de ces probabilités sur plusieurs ordres de grandeur : ainsi, la différence entre valeur exacte et valeur approchée ne dépasse pas 20 % dans des régions où la probabilité a décru d'un facteur 10^6 par rapport à sa valeur maximale. Nous avons choisi pour les figures 1 et 2 une valeur de N beaucoup moins grande ($N = 100$), pour que des écarts puissent être perçus entre $P(N, n)$ et $w_G(n)$: NP_+P_- atteint à peine 9 pour la figure 1, et 25 pour la figure 2. L'approximation gaussienne est visiblement encore raisonnable dans ces deux cas; dans le second, même, les différences n'apparaissent pas sur une figure de dimension normale. Cette comparaison numérique et graphique montre la *précision remarquable* de l'approximation gaussienne pour la distribution binomiale.

COMPLÉMENT I.D

Marche au hasard et mouvement brownien

I. Marche au hasard à une dimension
   1. Pas de longueur constante
   2. Pas de longueur variable

II. Représentation graphique de la marche au hasard à une dimension
   1. La représentation graphique
   2. Théorème de réflexion
   3. Application

III. Marche au hasard à trois dimensions et mouvement brownien
   1. Marche au hasard isotrope à trois dimensions
   2. Un modèle simple pour le mouvement brownien
   3. Vérification expérimentale

Le mouvement brownien a été découvert en 1827 par le botaniste anglais R. Brown. Observant au microscope des grains de pollen en suspension dans l'eau, Brown constata que les grains étaient animés d'un mouvement désordonné incessant : « ils vont et viennent en tournoyant, montent, descendent, remontent encore sans tendre aucunement vers le repos » (Jean Perrin). L'hypothèse qu'il s'agissait là « d'animalcules rudimentaires » fut rapidement abandonnée et le phénomène resta inexpliqué.

En 1905 Einstein[1] émit l'idée que ce mouvement était provoqué par les chocs des molécules du liquide sur les grains en suspension. Prenant pour hypothèse que les changements de direction dus à ces chocs sont aléatoires, il proposa une théorie du mouvement brownien dont les conséquences expérimentales furent vérifiées en 1908 par Jean Perrin. L'existence du mouvement brownien constitua dès lors un argument décisif pour les physiciens qui tentaient de convaincre leurs collègues de la nature atomique de la matière. Jean Perrin écrit : « ... il devient donc difficile de nier la réalité objective des molécules... Le mouvement moléculaire nous est rendu visible. Le mouvement brownien en est l'image fidèle, ou mieux, il est déjà un mouvement moléculaire ».

La théorie proposée par Einstein consiste à ramener l'étude du mouvement brownien à ce que l'on appelle aujourd'hui de façon imagée un problème de « marche au hasard ». Ce problème, dont il existe bien d'autres exemples, a donné lieu à d'importants développements mathématiques : on rencontre des processus de marche au hasard à une, deux, trois dimensions; le mouvement peut avoir lieu dans un espace infini ou borné; les « pas » peuvent être de longueur constante ou variable, toutes les directions peuvent être également probables ou certaines d'entre elles privilégiées; les situations sont donc plus ou moins complexes. Nous nous limiterons ici à quelques exemples[2].

---

1. M. Smoluchovski a publié indépendamment en 1906 une théorie similaire quoique moins détaillée.

2. Les techniques et résultats de la marche au hasard sont utilisés dans divers domaines de la physique, lorsque l'on est amené à considérer une grandeur physique comme variable aléatoire prenant des valeurs successives *indépendantes* les unes des autres, au moins en première approximation : diffusion de particules dans un milieu constitué de particules différentes, sédimentation de poussières dans un fluide, ou même évaluation du champ de gravitation produit par les étoiles (c'est ici la répartition des étoiles qui est traitée comme aléatoire — voir référence 1). Mais la marche au hasard peut également intervenir dans des problèmes où aucune donnée n'est *a priori* aléatoire, comme par exemple en théorie du potentiel (référence 2).

I. MARCHE AU HASARD A UNE DIMENSION

1. PAS DE LONGUEUR CONSTANTE

Nous considérons tout d'abord le cas le plus simple. Un marcheur se déplace, avec des pas de longueur constante ℓ, le long d'une rue infinie que nous schématisons par une droite orientée. Un pas peut être dirigé dans le sens positif ou négatif avec des probabilités respectives P_+ et P_-. Les pas sont statistiquement indépendants : la direction d'un pas est indépendante de celle du précédent; on dit que le marcheur a «perdu la mémoire» de son mouvement antérieur, ou encore que le processus est markovien. Quelle est la probabilité $\mathcal{P}(N, s)$ pour que, après avoir effectué N pas, le marcheur se trouve au point A d'abscisse $s\ell$ (s entier, $|s| \leqslant N$) par rapport au point de départ D?

Ce problème est identique à celui de pile ou face : à chaque pas le choix de la direction positive ou négative pourrait être déterminé par le résultat (pile ou face) du lancer d'une pièce de monnaie. La distribution $\mathcal{P}(N, s)$ sera donc binomiale.

Appelons *trajet* l'ensemble ordonné des N positions successives du marcheur. Tout trajet comprenant n_+ pas dans la direction positive (et $n_- = (N - n_+)$ dans la direction négative) amène le marcheur au point d'abscisse

$$n_+\ell - (N - n_+)\ell = (2n_+ - N)\ell \,. \tag{D.1}$$

Il existe donc de nombreux trajets différents qui conduisent en N pas de D au point d'abscisse $s\ell$: ce sont tous ceux qui comportent un nombre de pas n_+ dans la direction positive tel que

$$(2n_+ - N)\ell = s\ell \,, \tag{D.2.a}$$

c'est-à-dire[3]

$$n_+ = \frac{N+s}{2} \,. \tag{D.2.b}$$

Ces trajets diffèrent entre eux par la place des n_+ pas positifs dans la séquence des N pas; leur nombre est égal à $C_N^{n_+}$. Ils sont tous également probables, la probabilité de chacun d'eux étant

$$(P_+)^{n_+}.(P_-)^{n_-} = (P_+)^{(N+s)/2}.(P_-)^{(N-s)/2} \,. \tag{D.3}$$

On a donc

$$\begin{aligned}\mathcal{P}(N, s) &= C_N^{(N+s)/2}.(P_+)^{(N+s)/2}.(P_-)^{(N-s)/2} \\ &= \frac{N!}{\left(\frac{N+s}{2}\right)!\left(\frac{N-s}{2}\right)!}(P_+)^{(N+s)/2}.(P_-)^{(N-s)/2} \,.\end{aligned} \tag{D.4}$$

En utilisant les résultats (44) et (47) de l'appendice I ainsi que la relation (D.2.b) entre s et n_+, on trouve facilement la valeur moyenne et l'écart quadratique moyen de s :

$$\bar{s} = N(P_+ - P_-) \,, \tag{D.5}$$

$$\Delta s = \left(\overline{s^2} - \bar{s}^2\right)^{1/2} = \sqrt{4NP_+P_-} \,. \tag{D.6}$$

Imaginons qu'un grand nombre de marcheurs partent ensemble de la même origine et prenons le cas particulier où $P_+ = P_- = 1/2$. Après N pas les marcheurs se distribuent le long de la rue de façon symétrique par rapport à leur point de départ, ce qui explique que la valeur moyenne de leur abscisse soit nulle. Mais chacun des marcheurs tend à s'éloigner de l'origine : l'écart quadratique moyen n'est pas nul; il croît comme la racine carrée du nombre N de pas effectués.

---

3. L'équation (D.2.b) n'a de solution que si N et s ont même parité (puisque le nombre n_+ doit être entier). Dans le cas contraire, $\mathcal{P}(N, s) = 0$.

2. PAS DE LONGUEUR VARIABLE

Considérons maintenant le cas où les pas successifs, toujours indépendants, n'ont pas une longueur constante : soit $w(x)\,dx$ la probabilité pour qu'un pas donné accroisse l'abscisse du marcheur d'une quantité comprise entre x et $x+dx$. Quelle est la densité de probabilité $W_N(X)$ pour que, après N pas, le marcheur se trouve à l'abscisse X?

Si N est assez grand, on peut utiliser le théorème de la limite centrale[4] :

$$W_N(X)=\frac{1}{\sqrt{2\pi N\sigma^2}}\exp\left[-\frac{(X-N\bar{x})^2}{2N\sigma^2}\right], \tag{D.7}$$

où $\bar{x}$ et σ sont respectivement la valeur moyenne et l'écart quadratique moyen de la variable aléatoire x.

II. REPRÉSENTATION GRAPHIQUE DE LA MARCHE AU HASARD A UNE DIMENSION

Dans le cas de *pas de longueur constante,* une représentation graphique facilite beaucoup l'analyse et la résolution des problèmes de marche au hasard. Nous allons maintenant décrire cette représentation, l'utiliser pour établir un théorème important et présenter une application.

1. LA REPRÉSENTATION GRAPHIQUE

Représentons le trajet du marcheur par un diagramme donnant sa position en fonction du temps : nous portons en abscisse le nombre N de pas effectués depuis un instant initial et en ordonnée l'entier s qui repère le marcheur par rapport à son point de départ. Le trajet parcouru est alors représenté par une ligne brisée qui joint les points associés aux positions successives, et que nous appelons un *«chemin».* Lorsqu'un chemin passe par un point M de coordonnées (N, s), cela signifie que le marcheur s'est trouvé, après N pas, au point d'abscisse $s\ell$ (voir figure 1). A chaque trajet est associé un chemin et un seul et vice versa.

---

4. La distribution de probabilité $w(X)$ de la somme

$$X=\sum_{j=1}^{N} x_j \tag{1}$$

de N variables aléatoires *indépendantes* x_j est, si N est grand devant 1, une gaussienne centrée en

$$\bar{X}=\sum_{j=1}^{N} \bar{x}_j \tag{2}$$

et de largeur ΔX telle que

$$(\Delta X)^2=\sum_{j=1}^{N} (\Delta x_j)^2 \tag{3}$$

(voir exercice I.12).

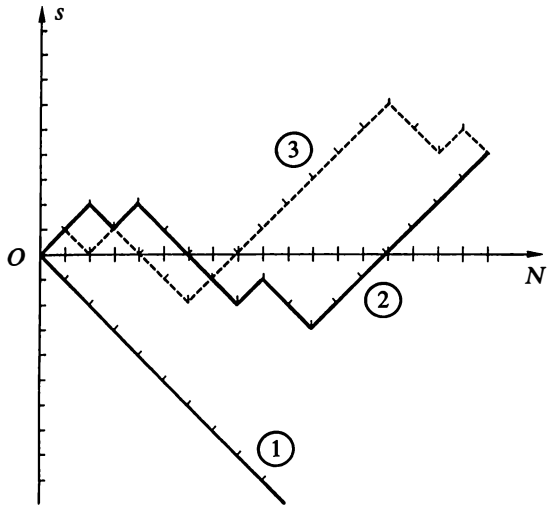


FIGURE 1
Représentation graphique de la marche au hasard : en abscisse, le nombre de pas effectués depuis le départ; en ordonnée, l'abscisse du marcheur (ℓ étant pris pour unité). Chemin 1 : tous les pas sont dans la direction négative. Chemins 2 et 3 : deux trajets différents qui mènent en 18 pas de l'origine D au point situé à l'abscisse 4ℓ.

2. THÉORÈME DE RÉFLEXION

Considérons, sur le graphique représentant une marche au hasard, une droite horizontale Δ et deux points A et B situés du même côté de cette droite (fig. 2). Certains des chemins qui vont de A à B traversent ou touchent Δ, d'autres au contraire ne l'atteignent pas.

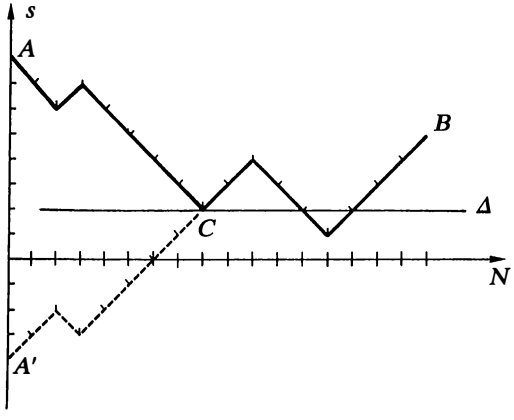


FIGURE 2
Deux chemins associés pour la démonstration du théorème de réflexion.

Nous allons montrer le théorème suivant : *le nombre de chemins qui vont de A à B en coupant ou en touchant la droite Δ est égal au nombre total de chemins qui vont du point A', symétrique de A par rapport à Δ, au point B.*

Soit un chemin qui va de A à B en coupant ou en touchant Δ et soit C son premier contact avec Δ. Nous lui associons un chemin joignant A' à B, construit de la manière suivante : de C à B, le nouveau chemin coïncide avec l'ancien; pour la partie antérieure à C, les deux chemins sont symétriques par rapport à Δ. Inversement, tout chemin qui va de A' à B coupe nécessairement Δ et nous pouvons lui associer, de la même façon, un chemin qui va de A à B en touchant Δ. Cela démontre le résultat énoncé ci-dessus, que l'on appelle le théorème de réflexion[5].

3. APPLICATION

Cherchons la probabilité $\tilde{\mathcal{P}}(N, s)$ pour que, après N pas de longueur constante ℓ dirigés dans le sens positif ou le sens négatif avec des probabilités égales, un marcheur parti de l'origine *passe pour la première fois* au point d'abscisse $s\ell$.

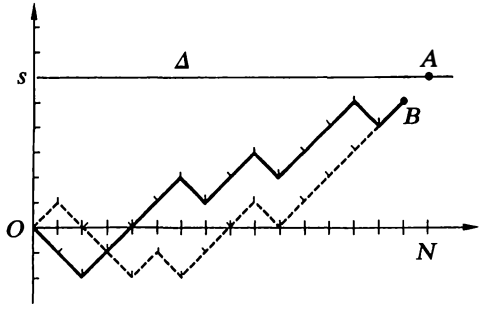


FIGURE 3
Premier passage en s après N pas : tous les chemins favorables passent par B et on peut les prolonger de façon unique pour atteindre A.

Dans le langage de la représentation graphique, nous cherchons le nombre de chemins qui vont du point $O(0, 0)$ au point $A(N, s)$ sans avoir touché auparavant la droite horizontale Δ d'ordonnée s (fig. 3). Tous ces chemins passent forcément par le point $B(N-1, s-1)$. Il y a au total $C_{N-1}^{(N+s-2)/2}$ chemins qui vont de O à B; parmi ceux-ci, le nombre de ceux qui touchent ou coupent Δ est égal, d'après le théorème de réflexion, au nombre de chemins qui vont de $O'(0, 2s)$ à B, soit[6] $C_{N-1}^{(N+s)/2}$; nous obtenons par soustraction le nombre de chemins qui nous intéresse. Comme ils comportent tous N pas, la probabilité de chacun d'eux est[7] $(1/2)^N$, de sorte que la probabilité $\tilde{\mathcal{P}}(N, s)$ cherchée vaut

$$\tilde{\mathcal{P}}(N, s) = \left(\frac{1}{2}\right)^N \left\{C_{N-1}^{(N+s-2)/2} - C_{N-1}^{(N+s)/2}\right\}$$

$$= \left(\frac{1}{2}\right)^N \frac{s}{N} C_N^{(N+s)/2} . \qquad \text{(D.8)}$$

5. Attention! Les *probabilités* d'aller de A à B en touchant Δ et d'aller de A' à B ne sont pas égales car un chemin qui va de A' à B ne comporte pas le même nombre de pas dans la direction positive qu'un chemin qui va de A à B (ces probabilités ne sont égales que dans le cas où $P_+ = P_- = 1/2$).

6. Le nombre de chemins qui vont de $O'(0, 2s)$ à $B(N-1, s-1)$ est en effet égal au nombre des chemins qui vont de $O(0, 0)$ à $B'(N-1, s-1-2s)$ ou encore de $O(0, 0)$ à $B''(N-1, s+1)$ (voir figure 3).

7. Si les probabilités des pas positifs et négatifs ne sont pas égales mais valent P_+ et P_-, chacun de ces chemins a une probabilité $(P_+)^{(N+s)/2}(P_-)^{(N-s)/2}$.

On peut montrer[8] que la somme sur toutes les valeurs de N de la probabilité $\tilde{\mathcal{P}}(N, s)$ vaut 1 :

$$\sum_{N=s}^{\infty} \left(\frac{1}{2}\right)^N \frac{s}{N} C_N^{(N+s)/2} = 1 \quad \forall s \,, \tag{D.9}$$

c'est-à-dire que le marcheur finira toujours par passer en un point de la droite déterminé à l'avance, quel que soit ce point[9]. En revanche la valeur moyenne du nombre de pas nécessaires pour arriver au point considéré n'est pas déterminée (la série correspondante ne converge pas).

III. MARCHE AU HASARD A TROIS DIMENSIONS ET MOUVEMENT BROWNIEN

1. MARCHE AU HASARD ISOTROPE A TROIS DIMENSIONS

Un mobile se déplace dans l'espace à trois dimensions par une succession de « pas »; ceux-ci sont statistiquement indépendants et pour chacun d'eux toutes les directions sont également probables. La longueur ℓ des pas est variable et caractérisée par une densité de probabilité $w(\ell)$ indépendante de la direction du pas. Nous voulons étudier la position du mobile lorsqu'il a effectué N pas; nous supposerons N grand.

La probabilité pour que le premier pas amène le mobile de son point de départ (que nous prenons pour origine) à un point situé à l'intérieur d'un volume élémentaire entourant le point M, de coordonnées sphériques (r, θ, φ), vaut

$$\mathrm{d}P(r, \theta, \varphi) = w(r)\, \mathrm{d}r \frac{\mathrm{d}\Omega}{4\pi} \,, \tag{D.10}$$

où $$\mathrm{d}\Omega = \sin\theta \, \mathrm{d}\theta \, \mathrm{d}\varphi \tag{D.11}$$

est l'angle solide élémentaire autour de la direction (θ, φ). En effet $w(r)\,\mathrm{d}r$ est la probabilité pour que la longueur ℓ du pas soit r à $\mathrm{d}r$ près et $\mathrm{d}\Omega/4\pi$ est la probabilité pour que sa direction soit (θ, φ) à $\mathrm{d}\Omega$ près.

Le pas suivant est bien sûr caractérisé par une distribution analogue, sa longueur et sa direction étant indépendantes des précédentes; mais son point de départ est l'aboutissement du premier pas. Pour résoudre cette difficulté, il suffit de *projeter sur trois axes cartésiens Ox, Oy, Oz*. La position du mobile après N pas est alors repérée par les coordonnées X_N, Y_N, Z_N telles que

$$X_N = \sum_{i=1}^{N} x_i \,, \tag{D.12}$$

où x_i est la projection sur Ox du i-ième pas (des formules analogues donnant Y_N et Z_N).

8. Voir référence 4, p. 350.

9. Dans le cas où les probabilités d'aller dans les directions positive et négative ne sont pas égales, on trouve que la probabilité de passer en un point d'abscisse positive vaut 1 tant que $P_+ > P_-$ si le nombre N de pas n'est pas limité. On trouve un résultat analogue pour la marche au hasard à 2 dimensions : dans le cas où toutes les directions sont également probables, quel que soit le point du plan, le marcheur a une probabilité égale à 1 d'y arriver. Autrement dit « tous les chemins mènent à Rome ». Cela n'est pas vrai à 3 dimensions.

On déduit facilement de (D.12) la valeur moyenne et l'écart quadratique moyen de la position après N pas; il suffit pour cela de connaître la valeur moyenne et l'écart quadratique moyen de la projection d'un pas selon chacun des trois axes. Le problème étant isotrope, on a nécessairement

$$\bar{x}=\bar{y}=\bar{z}=0\ , \tag{D.13}$$

$$\overline{x^2}=\overline{y^2}=\overline{z^2}=\frac{1}{3}\,\overline{\ell^2}\ , \tag{D.14}$$

avec $\overline{\ell^2}\,\widehat{=}\int_0^{+\infty}\ell^2 w(\ell)\,\mathrm{d}\ell$. (D.15)

On obtient ainsi[10]

$$\bar{X}_N=\bar{Y}_N=\bar{Z}_N=0\ , \tag{D.16}$$

$$(\Delta X_N)^2=(\Delta Y_N)^2=(\Delta Z_N)^2=\frac{N}{3}\,\overline{\ell^2}\ . \tag{D.17}$$

Là encore les valeurs moyennes des trois coordonnées sont nulles (puisque toutes les directions sont également probables) et les écarts quadratiques moyens varient comme $\sqrt{N}$; en particulier la valeur moyenne du carré de la distance à l'origine $R_N^2=X_N^2+Y_N^2+Z_N^2$ est tout simplement égale à $N\overline{\ell^2}$.

Si N est suffisamment grand, nous pouvons utiliser le théorème de la limite centrale (note 4) : la distribution de probabilité $W_N(X)$ de l'abscisse X après N pas est une gaussienne :

$$W_N(X)=\sqrt{\frac{3}{2\pi N\overline{\ell^2}}}\exp\left[-\frac{3X^2}{2N\overline{\ell^2}}\right],\quad N\gg 1\ . \tag{D.18}$$

Les distributions suivant les autres axes sont identiques.

---

10. On sait que la valeur moyenne d'une somme est égale à la somme des valeurs moyennes :

$$\bar{X}_N=\sum_{i=1}^{N}\bar{x}_i=N\bar{x}=0\ , \tag{1}$$

ainsi que

$$\bar{Y}_N=\bar{Z}_N=0\ . \tag{2}$$

Le calcul de l'écart quadratique moyen se ramène donc à celui de $\overline{X_N^2}$:

$$(\Delta X_N)^2=\overline{X_N^2}-(\bar{X}_N)^2=\overline{X_N^2}\ . \tag{3}$$

Pour le calculer, écrivons

$$X_N^2=\left(\sum_{i=1}^{N}x_i\right)\left(\sum_{j=1}^{N}x_j\right)=\sum_{i,j=1}^{N}x_ix_j=\sum_{i=1}^{N}x_i^2+\sum_{\substack{i,j=1\\ j\neq i}}^{N}x_ix_j\ . \tag{4}$$

Pour $i\neq j$, les grandeurs x_i et x_j sont statistiquement indépendantes; la valeur moyenne de leur produit est donc égale au produit de leurs valeurs moyennes (cf. appendice I, § I.2.b), qui sont ici nulles. On trouve finalement

$$(\Delta X_N)^2=\overline{X_N^2}=\sum_{i=1}^{N}\overline{x_i^2}=\frac{N}{3}\,\overline{\ell^2}\ . \tag{5}$$

2. UN MODÈLE SIMPLE POUR LE MOUVEMENT BROWNIEN

La marche au hasard peut servir à modéliser le mouvement d'une particule brownienne dans un gaz ou un liquide. On suppose que la particule se déplace librement et n'interagit avec les molécules du gaz ou du liquide que lors de collisions instantanées qui la renvoient au hasard dans une direction quelconque avec une probabilité uniformément répartie dans toutes les directions : sa trajectoire est ainsi une ligne brisée dont chaque segment correspond à la trajectoire entre deux collisions[11].

Au bout d'un temps t grand devant le temps moyen τ qui sépare deux collisions successives, la particule a subi un grand nombre N de chocs, évidemment proportionnel à t ($N = t/\tau$). L'abscisse moyenne de la particule dans une direction fixée est nulle avec un écart quadratique donné par (D.17), que l'on peut écrire sous la forme

$$(\Delta X)^2 = 2Dt\,, \tag{D.19}$$

où D est le *coefficient de diffusion*[12] des particules en suspension dans le liquide ou dans le gaz considéré. Dans notre modèle simple,

$$D = \frac{1}{6}\frac{\overline{\ell^2}}{\tau}\,. \tag{D.20}$$

3. VÉRIFICATION EXPÉRIMENTALE

Pour observer le mouvement brownien, il faut étudier des particules assez grosses pour être visibles au microscope mais assez légères pour que les chocs des molécules leur communiquent un mouvement perceptible.

Pour vérifier les prédictions de la théorie du mouvement brownien, Jean Perrin observa au microscope le mouvement de grains calibrés[13] en suspension dans un liquide. Il suivait un grain et en pointait la position toutes les trente secondes. Cette observation a été faite pour 500 grains. On voit sur la figure 4 les positions consécutives de trois d'entre eux et sur la figure 5 les déplacements, ramenés à une même origine, des 500 grains après un temps t suffisamment long (dans cette expérience, c'est bien sûr la projection du mouvement sur un plan perpendiculaire à l'axe du microscope que l'on observe).

---

11. La théorie du mouvement brownien est plus complexe que l'esquisse que nous en donnons ici. Voir référence 3, p. 560 et exercices IV.5 à IV.7.

12. Le coefficient de diffusion mesure *la réponse du système à un gradient de concentration.* En l'absence de champ extérieur, l'équilibre correspond à une répartition homogène des particules en suspension. Si l'on crée une inhomogénéité dans cette répartition, il apparaît un flux de particules qui tend à rétablir la situation d'équilibre. Désignons par c la concentration des particules considérées (nombre de particules par unité de volume). Lorsqu'on maintient un gradient de concentration $\partial c/\partial z$ dans la direction Oz, il s'établit un flux de particules J_z le long de Oz (nombre de particules franchissant par unité de temps une surface unité perpendiculaire à Oz). Dans la limite où le gradient est assez faible, on observe une « réponse linéaire », c'est-à-dire que J_z est proportionnel à $\partial c/\partial z$. Le coefficient de diffusion est alors défini par

$$J_z \widehat{=} -D\frac{\partial c}{\partial z}$$

(voir référence 3, p. 483).

13. Les expériences de Jean Perrin furent réalisées avec des grains de gomme gutte (rayons de 0,2 μm à 0,5 μm et masses de $4{,}8 \times 10^{-14}$ g à $6{,}00 \times 10^{-13}$ g), ainsi que des grains de mastic (rayons de 0,5 μm à 5,5 μm et masses de $6{,}5 \times 10^{-13}$ g à $7{,}5 \times 10^{-10}$ g).

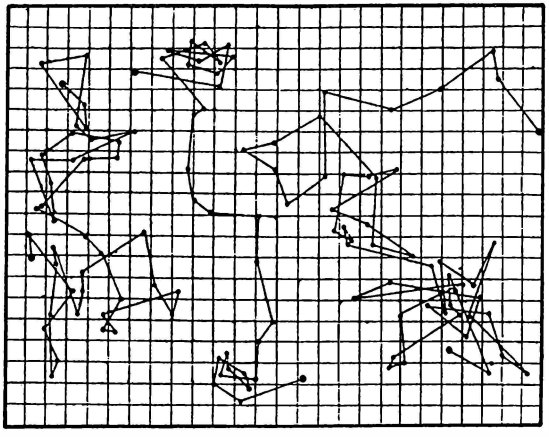


FIGURE 4
Positions successives de 3 grains en suspension dans un liquide. Les pointages ont été réalisés toutes les 30 secondes par J. Perrin (16 divisions représentent 50 μm).

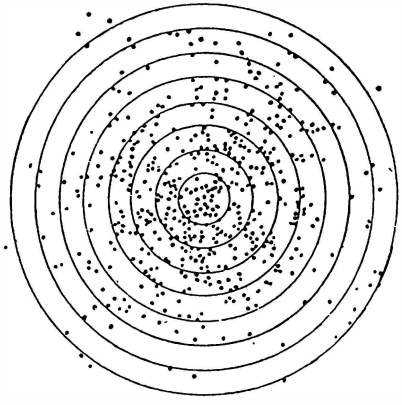


FIGURE 5
Positions de 500 grains à un instant t après leur départ d'une origine supposée commune.

Effectivement Perrin trouva que l'écart quadratique moyen de l'abscisse X des grains augmentait proportionnellement à la racine carrée du temps. De plus, le nombre de points situés dans les divers anneaux dessinés sur la figure 5 était en excellent accord avec les probabilités calculées[14]. Puis il compara les valeurs expérimentale et

14. A partir de la formule (D.18) on peut trouver la probabilité $\omega_N(\rho, \varphi)\mathrm{d}\rho\ \mathrm{d}\varphi$ pour que la projection du mobile sur un plan perpendiculaire à l'axe Oz se trouve dans la surface élémentaire entourant le point de coordonnées polaires (ρ, φ) :

$$\omega_N(\rho, \varphi)\mathrm{d}\rho\ \mathrm{d}\varphi = W_N(X)\mathrm{d}X \,.\, W_N(Y)\mathrm{d}Y$$
$$= \frac{3}{2\pi N\bar{\ell}^2} \exp\left[-\frac{3(X^2+Y^2)}{2N\bar{\ell}^2}\right]\mathrm{d}X\ \mathrm{d}Y$$
$$= \frac{3}{2\pi N\bar{\ell}^2} \exp\left[-\frac{3\rho^2}{2N\bar{\ell}^2}\right]\rho\ \mathrm{d}\rho\ \mathrm{d}\varphi\ . \qquad (1)$$

.../...

théorique[15] de la quantité $\bar{\rho}/\sqrt{\overline{\rho^2}}$ (ρ étant la distance au point de départ). Il trouva 0,894; compte tenu du nombre d'observations faites, ceci est en excellent accord avec la valeur théorique 0,886.

Une dernière vérification expérimentale réalisée avec succès par Perrin et Brillouin fut considérée à l'époque comme déterminante. Ils mesurèrent le coefficient de diffusion D qui apparaît dans la formule (D.19) et le comparèrent à son expression théorique[16]. Ils en déduisirent alors une valeur de la constante de Boltzmann qui, à l'aide de la constante des gaz parfaits

$$R \widehat{=} kN_A = 8{,}31 \text{ joule/degré}\,, \tag{D.21}$$

donne le nombre d'Avogadro N_A. Ils trouvèrent

$$5{,}5 \times 10^{23} < N_A < 8{,}0 \times 10^{23}\,, \tag{D.22}$$

en très bon accord avec les valeurs déterminées par d'autres méthodes.

Références

1. Article de S. Chandrasekar dans A. Wax, *Selected Papers on Noise and Stochastic Processes*, Dover, New York (1954).
2. Reuben Hersh et Richard J. Griego, « Brownian Motion and Potential Theory », *Scientific American*, mars 1969, p. 66.
3. F. Reif, *Fundamentals of Statistical and Thermal Physics*, McGraw-Hill, New York (1965), chapitres 12 et 15.
4. W. Feller, *An Introduction to Probability Theory and Its Applications*, John Wiley, New York (1950).

On en déduit la densité de probabilité $w_N(\rho)$ pour que le mobile se trouve à une distance ρ de l'origine :

$$w_N(\rho) = \int_0^{2\pi} \mathrm{d}\varphi\, \omega_N(\rho, \varphi) = \frac{3}{N\overline{\ell^2}}\, \rho\, \exp\left[-3\rho^2/2N\overline{\ell^2}\right]. \tag{2}$$

15. Les formules de la note 14 donnent facilement la valeur moyenne de ρ et son écart quadratique moyen :

$$\bar{\rho} = \int_0^{\infty} \rho w_N(\rho)\, \mathrm{d}\rho = \sqrt{\frac{\pi N\overline{\ell^2}}{6}}\,, \tag{1}$$

$$\overline{\rho^2} = \int_0^{\infty} \rho^2 w_N(\rho)\, \mathrm{d}\rho = \frac{2}{3} N\overline{\ell^2}\,, \tag{2}$$

d'où $$\frac{\bar{\rho}}{(\overline{\rho^2})^{1/2}} = \frac{\sqrt{\pi}}{2} \simeq 0{,}886\,. \tag{3}$$

16. Pour des particules sphériques, le coefficient de diffusion est donné par la loi de Stokes :

$$D = \frac{kT}{6\pi\eta a}\,,$$

où k est la constante de Boltzmann, T la température absolue, η le coefficient de viscosité du liquide et a le rayon des particules browniennes (voir exercice IV.5).

COMPLÉMENT I.E

Densité d'états d'un gaz parfait

I. Volume d'une boule et surface d'une sphère dans un espace de dimension n

II. Calcul de la densité d'états d'un gaz parfait

Le calcul de la densité d'états pour un système de N particules libres, indépendantes et discernables (gaz parfait) a été effectué au chapitre I, paragraphe III.B.2.b. Le résultat

$$\rho_N(E) = C(N)\, V^N \left(\frac{E}{N}\right)^{\frac{3N}{2}-1} \tag{E.1}$$

laissait indéterminé le coefficient numérique $C(N)$. Nous nous proposons ici de retrouver ce résultat par une méthode différente, et de déterminer explicitement le coefficient $C(N)$.

Nous aurons besoin, dans ce calcul, de connaître l'expression du volume et de la surface d'une sphère dans un espace de dimension quelconque n; nous commencerons donc par établir ces expressions.

I. VOLUME D'UNE BOULE ET SURFACE D'UNE SPHÈRE DANS UN ESPACE DE DIMENSION *n*

Soit donc un espace euclidien à n dimensions, dont les points sont repérés dans un système de coordonnées orthonormé : x_i $(i = 1, 2, \ldots, n)$. La sphère de rayon r est définie par son équation :

$$x_1^2 + x_2^2 + \ldots + x_n^2 = r^2 \ . \tag{E.2}$$

Le volume de la boule délimitée par cette sphère est nécessairement de la forme

$$V_n(r) = \alpha_n r^n \ , \tag{E.3}$$

où α_n est un coefficient qui dépend de la dimension n de l'espace, mais pas du rayon r de la sphère. La surface de la sphère s'en déduit par dérivation :

$$S_n(r) = n\alpha_n r^{n-1} \ . \tag{E.4}$$

Pour déterminer α_n, nous calculons l'intégrale

$$I_n = \int_{-\infty}^{+\infty} \ldots \int_{-\infty}^{+\infty} \mathrm{d}x_1\, \mathrm{d}x_2 \ldots \mathrm{d}x_n\, \mathrm{e}^{-(x_1^2 + x_2^2 + \ldots x_n^2)} \tag{E.5}$$

par deux méthodes différentes. La première consiste à remarquer que cette intégrale est le produit de n intégrales identiques à

$$I_1 = \int_{-\infty}^{+\infty} \mathrm{d}x\ \mathrm{e}^{-x^2} = \sqrt{\pi} \ , \tag{E.6}$$

de sorte que

$$I_n = (\pi)^{n/2} \ . \tag{E.7}$$

Mais on peut aussi, pour calculer I_n, découper l'espace en couches sphériques concentriques puisque la fonction à intégrer ne dépend que de r. Le volume élémentaire entre deux couches de rayons r et $r+\mathrm{d}r$ est $\mathrm{d}V = S_n(r)\,\mathrm{d}r$, et donc

$$I_n = \int_0^\infty \mathrm{d}r\; n\alpha_n r^{n-1} \mathrm{e}^{-r^2} . \tag{E.8}$$

Le changement de variable $u = r^2$ donne

$$I_n = \frac{1}{2}\int_0^\infty \mathrm{d}u\; n\alpha_n u^{n/2-1} \mathrm{e}^{-u} . \tag{E.9}$$

Introduisant la fonction[1]

$$\Gamma(x) \widehat{=} \int_0^\infty \mathrm{e}^{-t} t^{x-1}\,\mathrm{d}t , \tag{E.10}$$

on obtient

$$I_n = \frac{1}{2}\, n\alpha_n \Gamma\left(\frac{n}{2}\right) = \alpha_n \Gamma\left(\frac{n}{2}+1\right) . \tag{E.11}$$

Il ne reste plus qu'à identifier (E.7) et (E.11) pour déterminer α_n :

$$\alpha_n = \frac{\pi^{n/2}}{\Gamma\left(\frac{n}{2}+1\right)} . \tag{E.12}$$

Notons que, pour $n \gg 1$, $\Gamma(n/2+1)$ peut être approximé par la formule de Stirling

$$\Gamma(x+1) \simeq x^x \mathrm{e}^{-x}\sqrt{2\pi x} ; \tag{E.13}$$

ainsi $$\alpha_n \simeq \left(\frac{2\pi \mathrm{e}}{n}\right)^{n/2} \frac{1}{\sqrt{\pi n}} \quad \text{si} \quad n \gg 1 . \tag{E.14}$$

II. CALCUL DE LA DENSITÉ D'ÉTATS D'UN GAZ PARFAIT

Nous considérons N particules libres, indépendantes et discernables enfermées dans une boîte de taille macroscopique. Les N particules, de masse m, sont identiques entre elles, et la boîte, de volume V, est parallélépipédique, d'arêtes L_x, L_y et L_z[2].

1. La fonction $\Gamma(x)$ est bien connue (voir par exemple L. Schwartz, *Méthodes mathématiques pour les sciences physiques*, Hermann, Paris (1961), p. 344). De sa définition (E.10), on déduit facilement, en intégrant par parties,

$$\Gamma(x+1) = x\Gamma(x) . \tag{1}$$

Comme de plus

$$\Gamma(1) = 1 , \tag{2}$$

on trouve, pour n entier,

$$\Gamma(n+1) = n! . \tag{3}$$

2. En fait, la forme précise du récipient est sans importance tant que la longueur d'onde de de Broglie d'une particule est négligeable devant les dimensions du récipient (voir complément I.A, § III.2).

L'hamiltonien H du système est alors la somme de N termes se rapportant chacun à une particule :

$$H = \sum_{i=1}^{N} \frac{\vec{P}_i^2}{2m} . \tag{E.15}$$

Nous choisissons les conditions aux limites périodiques. Un état stationnaire du système est donc caractérisé par la donnée de N vecteurs à trois dimensions

$$\{\vec{k}_1, \vec{k}_2, \ldots, \vec{k}_N\} , \tag{E.16}$$

$\hbar\vec{k}_i$ étant l'impulsion de la particule (i). L'énergie de cet état vaut

$$E = \sum_{i=1}^{N} \frac{\hbar^2 \vec{k}_i^{\,2}}{2m} \tag{E.17}$$

et les composantes k_{ix}, k_{iy}, k_{iz} de chacun des vecteurs $\vec{k}_i$ doivent être des multiples entiers (positifs, négatifs ou nuls) de $2\pi/L_x$, $2\pi/L_y$ et $2\pi/L_z$ respectivement (voir I, § III.B.2.b).

On peut donc considérer que fixer l'état du système des N particules revient à fixer un vecteur $\vec{K}$ dans un *espace à 3N dimensions*, ce vecteur étant défini par ses composantes :

$$\vec{K} \widehat{=} \{k_{1x}, k_{1y}, k_{1z}, \ldots, k_{Ny}, k_{Nz}\} . \tag{E.18}$$

Les seuls vecteurs $\vec{K}$ admissibles ont leur extrémité à l'un des nœuds d'un réseau de cet espace à $3N$ dimensions.

Le raisonnement est alors identique à celui de la note 19 du chapitre I. L'énergie E du système des N particules s'écrit

$$E = \frac{\hbar^2 \vec{K}^{\,2}}{2m} , \tag{E.19}$$

le volume de la maille élémentaire étant

$$v_{3N} = \left(\frac{2\pi}{L_x} \cdot \frac{2\pi}{L_y} \cdot \frac{2\pi}{L_z}\right)^N = \frac{(2\pi)^{3N}}{V^N} ; \tag{E.20}$$

le nombre d'états $\rho_N(E)\,\mathrm{d}E$ d'énergie comprise entre E et $E + \mathrm{d}E$ est donc

$$\rho_N(E)\,\mathrm{d}E = \frac{S_{3N}(K)\,\mathrm{d}K}{v_{3N}} = \frac{3N\alpha_{3N} V^N K^{3N-1}}{(2\pi)^{3N}}\,\mathrm{d}K . \tag{E.21}$$

α_{3N} est donné par (E.12), K et $\mathrm{d}K$ s'expriment en fonction de E et $\mathrm{d}E$ à l'aide de (E.19); on obtient finalement

$$\rho_N(E) = \frac{3N}{2} \left(\frac{m}{2\pi\hbar^2}\right)^{3N/2} \frac{V^N}{\Gamma\left(\frac{3N}{2}+1\right)} E^{3N/2-1} . \tag{E.22}$$

Ce résultat est bien identique à (E.1), avec

$$C(N) = \frac{3}{2} \left(\frac{mN}{2\pi\hbar^2}\right)^{3N/2} \cdot \frac{1}{\Gamma\left(\frac{3N}{2}+1\right)} . \tag{E.23}$$

Dans tous les cas pratiques $N \gg 1$; on peut donc utiliser la formule de Stirling (E.13) pour calculer $C(N)$:

$$C(N) \simeq \frac{1}{2} \sqrt{\frac{3}{N\pi}} \left(\frac{m\,\mathrm{e}}{3\pi\hbar^2}\right)^{3N/2} . \tag{E.24}$$

Attention : le résultat ci-dessus suppose les N particules discernables. Pour des particules identiques, et donc indiscernables, et si l'approximation de Maxwell-Boltzmann est valable, on obtient un résultat approché en divisant $C(N)$ par N ! (Complément II-C et chapitre III, § IV-A-3).

COMPLÉMENT I.F

États électroniques dans les solides cristallins

I. Description du réseau cristallin
   1. Réseau de Bravais
   2. Maille élémentaire
   3. Réseau réciproque

II. États électroniques dans un cristal
   1. Potentiel périodique et théorème de Bloch. Bandes d'énergie
   2. Les conditions aux limites de Born-von Karman
   3. Densité d'états. Expression générale

III. Exemples simples d'évaluation de la densité d'états
   1. Le modèle des électrons libres
   2. Le modèle des liaisons fortes

Nous nous proposons d'introduire ici quelques notions concernant les solides cristallins et plus particulièrement leurs états électroniques. Nous nous limiterons aux considérations élémentaires nécessaires à la compréhension de la suite de ce cours. Les exemples que nous donnerons seront toujours les plus simples[1].

I. DESCRIPTION DU RÉSEAU CRISTALLIN

1. RÉSEAU DE BRAVAIS

Un cristal est un ensemble de *motifs atomiques identiques disposés régulièrement dans l'espace* de façon à former un réseau périodique à trois dimensions. Le motif élémentaire qui se répète ainsi régulièrement peut être constitué par un atome unique, un groupe d'atomes, un ou plusieurs ions ou molécules.

On appelle *réseau de Bravais du cristal* un ensemble de points qui reproduit la géométrie de la structure périodique sous-jacente, abstraction faite du motif élémentaire de base : on choisit dans ce motif un point particulier; le réseau de Bravais est alors l'ensemble des positions qu'occupe ce point particulier dans la succession de motifs identiques constituant le cristal. Le réseau de Bravais d'un cristal déterminé peut être engendré par un système de trois vecteurs indépendants $\vec{a}_1$, $\vec{a}_2$, $\vec{a}_3$. Ses points sont repérés par les vecteurs $\vec{R}$ de la forme

$$\vec{R} = n_1\vec{a}_1 + n_2\vec{a}_2 + n_3\vec{a}_3 \ , \qquad \text{(F.1)}$$

où n_1, n_2, n_3 sont trois entiers positifs, négatifs ou nuls quelconques[2]. En chacun de

1. Pour plus de détails nous renvoyons aux ouvrages spécialisés cités en référence.

2. Le système $(\vec{a}_1, \vec{a}_2, \vec{a}_3)$ des vecteurs élémentaires non coplanaires qui engendrent un réseau de Bravais donné n'est pas unique. Il existe une infinité de choix différents possibles.

ces points, appelés *nœuds* du réseau, s'attache le motif élémentaire constitutif[3], [4].

La figure 1 représente le plus simple des réseaux de Bravais, le réseau cubique simple; il est engendré par trois vecteurs orthogonaux de même module.

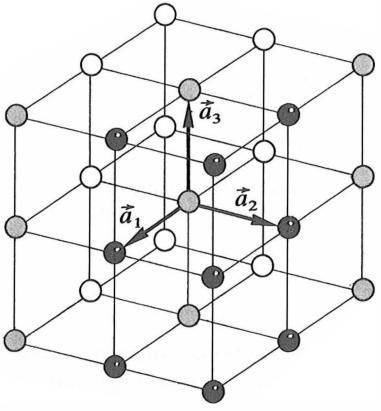


FIGURE 1
Structure du réseau cubique simple. Il peut être engendré par les trois vecteurs $\vec{a}_1$, $\vec{a}_2$, $\vec{a}_3$ orthogonaux et de même module.

2. MAILLE ÉLÉMENTAIRE

On appelle maille élémentaire d'un cristal un domaine qui, translaté par l'ensemble des vecteurs du réseau de Bravais, remplit tout l'espace sans recouvrement ni omission. Chaque maille élémentaire contient ainsi un seul nœud du réseau de Bravais.

Pour définir un solide cristallin, il suffit donc de se donner le réseau de Bravais correspondant et la position des atomes, ions ou molécules constitutifs à l'intérieur de la maille élémentaire.

Il y a une infinité de choix possibles pour la maille élémentaire d'un cristal donné (fig. 2), mais toutes les mailles élémentaires possibles occupent des volumes égaux.

---

3. Tel qu'il a été défini, un réseau de Bravais, et donc le cristal qu'il représente, s'étend à l'infini. Les solides réels sont finis, mais leurs dimensions sont telles que la plupart des atomes constitutifs sont trop éloignés de la surface pour en ressentir la présence; l'approximation du cristal infini est donc souvent suffisante. S'il est cependant nécessaire de tenir compte du fait que le cristal est fini, on lui attribue par commodité la forme la plus simple possible en limitant le réseau de Bravais aux vecteurs $\vec{R}$ (formule (F.1)) vérifiant, par exemple,

$$0 \leqslant n_1 < N_1 ,$$

$$0 \leqslant n_2 < N_2 ,$$

$$0 \leqslant n_3 < N_3 ,$$

où N_1, N_2, N_3 sont trois nombres entiers très grands dont le produit $N_1 N_2 N_3$ est égal au nombre N de nœuds dans le cristal réel considéré.

4. Un cristal réel présente aussi des défauts, qui ne sont pas pris en compte par le réseau de Bravais. Nous supposons que ces défauts sont rares et n'affectent pas les propriétés du cristal. Nous nous limitons donc aux *cristaux parfaits*. Mais même dans ce cas, du fait de l'agitation thermique, les atomes ou ions constitutifs du cristal n'occupent pas exactement les positions définies à partir du réseau de Bravais; ils s'en écartent de façon aléatoire, mais peu si la température est bien inférieure à la température de fusion du cristal. Nous négligeons ici cet effet sur lequel nous reviendrons au chapitre III (complément III.E).

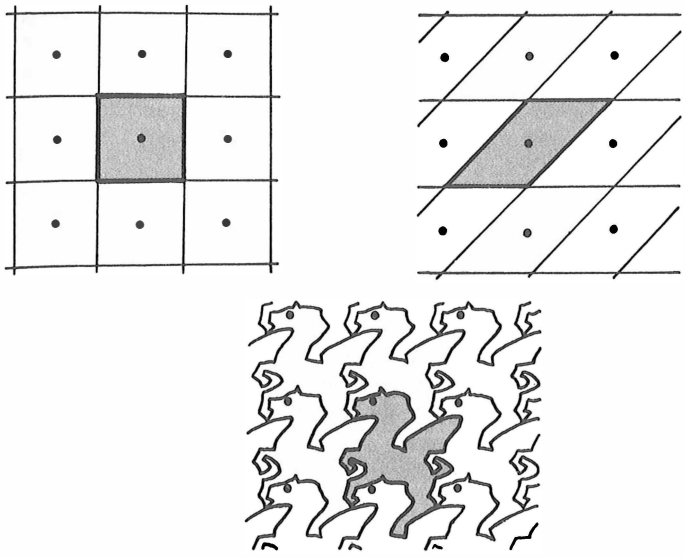


FIGURE 2
Réseau carré à deux dimensions avec trois choix possibles de la maille élémentaire (le troisième est repris de M. C. Escher).

Il est commode de choisir une maille élémentaire qui possède les mêmes éléments de symétrie que le réseau de Bravais. La plus simple est la *maille de Wigner-Seitz :* la maille élémentaire de Wigner-Seitz autour d'un nœud donné est constituée par l'ensemble des points de l'espace qui sont plus proches de ce nœud que de tout autre nœud du réseau. C'est donc le plus petit polyèdre limité par les plans médiateurs des segments qui joignent le nœud considéré à tous les autres (fig. 3). On vérifie aisément que la maille de Wigner-Seitz d'un cristal cubique simple est un cube.

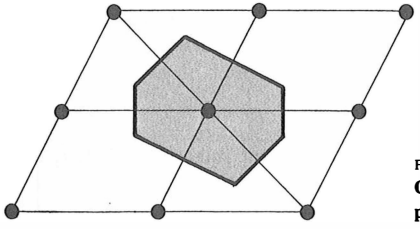


FIGURE 3
Construction de la maille élémentaire de Wigner-Seitz pour un réseau de Bravais à deux dimensions.

3. RÉSEAU RÉCIPROQUE

Étant donné un réseau de Bravais, son *réseau réciproque* est constitué par les extrémités des vecteurs $\vec{K}$ tels que

$$\exp(i\vec{K}\,.\,\vec{R}\,)=1 \qquad \text{(F.2)}$$

pour tous les $\vec{R}$ associés au réseau initial (appelé, par opposition, *réseau direct*). Le réseau réciproque est lui-même un réseau de Bravais. En effet, si $(\vec{a}_1, \vec{a}_2, \vec{a}_3)$ est le système de vecteurs élémentaires qui engendrent le réseau direct, tous les vecteurs du réseau réciproque sont de la forme

$$\vec{K} = m_1\vec{b}_1 + m_2\vec{b}_2 + m_3\vec{b}_3\ , \qquad \text{(F.3)}$$

où m_1, m_2, m_3 sont trois entiers positifs, négatifs ou nuls, et $\vec{b}_1$, $\vec{b}_2$, $\vec{b}_3$ les trois vecteurs non coplanaires définis par

$$\vec{b}_1 \widehat{=} 2\pi \frac{\vec{a}_2 \wedge \vec{a}_3}{\vec{a}_1 . (\vec{a}_2 \wedge \vec{a}_3)} , \tag{F.4.a}$$

$$\vec{b}_2 \widehat{=} 2\pi \frac{\vec{a}_3 \wedge \vec{a}_1}{\vec{a}_1 . (\vec{a}_2 \wedge \vec{a}_3)} , \tag{F.4.b}$$

$$\vec{b}_3 \widehat{=} 2\pi \frac{\vec{a}_1 \wedge \vec{a}_2}{\vec{a}_1 . (\vec{a}_2 \wedge \vec{a}_3)} . \tag{F.4.c}$$

Ainsi le réseau réciproque d'un cristal cubique simple est lui-même cubique simple. Il possède une maille élémentaire cubique de côté $2\pi/a$ (a étant le côté de la maille élémentaire cubique du réseau direct).

D'une façon générale, si v est le volume de la maille élémentaire du réseau direct, le volume de la maille élémentaire du réseau réciproque vaut[5], dans l'espace des $\vec{K}$, $(2\pi)^3/v$.

La maille élémentaire de Wigner-Seitz du réseau réciproque est appelée *première zone de Brillouin* du cristal.

II. ÉTATS ÉLECTRONIQUES DANS UN CRISTAL

1. POTENTIEL PÉRIODIQUE ET THÉORÈME DE BLOCH. BANDES D'ÉNERGIE

Les électrons d'un solide cristallin sont soumis au potentiel périodique créé par l'ensemble des noyaux atomiques. Ils interagissent également les uns avec les autres; cependant, on considère généralement ces électrons comme des *particules indépendantes,* chacune étant soumise à un « potentiel moyen » créé par les noyaux et les autres électrons du cristal[6]. Ce potentiel moyen reprend les mêmes valeurs dans chaque maille élémentaire, de sorte que chaque électron évolue dans un *potentiel périodique* $V(\vec{r})$:

$$V(\vec{r}) = V(\vec{r} + \vec{R}) . \tag{F.5}$$

Le *théorème de Bloch,* que nous admettrons sans démonstration (cf. références 1, chapitre 8, et 2, chapitre 9), donne la forme générale des fonctions d'onde spatiales[7] stationnaires $\psi(\vec{r})$, solutions de l'équation de Schrödinger

$$\left[-\frac{\hbar^2}{2m} \Delta + V(\vec{r}) \right] \psi(\vec{r}) = \varepsilon \psi(\vec{r}) \tag{F.6}$$

---

5. Pour démontrer ce résultat, il suffit de calculer le produit mixte $(\vec{b}_1 \wedge \vec{b}_2) . \vec{b}_3$ à l'aide des formules (F.4) et de la formule du double produit vectoriel

$$\vec{V}_1 \wedge (\vec{V}_2 \wedge \vec{V}_3) = (\vec{V}_1 . \vec{V}_3) \vec{V}_2 - (\vec{V}_1 . \vec{V}_2) \vec{V}_3 .$$

6. Il n'est pas évident que l'on puisse traiter correctement l'interaction coulombienne entre électrons par l'intermédiaire d'un potentiel moyen. L'approximation est en fait très bonne et de nombreuses propriétés électroniques des solides s'expliquent simplement dans le modèle des électrons indépendants.

Il faut cependant noter que ce potentiel moyen, qui permet de calculer les fonctions d'onde électroniques, doit être déterminé à partir de ces mêmes fonctions d'onde : un électron donné ressent en moyenne (outre l'attraction des noyaux) la distribution de charge créée par l'ensemble des autres électrons, distribution qui dépend de leurs fonctions d'onde. Le problème doit donc être résolu de façon « autocohérente ».

7. L'hamiltonien de (F.6) ne dépend pas du spin. A cette approximation, l'état d'un électron est le produit tensoriel de son état spatial et de son état de spin (appendice II, § III.2.b).

pour un électron dans le potentiel périodique $V(\vec{r})$. Ces fonctions d'onde peuvent toujours être choisies comme le *produit d'une onde plane par une fonction périodique* dont la périodicité est celle du réseau de Bravais du cristal[8] :

$$\psi_{\vec{k}}(\vec{r}) = e^{i\vec{k}\cdot\vec{r}} u_{\vec{k}}(\vec{r}), \tag{F.7.a}$$

avec $$u_{\vec{k}}(\vec{r}+\vec{R}) = u_{\vec{k}}(\vec{r}). \tag{F.7.b}$$

Remarquons qu'on peut se restreindre à choisir $\vec{k}$ dans la première zone de Brillouin. En effet tout $\vec{k}'$ extérieur à celle-ci peut s'écrire

$$\vec{k}' = \vec{k} + \vec{K}, \tag{F.8}$$

où $\vec{K}$ est un vecteur du réseau réciproque et $\vec{k}$ un vecteur de la première zone. Une fonction d'onde produit de l'onde plane $\exp[i\vec{k}'.\vec{r}]$ par la fonction périodique $u(\vec{r})$ est aussi égale au produit de $\exp[i\vec{k}.\vec{r}]$ par la fonction $u(\vec{r})\exp[i\vec{K}.\vec{r}]$, périodique d'après (F.2) et (F.7.b).

Pour un cristal infini, $\vec{k}$ peut prendre n'importe quelle valeur à l'intérieur de la première zone de Brillouin; les états de $\vec{k}$ voisins ont des propriétés voisines, leurs énergies $\varepsilon(\vec{k})$ varient continûment avec $\vec{k}$. Quand $\vec{k}$ décrit la première zone de Brillouin, $\varepsilon(\vec{k})$ couvre un domaine d'énergie qu'on appelle une *bande d'énergie permise.* Un cristal possède plusieurs bandes d'énergie permises[9], souvent séparées par des *bandes interdites*, domaines d'énergie pour lesquels il n'existe pas d'état (c'est-à-dire de solution de l'équation de Schrödinger (F.6)).

---

8. Ainsi la densité de probabilité quantique $|\psi_{\vec{k}}(\vec{r})|^2 = |u_{\vec{k}}(\vec{r})|^2$ possède la symétrie du cristal : comme il se doit, $|\psi_{\vec{k}}(\vec{r})|^2$ reprend les mêmes valeurs dans chaque maille élémentaire. La fonction d'onde $\psi_{\vec{k}}(\vec{r})$, elle, varie d'une maille à l'autre, mais seulement par l'intermédiaire du facteur de phase $e^{i\vec{k}.\vec{r}}$ (c'est là une propriété générale des ondes dans les milieux périodiques).

9. Dans un atome isolé, les divers niveaux que peut occuper un électron ont des énergies bien séparées (par des intervalles de l'ordre de l'électron-volt). Si l'on approche deux atomes identiques, les électrons peuvent « sauter » de l'un à l'autre, et ceci donne lieu à un phénomène de *résonance quantique* qui dédouble les niveaux : chaque état atomique se couple ainsi à son homologue de l'autre atome, et ce couplage engendre deux états de type moléculaire (c'est-à-dire englobant les deux noyaux à la fois), ayant des énergies voisines mais différentes. Dans un cristal, le couplage et la résonance concernent un grand nombre N d'atomes, de sorte que *chacun des états atomiques initiaux* donne naissance à un groupe de N états d'énergies distinctes mais très rapprochées, constituant une *bande permise* (les N niveaux, où N est de l'ordre du nombre d'Avogadro, sont concentrés dans une bande dont la largeur est de l'ordre de l'électron-volt). D'un point de vue

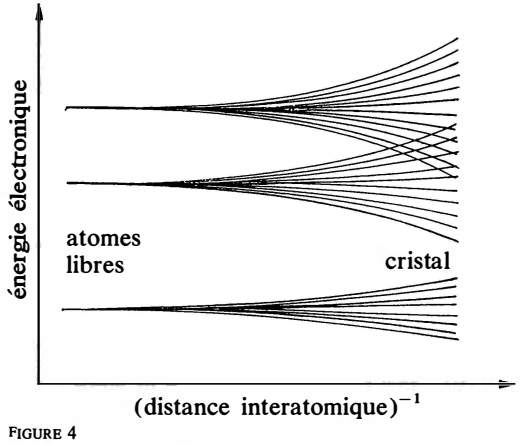


FIGURE 4

Correspondance schématique entre niveaux d'énergie dans un atome isolé et dans un cristal.

Ainsi la fonction d'onde spatiale d'un électron dans un cristal sera spécifiée par son vecteur d'onde $\vec{k}$ (choisi dans la première zone de Brillouin) et la bande n à laquelle elle appartient. On écrira cette fonction

$$\psi_{n,\vec{k}}(\vec{r}) = e^{i\vec{k}\cdot\vec{r}} u_{n,\vec{k}}(\vec{r}) , \tag{F.9}$$

et on notera $\varepsilon_n(\vec{k})$ l'énergie correspondante.

2. LES CONDITIONS AUX LIMITES DE BORN-VON KARMAN

Comme dans le cas des particules libres, la taille finie du cristal limite à un ensemble discret les valeurs de $\vec{k}$; ce sont les conditions aux limites imposées à la surface du cristal qui déterminent ces valeurs permises de $\vec{k}$. On choisit d'ordinaire les conditions de Born-von Karman, qui généralisent les conditions périodiques étudiées à l'appendice II, paragraphe II.4.c pour une particule dans une boîte parallélépipédique.

Pour un cristal, il est pratique d'utiliser un volume qui contient un nombre entier de mailles élémentaires (cf. note 3); on impose donc

$$\psi_{n,\vec{k}}(\vec{r} + N_i\vec{a}_i) = \psi_{n,\vec{k}}(\vec{r}) \quad (i = 1, 2, 3) , \tag{F.10}$$

où $\vec{a}_1$, $\vec{a}_2$, $\vec{a}_3$ sont trois vecteurs élémentaires définissant le réseau de Bravais et N_1, N_2, N_3 trois entiers très grands, d'ordre $N^{1/3}$ ($N = N_1N_2N_3$ est le nombre total de mailles élémentaires dans le cristal fini considéré).

La relation (F.10) combinée avec la forme (F.7) des fonctions d'onde donne

$$\exp[iN_j\vec{k}\,.\,\vec{a}_j] = 1 \quad (j = 1, 2, 3) ; \tag{F.11}$$

c'est-à-dire que *les valeurs de $\vec{k}$ permises sont discrètes* et de la forme

$$\vec{k} = \frac{p_1}{N_1}\vec{b}_1 + \frac{p_2}{N_2}\vec{b}_2 + \frac{p_3}{N_3}\vec{b}_3 , \tag{F.12}$$

où les p_i sont des entiers positifs, négatifs ou nuls, et les $\vec{b}_i$ sont définis en (F.4).

Comme au paragraphe III.B.2 du chapitre I, calculons le volume élémentaire $\Delta V_{\vec{k}}$ de l'espace réciproque associé à chaque valeur permise de $\vec{k}$. C'est le volume du parallélépipède construit sur les vecteurs $\vec{b}_1/N_1$, $\vec{b}_2/N_2$, $\vec{b}_3/N_3$:

$$\Delta V_{\vec{k}} = \frac{1}{N}\vec{b}_1\,.\,(\vec{b}_2 \wedge \vec{b}_3) . \tag{F.13}$$

---

plus technique, en reportant la forme (F.7) des fonctions d'onde dans l'équation de Schrödinger, on obtient

$$\left[\frac{\hbar^2}{2m}(-i\vec{\nabla} + \vec{k})^2 + V(\vec{r})\right] u_{\vec{k}}(\vec{r}) = \varepsilon(\vec{k})\, u_{\vec{k}}(\vec{r}) , \tag{1}$$

avec la condition

$$u_{\vec{k}}(\vec{r}) = u_{\vec{k}}(\vec{r} + \vec{R}) . \tag{2}$$

Pour $\vec{k}$ *fixé*, on cherche l'ensemble des fonctions propres $u_{\vec{k}}(\vec{r})$ et des valeurs propres $\varepsilon(\vec{k})$ d'un opérateur hermitique, avec (par suite de la relation (2)) des conditions aux limites périodiques dans un domaine restreint à la maille élémentaire. On montre de façon générale que, dans un tel problème, les valeurs propres forment un *spectre discret* $\varepsilon_n(\vec{k})$ ($n = 1, 2, \ldots$). Lorsque, *pour n fixé*, $\vec{k}$ parcourt la première zone de Brillouin, $\varepsilon_n(\vec{k})$ engendre une *bande permise* repérée par l'entier n (fig. 4). Si l'on tient compte du fait que le nombre N d'atomes du cristal est fini, cette bande permise comprend N niveaux distincts (voir paragraphe 2).

Or le produit mixte $\vec{b}_1.(\vec{b}_2 \wedge \vec{b}_3)$ est égal au volume de la maille élémentaire du réseau réciproque, c'est-à-dire au volume de la première zone de Brillouin; *le nombre de vecteurs d'onde $\vec{k}$ permis (dans la première zone de Brillouin) est donc simplement égal au nombre N de nœuds dans le cristal*[10].

De plus, si v est le volume de la maille élémentaire du réseau direct, celui de la première zone de Brillouin vaut, dans l'espace réciproque, $(2\pi)^3/v$. Il s'ensuit que

$$\Delta V_{\vec{k}} = \frac{(2\pi)^3}{Nv} ; \tag{F.14}$$

or Nv est égal au volume V du cristal, et

$$\Delta V_{\vec{k}} = \frac{(2\pi)^3}{V} . \tag{F.15}$$

Le volume élémentaire par valeur permise de $\vec{k}$ est donc le même que pour une particule libre enfermée dans une boîte de même volume que le cristal.

3. DENSITÉ D'ÉTATS. EXPRESSION GÉNÉRALE

De nombreuses propriétés macroscopiques du cristal s'obtiennent en calculant des moyennes sur les états électroniques (chapitre I, § III.B). Dans le modèle des électrons indépendants, on se ramène souvent à des moyennes de grandeurs A ne faisant intervenir qu'un seul état électronique, de la forme

$$\overline{A} = \sum_{\sigma=\pm 1/2} \sum_{n} \sum_{\vec{k}\in Z.B.} A_{n,\sigma}(\vec{k})\, P_{n,\sigma}(\vec{k}) , \tag{F.16}$$

où $P_{n,\sigma}(\vec{k})$ est la probabilité pour qu'un électron soit dans l'état de la bande n caractérisé par le vecteur d'onde $\vec{k}$ et la projection σ du spin et $A_{n,\sigma}(\vec{k})$ la valeur prise dans cet état par la grandeur A; la somme sur $\vec{k}$ est restreinte aux valeurs permises de $\vec{k}$ appartenant à la première zone de Brillouin.

Pour des cristaux de taille macroscopique, les valeurs permises de $\vec{k}$ sont très voisines et la somme sur $\vec{k}$ peut être remplacée par une intégrale; compte tenu de (F.15),

$$\overline{A} = \sum_{\sigma=\pm 1/2} \sum_{n} \frac{V}{(2\pi)^3} \int_{Z.B.} \mathrm{d}^3k\, A_{n,\sigma}(\vec{k}) P_{n,\sigma}(\vec{k}) . \tag{F.17}$$

Souvent, la quantité à intégrer est indépendante de σ et ne dépend de n et de $\vec{k}$ que par l'intermédiaire de l'énergie $\varepsilon_n(\vec{k})$:

$$A_{n,\sigma}(\vec{k})\, P_{n,\sigma}(\vec{k}) \equiv F[\varepsilon_n(\vec{k})] . \tag{F.18}$$

Si c'est le cas, l'intégration sur $\vec{k}$ peut être remplacée par une intégration sur l'énergie : en introduisant la densité d'états électroniques pour la bande n, définie par[11]

$$\rho_n(\varepsilon) \mathrel{\widehat{=}} \frac{2V}{(2\pi)^3} \int_{Z.B.} \mathrm{d}^3k\, \delta[\varepsilon - \varepsilon_n(\vec{k})] , \tag{F.19}$$

---

10. Ce résultat était à prévoir; il correspond simplement à la conservation du nombre de degrés de liberté (cf. note précédente).

11. Le facteur 2 qui apparaît dans la définition de $\rho_n(\varepsilon)$ provient des deux états de spin de l'électron. Si l'hamiltonien dépend du spin (en présence d'un champ magnétique, par exemple), il y a deux niveaux d'énergie distincts pour $\vec{k}$ et n donnés : $\varepsilon_{n,\sigma}(\vec{k})$ avec $\sigma = \pm 1/2$. On définit alors une *densité d'états pour chaque direction de spin* (voir par exemple complément VI.A, § II.2.c) :

$$\rho_{n,\sigma}(\varepsilon) = \frac{V}{(2\pi)^3} \int_{Z.B.} \mathrm{d}^3k\, \delta[\varepsilon - \varepsilon_{n,\sigma}(\vec{k})] .$$

il vient en effet

$$\overline{A} = \sum_n \int \mathrm{d}\varepsilon \, \rho_n(\varepsilon) F(\varepsilon) \; ; \tag{F.20}$$

la densité d'états totale pour les électrons du cristal est

$$\rho(\varepsilon) \widehat{=} \sum_n \rho_n(\varepsilon) \; , \tag{F.21}$$

et (F.20) s'écrit sous la forme usuelle

$$\overline{A} = \int \mathrm{d}\varepsilon \, \rho(\varepsilon) F(\varepsilon) \; . \tag{F.22}$$

On peut transformer l'expression (F.19) de $\rho_n(\varepsilon)$ en effectuant l'intégration par rapport à l'une des composantes du vecteur $\vec{k}$; l'intégrale de volume se transforme alors en intégrale de surface[12],[13] :

$$\rho_n(\varepsilon) = \frac{V}{4\pi^3} \int_{S_n(\varepsilon)} \frac{\mathrm{d}S}{|\vec{\nabla} \varepsilon_n(\vec{k})|} \; , \tag{F.23}$$

où $S_n(\varepsilon)$ est la surface définie dans l'espace des $\vec{k}$ par l'équation $\varepsilon_n(\vec{k}) = \varepsilon$, et limitée à la première zone de Brillouin.

III. EXEMPLES SIMPLES D'ÉVALUATION DE LA DENSITÉ D'ÉTATS

La connaissance de la structure de bandes, c'est-à-dire des fonctions $\varepsilon_n(\vec{k})$ pour $n = 1, 2, \ldots$, permet donc de calculer la densité d'états électroniques d'un solide. La difficulté principale consiste à déterminer *conjointement* les fonctions d'onde $\psi_{n,\vec{k}}$, les énergies correspondantes $\varepsilon_n(\vec{k})$ et le potentiel périodique $V(\vec{r})$ qui leur est associé (cf. note 6). Ce problème est difficile et sort largement du cadre de cet ouvrage. Nous nous contenterons de deux modèles très simples qui rendent compte, au moins qualitativement, de nombreux cas pratiques.

1. LE MODÈLE DES ÉLECTRONS LIBRES

Dans les éléments métalliques, les électrons périphériques sont peu liés au noyau et évoluent sur des « orbites » de grand rayon. Dans le cristal ces électrons ont des fonctions d'onde qui s'étendent sur l'ensemble du solide. Ils ne ressentent que faiblement le potentiel périodique du réseau, sont très mobiles et se déplacent très facilement lorsqu'on les soumet à un champ électrique. Ce sont eux qui sont responsables des propriétés électriques des métaux; on les appelle *électrons de conduction,* et leurs états correspondent à une bande permise appelée *bande de conduction.*

12. Le passage de (F.19) à (F.23) est explicité dans la note 7 du complément I.B.

13. La quantité à intégrer dans (F.23) diverge aux points où $\vec{\nabla} \varepsilon_n(\vec{k})$ s'annule. A trois dimensions, ces singularités sont en général intégrables et conduisent à des valeurs finies de la densité d'états. Cependant la dérivée $\mathrm{d}\rho_n/\mathrm{d}\varepsilon$ diverge en ces points, que l'on appelle « singularités de van Hove ». C'est ce qui se passe en particulier aux bords inférieur et supérieur d'une bande permise : la densité d'états s'y annule avec une dérivée infinie (voir références 1, chapitre 8 et 2, complément C).

Pour ces électrons de conduction des métaux (et plus particulièrement des métaux alcalins et alcalinoterreux), une assez bonne approximation consiste à les supposer entièrement libres à l'intérieur du cristal, c'est-à-dire à négliger totalement le potentiel périodique du réseau[14]. Les niveaux d'énergie sont alors

$$\varepsilon_c(\vec{k}) = \frac{\hbar^2 \vec{k}^2}{2m}, \tag{F.24}$$

et les fonctions d'onde correspondantes des ondes planes. Dans ce modèle la notion de zone de Brillouin est sans objet puisqu'on néglige l'effet de la structure périodique du réseau cristallin[15].

La densité d'états est donc celle calculée au paragraphe III.B.2.a du chapitre I pour une particule libre dans une boîte de volume V (au facteur de spin près)[16] :

$$\rho_c(\varepsilon) = \frac{V\sqrt{2}}{\pi^2\hbar^3} m^{3/2} \varepsilon^{1/2}. \tag{F.25}$$

Elle est représentée sur la figure 5.

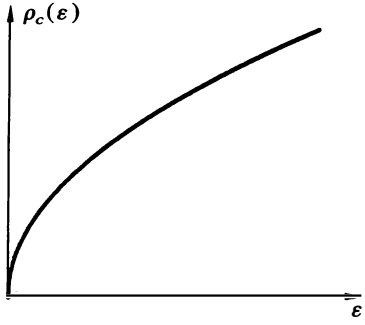


FIGURE 5
Densité d'états pour les électrons de la bande de conduction d'un métal, dans le modèle des électrons libres.

Il est intéressant de noter que l'on obtient un résultat semblable pour une *bande quelconque* si on ne s'intéresse qu'*au voisinage d'un de ses bords.* En effet le développement de $\varepsilon_n(\vec{k})$, par exemple au voisinage de son minimum, s'écrit (avec un choix judicieux des coordonnées)

$$\varepsilon_n(\vec{k}) = \varepsilon_0 + \frac{\hbar^2}{2}\left[\frac{k_x^2}{m_x} + \frac{k_y^2}{m_y} + \frac{k_z^2}{m_z}\right] + \dots, \tag{F.26}$$

14. Évidemment, il existe dans les métaux d'autres bandes, provenant des couches atomiques profondes (cf. note 9). Les électrons qui occupent de tels états sont beaucoup plus sensibles au potentiel du réseau, qui n'est alors en aucun cas négligeable.

15. Le modèle peut être amélioré en traitant $V(\vec{r})$ comme une perturbation : c'est le modèle des *électrons presque libres* (voir référence 1, chapitre 9).

16. Les expressions (F.24) et (F.25) supposent que l'on repère les énergies à partir du bas de la bande de conduction. Si le zéro des énergies est placé ailleurs, (F.24) devient

$$\varepsilon_c(\vec{k}) = \frac{\hbar^2 \vec{k}^2}{2m} + \varepsilon_0, \tag{1}$$

et (F.25)

$$\rho_c(\varepsilon) = \frac{V\sqrt{2}}{\pi^2\hbar^3} m^{3/2} (\varepsilon - \varepsilon_0)^{1/2}, \tag{2}$$

où ε_0 est l'énergie minimale des états de la bande de conduction.

où m_x, m_y, m_z sont trois coefficients positifs, ayant les dimensions d'une masse (ils forment ce qu'on appelle le «tenseur de masse effective»). La densité d'états correspondante, au voisinage de ε_0, est alors approximativement égale à

$$\rho_n(\varepsilon) = \frac{V}{\pi^2\hbar^3}(2m_x m_y m_z)^{1/2}(\varepsilon - \varepsilon_0)^{1/2} . \tag{F.27}$$

2. LE MODÈLE DES LIAISONS FORTES

Les électrons des couches profondes d'un atome restent *fortement liés* au noyau de cet atome particulier lorsqu'il fait partie d'un cristal : la probabilité que ces électrons «sautent» vers un noyau voisin est très faible, et leurs fonctions d'onde sont essentiellement localisées autour d'un site atomique. Il en est de même pour les électrons de valence lorsqu'il s'agit d'éléments isolants.

Pour de tels électrons le modèle des liaisons fortes est une bonne approximation; contrairement au précédent, il prend en compte comme effet principal le potentiel attractif des noyaux. Les fonctions d'onde sont écrites comme des combinaisons linéaires d'orbitales atomiques et l'écart entre le potentiel périodique du réseau et le potentiel atomique de chaque noyau est traité en perturbation[17]. On détermine ainsi une bande d'états pour chaque état atomique de départ.

---

17. Considérons tout d'abord un électron dans le champ d'un *atome isolé* caractérisé par le potentiel $U_{at}(\vec{r})$ centré en $\vec{r} = \vec{0}$; soit $\varphi_{at}(\vec{r})$ une fonction d'onde de cet électron et ε_{at} l'énergie correspondante. Nous supposons que $\varphi_{at}(\vec{r})$ est fortement localisée (elle décroît très vite lorsque r croît, sur une distance caractéristique inférieure à la distance interatomique dans le cristal). La fonction d'onde $\psi_{\vec{k}}(\vec{r})$ d'un électron dans le cristal est alors approximée par une *combinaison linéaire d'orbitales atomiques* centrées sur les noyaux du cristal; si ces divers noyaux sont repérés par les vecteurs $\vec{R}_j$ ($j = 1, 2, 3, 4, \ldots$), on obtient en utilisant le théorème de Bloch

$$\psi_{\vec{k}}(\vec{r}) = \sum_j e^{i\vec{k}.\vec{R}_j}\varphi_{at}(\vec{r} - \vec{R}_j) . \tag{1}$$

En effet $\psi_{\vec{k}}(\vec{r})$ est bien de la forme (F.7), à condition que

$$u_{\vec{k}}(\vec{r}) = \psi_{\vec{k}}(\vec{r})e^{-i\vec{k}.\vec{r}} = \sum_j e^{-i\vec{k}.(\vec{r}-\vec{R}_j)}\varphi_{at}(\vec{r} - \vec{R}_j) \tag{2}$$

vérifie la condition de périodicité (F.7.b). Or

$$u_{\vec{k}}(\vec{r} + \vec{R}) = \sum_j e^{-i\vec{k}.(\vec{r}+\vec{R}-\vec{R}_j)}\varphi_{at}(\vec{r} + \vec{R} - \vec{R}_j) . \tag{3}$$

Si $\vec{R}$ est de la forme (F.1), $(\vec{R}_j - \vec{R})$ repère un nœud du réseau de Bravais tout comme $\vec{R}_j$; si le cristal est infini, l'ensemble des vecteurs $(\vec{R}_j - \vec{R})$ coïncide avec l'ensemble des vecteurs $\vec{R}_j$, de sorte que

$$u_{\vec{k}}(\vec{r} + \vec{R}) = u_{\vec{k}}(\vec{r}) . \tag{4}$$

La dimension finie du cristal est ensuite prise en compte en imposant les conditions de Born-von Karman qui limitent les valeurs de $\vec{k}$ permises. En pratique, la faible portée des fonctions d'onde atomiques réduit la somme (1) à quelques termes : pour une valeur donnée de $\vec{r}$, seuls les noyaux les plus proches contribuent à cette somme.

A partir des fonctions d'onde approchées (1), on détermine les énergies $\varepsilon(\vec{k})$ correspondantes en calculant la valeur moyenne quantique de l'hamiltonien

$$H = -\frac{\hbar^2}{2m}\Delta + V(\vec{r}) \tag{5}$$

.../...

Pour un cristal de structure *cubique simple* et pour des *fonctions atomiques à symétrie sphérique* (orbitales s) on trouve[18]

$$\varepsilon(\vec{k}) = \varepsilon_0 + \varepsilon_1(\cos k_x a + \cos k_y a + \cos k_z a)\,, \tag{F.28}$$

d'un électron dans le solide :

$$\varepsilon(\vec{k}) = \frac{\int \mathrm{d}^3 r\, \psi_{\vec{k}}^*(\vec{r})\, H \psi_{\vec{k}}(\vec{r})}{\int \mathrm{d}^3 r\, \psi_{\vec{k}}^*(\vec{r})\, \psi_{\vec{k}}(\vec{r})}\,, \tag{6}$$

c'est-à-dire

$$\varepsilon(\vec{k}) = \frac{\sum\limits_{i,\,j} \mathrm{e}^{i\vec{k}\cdot(\vec{R}_j - \vec{R}_i)} \int \mathrm{d}^3 r\, \varphi_{\mathrm{at}}^*(\vec{r} - \vec{R}_i)\, H \varphi_{\mathrm{at}}(\vec{r} - \vec{R}_j)}{\sum\limits_{i,\,j} \mathrm{e}^{i\vec{k}\cdot(\vec{R}_j - \vec{R}_i)} \int \mathrm{d}^3 r\, \varphi_{\mathrm{at}}^*(\vec{r} - \vec{R}_i)\, \varphi_{\mathrm{at}}(\vec{r} - \vec{R}_j)}\,. \tag{7}$$

Au voisinage de chaque ion du réseau, $V(\vec{r})$ coïncide avec le potentiel atomique $U_{\mathrm{at}}(\vec{r} - \vec{R}_i)$ de cet ion; ce n'est qu'entre les ions que U_{at} et V diffèrent sensiblement, mais dans une région où la probabilité de présence de l'électron est faible, ce qui permet de traiter l'écart entre U_{at} et V comme une perturbation. On remplace alors dans (7) H par

$$H = H_{\mathrm{at}}^j + (H - H_{\mathrm{at}}^j) = H_{\mathrm{at}}^j + V(\vec{r}) - U_{\mathrm{at}}(\vec{r} - \vec{R}_j)\,, \tag{8}$$

où H_{at}^j est l'hamiltonien pour un atome isolé placé au nœud j du réseau, et on ne garde que les intégrales qui font intervenir des fonctions atomiques centrées sur le même site ou sur deux sites voisins (les autres sont négligeables). On obtient ainsi, en utilisant le fait que les fonctions d'onde atomiques sont normées,

$$\varepsilon(\vec{k}) = \varepsilon_{\mathrm{at}} + \frac{\alpha + \sum\limits_{\mathrm{p.v.}} \beta(\vec{R})\, \mathrm{e}^{i\vec{k}\cdot\vec{R}}}{1 + \sum\limits_{\mathrm{p.v}} \gamma(\vec{R})\, \mathrm{e}^{i\vec{k}\cdot\vec{R}}}\,, \tag{9}$$

où les sommes ne portent que sur les vecteurs $\vec{R}$ du réseau de Bravais qui connectent l'origine à ses plus proches voisins (p.v.), et

$$\alpha \mathrel{\widehat{=}} \int \mathrm{d}^3 r\, \varphi_{\mathrm{at}}^*(\vec{r})\,[V(\vec{r}) - U_{\mathrm{at}}(\vec{r})]\, \varphi_{\mathrm{at}}(\vec{r})\,, \tag{10}$$

$$\beta(\vec{R}) \mathrel{\widehat{=}} \int \mathrm{d}^3 r\, \varphi_{\mathrm{at}}^*(\vec{r})\,[V(\vec{r}) - U_{\mathrm{at}}(\vec{r} - \vec{R})]\, \varphi_{\mathrm{at}}(\vec{r} - \vec{R})\,, \tag{11}$$

$$\gamma(\vec{R}) \mathrel{\widehat{=}} \int \mathrm{d}^3 r\, \varphi_{\mathrm{at}}^*(\vec{r})\, \varphi_{\mathrm{at}}(\vec{r} - \vec{R})\,. \tag{12}$$

Les valeurs de α et $\beta(\vec{R})$ sont petites devant $\varepsilon_{\mathrm{at}}$; de même le terme en $\gamma(\vec{R})$ est petit devant l'unité et peut être négligé. Finalement, si l'on pose $\varepsilon_0 = \varepsilon_{\mathrm{at}} + \alpha$, on trouve

$$\varepsilon(\vec{k}) \simeq \varepsilon_0 + \sum_{\mathrm{p.v.}} \beta(\vec{R})\, \mathrm{e}^{i\vec{k}\cdot\vec{R}}\,. \tag{13}$$

18. Pour un cristal cubique simple, il y a six proches voisins de l'origine, tels que

$$\vec{R} = (\pm a, 0, 0),\ (0, \pm a, 0),\ (0, 0, \pm a)\,, \tag{1}$$

et pour des fonctions d'onde à symétrie sphérique, le coefficient $\beta(\vec{R})$ défini dans la note 17 est le même pour ces six $\vec{R}$. Dans ce cas, l'expression (13) de la note précédente se réduit à

$$\begin{aligned} \varepsilon(\vec{k}) &= \varepsilon_0 + \beta \sum_{\mathrm{p.v.}} \mathrm{e}^{i\vec{k}\cdot\vec{R}} \\ &= \varepsilon_0 + 2\beta\,[\cos k_x a + \cos k_y a + \cos k_z a]\,. \end{aligned} \tag{2}$$

où ε_0 et ε_1 sont deux constantes, et a le côté de la maille élémentaire cubique sur laquelle est construit le cristal.

La bande permise correspondante s'étale donc entre les valeurs extrêmes $(\varepsilon_0 - 3\varepsilon_1)$ et $(\varepsilon_0 + 3\varepsilon_1)$ de l'énergie.

Pour calculer la densité d'états dans cette bande, nous utilisons la formule générale (F.19) dans laquelle nous remplaçons la fonction de Dirac par sa représentation intégrale

$$\delta(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \mathrm{d}t \; \mathrm{e}^{ixt} \; . \tag{F.29}$$

La première zone de Brillouin étant un cube de côté $2\pi/a$, on obtient

$$\begin{aligned} \rho(\varepsilon) &= \frac{V}{8\pi^4} \int_{-\pi/a}^{\pi/a} \mathrm{d}k_x \int_{-\pi/a}^{\pi/a} \mathrm{d}k_y \int_{-\pi/a}^{\pi/a} \mathrm{d}k_z \int_{-\infty}^{+\infty} \mathrm{d}t \; \mathrm{e}^{i[\varepsilon - \varepsilon(\vec{k})]t} \\ &= \frac{V}{\pi} \int_{-\infty}^{+\infty} \mathrm{d}t \; \mathrm{e}^{i(\varepsilon - \varepsilon_0)t} \left[\frac{1}{2\pi} \int_{-\pi/a}^{\pi/a} \mathrm{d}k \; \mathrm{e}^{-i\varepsilon_1 \cos ka} \right]^3 . \end{aligned} \tag{F.30}$$

La quantité entre crochets fait intervenir la fonction de Bessel d'ordre zéro

$$J_0(x) = \frac{1}{\pi} \int_0^{\pi} \mathrm{d}q \; \mathrm{e}^{ix \cos q} = \frac{1}{\pi} \int_0^{\pi} \mathrm{d}q \; \cos(x \cos q) \; , \tag{F.31}$$

à l'aide de laquelle (F.30) s'écrit

$$\rho(\varepsilon) = \frac{V}{\pi a^3} \int_{-\infty}^{+\infty} \mathrm{d}t \; \mathrm{e}^{i(\varepsilon - \varepsilon_0)t} J_0^3(\varepsilon_1 t) = \frac{V}{\pi a^3} \int_{-\infty}^{+\infty} \mathrm{d}t \left[\cos(\varepsilon - \varepsilon_0) t \right] J_0^3(\varepsilon_1 t) \; . \tag{F.32}$$

Il reste une intégrale, à calculer numériquement, pour déterminer la densité d'états cherchée. Le résultat est représenté sur la figure 6[19], [20].

---

19. La densité d'états n'est différente de zéro que pour les valeurs de l'énergie comprises entre les limites de la bande : $(\varepsilon_0 - 3\varepsilon_1)$ et $(\varepsilon_0 + 3\varepsilon_1)$. Elle présente des singularités de van Hove (cf. note 13) aux points où $\vec{\nabla}\varepsilon(\vec{k})$ s'annule, c'est-à-dire pour les valeurs de $\vec{k}$ de la première zone de Brillouin telles que

$$\begin{cases} k_x = 0 \quad \text{ou} \quad \pi/a \\ k_y = 0 \quad \text{ou} \quad \pi/a \\ k_z = 0 \quad \text{ou} \quad \pi/a \; , \end{cases}$$

ce qui correspond aux 4 valeurs de l'énergie :

$$\varepsilon_0 - 3\varepsilon_1, \quad \varepsilon_0 - \varepsilon_1, \quad \varepsilon_0 + \varepsilon_1, \quad \varepsilon_0 + 3\varepsilon_1 \; .$$

20. Comme il se doit (cf. note 9), l'intégrale de la densité d'états dans une bande égale le nombre N de sites dans le cristal multiplié par 2 (à cause des deux états de spin possibles pour chaque état atomique de départ). En effet

$$\Phi \widehat{=} \int_{-\infty}^{+\infty} \rho(\varepsilon) \, \mathrm{d}\varepsilon = \frac{V}{\pi a^3} \int_{-\infty}^{+\infty} \mathrm{d}\varepsilon \int_{-\infty}^{+\infty} \mathrm{d}t \; \mathrm{e}^{i(\varepsilon - \varepsilon_0)t} J_0^3(\varepsilon_1 t) \; . \tag{1}$$

Changeons l'ordre des intégrations et utilisons la relation (F.29); il vient

$$\Phi = \frac{V}{\pi a^3} \int_{-\infty}^{+\infty} \mathrm{d}t \; \mathrm{e}^{-i\varepsilon_0 t} J_0^3(\varepsilon_1 t) \times 2\pi\delta(t) = \frac{2V}{a^3} J_0^3(0) \; . \tag{2}$$

On vérifie immédiatement sur la relation (F.31) que $J_0(0) = 1$, d'où

$$\Phi = \frac{2V}{a^3} = 2N \; . \tag{3}$$

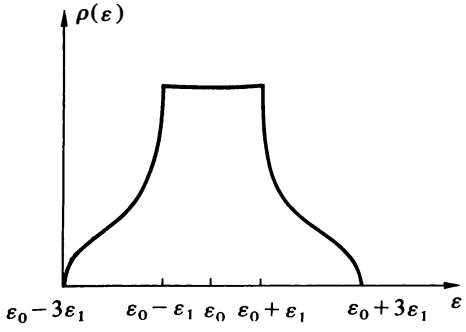


FIGURE 6
Densité d'états pour un cristal cubique simple, dans le modèle des liaisons fortes.

Références

1. N.W. Ashcroft et N.D. Mermin, *Solid State Physics,* Holt-Saunders, New York (1981).

2. C. Kittel, *Introduction à la physique de l'état solide,* Dunod, Paris (1972).

COMPLÉMENT I.G

Information et entropie

I. Information et mesure de l'information (distributions discrètes de probabilités)
  1. Information. Manque d'information
  2. Mesure du manque d'information : l'entropie statistique
    a. Hypothèses
    b. Expression de l'entropie statistique pour un ensemble d'événements équiprobables
    c. Expression de l'entropie statistique pour une distribution quelconque
  3. Exemples simples

II. Propriétés de l'entropie statistique
  1. Inégalité comparant deux distributions
  2. Inégalité de « sous-additivité »
  3. Inégalité de concavité

III. Mesure de l'information manquante dans le cas d'une distribution continue de probabilité
  1. Discrétisation du problème
  2. Difficulté du passage à la limite continue
    a. Divergence de la formule fondamentale
    b. Les deux solutions possibles
  3. Choix de la variable pertinente

IV. Discussion : lien entre information et mécanique statistique

I. INFORMATION ET MESURE DE L'INFORMATION (DISTRIBUTIONS DISCRÈTES DE PROBABILITÉS)

Nous avons donné dans le chapitre I la définition et les principales propriétés de l'entropie statistique, qui mesure l'information manquante associée à une distribution de probabilités. Nous reprenons ici la question pour montrer que, moyennant un petit nombre d'hypothèses plausibles, l'entropie statistique est *la seule fonction susceptible de mesurer l'information manquante.*

1. INFORMATION. MANQUE D'INFORMATION

Prenons un exemple simple. Supposons que nous cherchions un objet dont nous savons seulement qu'il se trouve dans l'un des M tiroirs de la maison. Cette *information* est clairement insuffisante pour nous permettre de localiser l'objet à coup sûr, sans tâtonnements. Notons cependant qu'elle n'est pas sans valeur : parmi tous les endroits où pourrait *a priori* se trouver l'objet (dans un autre appartement, au bureau des objets trouvés, ...), nous n'avons à envisager que les M tiroirs.

Pour mesurer quantitativement l'information, il faut en premier lieu fixer un *point de référence.* Il est naturel de choisir pour cela, dans chaque cas particulier, la situation où l'on posséderait une *information complète :* dans l'exemple précédent, notre information serait complète si nous savions lequel des M tiroirs contient l'objet cherché. Par rapport à cette situation de référence, les autres situations font apparaître un certain *manque d'information;* c'est celui-ci que l'on va évaluer.

Si l'on sait mesurer le manque d'information (ou *« information manquante »*), on sait par là même mesurer l'information fournie par une expérience : il suffit de calculer la différence entre le manque d'information initial et le manque d'information restant après l'expérience. Ainsi, dans notre exemple, l'ouverture d'un tiroir apporte une certaine information : si l'objet se trouve dans ce tiroir, nous acquérons d'un coup toute l'information qui nous manquait; mais même si le tiroir que nous avons ouvert est vide, l'information acquise n'est pas nulle : le manque d'information a été réduit, puisque nous n'avons plus que $(M-1)$ tiroirs à explorer.

2. MESURE DU MANQUE D'INFORMATION : L'ENTROPIE STATISTIQUE

Considérons une situation dans laquelle M événements e_m peuvent se produire, avec des probabilités P_m $(m = 1, 2, \ldots, M)$. Nous voulons déterminer la fonction $S(P_1, P_2, \ldots, P_M)$ qui va mesurer le manque d'information lié à la distribution de probabilités $\{P_m\}$, à partir de quelques propriétés qu'il paraît naturel de lui imposer.

a. Hypothèses

(i) La situation de référence, dans laquelle l'information est complète et donc le manque d'information nul, est celle où l'une des probabilités vaut 1, toutes les autres étant nulles :

$$S(0, 0, \ldots, 1, \ldots, 0) = 0 \ . \tag{G.1}$$

(ii) Dans toutes les autres situations, nous prendrons S *positive.* Nous supposerons que c'est une fonction *continue* de ses diverses variables, et qu'elle est *symétrique,* c'est-à-dire inchangée dans une permutation quelconque de ses variables : par exemple,

$$S(P_2, P_1, P_3, \ldots, P_M) = S(P_1, P_2, P_3, \ldots, P_M) \ . \tag{G.2}$$

(iii) Les *événements impossibles* apportent une *contribution nulle* au manque d'information, et donc à S : si par exemple on rajoute un $(M+1)$-ième événement de probabilité nulle aux M événements considérés, le manque d'information est clairement inchangé dans cette opération, qui ne modifie pas en fait la situation initiale :

$$S(P_1, P_2, \ldots, P_M, 0) = S(P_1, P_2, \ldots, P_M) \ . \tag{G.3}$$

(iv) Dans le cas d'*événements équiprobables,* il est évident que le manque d'information croît avec le nombre de ces événements :

$$S\left(P_1 = P_2 = \ldots = P_M = \frac{1}{M}\right) \nearrow \text{ quand } M \nearrow \ . \tag{G.4}$$

(v) Il faut enfin imposer une propriété d'*additivité.*

Considérons tout d'abord *deux ensembles d'événements indépendants :* e_m $(m = 1, 2, \ldots, M)$, de probabilités P_m, et $e'_{m'}$ $(m' = 1, 2, \ldots, M')$, de probabilités $P'_{m'}$. Envisageons une expérience «double», dans laquelle se produisent à la fois un événement du premier ensemble et un événement du deuxième. La probabilité pour que se réalisent e_m et $e'_{m'}$ est simplement le produit $P_m \, . \, P'_{m'}$. Le manque d'information sur l'ensemble des $M \, . \, M'$ événements doubles (e_m et $e'_{m'}$) ne peut être dans ce cas que la somme des manques d'information apportés par les deux ensembles pris séparément. On va donc imposer la condition

$$S\left(\{P_m \, . \, P'_{m'};\ m = 1, 2, \ldots, M;\ m' = 1, 2, \ldots, M'\}\right) = S(P_1, P_2, \ldots, P_M) + S(P'_1, P'_2, \ldots, P'_{M'}) \ . \tag{G.5}$$

Mais nous aurons besoin d'une *condition d'additivité plus générale.* Reprenons d'abord l'exemple du paragraphe 1, en le compliquant un peu : chacun des tiroirs contient des boîtes, et c'est dans une de ces boîtes que se cache l'objet. Nous repérons les tiroirs par leur numéro m $(m = 1, 2, \ldots, M)$, et supposons connues les probabilités P_m pour que l'objet se trouve dans les divers tiroirs (probabilités qui ne sont plus nécessairement égales). En outre, le tiroir (m) contient M'_m boîtes (m') $(m' = 1, 2, \ldots, M'_m)$, et l'on se donne les probabilités $P'^{(m)}_{m'}$ pour que, *si l'objet se trouve dans le tiroir* (m), ce soit la boîte (m') qui le recèle. Pour formuler le problème en des termes plus généraux, considérons un ensemble de M *événements «primaires»* e_m (correspondant aux tiroirs), de probabilités P_m, normalisées comme à l'habitude :

$$\sum_{m=1}^{M} P_m = 1 \ . \tag{G.6}$$

A chacun de ces événements est associé un ensemble d'*événements «secondaires»* (les boîtes) $e'^{(m)}_{m'}$ ($m' = 1, 2, \ldots, M'_m$); la *probabilité conditionnelle* pour que l'événement «secondaire» $e'^{(m)}_{m'}$ se produise lorsque l'événement «primaire» e_m qui le commande est déjà réalisé est notée $P'^{(m)}_{m'}$; ces probabilités conditionnelles sont elles aussi normalisées, séparément pour chaque valeur de m :

$$\sum_{m'=1}^{M'_m} P'^{(m)}_{m'} = 1 \quad \text{pour tout } m. \tag{G.7}$$

En définitive, on étudie donc un ensemble de

$$N = \sum_{m=1}^{M} M'_m \tag{G.8}$$

événements «doubles» (constitués chacun d'un événement e_m suivi d'un des événements $e'^{(m)}_{m'}$ associés), dont les probabilités valent

$$P_{m,m'} = P_m \cdot P'^{(m)}_{m'} . \tag{G.9}$$

Dans un cas de ce genre, nous demandons à la fonction S de vérifier la condition

$$S(\{P_{m,m'};\ m = 1, 2, \ldots, M;\ m' = 1, 2, \ldots, M'_m\}) = S(P_1, P_2, \ldots, P_M) + \sum_{m=1}^{M} P_m S(P'^{(m)}_1, P'^{(m)}_2, \ldots, P'^{(m)}_{M'_m}). \tag{G.10}$$

Cette égalité se comprend facilement, sauf peut-être la présence du facteur P_m (probabilité de l'événement e_m qui «commande» l'ensemble des événements $e'^{(m)}_{m'}$) pondérant le manque d'information associé à la distribution de probabilités $\{P'^{(m)}_{m'},\ m' = 1, 2, \ldots, M'_m\}$. Pourtant, si l'on y réfléchit, cette pondération est nécessaire : s'il se trouve que, pour une valeur donnée de m, P_m est nulle, tous les événements $e'^{(m)}_{m'}$ correspondant à cette valeur de m sont impossibles, et le manque d'information associé à leurs probabilités conditionnelles ne doit pas contribuer au manque d'information global (propriété (iii) ci-dessus); il doit au contraire contribuer à plein dans le cas où $P_m = 1$[1].

La condition générale d'additivité (G.10) contient (G.5) comme cas particulier. Supposons en effet que l'ensemble des événements $e'^{(m)}_{m'}$, et leurs probabilités conditionnelles $P'^{(m)}_{m'}$, soient indépendants de m. Ceci revient à dire qu'on a alors deux ensembles d'événements indépendants, $\{e_m\}$ et $\{e'_{m'}\}$, de probabilités $\{P_m\}$ et $\{P'_{m'}\}$. La formule (G.9) devient dans ce cas

$$P_{m,m'} = P_m \cdot P'_{m'} , \tag{G.11}$$

et l'égalité (G.10) se réduit à (G.5).

b. Expression de l'entropie statistique pour un ensemble d'événements équiprobables

Supposons d'abord que les M événements e_m considérés soient équiprobables :

$$P_1 = P_2 = \ldots = P_M = \frac{1}{M} . \tag{G.12}$$

Nous utiliserons dans ce cas une notation simplifiée :

$$I(M) \widehat{=} S\left(P_1 = P_2 = \ldots = P_M = \frac{1}{M}\right) . \tag{G.13}$$

---

1. Cet argument n'exige pas que le manque d'information $S(\{P'^{(m)}_{m'}\})$ soit multiplié par P_m elle-même; il pourrait l'être par une fonction $f(P_m)$ positive quelconque, pourvu que $f(0) = 0$ et $f(1) = 1$. Mais il faut aussi, pour que la relation (G.10) redonne (G.5) dans le cas de distributions indépendantes, que

$$\sum_{m=1}^{M} f(P_m) = 1 ,$$

ce qui n'est possible que si $f(P_m) \equiv P_m$.

L'hypothèse (iv) du paragraphe a indique que $I(M)$ est une fonction croissante de M :

$$I(M') > I(M) \quad \text{si} \quad M' > M . \tag{G.14}$$

Si l'on prend *deux ensembles indépendants d'événements équiprobables* (M événements de probabilités $1/M$ d'une part, et M' événements de probabilités $1/M'$ d'autre part), l'hypothèse (G.5) entraîne la relation fonctionnelle

$$I(M.M') = I(M) + I(M') . \tag{G.15}$$

Cette relation est *caractéristique de la fonction logarithme*[2], c'est-à-dire qu'elle implique

$$I(M) = \lambda \ln M , \tag{G.16}$$

où λ est une constante, positive puisque I doit l'être (hypothèse (ii)).

c. Expression de l'entropie statistique pour une distribution quelconque

Soit une distribution de probabilités $\{P_m ; m = 1, 2, \ldots, M\}$ quelconque, normalisée. Pour évaluer le manque d'information correspondant, nous allons imaginer une situation où nous puissions utiliser la propriété d'additivité (G.10) et la formule (G.16) que nous venons d'obtenir pour les distributions équiprobables.

Le raisonnement suppose que les M probabilités P_m sont des *nombres rationnels.* Du point de vue physique, ceci n'est guère une restriction car il n'est pas possible de distinguer entre un nombre irrationnel ou transcendant et son approximation rationnelle si celle-ci est suffisamment précise. Du point de vue mathématique, la *continuité* postulée pour la fonction S (propriété (ii)) permet d'étendre à tous les réels le résultat que nous allons démontrer.

Supposons donc que les P_m sont des rationnels. En les réduisant au même dénominateur, on peut les mettre sous la forme

$$P_m = \frac{Q_m}{Q} , \tag{G.17}$$

où Q_m et Q sont des *entiers* auxquels la condition de normalisation impose la relation

$$\sum_{m=1}^{M} Q_m = Q . \tag{G.18}$$

Considérons alors un ensemble de Q *événements équiprobables.* Nous le subdivisons en M *sous-ensembles*[3], repérés par m, de sorte que le sous-ensemble (m) comprenne Q_m événements. Nous imaginons que les expériences au cours desquelles se produisent les Q événements équiprobables peuvent être réalisées en deux temps. Dans une première phase, on choisit de façon aléatoire l'un des sous-ensembles (m), auxquels sont affectées les *probabilités initiales* P_m; ainsi, les sous-ensembles (m) constituent les M événements e_m que nous voulons étudier. Ensuite, une fois le sous-ensemble choisi, nous tirons au hasard l'un des événements qui le composent, et qui sont considérés ici comme les événements $e'^{(m)}_{m'}$ du paragraphe a.(v); nous prenons ces événements *équiprobables dans le sous-ensemble*, c'est-à-dire tels que leurs probabilités conditionnelles soient égales :

$$P'^{(m)}_{m'} = \frac{1}{Q_m} ; \quad m' = 1, 2, \ldots, Q_m . \tag{G.19}$$

---

2. Le fait que la relation est vraie seulement pour M *et* M' *entiers* pose un petit problème technique mais ne change pas le résultat (réf. 1, p. 95).

3. Notons que la relation (G.18) implique

$$Q \geq M .$$

Finalement, grâce à la relation (G.17), on retombe ainsi sur les probabilités égales des Q événements initiaux : l'égalité (G.9), compte tenu de (G.19), donne ici

$$P_{m,m'}=\frac{1}{Q}\,. \tag{G.20}$$

Il suffit maintenant d'appliquer la relation d'additivité (G.10) au cas (G.20), (G.19), (G.17) :

$$I(Q)=S(P_1, P_2, \ldots, P_M)+\sum_{m=1}^{M} P_m I(Q_m)\,. \tag{G.21}$$

L'expression (G.16) de la fonction $I(M)$ permet d'en tirer aussitôt[4] la *formule de Shannon :*

$$\boxed{S(P_1, P_2, \ldots, P_M)=-\lambda \sum_{m=1}^{M} P_m \ln P_m}\,. \tag{G.22}$$

3. EXEMPLES SIMPLES

En mécanique statistique, on choisit pour λ *la constante de Boltzmann k*, ce qui permet d'identifier l'entropie statistique associée à l'état macroscopique d'un système à l'équilibre avec l'entropie thermodynamique introduite initialement par Clausius (1850). L'entropie thermodynamique s'interprète ainsi comme *la mesure du manque d'information concernant l'état microscopique* d'un système macroscopique à l'équilibre.

En théorie de l'information, on a coutume de prendre

$$\lambda=\frac{1}{\ln 2}\,, \tag{G.23}$$

de sorte que le manque d'information associé à un ensemble de 2 événements équiprobables (tels que pile ou face) ait pour valeur 1. Cette unité s'appelle le *bit* (contraction des mots anglais « binary digit », qui signifient « chiffre binaire »). Le manque d'information lié au jeu de dés est supérieur à celui de pile ou face, puisque le nombre d'événements possibles est supérieur : pour un dé, il vaut

$$I(6)=\frac{\ln 6}{\ln 2}\simeq 2{,}58 \text{ bit}\,, \tag{G.24}$$

et pour un ensemble de n dés il est n fois plus grand :

$$I(6^n)=n\,I(6)\,. \tag{G.25}$$

Évaluons l'information que peut *a priori* apporter la lecture d'une page de livre, c'est-à-dire l'information manquante associée à l'ensemble de tous les textes d'une page *a priori* possibles. Une telle page comporte environ 4 000 signes, si l'on entend par « signe » l'une des 26 lettres de l'alphabet ou un espace blanc. Il est donc possible *a priori* de composer

$$M=27^{4\,000} \tag{G.26}$$

pages différentes, ce qui correspond à une information manquante

$$I(27^{4\,000})=4\,000\,I(27)\simeq 19\,000 \text{ bit}\,, \tag{G.27}$$

soit environ 4,7 bit par signe. Mais ce calcul suppose tous les signes équiprobables,

---

4. En effet,

$$S(P_1, P_2, \ldots, P_M)=\lambda \ln Q-\sum_{m=1}^{M} P_m\lambda \ln Q_m=-\lambda \sum_{m=1}^{M} P_m(\ln Q_m-\ln Q)$$

(où l'on a utilisé la condition de normalisation des P_m). Compte tenu de l'égalité (G.17), on aboutit à la formule (G.22).

ce qui est loin d'être vrai. Nous avons donc surestimé l'information manquante; si l'on tient compte des probabilités, dans la langue française, des 27 signes définis ci-dessus, la formule (G.22) donne environ 4 bit par signe. Il s'agit là encore d'une évaluation par excès, car on y néglige les corrélations entre signes (cf. paragraphe II.2) : la lettre q n'est jamais suivie de k, mais souvent de u. Lorsque, par un calcul plus compliqué, on prend ces corrélations en compte, on trouve que l'information par signe est en réalité nettement plus faible (comprise entre 1 et 2 bit).

II. PROPRIÉTÉS DE L'ENTROPIE STATISTIQUE

Outre les propriétés que nous avons décrites au paragraphe II.B du chapitre I, et dont certaines nous ont servi ici d'hypothèses, l'entropie statistique vérifie trois inégalités importantes.

1. INÉGALITÉ COMPARANT DEUX DISTRIBUTIONS

Considérons un ensemble d'événements e_m $(m = 1, 2, .., M)$ et deux distributions de probabilités possibles pour ces événements, $\{P_m\}$ et $\{P'_m\}$, séparément normalisées. On démontre la relation[5]

$$S(P_m) \leqslant -\lambda \sum_{m=1}^{M} P_m \ln P'_m , \tag{G.28}$$

5. La démonstration est fondée sur l'inégalité bien connue (voir figure 1) :

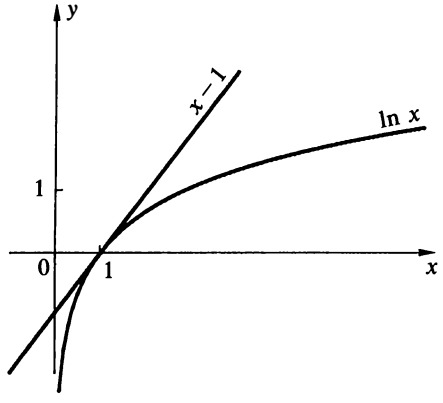


FIGURE 1
Comparaison graphique entre la fonction ln x et la fonction $x - 1$.

$$\ln x \leqslant x - 1 , \tag{1}$$

qui devient égalité si et seulement si $x = 1$. On peut en effet écrire

$$S(P_m) + \lambda \sum_{m=1}^{M} P_m \ln P'_m = \lambda \sum_{m=1}^{M} P_m \ln \frac{P'_m}{P_m} . \tag{2}$$

Appliquons l'inégalité (1) au second membre :

$$S(P_m) + \lambda \sum_{m=1}^{M} P_m \ln P'_m \leqslant \lambda \sum_{m=1}^{M} P_m \left(\frac{P'_m}{P_m} - 1 \right) , \tag{3}$$

avec égalité si et seulement si

$$\frac{P'_m}{P_m} = 1 \quad \text{pour tout } m . \tag{4}$$

On en déduit le résultat (G.28), car

$$\sum_{m=1}^{M} P_m \left(\frac{P'_m}{P_m} - 1 \right) = \sum_{m=1}^{M} P'_m - \sum_{m=1}^{M} P_m = 0 \tag{5}$$

à cause de la normalisation des deux distributions.

dans laquelle l'égalité a lieu si et seulement si

$$P'_m = P_m \quad \text{pour tout } m. \tag{G.29}$$

Quoique difficile à interpréter intuitivement, cette inégalité est utile dans bien des démonstrations.

2. INÉGALITÉ DE « SOUS-ADDITIVITÉ »

Considérons deux ensembles d'événements, $\{e_m;\ m = 1, 2, \ldots, M\}$ et $\{e'_{m'};\ m' = 1, 2, \ldots, M'\}$, et envisageons des *expériences doubles*, dans chacune desquelles se produisent un événement e_m et un événement $e'_{m'}$. Soit $P_{m,m'}$ la probabilité de l'événement composé $(e_m, e'_{m'})$. A cette distribution de probabilités $\{P_{m,m'};\ m = 1, 2, \ldots, M;\ m' = 1, 2, \ldots, M'\}$ est associée l'entropie statistique $S(P_{m,m'})$.

La relation d'additivité (G.10) est évidemment toujours valable[6]. Mais nous allons prendre ici un point de vue légèrement différent.

Supposons d'abord que, dans chacune de ces expériences doubles, nous regardions seulement quel événement du premier ensemble s'y réalise. La probabilité P_m de l'événement e_m, quel que soit par ailleurs l'événement $e'_{m'}$ qui l'accompagne, vaut

$$P_m = \sum_{m'=1}^{M'} P_{m,m'} \,. \tag{G.30}$$

De même, si nous observons seulement les événements $e'_{m'}$, ils ont pour probabilités

$$P'_{m'} = \sum_{m=1}^{M} P_{m,m'} \,. \tag{G.31}$$

Aux distributions de probabilités $\{P_m\}$ et $\{P'_{m'}\}$ sont associées des entropies statistiques $S(P_m)$ et $S(P'_{m'})$ respectivement.

On a alors de façon générale[7]

$$S(P_{m,m'}) \leqslant S(P_m) + S(P'_{m'}) \,, \tag{G.32}$$

---

6. On peut en effet, après avoir introduit les probabilités P_m par la formule (G.30), calculer les probabilités conditionnelles $P'^{(m)}_{m'}$ à partir de la relation (G.9) :

$$P'^{(m)}_{m'} = \frac{P_{m,m'}}{P_m} \,. \tag{1}$$

L'égalité (G.10) est alors vérifiée, car on a

$$S(P'^{(m)}_{m'}) \mathrel{\widehat{=}} -\lambda \sum_{m'=1}^{M'} P'^{(m)}_{m'} \ln P'^{(m)}_{m'} = -\frac{\lambda}{P_m} \sum_{m'=1}^{M'} P_{m,m'} \left[\ln P_{m,m'} - \ln P_m\right] \,, \tag{2}$$

de sorte que

$$\sum_{m=1}^{M} P_m S(P'^{(m)}_{m'}) = -\lambda \sum_{m=1}^{M} \sum_{m'=1}^{M'} P_{m,m'} \left[\ln P_{m,m'} - \ln P_m\right] = S(P_{m,m'}) - S(P_m) \tag{3}$$

d'après l'expression (G.30) de P_m.

7. On peut en effet écrire, à partir de (G.30) et (G.31),

$$S(P_m) = -\lambda \sum_{m=1}^{M} P_m \ln P_m = -\lambda \sum_{m,m'} P_{m,m'} \ln P_m \,, \tag{1}$$

$$S(P'_{m'}) = -\lambda \sum_{m'=1}^{M'} P'_{m'} \ln P'_{m'} = -\lambda \sum_{m,m'} P_{m,m'} \ln P'_{m'} \,, \tag{2}$$

et donc

$$S(P_m) + S(P'_{m'}) = -\lambda \sum_{m,m'} P_{m,m'} \ln (P_m \,.\, P'_{m'}) \,. \tag{3}$$

.../...

l'égalité n'ayant lieu que si les deux ensembles d'événements sont *indépendants*, c'est-à-dire si

$$P_{m,m'} = P_m \cdot P'_{m'} \,. \tag{G.33}$$

L'inégalité de sous-additivité (G.32) se comprend intuitivement. Supposons en effet que, par une expérience où nous observons seulement l'événement appartenant à l'ensemble $\{e_m\}$, nous acquérions l'information $S(P_m)$. Si les deux ensembles d'événements sont corrélés, le manque d'information restant,

$$S(P_{m,m'}) - S(P_m) \,, \tag{G.34}$$

est inférieur à $S(P'_{m'})$. C'est que la connaissance de l'événement e_m qui s'est produit nous fournit, à travers les corrélations, une certaine information sur l'événement du deuxième ensemble qui l'accompagnait, même si ce dernier n'a pas été observé. Plus précisément, la probabilité pour que ce deuxième événement soit $e'_{m'}$ n'est pas $P'_{m'}$ donnée par (G.31) mais $P_{m,m'}$, où m est ici connu. Il peut même y avoir des cas où les corrélations sont si fortes que la donnée de m détermine à coup sûr la valeur de m' qui lui est associée; dans de tels cas, l'observation de l'événement e_m fournit à elle seule toute l'information qui manquait initialement[8].

Comme $\{P_m \cdot P'_{m'}\}$ est une distribution de probabilités possible pour les événements doubles $(e_m, e'_{m'})$ (distribution valable dans le cas où les deux ensembles d'événements sont indépendants), on peut utiliser l'inégalité (G.28), qui fournit le résultat :

$$S(P_m) + S(P'_{m'}) \geqslant S(P_{m,m'}) \,, \tag{4}$$

l'égalité se produisant si et seulement si

$$P_{m,m'} = P_m \cdot P'_{m'} \,. \tag{5}$$

8. Supposons par exemple que, M' et M étant égaux, la probabilité de l'événement double $(e_m, e'_{m'})$ soit nulle pour $m' \neq m$ et égale à une constante (indépendante de m) pour $m' = m$; autrement dit, supposons que

$$P_{m,m'} = \frac{1}{M}\,\delta_{mm'} \,, \tag{1}$$

où $\delta_{mm'}$ est le symbole de Kronecker (le facteur $1/M$ assure la normalisation de la distribution). Il est facile d'imaginer des situations de ce type : par exemple, si les événements e_m sont des noms de famille et les événements $e'_{m'}$ des prénoms, le tirage au sort d'un ensemble {nom, prénom} dans une assemblée de M personnes de patronymes tous différents possède ces propriétés.

Les formules (G.30) et (G.31) donnent ici

$$P_m = P'_{m'} = \frac{1}{M} \,, \tag{2}$$

de sorte que

$$S(P_m) = S(P'_{m'}) = \lambda \ln M \,. \tag{3}$$

Quant au manque d'information global, il vaut

$$S(P_{m,m'}) = -\lambda \sum_{m,m'} \frac{1}{M}\,\delta_{mm'} \ln\left(\frac{1}{M}\,\delta_{mm'}\right) = -\lambda \sum_{m} \frac{1}{M} \ln \frac{1}{M} = \lambda \ln M \,. \tag{4}$$

L'inégalité (G.32) est bien satisfaite : son premier membre est égal à la moitié du second. Mais ici, si l'on acquiert l'information $S(P_m)$, le manque d'information restant est nul :

$$S(P_{m,m'}) - S(P_m) = 0 \,. \tag{5}$$

On peut formuler l'inégalité de sous-additivité de façon légèrement différente. Considérons un système physique, et soient A et B deux observables (ou deux ensembles d'observables) pouvant être mesurées simultanément. Si $S(A, B)$ est l'information apportée par une telle mesure simultanée, $S(A)$ celle que l'on acquiert en mesurant seulement A et $S(B)$ celle que fournit la mesure de B seule, la relation (G.32) s'écrit

$$S(A, B) \leqslant S(A) + S(B) , \tag{G.35}$$

l'égalité n'étant réalisée que si les observables A et B sont statistiquement indépendantes. On introduit parfois l'*information conditionnelle* $S_A(B)$ par

$$S_A(B) \widehat{=} S(A, B) - S(A) ; \tag{G.36}$$

c'est l'information que peut encore apporter une mesure de B après que A ait été mesurée. L'inégalité de sous-additivité prend alors la forme

$$S_A(B) \leqslant S(B) . \tag{G.37}$$

3. INÉGALITÉ DE CONCAVITÉ

Soit encore $\{P_m\}$ et $\{P'_m\}$ deux distributions de probabilités relatives au même ensemble d'événements $\{e_m ; m = 1, 2, \ldots, M\}$, et considérons une troisième distribution, obtenue par combinaison linéaire des deux premières[9] :

$$P''_m = \alpha P_m + (1 - \alpha) P'_m , \tag{G.38}$$

où $$0 \leqslant \alpha \leqslant 1 . \tag{G.39}$$

---

9. Prenons un exemple. Des billes de M couleurs différentes sont enfermées au départ dans deux sacs distincts : le premier contient N billes au total, dont N_1 de la couleur (1), N_2 de la couleur (2), ..., N_M de la couleur (M); pour le second sac, ces nombres sont respectivement N' et N'_m $(m = 1, 2, \ldots, M)$. La probabilité de tirer une bille de couleur (m) vaut pour le premier sac

$$P_m = \frac{N_m}{N} , \tag{1}$$

et pour le second

$$P'_m = \frac{N'_m}{N'} . \tag{2}$$

Réunissons maintenant toutes ces billes dans un sac unique. Il y en a $(N + N')$ au total, dont $(N_m + N'_m)$ de couleur (m), de sorte que la probabilité de tirer une bille de couleur (m) vaut ici

$$P''_m = \frac{N_m + N'_m}{N + N'} = P_m \frac{N}{N + N'} + P'_m \frac{N'}{N + N'} . \tag{3}$$

En choisissant N et N', on peut faire varier

$$\alpha = \frac{N}{N + N'} \tag{4}$$

entre 0 et 1.

Le mélange des billes dans un seul sac augmente le manque d'information : si l'on a par exemple un sac de billes toutes blanches et un autre sac de billes toutes rouges, le manque d'information est nul, puisqu'il y a certitude sur la couleur de la bille que l'on va tirer de l'un ou l'autre sac; si maintenant on mélange ces billes dans un sac unique, on ne sait plus avec certitude la couleur des billes que l'on va en tirer, ce qui se traduit par une entropie statistique non nulle. De même, nous verrons au complément II.C (§ I.3.b) que le mélange de deux gaz de

.../...

On montre l'inégalité suivante, qui exprime que la fonction entropie statistique est *partout concave*[10] :

$$S\left(\alpha P_m + (1-\alpha) P'_m\right) \geqslant \alpha S(P_m) + (1-\alpha) S(P'_m) , \tag{G.40}$$

l'égalité n'ayant lieu que pour $\alpha = 0$ ou $\alpha = 1$.

---

nature différente s'accompagne d'une augmentation de l'entropie de l'ensemble; l'existence de cette entropie supplémentaire, dite *entropie de mélange,* permet de comprendre, dans le cadre du postulat fondamental (II, § I.B), pourquoi deux gaz mélangés ne se séparent pas spontanément l'un de l'autre.

De façon plus générale, les propriétés de concavité de l'entropie sont à la base de l'explication théorique de la *stabilité* des systèmes macroscopiques : le maximum de l'entropie, qui caractérise l'état d'équilibre macroscopique pour un système isolé (cf. II, § I.B.3), est un point stationnaire où la concavité de la fonction entropie est tournée vers le bas (une concavité en sens inverse indiquerait que le point stationnaire correspond à un état d'équilibre instable du système). Des illustrations concrètes de cet argument peuvent être trouvées au chapitre II (§ III.A.2), au chapitre III (§ II.D.3) et dans le complément III.G (§ III).

10. Cette propriété est vérifiée séparément par chacun des termes de la somme donnant S. Si en effet l'on trace la fonction

$$y = -x \ln x \tag{1}$$

pour x compris entre 0 et 1 (fig. 2), on voit que sa concavité est partout tournée vers le bas : sa dérivée seconde,

$$y'' = -\frac{1}{x} , \tag{2}$$

est constamment négative. Considérons alors le point M, de coordonnées

$$x_M \mathrel{\widehat{=}} \alpha x_1 + (1-\alpha) x_2 , \tag{3.a}$$

$$y_M \mathrel{\widehat{=}} -\alpha x_1 \ln x_1 - (1-\alpha) x_2 \ln x_2 . \tag{3.b}$$

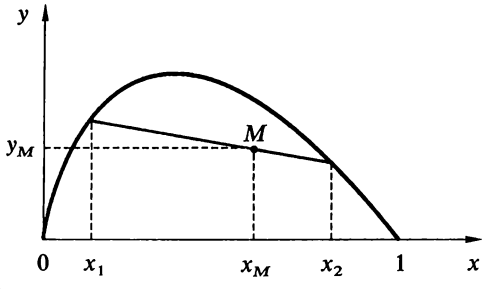


FIGURE 2

Courbe représentative de la fonction $y = -x \ln x$, pour $0 \leqslant x \leqslant 1$. La corde joignant deux points de cette courbe est toujours située au-dessous de l'arc de courbe correspondant.

Quand α varie entre 0 et 1, M se déplace sur la *corde* joignant les points de la courbe d'abscisses x_1 et x_2. La propriété de concavité de la fonction $y = -x \ln x$:

$$-\left[\alpha x_1 + (1-\alpha) x_2\right] \ln\left[\alpha x_1 + (1-\alpha) x_2\right] \geqslant -\alpha x_1 \ln x_1 - (1-\alpha) x_2 \ln x_2 \tag{4}$$

exprime donc que sa courbe représentative est partout située au-dessus d'une corde quelconque.

III. MESURE DE L'INFORMATION MANQUANTE DANS LE CAS D'UNE DISTRIBUTION CONTINUE DE PROBABILITÉ

Considérons une variable aléatoire x, pouvant prendre *a priori* n'importe quelle valeur réelle comprise entre deux bornes a et b $(a<b)$; soit $w(x)$ la densité de probabilité correspondante, que nous supposons continue et normalisée selon

$$\int_a^b w(x)\,\mathrm{d}x = 1 \,. \tag{G.41}$$

Comment mesurer l'information manquante associée à cette distribution de probabilité?

1. DISCRÉTISATION DU PROBLÈME

Essayons de nous ramener au cas déjà connu de probabilités discrètes. Découpons pour cela l'intervalle $[a, b]$ en un nombre M de segments d'égale longueur δx :

$$\delta x = \frac{b-a}{M} \,. \tag{G.42}$$

Le m-ième de ces segments est l'ensemble des points d'abscisse x telle que

$$a_m \leqslant x < b_m \,, \tag{G.43}$$

avec $$a_m = a + (m-1)\frac{b-a}{M} \,, \tag{G.44.a}$$

$$b_m = a + m\frac{b-a}{M} \,. \tag{G.44.b}$$

La probabilité δP_m pour que la variable aléatoire prenne une des valeurs (G.43) se calcule aisément :

$$\delta P_m = \int_{a_m}^{b_m} w(x)\,\mathrm{d}x \,. \tag{G.45}$$

La fonction $w(x)$ est continue; il existe donc une valeur x_m de la variable, comprise entre a_m et b_m, telle que

$$\delta P_m = \frac{b-a}{M}\, w(x_m) \,. \tag{G.46}$$

Nous voilà effectivement ramenés à un problème de probabilités discrètes : l'événement e_m, défini comme « la variable x est comprise entre a_m et b_m », a la probabilité δP_m de se réaliser; ces événements sont en nombre fini M et, si l'on rapproche les formules (G.45) et (G.41), on constate que l'ensemble des δP_m est normalisé :

$$\sum_{m=1}^{M} \delta P_m = 1 \,. \tag{G.47}$$

En prenant M suffisamment grand, on peut espérer reproduire avec une bonne approximation la distribution continue dont on est parti.

Mais, pour ce qui est de l'entropie statistique, ce procédé se heurte à *deux difficultés fondamentales,* que nous allons analyser successivement.

2. DIFFICULTÉ DU PASSAGE A LA LIMITE CONTINUE

a. Divergence de la formule fondamentale

L'information manquante correspondant à la distribution de probabilités discrètes $\{\delta P_m;\ m = 1, 2, \ldots, M\}$ est mesurée par l'entropie statistique

$$S(\delta P_m) = -\lambda \sum_{m=1}^{M} \delta P_m \ln \delta P_m \,, \tag{G.48}$$

qui vaut ici

$$S(\delta P_m) = -\lambda \left[\sum_{m=1}^{M} \frac{b-a}{M} w(x_m) \ln w(x_m) \right] - \lambda \ln \frac{b-a}{M} \tag{G.49}$$

(on a utilisé la propriété de normalisation (G.47)). Pour passer à la limite continue, nous voulons faire tendre M vers l'infini. Le premier terme de (G.49) tend alors, par définition de l'intégrale, vers

$$S_t \widehat{=} -\lambda \int_a^b \mathrm{d}x\, w(x) \ln w(x) \,, \tag{G.50}$$

mais le second tend vers l'infini. *Le passage à la limite est donc impossible.*

b. Les deux solutions possibles

Pour contourner cette difficulté, deux attitudes sont possibles.

La première s'appuie sur la remarque que *le terme qui devient infini ne dépend pas de la distribution de probabilité* étudiée. Si donc on renonce à une mesure absolue de l'information manquante pour s'intéresser seulement aux *différences d'informations manquantes* entre les diverses distributions continues de probabilité $w(x)$, on peut retenir l'expression (G.50) de S_t comme mesure de cette information manquante relative. L'existence de cette « constante infinie » que l'on retranche ici n'est pas surprenante : il faudrait une précision infinie pour localiser exactement un point mathématique sur le segment $[a, b]$; il est donc normal que le manque d'information associé à une distribution continue de probabilité soit infini.

Le choix de l'entropie statistique tronquée S_t pour mesurer l'information manquante relative présente cependant certains inconvénients. En premier lieu, se pose un problème de dimensions : $w(x)$ a les dimensions de $1/x$ (cf. (G.41)), de sorte que son logarithme dépend de l'unité choisie pour mesurer x (sauf si x est un nombre pur)[11]. En outre, S_t n'est plus constamment positive comme l'était l'entropie statistique dans le cas discret. Elle tend même vers $-\infty$ lorsque $w(x)$ tend vers une « fonction » de Dirac $\delta(x - x_0)$, c'est-à-dire lorsque la valeur de la variable aléatoire est donnée

11. La raison de cette difficulté réside dans la forme de la « constante infinie » que l'on a supprimée : le second terme de (G.49) comporte lui aussi le logarithme d'une grandeur ayant des dimensions. La formule (G.49) complète ne souffre pas de ce défaut; il apparaît lorsqu'on tronque l'expression de S.

avec une précision parfaite[12]. Les autres propriétés de l'entropie statistique sont cependant préservées.

Mais la formule (G.49) permet un autre choix, que nous ferons dans la suite (cf. complément II.F). Il se fonde sur l'idée qu'un point mathématique, c'est-à-dire une valeur de x définie avec une précision parfaite, n'existe jamais vraiment en physique. On doit donc pouvoir déterminer, dans chaque problème particulier, un intervalle minimum $\delta_0 x$ caractérisant la meilleure précision possible sur une mesure de la variable x. On peut dès lors considérer que le passage à la limite continue est en fait illusoire, et que la réalité est plus proche de l'« approximation » discrète que nous avons construite : on choisira M de telle sorte que

$$\frac{b-a}{M} = \delta_0 x \,, \tag{G.51}$$

et on prendra pour entropie statistique associée à la distribution $w(x)$ l'expression complète (G.49). Si $\delta_0 x$ correspond vraiment à la meilleure précision possible sur les mesures de x, les variations de $w(x)$ sur un intervalle de largeur $\delta_0 x$ sont nécessairement négligeables; il en résulte que la somme sur m de la formule (G.49) est impossible à distinguer en pratique, pour M fixé selon (G.51), de l'intégrale qui en est la limite continue. En définitive, dans ce second point de vue, le manque d'information lié à une densité de probabilité $w(x)$ s'écrit

$$S = -\lambda \int_a^b \mathrm{d}x\, w(x) \ln w(x) - \lambda \ln \delta_0 x \,, \tag{G.52}$$

c'est-à-dire

$$\boxed{S = -\lambda \int_a^b \mathrm{d}x\, w(x) \ln\left[w(x)\, \delta_0 x \right]} \,. \tag{G.53}$$

Cette fois, l'entropie statistique est positive (ceci apparaît clairement sur la formule de départ (G.48), puisque les δP_m sont de véritables probabilités, comprises

---

12. La « fonction » $\delta(x - x_0)$ peut être considérée comme la limite d'une fonction « créneau » dont la largeur tend vers zéro et la hauteur vers l'infini (de telle sorte que son intégrale reste égale à 1) :

$$\delta(x - x_0) = \lim_{\varepsilon \to 0_+} \delta_\varepsilon (x - x_0) \,, \tag{1}$$

où $\quad \delta_\varepsilon (x - x_0) = 0 \quad$ pour $\quad |x - x_0| > \varepsilon$

$$= \frac{1}{2\varepsilon} \quad \text{pour} \quad x_0 - \varepsilon < x < x_0 + \varepsilon \,. \tag{2}$$

Pour

$$w(x) = \delta_\varepsilon (x - x_0) \,, \tag{3}$$

l'entropie statistique tronquée S_t vaut

$$S_t = -\lambda \int_{x_0 - \varepsilon}^{x_0 + \varepsilon} \mathrm{d}x\, \frac{1}{2\varepsilon} \ln \frac{1}{2\varepsilon} = \lambda \ln 2\varepsilon \,. \tag{4}$$

Elle tend vers $-\infty$ quand ε tend vers zéro.

entre 0 et 1). Elle s'annule lorsque le manque d'information est minimal, c'est-à-dire lorsque la valeur x_0 de la variable x est connue avec une incertitude égale à $\delta_0 x$[13]. Pour une distribution uniforme (w constante sur l'intervalle $[a, b]$), elle vaut

$$S = \lambda \ln \frac{b-a}{\delta_0 x} . \tag{G.54}$$

3. CHOIX DE LA VARIABLE PERTINENTE

La mesure de l'information manquante dans le cas d'une distribution continue de probabilité fait apparaître un autre problème délicat. En effet, dans l'un et l'autre points de vue décrits au paragraphe 2.b, l'intervalle $[a, b]$ sur lequel la variable aléatoire x prend ses valeurs a été divisé en *segments égaux*. Supposons maintenant que nous changions de variable :

$$y = f(x) , \tag{G.55}$$

où la fonction f est continue et monotone sur l'intervalle $[a, b]$. La nouvelle variable aléatoire y est distribuée sur l'intervalle $[f(a), f(b)]$ avec la densité de probabilité (complément I.B, § I.1)

$$\tilde{w}(y) = w\left[f^{-1}(y)\right] \frac{\mathrm{d}f^{-1}}{\mathrm{d}y} , \tag{G.56}$$

où f^{-1} est la fonction inverse de f. Si nous appliquons à $\tilde{w}(y)$ la méthode que nous avons décrite ci-dessus (en découpant l'intervalle $[f(a), f(b)]$ en segments égaux), nous obtenons une *valeur différente de l'entropie statistique*[14], que nous ayons choisi l'expression tronquée (G.50) ou la formule discrétisée (G.53).

---

13. La densité de probabilité $w(x)$ peut dans ce cas être prise sous la forme

$$\begin{aligned} w(x) &= \frac{1}{\delta_0 x} \quad \text{si} \quad x_0 - \frac{\delta_0 x}{2} < x < x_0 + \frac{\delta_0 x}{2} \\ &= 0 \quad \text{si} \quad |x - x_0| > \frac{\delta_0 x}{2} . \end{aligned} \tag{1}$$

La formule (G.53) donne alors

$$S = -\lambda \int_{x_0 - \delta_0 x/2}^{x_0 + \delta_0 x/2} \mathrm{d}x \, \frac{1}{\delta_0 x} \ln \frac{\delta_0 x}{\delta_0 x} = 0 . \tag{2}$$

14. Dans la formule (G.50), le changement de variable (G.55) donne

$$S_t = -\lambda \int_{f(a)}^{f(b)} \frac{\mathrm{d}f^{-1}}{\mathrm{d}y} \, \mathrm{d}y \; w\left[f^{-1}(y)\right] \ln w\left[f^{-1}(y)\right] , \tag{1}$$

c'est-à-dire, compte tenu de (G.56),

$$S_t = -\lambda \int_{f(a)}^{f(b)} \mathrm{d}y \; \tilde{w}(y) \ln \tilde{w}(y) + \lambda \int_{f(a)}^{f(b)} \mathrm{d}y \; \tilde{w}(y) \ln \frac{\mathrm{d}f^{-1}}{\mathrm{d}y} . \tag{2}$$

Cette expression diffère de celle qu'on aurait obtenue en utilisant directement la variable y :

$$\tilde{S}_t = -\lambda \int_{f(a)}^{f(b)} \mathrm{d}y \; \tilde{w}(y) \ln \tilde{w}(y) . \tag{3}$$

Ce calcul s'applique également à la formule (G.53), qui est en outre directement affectée par le fait que des intervalles égaux en x ne sont plus en général égaux en y :

$$\delta_0 y = \frac{\mathrm{d}f}{\mathrm{d}x} \delta_0 x , \tag{4}$$

où la dérivée $\mathrm{d}f/\mathrm{d}x$ dépend en général de x; pour $\delta_0 x$ indépendant de x sur l'intervalle $[a, b]$, $\delta_0 y$ n'est pas en général indépendant de y sur l'intervalle $[f(a), f(b)]$.

Ainsi, dans le cas d'une distribution continue de probabilité, *il faut choisir, parmi les variables* a priori *possibles, celle que l'on utilisera pour calculer l'entropie statistique* mesurant le manque d'information associé à la distribution. Ce choix devra être fondé sur des critères physiques.

IV. DISCUSSION : LIEN ENTRE INFORMATION ET MÉCANIQUE STATISTIQUE

Nous avons, dans ce complément, trouvé l'expression de l'entropie statistique, mesure de l'information manquante, à partir d'hypothèses « naturelles » sur la signification et les propriétés de l'information. Notre but est d'utiliser cette expression en mécanique statistique. Nous allons discuter succinctement, dans la perspective de cette application, la démarche du paragraphe I.

Il apparaît clairement, dans la démonstration du paragraphe I.2.c, que c'est la *propriété d'additivité* qui est essentielle pour aboutir à la formule (G.22) donnant l'entropie statistique. On peut envisager de remplacer cette hypothèse par une autre moins restrictive, et d'admettre ainsi éventuellement d'autres fonctions des probabilités pour mesurer l'information. Mais ces développements de la théorie de l'information restent étrangers à la mécanique statistique, où l'additivité de l'entropie est fondamentale du point de vue physique (voir chapitre II et complément II.A).

Dans le même ordre d'idées, on peut souligner que le point de vue adopté ici suppose implicitement que *l'information dépend seulement des probabilités* des événements possibles. Or ceci n'est pas forcément réaliste : l'intérêt d'un événement peut dépendre de bien autre chose que de sa seule probabilité[15]. Si c'est le cas, on affectera à chaque événement e_m, outre sa probabilité P_m, une *« utilité »* u_m, et on mesurera l'information manquante par

$$S_u \widehat{=} -\lambda \sum_{m=1}^{M} u_m P_m \ln P_m \,. \qquad \text{(G.57)}$$

Ce type de généralisation de l'entropie statistique n'intéresse pas non plus la mécanique statistique, pour laquelle les divers états microscopiques possibles pour un système ont tous même « valeur ».

Même dans le cadre de la formule de Shannon (G.22), des discussions ont eu lieu sur la relation entre l'information et l'entropie au sens de la thermodynamique. Le problème peut être posé en comparant les propriétés thermodynamiques d'un système macroscopique et la mesure de l'information manquante dans une des situations simples qui servent à la compréhension intuitive des notions probabilistes.

Pour un gaz par exemple, l'entropie thermodynamique est une propriété objective, indépendante de l'observateur : sa valeur est une fonction bien définie des paramètres (énergie, volume, nombre de molécules) qui caractérisent l'état macroscopique du gaz; si par exemple, tout en le maintenant isolé, on offre au gaz un volume supérieur à celui qu'il occupe initialement (II, § II.B.1), il se répand dans le volume total disponible, parce que son état d'équilibre macroscopique correspond toujours à la valeur maximale que peut prendre son entropie, compte tenu des contraintes extérieures.

---

15. Considérons par exemple un joueur de bridge qui doit réaliser un contrat résultant des annonces de son équipe. Il connaît son jeu et celui du mort, et attend l'entame de l'adversaire. *A priori,* les 26 cartes détenues par l'équipe adverse ont pour lui des probabilités égales d'être jouées (on pourrait, sans changer le fond de l'argument, raffiner cette estimation pour tenir compte du fait que le jeu de la carte ne se fait pas au hasard). Un joueur averti tire pourtant de la première carte jouée de précieuses informations, qui dépendent beaucoup de sa nature; assez souvent même, le contrat peut être gagné si l'entame a lieu dans telle couleur, ou perdu irrémédiablement si c'est dans telle autre.

Considérons au contraire des cartes, disposées l'une à côté de l'autre sur une table, face tournée vers le bas. On peut calculer l'entropie statistique associée à cet état[16]. Mais il peut se faire que la table soit transparente; dans ce cas, un observateur malin peut par exemple disposer un miroir au-dessous du jeu de cartes et obtenir ainsi tout ou partie de l'information qui manque lorsqu'on ne voit que le dos des cartes. Ainsi, l'entropie statistique n'est plus ici une propriété intrinsèque du système étudié, dépendant seulement de son état; la position de l'observateur par rapport au système influe aussi sur sa valeur.

La question est évidemment de savoir si, dans le cas du gaz, un observateur « malin » pourrait obtenir une partie de l'information comptabilisée dans l'entropie. Cette question a été posée, dans le cadre de la thermodynamique macroscopique, dès qu'a été reconnue l'existence d'une réalité microscopique sous-jacente (molécules). L'exemple le plus célèbre de ce type de problème est celui du *« démon de Maxwell »* (réf. 3, p. 157), capable de mettre en défaut le second principe de la thermodynamique en transférant, sans dépense de travail, de la chaleur d'un corps froid à un corps chaud; mais pour ce faire, le « démon » doit *voir les molécules,* pour sélectionner celles dont la vitesse d'agitation thermique est supérieure à la moyenne; il doit donc pouvoir acquérir et utiliser une information détaillée sur l'état microscopique d'un système macroscopique. Il serait prématuré de poursuivre ici la discussion de ce problème. Nous la reprendrons après l'énoncé du postulat fondamental de la mécanique statistique (II, § I.B), et surtout lorsque nous analyserons l'irréversibilité de l'évolution des systèmes à l'échelle macroscopique (IV, § II.B.4).

Références

1. R. Balian, *Du microscopique au macroscopique,* cours de Physique Statistique de l'École Polytechnique, Ellipses, Paris (1982), chapitre III.
2. A. Katz, *Principles of Statistical Mechanics,* Freeman, San Francisco (1967), chapitres 2, 4 et 5.
3. L. Brillouin, *La science et la théorie de l'information,* Masson, Paris (1959).
4. S. Guiasu et R. Theodorescu, *Incertitude et information,* Presses de l'Université Laval, Québec (1971).

16. Pour un jeu de 52 cartes, il y a 52! configurations possibles, toutes *a priori* équiprobables. L'entropie statistique correspondante vaut donc

$$I(52!) = \frac{\ln 52!}{\ln 2}\,, \tag{1}$$

que l'on calcule à l'aide de la formule de Stirling :

$$\ln n! \simeq n \ln n - n + \frac{1}{2}\ln(2\pi n)\,. \tag{2}$$

On trouve ainsi

$$I(52!) \simeq 225{,}6 \text{ bit}\,. \tag{3}$$

COMPLÉMENT I.H

L'opérateur-densité

I. Définition
   1. Mélange statistique d'états purs
   2. Caractérisation d'un mélange par un opérateur-densité

II. Principales propriétés
   1. Hermiticité, trace, positivité
   2. Valeurs et vecteurs propres
   3. Matrice-densité
   4. Évolution dans le temps
   5. Opérateur-densité décrivant un état pur

III. Opérateur-densité et distribution statistique des états stationnaires
   1. Système à l'équilibre macroscopique
   2. États hors d'équilibre

IV. Entropie statistique d'un état décrit par un opérateur-densité
   1. Expression de l'entropie statistique
   2. Une inégalité utile
   3. Propriétés de l'entropie statistique
      a. Positivité. Minimum
      b. Maximum sur un sous-espace de dimension finie
      c. Additivité
      d. Concavité

V. Conclusion

L'état quantique macroscopique d'un système est défini comme un mélange statistique d'états purs microscopiques. Ce complément est destiné à introduire et étudier succinctement l'outil théorique qui permet de caractériser le plus commodément un tel état macroscopique et de calculer ses propriétés.

I. DÉFINITION

1. MÉLANGE STATISTIQUE D'ÉTATS PURS

Reprenons la situation décrite au paragraphe III.A.2 du chapitre I : le système étudié[1] peut *a priori* se trouver dans l'un quelconque des états (purs) d'un ensemble

$$\{|\psi_m\rangle;\ m=1,\ 2,\ \ldots\}\ , \tag{H.1}$$

chacun de ces états ayant une certaine probabilité d'être réalisé :

$$\{P_m;\ m=1,\ 2,\ \ldots\}\ , \tag{H.2}$$

avec $$\sum_m P_m = 1\ . \tag{H.3}$$

Nous supposons les kets $|\psi_m\rangle$ normés :

$$\langle\psi_m|\psi_m\rangle = 1;\quad m=1,\ 2,\ \ldots\ . \tag{H.4}$$

1. Ce système peut être aussi bien microscopique que macroscopique : pour un système macroscopique, l'impossibilité de préciser l'état pur dans lequel il se trouve à un instant donné impose l'utilisation de mélanges statistiques; mais un système microscopique peut lui aussi être placé dans une situation telle que son état quantique ne soit pas un état pur, mais un mélange statistique de tels états (voir par exemple III, § I.B.1).

Mais ils n'ont aucune raison d'être orthogonaux entre eux (cf. note 7 ci-dessous).

Les prédictions physiques que l'on peut formuler dans une telle situation sont obtenues par les méthodes habituelles du calcul des probabilités : si $G(|\psi_m\rangle)$ est la valeur ou l'expression de la grandeur ou de la propriété G lorsque le système se trouve avec certitude dans l'état $|\psi_m\rangle$, la valeur moyenne de G dans le mélange statistique des états $|\psi_m\rangle$ affectés des probabilités P_m est donnée par

$$\overline{G} = \sum_m P_m G(|\psi_m\rangle) \,. \tag{H.5}$$

Cette formule générale permet de calculer également l'écart quadratique moyen (ou tout autre moment) de la distribution statistique d'une grandeur ou propriété quelconque.

2. CARACTÉRISATION D'UN MÉLANGE PAR UN OPÉRATEUR-DENSITÉ

L'opérateur-densité D associé au mélange statistique $\{|\psi_m\rangle, P_m\}$ est défini par[2]

$$\boxed{D \mathrel{\widehat{=}} \sum_m |\psi_m\rangle P_m \langle\psi_m|} \,. \tag{H.6}$$

Les prédictions physiques (formules (H.5)) se calculent de façon directe et simple à partir de l'opérateur-densité.

Supposons par exemple que nous ayons l'intention de mesurer une grandeur physique A et que nous voulions évaluer la probabilité $P(a)$ pour que le résultat soit la valeur propre a de l'observable A. Dans le cas général, cette valeur propre est g fois dégénérée, c'est-à-dire qu'il existe g vecteurs $|u_k\rangle$ linéairement indépendants (normés et orthogonaux entre eux) tels que

$$A|u_k\rangle = a|u_k\rangle; \quad k = 1, 2, \ldots g \,. \tag{H.7}$$

La formule générale (H.5) donne ici

$$P(a) = \sum_m P_m P(a; |\psi_m\rangle) \,, \tag{H.8}$$

où

$$P(a; |\psi_m\rangle) = \sum_{k=1}^{g} |\langle u_k|\psi_m\rangle|^2 \tag{H.9}$$

est la probabilité d'obtenir le résultat a lorsque le système est à coup sûr dans l'état $|\psi_m\rangle$. Si l'on reporte (H.9) dans (H.8) on obtient, en utilisant les propriétés classiques du produit scalaire (C.D.L., p. 111),

$$\begin{aligned} P(a) &= \sum_m \sum_{k=1}^{g} P_m \langle u_k|\psi_m\rangle\langle\psi_m|u_k\rangle = \sum_{k=1}^{g} \langle u_k| \left[\sum_m P_m |\psi_m\rangle\langle\psi_m| \right] |u_k\rangle \\ &= \sum_{k=1}^{g} \langle u_k|D|u_k\rangle \,. \end{aligned} \tag{H.10}$$

---

2. L'objet mathématique $|\psi_m\rangle\langle\psi_m|$, obtenu par juxtaposition d'un ket et du bra conjugué (dans cet ordre), est un *opérateur :* à tout ket $|\varphi\rangle$ il fait correspondre un autre ket,

$$|\psi_m\rangle\langle\psi_m|\varphi\rangle \,, \tag{1}$$

produit du ket $|\psi_m\rangle$ par le nombre complexe $\langle\psi_m|\varphi\rangle$. Plus précisément,

$$|\psi_m\rangle\langle\psi_m| \mathrel{\widehat{=}} \mathcal{T}(|\psi_m\rangle) \tag{2}$$

est le *projecteur* sur l'état $|\psi_m\rangle$ (C.D.L., p. 115).

Dans la formule (H.6), le nombre P_m peut être placé aussi bien avant le ket $|\psi_m\rangle$ ou après le bra $\langle\psi_m|$. L'important est l'ordre de succession du ket et du bra.

On peut écrire ce résultat sous la forme[3],[4]

$$\boxed{P(a)=\mathrm{Tr}\left[\mathcal{P}(\mathcal{E}_a)D\right]}\ , \tag{H.11}$$

où $$\mathcal{P}(\mathcal{E}_a)\widehat{=}\sum_{k=1}^{g}|u_k\rangle\langle u_k| \tag{H.12}$$

est le projecteur sur le sous-espace de dégénérescence $\mathcal{E}_a$ associé à la valeur propre a (c'est-à-dire le sous-espace engendré par les g vecteurs $|u_k\rangle$).

On trouve de même que la valeur moyenne $\overline{A}$ d'une grandeur physique A dans l'état macroscopique (H.1)−(H.2) est donnée par[5]

$$\boxed{\overline{A}=Tr(AD)}\ . \tag{H.13}$$

---

3. Le symbole Tr A désigne la *trace de l'opérateur A*, définie comme la *somme de ses éléments de matrice diagonaux : si* $\{|v_i\rangle\}$ est une base quelconque de l'espace des états,

$$\mathrm{Tr}\ A\widehat{=}\sum_i\langle v_i|A|v_i\rangle\ . \tag{1}$$

On montre (C.D.L., p. 167) que la trace ainsi définie est une caractéristique de l'opérateur A, *indépendante de la base choisie* pour la calculer.

La trace d'un produit d'opérateurs est invariante par permutation circulaire de ces opérateurs (C.D.L., p. 167) :

$$\mathrm{Tr}(ABC)=\mathrm{Tr}(CAB)=\mathrm{Tr}(BCA)\ . \tag{2}$$

4. Partant de l'expression (H.10), introduisons la relation de fermeture associée à la base $\{|v_i\rangle\}$ choisie dans l'espace des états :

$$\sum_i|v_i\rangle\langle v_i|=\mathbb{1}\ . \tag{1}$$

Il vient

$$P(a)=\sum_{k=1}^{g}\sum_i\langle u_k|D|v_i\rangle\langle v_i|u_k\rangle=\sum_{k=1}^{g}\sum_i\langle v_i|u_k\rangle\langle u_k|D|v_i\rangle \tag{2}$$

(la dernière étape consiste simplement à permuter deux nombres complexes). On reconnaît alors le projecteur $\mathcal{P}(\mathcal{E}_a)$ défini par (H.12) :

$$P(a)=\sum_i\langle v_i|\mathcal{P}(\mathcal{E}_a)D|v_i\rangle=\mathrm{Tr}\left[\mathcal{P}(\mathcal{E}_a)D\right]\ . \tag{3}$$

5. On a par définition (formule (H.5))

$$\overline{A}=\sum_m P_m\langle\psi_m|A|\psi_m\rangle\ . \tag{1}$$

Comme dans la note précédente, utilisons la relation de fermeture que vérifie la base $\{|v_i\rangle\}$ choisie :

$$\overline{A}=\sum_m\sum_i P_m\langle\psi_m|A|v_i\rangle\langle v_i|\psi_m\rangle\ . \tag{2}$$

On peut changer l'ordre des trois nombres figurant au second membre :

$$\overline{A}=\sum_i\sum_m\langle v_i|\psi_m\rangle P_m\langle\psi_m|A|v_i\rangle\ . \tag{3}$$

.../...

Comme toutes les prédictions physiques concernant un système quantique font intervenir des valeurs moyennes d'observables ou de projecteurs tels que (H.12), les formules (H.11) et (H.13) sont universelles. Ces deux formules ont même structure[6] : on y prend la trace du produit de l'opérateur-densité et de l'observable ou du projecteur associé à la propriété que l'on veut étudier.

L'opérateur-densité n'est pas seulement un outil commode permettant de donner une expression simple et compacte aux propriétés du système. On montre en effet[7] que *des mélanges statistiques* $\{|\psi_m\rangle, P_m; m = 1, 2, \ldots\}$ *différents peuvent conduire au même opérateur-densité.* Ces mélanges statistiques sont alors *équivalents,* puisque les prédictions physiques qu'on en déduit sont toutes identiques d'après les formules (H.11) et (H.13). Par conséquent, *c'est l'opérateur-densité, et non pas le mélange statistique à partir duquel on l'a défini, qui caractérise de façon univoque l'état macroscopique d'un système.*

II. PRINCIPALES PROPRIÉTÉS

1. HERMITICITÉ, TRACE, POSITIVITÉ

Les probabilités P_m étant des nombres réels, un opérateur-densité est toujours hermitique (C.D.L, p. 120) :

$$D^\dagger = D \,. \tag{H.14}$$

En utilisant l'expression (H.6) de l'opérateur-densité, on obtient finalement

$$\overline{A} = \sum_i \langle v_i | DA | v_i \rangle = \mathrm{Tr}(DA) = \mathrm{Tr}(AD) \,. \tag{4}$$

6. Ceci n'a rien d'étonnant : un projecteur est lui aussi une observable (voir C.D.L., p. 138).

7. Prenons l'exemple simple d'une particule de spin 1/2, en nous restreignant à ses états de spin (espace à 2 dimensions). Notons $|\psi(\theta, \varphi)\rangle$ le ket correspondant au cas où le spin de la particule pointe dans la direction caractérisée par les angles polaires θ et φ (ket propre, avec la valeur propre $+\hbar/2$, de la projection de l'observable de spin $\vec{S}$ sur la direction (θ, φ)). L'état de spin « non-polarisé » est alors décrit comme un mélange statistique de tous les états $|\psi(\theta, \varphi)\rangle$ possibles, avec des probabilités égales. L'opérateur-densité correspondant s'écrit donc

$$D = \frac{1}{4\pi} \int_0^{2\pi} \mathrm{d}\varphi \int_0^{\pi} \sin\theta \, \mathrm{d}\theta \, |\psi(\theta, \varphi)\rangle\langle\psi(\theta, \varphi)| \,. \tag{1}$$

En effet, au lieu d'une somme discrète sur m, on a ici une intégrale sur les paramètres continus θ et φ, l'angle solide élémentaire autour de la direction (θ, φ) ayant l'expression bien connue

$$\mathrm{d}\omega(\theta, \varphi) = \sin\theta \, \mathrm{d}\theta \, \mathrm{d}\varphi \,. \tag{2}$$

Il est facile de vérifier (C.D.L., p. 437) que ce même opérateur-densité est obtenu à partir d'un mélange statistique de deux états purs orthogonaux quelconques, avec des probabilités égales :

$$D = \frac{1}{2} |+\rangle\langle+| + \frac{1}{2} |-\rangle\langle-| \,, \tag{3}$$

où $|+\rangle$ et $|-\rangle$ sont par exemple les deux états propres orthogonaux (valeurs propres $+\hbar/2$ et $-\hbar/2$) de l'observable S_z, projection du spin $\vec{S}$ sur l'axe Oz. On aboutit encore au même résultat si l'on prend un mélange, avec des probabilités égales (1/4), des quatre états $|+\rangle$, $|-\rangle$, $|+\rangle_x$ et $|-\rangle_x$, où $|+\rangle_x$ et $|-\rangle_x$ sont les états propres de l'observable S_x, projection du spin $\vec{S}$ sur l'axe Ox (C.D.L., p. 393) :

$$|\pm\rangle_x = \frac{1}{\sqrt{2}} [\,|+\rangle \pm |-\rangle\,] \,. \tag{4}$$

D'autre part, la condition de normalisation des probabilités P_m se traduit ici par une normalisation de la trace de l'opérateur-densité associé[8] :

$$\sum_m P_m = 1 \implies \text{Tr } D = 1 \,. \tag{H.15}$$

Enfin, l'opérateur-densité est semi-défini positif, c'est-à-dire que

$$\langle \varphi | D | \varphi \rangle \geqslant 0 \quad \forall | \varphi \rangle \,. \tag{H.16}$$

Cette propriété découle de la positivité des probabilités P_m[9].

2. VALEURS ET VECTEURS PROPRES

Soit $|\varphi_j\rangle$ les vecteurs propres de l'opérateur-densité D, d_j les valeurs propres asssociées :

$$D|\varphi_j\rangle = d_j |\varphi_j\rangle \,. \tag{H.17}$$

L'opérateur D étant hermitique, ses valeurs propres d_j sont *réelles*, et l'on peut choisir ses vecteurs propres de façon qu'ils forment un ensemble orthonormé :

$$\langle \varphi_j | \varphi_{j'} \rangle = \delta_{jj'} \,; \tag{H.18}$$

cet ensemble constitue une *base* dans l'espace des états. On peut dès lors exprimer D en fonction de ses vecteurs propres et valeurs propres sous la forme (C.D.L, p. 138)

$$D = \sum_j d_j |\varphi_j\rangle\langle\varphi_j| \,. \tag{H.19}$$

En outre, la propriété de positivité (H.16) implique[10]

$$d_j \geqslant 0 \,, \tag{H.20}$$

---

8. Une base orthonormée $\{|v_i\rangle\}$ ayant été choisie dans l'espace des états, la trace de D vaut par définition (note 3)

$$\text{Tr } D = \sum_i \langle v_i | D | v_i \rangle = \sum_i \sum_m \langle v_i | \psi_m \rangle P_m \langle \psi_m | v_i \rangle \,. \tag{1}$$

Si l'on change l'ordre des facteurs dans ce produit de trois nombres, on peut écrire

$$\text{Tr } D = \sum_m P_m \sum_i \langle \psi_m | v_i \rangle \langle v_i | \psi_m \rangle \,. \tag{2}$$

La relation de fermeture de la base $\{|v_i\rangle\}$ (note 4, formule 1) donne ensuite

$$\text{Tr } D = \sum_m P_m \langle \psi_m | \psi_m \rangle \,, \tag{3}$$

c'est-à-dire, puisque les $|\psi_m\rangle$ sont supposés normés à 1,

$$\text{Tr } D = \sum_m P_m = 1 \,. \tag{4}$$

9. En effet,

$$\langle \varphi | D | \varphi \rangle = \sum_m \langle \varphi | \psi_m \rangle P_m \langle \psi_m | \varphi \rangle = \sum_m P_m |\langle \varphi | \psi_m \rangle|^2$$

apparaît comme une somme de termes tous positifs ou nuls.

10. En effet, $|\varphi_j\rangle$ étant normé à 1, on tire de (H.17)

$$d_j = \langle \varphi_j | D | \varphi_j \rangle \,;$$

d_j est donc positive ou nulle d'après (H.16).

et la condition de normalisation (H.15)[11]

$$\sum_j d_j = 1 \ . \tag{H.21}$$

L'ensemble des valeurs propres d_j constitue donc une distribution de *probabilités;* nous les noterons désormais P_j, et écrirons les relations (H.17), (H.19), (H.20) et (H.21) sous la forme

$$D|\varphi_j\rangle = P_j|\varphi_j\rangle \ , \tag{H.22}$$

$$D = \sum_j |\varphi_j > P_j < \varphi_j| \ , \tag{H.23}$$

$$0 \leqslant P_j \leqslant 1 \ , \tag{H.24.a}$$

$$\sum_j P_j = 1 \ . \tag{H.24.b}$$

On constate ainsi que l'état décrit par un opérateur-densité D donné peut être considéré comme un *mélange statistique des états propres* $|\varphi_j\rangle$ *de D, affectés des probabilités* P_j*, valeurs propres de* D[12]. Dans cette interprétation, les états purs entrant dans le mélange sont *orthogonaux,* de sorte qu'ils s'excluent les uns les autres : si le système est dans l'un des états $|\varphi_j\rangle$, la probabilité $|\langle\varphi_{j'}|\varphi_j\rangle|^2$ qu'on le trouve dans un autre état $|\varphi_{j'}\rangle$ du même ensemble est nulle. C'est pourquoi l'expression correspondante de D est particulièrement commode[13]. Inversement, si l'on part d'un mélange statistique d'états purs *orthogonaux,* l'opérateur-densité qui en découle admet ces états comme états propres et les probabilités associées comme valeurs propres[14].

---

11. Puisque la valeur de la trace peut être calculée dans n'importe quelle base, choisissons celle que constituent les $|\varphi_j\rangle$. Le calcul est alors très simple :

$$\mathrm{Tr}\ D = \sum_j \langle\varphi_j|D|\varphi_j\rangle = \sum_j d_j \ .$$

La relation (H.15) s'écrit donc sous la forme (H.21).

12. Cette démonstration prouve que tout opérateur hermitique, de trace unité et semi-défini positif est un opérateur-densité susceptible de décrire un mélange statistique d'états purs.

13. L'expression (H.23) n'est en général pas unique. Elle l'est seulement si aucune des valeurs propres P_j de D n'est dégénérée. Pour une valeur propre dégénérée g fois, en effet, on peut choisir librement les g vecteurs propres orthonormés qui lui correspondent; si l'on change ce système de g vecteurs, on change évidemment l'expression (H.23) de D, mais pas D lui-même car la somme des projecteurs sur ces g vecteurs est toujours égale au projecteur sur le sous-espace de dégénérescence qu'ils engendrent, et qui reste inchangé.

14. Supposons en effet que les états $|\psi_m\rangle$ du mélange défini en (H.1)-(H.2), qui sont déjà normés à 1, soient en outre orthogonaux :

$$\langle\psi_{m'}|\psi_m\rangle = \delta_{m'm} \ . \tag{1}$$

Appliquons alors l'opérateur-densité (H.6) à l'un d'entre eux, $|\psi_{m_0}\rangle$:

$$D|\psi_{m_0}\rangle = \sum_m |\psi_m\rangle P_m \langle\psi_m|\psi_{m_0}\rangle = \sum_m |\psi_m\rangle P_m \delta_{mm_0} = P_{m_0}|\psi_{m_0}\rangle \ . \tag{2}$$

Le ket $|\psi_{m_0}\rangle$ est donc vecteur propre de D avec la valeur propre P_{m_0}.

3. MATRICE-DENSITÉ

Lorsqu'on choisit une *base orthonormée* $\{|v_i\rangle\}$, *a priori* quelconque, dans l'espace des états, les kets sont représentés par des vecteurs-colonnes et les opérateurs par des matrices. Un opérateur-densité est alors caractérisé par une *matrice-densité*, d'éléments

$$D_{ij} \widehat{=} \langle v_i | D | v_j \rangle \ . \tag{H.25}$$

La matrice-densité est *hermitique :*

$$D^\dagger = D \implies D^*_{ij} = D_{ji} \ . \tag{H.26}$$

Ses éléments diagonaux sont non négatifs d'après (H.16) :

$$D_{ii} \widehat{=} \langle v_i | D | v_i \rangle \geqslant 0 \ , \tag{H.27}$$

et leur somme vaut 1 d'après (H.15) :

$$\mathrm{Tr}\ D = \sum_i D_{ii} = 1 \ . \tag{H.28}$$

L'élément diagonal D_{ii} est appelé *«population»* de l'état $|v_i\rangle$, l'élément non-diagonal D_{ij} *«cohérence»* entre les deux états $|v_i\rangle$ et $|v_j\rangle$.

Si la base choisie est celle des *états propres* $|\varphi_j\rangle$ de l'opérateur-densité considéré, la matrice-densité est *diagonale :* les cohérences sont toutes nulles et la population de l'état $|\varphi_j\rangle$ coïncide avec sa probabilité P_j[15].

4. ÉVOLUTION DANS LE TEMPS

Reprenons la définition (H.6) de l'opérateur-densité D. Chacun des états purs $|\psi_m\rangle$ évolue dans le temps suivant l'équation de Schrödinger :

$$i\hbar \frac{\mathrm{d}}{\mathrm{d}t} |\psi_m(t)\rangle = H |\psi_m(t)\rangle \ , \tag{H.29}$$

où H est l'hamiltonien du système. Les probabilités P_m caractérisant le mélange statistique sont fixées une fois pour toutes : si, à l'instant $t = 0$ par exemple, la probabilité pour que le système soit dans l'état $|\psi_m(0)\rangle$ est P_m, elle est aussi P_m pour que le système se trouve à l'instant t dans l'état $|\psi_m(t)\rangle$ donné par (H.29). On calcule alors facilement la dérivée par rapport au temps de l'opérateur $D(t)$ correspondant :

$$\frac{\mathrm{d}}{\mathrm{d}t} D(t) = \sum_m P_m \left[\frac{\mathrm{d}}{\mathrm{d}t} |\psi_m(t)\rangle \right] \langle \psi_m(t)| + \sum_m P_m |\psi_m(t)\rangle \frac{\mathrm{d}}{\mathrm{d}t} \langle \psi_m(t)| \ . \tag{H.30}$$

L'évolution au cours du temps de l'opérateur-densité obéit donc à l'équation[16]

$$i\hbar \frac{\mathrm{d}}{\mathrm{d}t} D(t) = [H, D(t)] \ . \tag{H.31}$$

---

15. Bien que les populations D_{ii} possèdent les propriétés d'un ensemble de probabilités (formules (H.27) et (H.28)), l'état décrit par l'opérateur-densité D n'est un mélange statistique des états $|v_i\rangle$, avec les probabilités D_{ii}, que *si les cohérences sont toutes nulles,* c'est-à-dire si les $|v_i\rangle$ sont tous vecteurs propres de D; en général,

$$D \neq \sum_i D_{ii} |v_i\rangle\langle v_i|$$

(les éléments de matrice non diagonaux du second membre sont toujours nuls, par suite de l'orthonormalisation des $|v_i\rangle$; ceux du premier membre ne le sont que si les $|v_i\rangle$ sont vecteurs propres de D).

16. En prenant le bra conjugué des deux membres de l'équation de Schrödinger (H.29) (C.D.L., p. 120), on trouve que $\langle \psi_m(t)|$ évolue selon

$$-i\hbar \frac{\mathrm{d}}{\mathrm{d}t} \langle \psi_m(t)| = \langle \psi_m(t)| H \ , \tag{1}$$

.../...

5. OPÉRATEUR-DENSITÉ DÉCRIVANT UN ÉTAT PUR

Considérons le cas particulier où le mélange statistique ne comporte qu'un seul ket, $|\psi\rangle$, de probabilité 1. Le système se trouve alors dans l'*état pur* caractérisé par le ket $|\psi\rangle$. La définition (H.6) donne dans ce cas

$$D = |\psi\rangle\langle\psi| \ . \tag{H.32}$$

Un état pur peut donc être également décrit par un opérateur-densité, qui est dans ce cas un *projecteur*[17].

L'opérateur-densité associé à un état pur possède les propriétés générales énumérées jusqu'ici. Il vérifie en outre les *relations particulières*[18]

$$D^2 = D \tag{H.33}$$

et donc

$$\mathrm{Tr}\ D^2 = 1 \ . \tag{H.34}$$

Inversement, un opérateur-densité vérifiant l'une ou l'autre des relations (H.33) et (H.34) décrit un état pur[19].

---

car H est hermitique. Il suffit de reporter cette expression et celle de la dérivée du ket $|\psi_m(t)\rangle$ dans (H.30) pour obtenir l'équation (H.31) :

$$i\hbar \frac{dD}{dt} = \sum_m P_m H |\psi_m(t)\rangle\langle\psi_m(t)| - \sum_m P_m |\psi_m(t)\rangle\langle\psi_m(t)| H = HD(t) - D(t)H = [H, D(t)] \ . \tag{2}$$

17. Considérons un état pur, décrit par le ket $|\psi\rangle$ normé. Tout ket $|\psi'\rangle$ de la forme

$$|\psi'\rangle = e^{i\theta}|\psi\rangle \ , \tag{1}$$

où θ est un nombre réel quelconque, décrit le même état physique que $|\psi\rangle$ (C.D.L., p. 219). Mais les opérateurs-densités associés à tous ces kets équivalents sont égaux :

$$|\psi'\rangle\langle\psi'| = e^{i\theta}|\psi\rangle e^{-i\theta}\langle\psi| = |\psi\rangle\langle\psi| \ . \tag{2}$$

A chaque état physique correspond donc un opérateur-densité unique.

18. La relation (H.33) découle immédiatement de (H.32) :

$$D^2 = |\psi\rangle\langle\psi|\psi\rangle\langle\psi| = |\psi\rangle\langle\psi|$$

(puisque $|\psi\rangle$ est normé à 1). Cette relation est caractéristique des projecteurs : deux projections successives sur le même vecteur donnent le même résultat qu'une seule.

19. Soit un opérateur-densité D, c'est-à-dire un opérateur possédant les propriétés (H.14), (H.15) et (H.16). Choisissons pour base, dans l'espace des états, un ensemble orthonormé $\{|\varphi_j\rangle\}$ de vecteurs propres de D. Si nous écrivons D sous la forme (H.23), son carré s'exprime simplement :

$$D^2 = \sum_j \sum_{j'} |\varphi_j\rangle P_j \langle\varphi_j|\varphi_{j'}\rangle P_{j'} \langle\varphi_{j'}| = \sum_j P_j^2 |\varphi_j\rangle\langle\varphi_j| \tag{1}$$

(par suite de la relation d'orthonormalisation (H.18) des $|\varphi_j\rangle$).

Supposons alors que D soit tel que

$$D^2 = D \ . \tag{2}$$

Les $|\varphi_j\rangle$ formant une base, ceci n'est possible que si

$$P_j^2 = P_j \quad \text{pour tout } j. \tag{3}$$

A cause de la normalisation des probabilités P_j (formule (H.24.b)), une et une seule d'entre elles vaut 1, toutes les autres étant nulles. D décrit donc un état pur.

Il en est de même si l'on sait seulement que

$$\mathrm{Tr}\ D^2 = 1 \ . \tag{4}$$

.../...

III. OPÉRATEUR-DENSITÉ ET DISTRIBUTION STATISTIQUE DES ÉTATS STATIONNAIRES

Considérons un système macroscopique. Nous le supposerons *isolé,* de sorte que son hamiltonien H est indépendant du temps (système conservatif).

1. SYSTÈME A L'ÉQUILIBRE MACROSCOPIQUE

L'opérateur-densité décrivant un état d'équilibre est *indépendant du temps :*

$$\frac{\mathrm{d}D}{\mathrm{d}t}=0\ . \tag{H.35}$$

L'équation d'évolution (H.31) montre qu'alors D *commute avec l'hamiltonien* H. Il existe donc (au moins) une base de l'espace des états constituée de *vecteurs propres communs à* H *et* D (C.D.L., p. 140); nous les noterons $|\varphi_\ell\rangle$:

$$H\,|\varphi_\ell\rangle=E_\ell|\varphi_\ell\rangle\ , \tag{H.36.a}$$

$$D\,|\varphi_\ell\rangle=P_\ell|\varphi_\ell\rangle\ . \tag{H.36.b}$$

Si les états stationnaires $|\varphi_\ell\rangle$ sont connus, la donnée de l'ensemble des P_ℓ est équivalente à celle de l'opérateur-densité D (cf. formule (H.23)). Ceci justifie donc le choix que nous avons fait (I, § III.A.2.b) de caractériser l'état macroscopique d'un système à l'équilibre par la *distribution statistique* $\{P_\ell\}$ *des états microscopiques stationnaires* $|\varphi_\ell\rangle$.

Le formalisme de l'opérateur-densité permet cependant une description plus générale et plus souple des états d'équilibre macroscopique que l'utilisation des états stationnaires $|\varphi_\ell\rangle$ et de leurs probabilités P_ℓ. On peut en effet définir et manipuler l'opérateur-densité sans avoir à choisir une base dans l'espace des états; et si l'on en choisit une, elle ne doit pas être nécessairement constituée d'états stationnaires. Ceci représente un avantage considérable dans la pratique, car on ne sait généralement pas, pour un système macroscopique, résoudre exactement l'équation aux valeurs propres de l'hamiltonien H. On sait le faire seulement pour quelques systèmes particulièrement simples, dans lesquels les interactions entre constituants microscopiques sont négligeables : gaz parfaits (III, § IV.B et chapitre VI), cristal paramagnétique parfait (complément II.B), ... Le problème devient inextricable dès que l'on doit tenir compte des interactions. On fait alors appel à des méthodes d'approximation, qu'il est beaucoup plus commode d'appliquer directement à l'opérateur-densité plutôt qu'à la détermination préalable des états stationnaires (cf. complément V.E, § II).

En effet, comme les $|\varphi_j\rangle$ sont orthonormés, D^2 est comme D diagonal dans cette base, de sorte que

$$\mathrm{Tr}\ D^2=\sum_j P_j^2\ . \tag{5}$$

On a donc

$$\sum_j P_j^2=\sum_j P_j=1\ . \tag{6}$$

Mais, les P_j étant tous inférieurs ou égaux à 1, on sait que

$$P_j^2\leqslant P_j \quad \text{pour tout } j. \tag{7}$$

Pour que la somme des P_j^2 ait même valeur que celle des P_j, il faut nécessairement que

$$P_j^2=P_j \quad \text{pour tout } j. \tag{8}$$

Comme plus haut, ceci implique que D se rapporte à un état pur.

En résumé, c'est le formalisme de l'opérateur-densité qu'il conviendrait d'utiliser en mécanique statistique, car c'est l'outil le mieux adapté, tant pour les développements théoriques que pour les applications. Mais la description en termes de probabilités des états stationnaires, plus simple et plus intuitive, est *fondamentalement équivalente* à celle que donne l'opérateur-densité.

2. ÉTATS HORS D'ÉQUILIBRE

Étudions maintenant le cas général où l'opérateur-densité dépend *a priori* du temps.

L'hamiltonien H, ses états propres $|\varphi_\ell\rangle$ et les énergies E_ℓ associées sont quant à eux indépendants du temps. Choisissons encore l'ensemble des $|\varphi_\ell\rangle$ comme base dans l'espace des états. L'équation d'évolution (H.31) indique alors[20] que les éléments de matrice diagonaux de l'opérateur-densité sont constants dans le temps :

$$\frac{\mathrm{d}}{\mathrm{d}t} D_{\ell\ell} = 0 \; ; \tag{H.37}$$

ses éléments de matrice non-diagonaux vérifient l'équation différentielle

$$i\hbar \frac{\mathrm{d}}{\mathrm{d}t} D_{\ell\ell'} = (E_\ell - E_{\ell'}) D_{\ell\ell'} \tag{H.38}$$

et sont donc de la forme

$$D_{\ell\ell'}(t) = c_{\ell\ell'} \mathrm{e}^{-i(E_\ell - E_{\ell'})t/\hbar} \; , \tag{H.39}$$

où les $c_{\ell\ell'}$ sont des constantes par rapport au temps.

Supposons que le système étudié soit *parfaitement isolé*. Son énergie est alors exactement fixée. Le mélange statistique qui décrit son état ne fait intervenir que des états purs correspondant à cette valeur de l'énergie, c'est-à-dire des états propres de H associés à une seule et même valeur propre. Dans ces conditions, les éléments de matrice de D qui pourraient dépendre du temps sont nuls, puisqu'ils devraient correspondre à des énergies E_ℓ et $E_{\ell'}$ différentes (formules (H.38)-(H.39)). Donc l'opérateur-densité D, solution de l'équation d'évolution (H.31) correspondant à une énergie fixée E, est représenté par une matrice-densité constante dans le temps, dont les seuls éléments non nuls relient des états propres de H associés à la valeur propre E. L'opérateur D étant hermitique, on peut diagonaliser cette matrice par un changement de base dans le sous-espace de dégénérescence de la valeur propre E; on est alors ramené à un ensemble de probabilités P_ℓ comme au paragraphe 1.

Mais *pour un système macroscopique, la réalité est tout autre.* Il n'est *jamais parfaitement isolé,* et son énergie n'est définie qu'avec *une incertitude δE non nulle* (voir complément IV.B). Il ne suit donc pas les équations (H.37)-(H.39); son évolution dans le temps est décrite par l'*« équation maîtresse »* que nous étudierons au chapitre IV. Effectivement, l'expérience montre qu'un système macroscopique « isolé », abandonné à lui-même dans un état hors d'équilibre, évolue au cours du temps pour atteindre finalement un état d'équilibre macroscopique, où ses propriétés ne dépendent plus du temps.

---

20. Appliquons les deux membres de (H.31) au ket $|\varphi_{\ell'}\rangle$ et multiplions le résultat par le bra $\langle\varphi_\ell|$:

$$i\hbar \langle\varphi_\ell| \frac{\mathrm{d}D}{\mathrm{d}t} |\varphi_{\ell'}\rangle = \langle\varphi_\ell| HD - DH |\varphi_{\ell'}\rangle \; . \tag{1}$$

Les kets de base étant indépendants du temps et vecteurs propres de l'opérateur hermitique H, cette égalité peut être écrite

$$i\hbar \frac{\mathrm{d}}{\mathrm{d}t} D_{\ell\ell'} = E_\ell D_{\ell\ell'} - E_{\ell'} D_{\ell\ell'} \; , \tag{2}$$

où $$D_{\ell\ell'} \widehat{=} \langle\varphi_\ell| D |\varphi_{\ell'}\rangle \; . \tag{3}$$

On en déduit aussitôt (H.37) et (H.38).

IV. ENTROPIE STATISTIQUE D'UN ÉTAT DÉCRIT PAR UN OPÉRATEUR-DENSITÉ

1. EXPRESSION DE L'ENTROPIE STATISTIQUE

Dans un mélange statistique d'états purs (§ I.1), les probabilités interviennent parce que l'état quantique (pur) du système n'est pas connu avec certitude[21]. Ceci traduit un *manque d'information* qu'il s'agit d'évaluer.

Pour ce faire, la première idée qui se présente est d'appliquer la définition usuelle de l'entropie statistique (I, § II.A) à la distribution de probabilités $\{P_m\}$ des états purs $|\psi_m\rangle$ dans le mélange. Mais nous savons que des mélanges statistiques différents, construits à partir d'un ensemble d'états $|\psi_m\rangle$ différents et de probabilités P_m différentes, peuvent décrire le même état physique (voir paragraphe I.2 et note 7); l'entropie statistique ne prend pas la même valeur pour les diverses distributions $\{P_m\}$ de ces mélanges équivalents[22]. Il faut donc préciser la définition de l'entropie statistique d'un mélange d'états purs de façon que sa valeur soit une propriété intrinsèque de l'état macroscopique considéré, c'est-à-dire ne dépende pas du mélange statistique particulier que l'on choisit pour représenter cet état; autrement dit, il faut que l'entropie statistique fasse directement intervenir l'opérateur-densité D caractérisant l'état macroscopique étudié.

La seule définition vérifiant cette condition et compatible avec la définition de l'entropie statistique posée antérieurement s'écrit[23]

$$\boxed{S(D) \widehat{=} -k \operatorname{Tr}(D \ln D)} \quad . \tag{H.40}$$

Explicitons en effet cette expression en prenant pour base l'ensemble $\{|\varphi_j\rangle\}$ des vecteurs propres de D; nous obtenons alors la formule familière

$$S = -k \sum_j P_j \ln P_j \, . \tag{H.41}$$

---

21. Les prédictions physiques données par la mécanique quantique s'expriment toujours en termes probabilistes. Ceci n'est cependant pas lié à un manque d'information si l'on connaît l'état pur dans lequel se trouve le système : d'après le premier postulat de la mécanique quantique (appendice II, § I.1), l'état d'un système est complètement caractérisé par la donnée d'un ket de l'espace des états; celui-ci contient toute l'information qu'il est possible d'obtenir.

22. L'état non polarisé d'un spin 1/2, par exemple (voir note 7), peut aussi bien être décrit par un mélange de deux états purs ayant chacun la probabilité 1/2, ce qui donnerait $S = k \ln 2$, que par un mélange de quatre états purs ayant chacun la probabilité 1/4, ce qui donnerait $S = k \ln 4$.

23. Le logarithme de l'opérateur D est défini comme l'opérateur ayant mêmes vecteurs propres que D, avec pour valeurs propres le logarithme de celles de D :

$$(\ln D)\,|\varphi_j\rangle = (\ln P_j)\,|\varphi_j\rangle \, . \tag{1}$$

L'équivalence entre les égalités (H.40) et (H.41) se démontre facilement. La trace étant indépendante de la base choisie pour la calculer, on peut prendre celle des $|\varphi_j\rangle$ pour expliciter (H.40) :

$$S(D) = -k \sum_j \langle \varphi_j | D \ln D | \varphi_j \rangle \, . \tag{2}$$

Mais $$(D \ln D)|\varphi_j\rangle = D \ln P_j |\varphi_j\rangle = P_j \ln P_j |\varphi_j\rangle \, . \tag{3}$$

On trouve donc aussitôt l'expression (H.41).

Mais cette dernière formule n'est valable que dans le cas où les kets constituant le mélange statistique sont *orthogonaux* (auquel cas il sont nécessairement vecteurs propres de D; cf. paragraphe II.2). Pour un mélange statistique quelconque, c'est la formule générale (H.40) qu'il convient d'utiliser.

2. UNE INÉGALITÉ UTILE

L'entropie statistique $S(D)$ associée à un opérateur-densité D déterminé vérifie, *pour tout opérateur-densité* D', l'inégalité suivante :

$$S(D) \leqslant -k \operatorname{Tr}(D \ln D') , \tag{H.42}$$

cette inégalité devenant une *égalité si et seulement si*

$$D' = D . \tag{H.43}$$

Cette relation est analogue à celle que nous avons écrite dans le complément I.G (formule (G.28)) pour des distributions de probabilités ordinaires. Elle est un peu plus délicate à démontrer, car D' *et* D *n'ont pas en général le même ensemble de vecteurs propres*[24].

---

24. Considérons les deux ensembles de vecteurs propres de D et D' :

$$D|\varphi_j\rangle = P_j|\varphi_j\rangle ; \quad D'|\varphi'_i\rangle = P'_i|\varphi'_i\rangle . \tag{1}$$

Nous les supposons tous deux orthonormés :

$$\langle\varphi_j|\varphi_{j'}\rangle = \delta_{jj'} ; \quad \langle\varphi'_i|\varphi'_{i'}\rangle = \delta_{ii'} . \tag{2}$$

Chacun d'eux constitue une base dans l'espace des états, et vérifie donc une relation de fermeture :

$$\sum_j |\varphi_j\rangle\langle\varphi_j| = \mathbb{1} ; \quad \sum_i |\varphi'_i\rangle\langle\varphi'_i| = \mathbb{1} . \tag{3}$$

Calculons alors

$$S(D) + k \operatorname{Tr}(D \ln D') = k\left[\operatorname{Tr}(D \ln D') - \operatorname{Tr}(D \ln D)\right] \tag{4}$$

en utilisant la base $\{|\varphi_j\rangle\}$:

$$S(D) + k \operatorname{Tr}(D \ln D') = k \sum_j P_j \langle\varphi_j|\ln D'|\varphi_j\rangle - k \sum_j P_j \ln P_j . \tag{5}$$

Pour expliciter l'élément de matrice de $\ln D'$, il faut introduire les $|\varphi'_i\rangle$, ce que l'on fait à l'aide de leur relation de fermeture :

$$\langle\varphi_j|\ln D'|\varphi_j\rangle = \sum_i \langle\varphi_j|\ln D'|\varphi'_i\rangle\langle\varphi'_i|\varphi_j\rangle = \sum_i (\ln P'_i)\langle\varphi_j|\varphi'_i\rangle\langle\varphi'_i|\varphi_j\rangle . \tag{6}$$

Il est commode de faire apparaître les mêmes produits scalaires de vecteurs dans le deuxième terme de l'expression (5) à calculer; ceci est facile car

$$\sum_i \langle\varphi_j|\varphi'_i\rangle\langle\varphi'_i|\varphi_j\rangle = \langle\varphi_j|\varphi_j\rangle = 1 . \tag{7}$$

On trouve ainsi

$$S(D) + k \operatorname{Tr}(D \ln D') = k \sum_j \sum_i P_j \langle\varphi_j|\varphi'_i\rangle\langle\varphi'_i|\varphi_j\rangle \times (\ln P'_i - \ln P_j) . \tag{8}$$

Utilisons alors, comme dans le complément I.G (note 5), l'inégalité

$$\ln x \leqslant x - 1 , \tag{9}$$

valable pour tout x réel et devenant une égalité seulement pour $x = 1$. On en déduit

$$P_j \ln \frac{P'_i}{P_j} \leqslant P_j \left(\frac{P'_i}{P_j} - 1\right) = P'_i - P_j , \tag{10}$$

.../...

3. PROPRIÉTÉS DE L'ENTROPIE STATISTIQUE

Nous allons retrouver à partir de l'expression (H.40) les propriétés de l'entropie statistique que nous avons énumérées au chapitre I, paragraphe II.B et dans le complément I.G, paragraphe II. Les difficultés créées par le fait que D est un opérateur sont résolues en utilisant l'inégalité (H.42).

a. Positivité. Minimum

Les valeurs propres d'un opérateur-densité quelconque sont des probabilités, comprises entre 0 et 1. Si donc on écrit l'entropie statistique $S(D)$ sous la forme (H.41), on voit aussitôt que

$$S(D) \geqslant 0 \quad \text{pour tout } D \,. \tag{H.44}$$

On constate de plus que $S(D)$ est *nulle*, c'est-à-dire *minimale*, si et seulement si l'une des valeurs propres de D est égale à 1, toutes les autres étant nulles, c'est-à-dire (§ II.5) *si et seulement si D décrit un état pur* (cf. note 19).

b. Maximum sur un sous-espace de dimension finie

Supposons que, par des arguments physiques[25], nous sachions que les seuls vecteurs d'état possibles pour le système dans la situation envisagée sont ceux d'un

---

l'égalité se produisant seulement pour $P'_i = P_j$. Reportons dans l'expression (8) :

$$S(D) + k \operatorname{Tr}(D \ln D') \leqslant k \sum_j \sum_i \langle \varphi_j | \varphi'_i \rangle \langle \varphi'_i | \varphi_j \rangle \times (P'_i - P_j) \,; \tag{11}$$

cette inégalité devient égalité seulement si tous les termes non nuls de la somme réalisent l'égalité dans (10), c'est-à-dire seulement si

$$\text{pour tout couple } (i, j), \text{ ou bien } \langle \varphi'_i | \varphi_j \rangle = 0 \,,$$
$$\text{ou bien } P'_i = P_j \,. \tag{12}$$

Le second membre de l'inégalité (11) est simple à évaluer :

$$\sum_j \sum_i \langle \varphi_j | \varphi'_i \rangle \langle \varphi'_i | \varphi_j \rangle P'_i = \sum_i P'_i \sum_j \langle \varphi'_i | \varphi_j \rangle \langle \varphi_j | \varphi'_i \rangle = \sum_i P'_i \langle \varphi'_i | \varphi'_i \rangle = \sum_i P'_i = 1 \,, \tag{13}$$

et de même pour l'autre double somme. Donc

$$S(D) + k \operatorname{Tr} D \ln D' \leqslant 0 \,. \tag{14}$$

D'après (12), la condition d'annulation s'écrit

$$\langle \varphi_j | \varphi'_i \rangle (P'_i - P_j) = 0 \quad \forall i, j \,. \tag{15}$$

En utilisant les relations (1), on peut la transformer en

$$\langle \varphi_j | (D' - D) | \varphi'_i \rangle = 0 \quad \forall i, j \,. \tag{16}$$

Comme l'ensemble des $|\varphi_j\rangle$ forme une base dans l'espace des états, l'égalité (16) implique la nullité des vecteurs $(D' - D)|\varphi'_i\rangle$:

$$(D' - D)|\varphi'_i\rangle = 0 \quad \forall i \,. \tag{17}$$

Mais les $|\varphi'_i\rangle$ forment eux aussi une base; c'est donc l'opérateur $D' - D$ lui-même qui est nul :

$$D' - D = 0 \,. \tag{18}$$

25. Par exemple, l'énergie d'un système isolé est fixée. Si l'on sait qu'elle vaut E avec une incertitude δE, seuls les kets correspondant à une énergie comprise entre E et $E + \delta E$ ont une probabilité non nulle de décrire l'état du système (voir chapitre II, § I.B.1).

sous-espace $\mathcal{E}_\alpha$, de dimension *finie* Ω, de l'espace des états $\mathcal{E}$. Ceci signifie que l'opérateur-densité D ne peut avoir d'éléments de matrice non nuls qu'entre vecteurs de ce sous-espace $\mathcal{E}_\alpha$.

Parmi les opérateurs-densités qui vérifient cette condition, il en est un qui est proportionnel à l'opérateur-identité $\mathbb{1}_\alpha$ dans $\mathcal{E}_\alpha$[26] : comme sa trace doit être égale à 1, il s'écrit

$$D_0 = \frac{1}{\Omega}\,\mathbb{1}_\alpha \,. \tag{H.45}$$

Il correspond à des *probabilités égales* pour les Ω kets d'une base quelconque du sous-espace $\mathcal{E}_\alpha$. On montre facilement, grâce à l'inégalité (H.42), que c'est lui qui *rend maximale l'entropie statistique*[27]. Ce maximum vaut

$$S(D_0) = k \ln \Omega \,. \tag{H.46}$$

c. Additivité

Considérons un système $\mathcal{S}$ constitué de deux sous-systèmes $\mathcal{S}_1$ et $\mathcal{S}_2$. L'espace des états $\mathcal{E}$ de $\mathcal{S}$ est le produit tensoriel des espaces d'états $\mathcal{E}_1$ et $\mathcal{E}_2$ associés à $\mathcal{S}_1$ et $\mathcal{S}_2$ (appendice II, § III).

Supposons que l'opérateur-densité D décrivant l'état de $\mathcal{S}$ soit un produit tensoriel d'un opérateur-densité $D(1)$ défini dans $\mathcal{E}_1$ et d'un opérateur-densité $D(2)$ défini dans $\mathcal{E}_2$[28] :

$$D = D(1) \otimes D(2) \,. \tag{H.47}$$

---

26. L'opérateur que nous notons ici $\mathbb{1}_\alpha$ laisse inchangé tout vecteur de $\mathcal{E}_\alpha$ et donne zéro si on l'applique à un vecteur n'ayant aucune composante dans $\mathcal{E}_\alpha$. Il s'identifie donc au *projecteur* $\mathcal{P}(\mathcal{E}_\alpha)$ *sur le sous-espace* $\mathcal{E}_\alpha$.

27. Écrivons l'inégalité (H.42) pour un opérateur-densité D quelconque (n'ayant des éléments de matrice qu'à l'intérieur du sous-espace $\mathcal{E}_\alpha$) et pour $D' = D_0$:

$$S(D) \leqslant -k \,\mathrm{Tr}(D \ln D_0) \,. \tag{1}$$

Choisissons, pour évaluer le second membre, une base orthonormée $\{|v_i\rangle\}$ quelconque de $\mathcal{E}_\alpha$:

$$S(D) \leqslant -k \sum_{i=1}^{\Omega} \langle v_i | D \ln D_0 | v_i \rangle \,. \tag{2}$$

Mais les $|v_i\rangle$ sont forcément vecteurs propres de D_0, à cause de sa forme très particulière (H.45) :

$$D_0 |v_i\rangle = \frac{1}{\Omega} |v_i\rangle \,, \tag{3}$$

de sorte que

$$\ln D_0 |v_i\rangle = -\ln \Omega |v_i\rangle \,. \tag{4}$$

Si l'on reporte dans (2), on trouve finalement

$$S(D) \leqslant k \ln \Omega \sum_{i=1}^{\Omega} \langle v_i | D | v_i \rangle \,, \tag{5}$$

c'est-à-dire, puisque la trace de D dans $\mathcal{E}_\alpha$ vaut 1,

$$S(D) \leqslant k \ln \Omega \,. \tag{6}$$

L'égalité a lieu pour $D = D_0$, et seulement dans ce cas.

28. Le produit tensoriel $A(1) \otimes A(2)$ d'un opérateur $A(1)$ agissant dans $\mathcal{E}_1$ et d'un opérateur $A(2)$ agissant dans $\mathcal{E}_2$ est un opérateur qui agit dans $\mathcal{E}$; son action sur un vecteur produit tensoriel obéit par définition à la règle

$$[A(1) \otimes A(2)]\,|\varphi(1)\rangle \otimes |\chi(2)\rangle \widehat{=} [A(1)|\varphi(1)\rangle] \otimes [A(2)|\chi(2)\rangle] \,.$$

Comme tout vecteur de $\mathcal{E}$ peut être écrit comme une combinaison linéaire de vecteurs produits tensoriels, ceci suffit à caractériser $A(1) \otimes A(2)$.

Ceci signifie physiquement que chacun des deux sous-systèmes est dans un état déterminé, *indépendant* de l'état de l'autre sous-système : on peut changer $D(2)$ sans modifier $D(1)$, et *vice versa*. Dans ce cas, l'entropie statistique associée à D est simplement la somme des entropies associées à $D(1)$ et $D(2)$ séparément[29] :

$$S(D) = S[D(1)] + S[D(2)] . \tag{H.48}$$

On a donc *additivité de l'entropie statistique* lorsque les sous-systèmes sont dans des *états indépendants* l'un de l'autre[30].

d. Concavité

Soit deux opérateurs-densités D_1 et D_2 relatifs au même système, et α un nombre réel tel que

$$0 \leqslant \alpha \leqslant 1 . \tag{H.49}$$

L'entropie statistique $S(D)$ est une fonction partout *concave*, c'est-à-dire qu'elle vérifie la relation[31]

$$S[\alpha D_1 + (1-\alpha) D_2] \geqslant \alpha S(D_1) + (1-\alpha) S(D_2) , \tag{H.50}$$

quels que soient D_1 et D_2; les deux membres sont égaux seulement si $D_1 = D_2$.

---

29. Pour calculer $S(D)$, on choisit la base orthonormée de $\mathcal{E}$ construite par produit tensoriel à partir d'une base orthonormée de $\mathcal{E}_1$ constituée de vecteurs propres $|\varphi_j(1)\rangle$ de $D(1)$ et d'une base orthonormée de $\mathcal{E}_2$ constituée de vecteurs propres $|\chi_i(2)\rangle$ de $D(2)$:

$$|w_{ji}\rangle \widehat{=} |\varphi_j(1)\rangle \otimes |\chi_i(2)\rangle , \tag{1}$$

avec
$$D(1)|\varphi_j(1)\rangle = P_j^{(1)} |\varphi_j(1)\rangle , \tag{2.a}$$

$$D(2)|\chi_i(2)\rangle = P_i^{(2)} |\chi_i(2)\rangle . \tag{2.b}$$

On a alors

$$D|w_{ji}\rangle = [D(1)|\varphi_j(1)\rangle] \otimes [D(2)|\chi_i(2)\rangle] = P_j^{(1)} P_i^{(2)} |\varphi_j(1)\rangle \otimes |\chi_i(2)\rangle = P_j^{(1)} P_i^{(2)} |w_{ji}\rangle , \tag{3}$$

et par conséquent

$$\ln D|w_{ji}\rangle = \ln[P_j^{(1)} P_i^{(2)}]|w_{ji}\rangle . \tag{4}$$

On trouve ainsi

$$\begin{aligned} S(D) &= -k \sum_{j,i} \langle w_{ji}| D \ln D| w_{ji}\rangle \\ &= -k \sum_{j,i} P_j^{(1)} P_i^{(2)} [\ln P_j^{(1)} + \ln P_i^{(2)}] \\ &= -k \sum_i P_i^{(2)} \sum_j P_j^{(1)} \ln P_j^{(1)} - k \sum_j P_j^{(1)} \sum_i P_i^{(2)} \ln P_i^{(2)} . \end{aligned} \tag{5}$$

Chacun des deux termes se factorise en une somme de probabilités qui vaut 1 multipliée par l'entropie statistique associée à l'autre distribution. On aboutit bien à la formule (H.48).

30. Il existe des opérateurs-densités définis dans $\mathcal{E}$ qui ne sont pas produits tensoriels d'un opérateur de $\mathcal{E}_1$ par un opérateur de $\mathcal{E}_2$. Ils décrivent des états du système global $\mathcal{S}$ dans lesquels les propriétés des sous-systèmes $\mathcal{S}_1$ et $\mathcal{S}_2$ sont *corrélées*. On peut dans ce cas démontrer une *inégalité de sous-additivité* analogue à la relation (G.32) du complément I.G (voir réf. 1, p. 101).

31. On utilise encore l'inégalité (H.42), en choisissant

$$D' = \alpha D_1 + (1-\alpha) D_2 , \tag{1}$$

et successivement $D = D_1$ puis $D = D_2$:

$$S(D_1) \leqslant -k \operatorname{Tr}\{D_1 \ln[\alpha D_1 + (1-\alpha) D_2]\} , \tag{2.a}$$

$$S(D_2) \leqslant -k \operatorname{Tr}\{D_2 \ln[\alpha D_1 + (1-\alpha) D_2]\} . \tag{2.b}$$

.../...

V. Conclusion

En mécanique quantique, l'état macroscopique d'un système est caractérisé de façon générale par un opérateur-densité défini dans l'espace des états du système. Nous avons introduit ici le formalisme correspondant, et présenté les principales propriétés des opérateurs-densités. Nous utiliserons ces notions et résultats dans le complément V.E.

Dans tout le reste du livre, nous raisonnons dans un cadre plus simple et plus intuitif : l'état macroscopique est décrit par la distribution statistique $\{P_\ell\}$ des états propres (ℓ) de l'hamiltonien. Nous avons montré ci-dessus (§ III) que cette description est fondamentalement équivalente à celle de l'opérateur-densité. Celle-ci est toutefois plus compacte, et surtout plus souple et plus puissante, car elle s'accommode d'un choix de base quelconque dans l'espace des états. Cette liberté de choix constitue un avantage important lorsqu'il s'agit par exemple de développer des méthodes d'approximation : le complément V.E (§ II) en fournira une illustration simple.

Références

1. R. Balian, *Du microscopique au macroscopique,* cours de Physique Statistique de l'École Polytechnique, Ellipses, Paris (1982), chapitres II et III.

2. C.D.L., complément E_{III}.

Comme α et $(1-\alpha)$ sont positifs et que l'on peut en outre les faire entrer dans la somme que représente la trace, on en déduit

$$\alpha S(D_1)+(1-\alpha)S(D_2)\leqslant -k\ \mathrm{Tr}\{[\alpha D_1+(1-\alpha)D_2]\ln[\alpha D_1+(1-\alpha)D_2]\}\ , \tag{3}$$

c'est-à-dire la formule (H.50). L'égalité a lieu si et seulement si

$$D_1=\alpha D_1+(1-\alpha)D_2 \tag{4.a}$$

et

$$D_2=\alpha D_1+(1-\alpha)D_2\ , \tag{4.b}$$

c'est-à-dire

$$D_1=D_2\ . \tag{5}$$

Exercices I

DÉNOMBREMENTS ET CALCULS DE PROBABILITÉS

I.1. On distribue N jetons dans deux sacs. Combien y a-t-il de répartitions différentes dans lesquelles le premier sac contient N_1 jetons

a) si l'on distingue les jetons les uns des autres par un numéro;

b) si on ne les distingue pas?

Quel est le nombre total de répartitions possibles des jetons entre les deux sacs (sans restriction sur le nombre de jetons par sac) dans les cas *a* et *b*?

I.2. On répartit N boules entre n boîtes numérotées, le nombre de boules par boîte n'étant pas limité. On caractérise une répartition par le nombre de boules présentes dans chaque boîte, sans distinguer les boules entre elles. Calculer le nombre de répartitions différentes. Application : de combien de façons différentes peut-on choisir n nombres entiers positifs ou nuls $q_1, q_2, \ldots, q_n$ de façon que leur somme soit égale à N?

Suggestion : on représente chaque distribution par un schéma sur lequel on fait figurer, alignées, les boules et les cloisons séparant les boîtes (lorsque deux cloisons se suivent, la boîte correspondante est vide); la permutation d'une cloison et d'une boule donne un schéma représentant une autre répartition des boules dans les boîtes, tandis que l'échange de deux boules entre elles ou de deux cloisons entre elles ne modifie pas la répartition.

Réponse : $(N+n-1)!/N!(n-1)!$

I.3. On lance simultanément 400 pièces de monnaie. Calculer la probabilité pour qu'exactement 220 d'entre elles montrent le côté pile.

I.4. Dans un jeu de pile ou face, deux joueurs lancent alternativement une pièce de monnaie. Le premier qui obtient pile a gagné.

Quelle est la probabilité pour que pile apparaisse pour la première fois au n-ième lancer? En déduire la probabilité pour que le gagnant soit le joueur qui entame la partie.

EXERCICES GÉNÉRAUX SUR LES PROBABILITÉS

I.5. Soit X et Y deux variables aléatoires indépendantes et S leur somme. Montrer que

$$\overline{S} = \overline{X} + \overline{Y}$$

et $$(\Delta S)^2 = (\Delta X)^2 + (\Delta Y)^2.$$

I.6. Soit u une variable aléatoire qui peut prendre un ensemble de valeurs *positives* u_i, les probabilités correspondantes étant notées P_i. On veut démontrer l'inégalité

$$\overline{\left(\frac{1}{u}\right)} \geqslant \frac{1}{\bar{u}} . \tag{1}$$

1) On considère la somme

$$S \widehat{=} \sum_i \sum_j P_i P_j \frac{1}{u_i} \frac{1}{u_j} (u_i - u_j)^2 . \tag{2}$$

Montrer que

$$S = 2\left[\bar{u} \times \overline{\left(\frac{1}{u}\right)} - 1\right] . \tag{3}$$

2) En déduire l'inégalité (1). Dans quels cas devient-elle une égalité?

I.7. Un objet ponctuel est assujetti à se trouver dans le plan xOy. La densité de probabilité $w(x, y)$ pour qu'il soit situé au point M, de coordonnées cartésiennes (x, y), est nulle en dehors du carré

$$\begin{cases} 0 < x < 1 \\ 0 < y < 1, \end{cases}$$

et uniforme sur ce carré.

Trouver l'expression de la densité de probabilité $\tilde{w}(r, \theta)$ lorsqu'on repère le point M par ses coordonnées polaires (r, θ). Calculer la probabilité de trouver l'objet à une distance de l'origine comprise entre r et $r + \mathrm{d}r$. Calculer la probabilité pour que l'angle entre la droite OM et l'axe Ox soit compris entre θ et $\theta + \mathrm{d}\theta$.

I.8. Dans la définition I.(II.2) de l'entropie statistique S, remplacer P_M en fonction de $P_1, P_2, \ldots, P_{M-1}$. Chercher ensuite le maximum de S comme fonction de ces $(M-1)$ variables indépendantes et retrouver le résultat I.(II.8).

I.9. Déterminer (par la méthode des multiplicateurs de Lagrange) la distribution de probabilités $\{P_\ell\}$ qui rend maximale l'entropie statistique S lorsque la valeur moyenne $\bar{y}$ d'une grandeur y est fixée. On montrera que la solution est de la forme

$$P_\ell = \frac{1}{A} \exp\left[-\lambda y_\ell\right] ,$$

et on écrira les relations permettant en principe de calculer les constantes A et λ.

I.10. Distribution de Poisson

1) Une variable aléatoire, dont les valeurs possibles sont les *entiers positifs ou nul* n, est distribuée selon la loi de Poisson si la probabilité pour qu'elle prenne la valeur n est de la forme

$$P_P(n; \lambda) \widehat{=} \frac{\lambda^n}{n!} \mathrm{e}^{-\lambda} , \tag{1}$$

où λ est un *paramètre réel positif.*

a) Vérifier que cette distribution est convenablement normalisée.

b) Montrer que la valeur moyenne $\bar{n}$ correspondante est égale à λ et l'écart quadratique moyen Δn à $\sqrt{\lambda}$.

c) Calculer le rapport $P_P(n;\lambda)/P_P(n-1;\lambda)$. En déduire le sens de variation de P_P en fonction de n, pour λ fixé.

2) On reprend la *distribution binomiale* $P(N, n)$ (appendice I, § II.1), qui donne la probabilité pour qu'un événement particulier (de probabilité P_+) se réalise n fois lorsqu'on effectue N expériences. On va montrer que $P(N, n)$ *est bien approximée par une distribution de Poisson si l'événement considéré est rare :*

$$P_+ \ll 1 , \tag{2.a}$$

et si $$N \gg 1 . \tag{2.b}$$

a) Montrer que

$$\frac{P(N, n)}{P(N, n-1)} = \frac{\bar{n}}{n}\left[1 + \frac{P_+}{P_-} - \frac{n-1}{NP_-}\right], \tag{3}$$

où $P_- = 1 - P_+$ et $\bar{n} = NP_+$.

b) Si les conditions (2) sont satisfaites et si $n \ll N$, on peut assimiler à 1 le crochet de la formule (3). Montrer qu'alors

$$\frac{P(N, n)}{P(N, 0)} \simeq \frac{\bar{n}^n}{n!} . \tag{4}$$

En déduire que

$$P(N, n) \simeq P_P(n; \bar{n}) . \tag{5}$$

c) On veut s'assurer que les conditions (2) sont suffisantes pour que l'approximation (5) soit valable dans tout le domaine où $P(N, n)$ est appréciable.

Montrer (en évaluant $N!$ et $(N-n)!$ par la formule de Stirling) que

$$\frac{P(N, n)}{P_P(n; \bar{n})} \simeq e^{\bar{n}-n}\left(\frac{1-P_+}{1-n/N}\right)^{N-n}\sqrt{\frac{N}{N-n}} . \tag{6}$$

Calculer ce rapport pour $n = \bar{n}$ et $n = 0$, puis comparer les valeurs les plus probables et les écarts quadratiques moyens de n donnés par $P(N, n)$ et $P_P(n; \bar{n})$.

I.11. Fonction caractéristique d'une distribution de probabilité. Soit une variable aléatoire x, de densité de probabilité $w(x)$. On définit la fonction caractéristique $\tilde{w}(k)$ de cette distribution comme la valeur moyenne de $\exp(ikx)$:

$$\tilde{w}(k) \mathrel{\widehat{=}} \int_{-\infty}^{+\infty} dx\, w(x) e^{ikx} . \tag{1}$$

A un facteur près, $\tilde{w}(k)$ est la transformée de Fourier de $w(x)$. On peut donc, si l'on connaît $\tilde{w}(k)$, calculer $w(x)$ par transformation de Fourier inverse :

$$w(x) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} dk\, \tilde{w}(k) e^{-ikx} . \tag{2}$$

1) Montrer que, si $w(x)$ décroît suffisamment vite lorsque $|x| \to \infty$, la fonction caractéristique peut être développée en série entière :

$$\tilde{w}(k) = 1 + i\bar{x}k - \frac{1}{2}\overline{x^2}k^2 + \ldots + \frac{i^n}{n!}\overline{x^n}k^n + \ldots , \tag{3}$$

où $\overline{x^n}$ est le «moment d'ordre n» de la distribution $w(x)$:

$$\overline{x^n} \widehat{=} \int_{-\infty}^{+\infty} \mathrm{d}x\, x^n w(x) . \tag{4}$$

2) Montrer que la fonction caractéristique $\widetilde{w}(k)$ associée à la somme $x_1 + x_2$ de deux variables aléatoires *indépendantes* x_1 et x_2 est le produit $\widetilde{w}_1(k) \cdot \widetilde{w}_2(k)$ des fonctions caractéristiques de ces deux variables. Généraliser à la somme de N variables aléatoires indépendantes.

3) Montrer que la fonction caractéristique d'une distribution de probabilité gaussienne est elle-même une gaussienne en k (cf. L. Schwartz, *Méthodes mathématiques pour les sciences physiques,* Hermann, Paris (1961), p. 204).

I.12. Théorème de la limite centrale. Soit $\{x_1, x_2, \ldots, x_N\}$ un ensemble de N variables aléatoires *indépendantes;* chacune d'elles est régie par une distribution de probabilité de densité $u_j(x_j)$ $(j = 1, 2, \ldots, N)$. On cherche la distribution statistique de la somme

$$S_N \widehat{=} \sum_{j=1}^{N} x_j \tag{1}$$

de ces N variables, dans le cas où $N \gg 1$.

1) *Préliminaires*

On suppose que la valeur moyenne m_j et l'écart quadratique moyen σ_j de chacune des variables x_j existent et sont connus.

a) Montrer que la valeur moyenne et l'écart quadratique moyen de la variable S_N sont donnés par

$$\overline{S}_N = Nm , \tag{2.a}$$

$$\Delta S_N = \sigma\sqrt{N} , \tag{2.b}$$

où m et σ s'obtiennent en prenant la moyenne des m_j et des σ_j^2 :

$$m \widehat{=} \frac{1}{N} \sum_{j=1}^{N} m_j , \tag{3.a}$$

$$\sigma^2 \widehat{=} \frac{1}{N} \sum_{j=1}^{N} \sigma_j^2 . \tag{3.b}$$

b) On définit la variable réduite

$$y_N \widehat{=} \frac{S_N - \overline{S}_N}{\Delta S_N} . \tag{4}$$

On va montrer plus loin (question 2) que, pour N grand, la densité de probabilité de cette variable est une gaussienne :

$$w(y_N) = \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-y_N^2/2} \quad \text{si} \quad N \gg 1 . \tag{5}$$

En admettant pour l'instant ce résultat, en déduire que la distribution de probabilité $W(S_N)$ de la variable S_N vaut

$$W(S_N) = \frac{1}{\sqrt{2\pi N\sigma^2}} \exp\left[-(S_N - Nm)^2/2N\sigma^2\right] \quad \text{si} \quad N \gg 1 . \tag{6}$$

2) *Démonstration du théorème*

a) Montrer que la fonction caractéristique (voir exercice I.11) $\tilde{w}_N(k)$ de la variable y_N définie en (4) s'écrit

$$\tilde{w}_N(k) = \exp\left[-ik\frac{m\sqrt{N}}{\sigma}\right] \times \prod_{j=1}^{N} \tilde{u}_j\left(\frac{k}{\sigma\sqrt{N}}\right), \tag{7}$$

où $\tilde{u}_j(k)$ est la fonction caractéristique correspondant à $u_j(x_j)$.

Pour analyser le comportement de $\tilde{w}_N(k)$ lorsque N devient grand, on va développer son logarithme (cf. note 6 du complément I.C).

b) En utilisant le résultat (3) de l'exercice I.11, montrer que

$$\ln\left[\tilde{u}_j\left(\frac{k}{\sigma\sqrt{N}}\right)\right] = im_j\frac{k}{\sigma\sqrt{N}} - \sigma_j^2\frac{k^2}{2N\sigma^2} + \mathcal{O}\left(\frac{1}{N^{3/2}}\right). \tag{8}$$

c) En déduire que

$$\ln\tilde{w}_N(k) = -\frac{k^2}{2} + \mathcal{O}\left(\frac{1}{\sqrt{N}}\right), \tag{9}$$

et que donc

$$\tilde{w}_N(k) = e^{-k^2/2}\left[1 + \mathcal{O}\left(\frac{1}{\sqrt{N}}\right)\right]. \tag{10}$$

d) Établir enfin (cf. question 3 de l'exercice I.11) le résultat final :

$$w(y_N) = \frac{1}{\sqrt{2\pi}}e^{-y_N^2/2}\left[1 + \mathcal{O}\left(\frac{1}{\sqrt{N}}\right)\right]. \tag{11}$$

Ce résultat, et l'expression (6) qui en découle pour la distribution statistique de la variable S_N, constituent le théorème de la limite centrale (pour une discussion de ses conditions de validité, voir A. Feller, *An Introduction to Probability Theory and Its Applications,* John Wiley, New York (1950), vol. II, p. 256).

PROBABILITÉS : APPLICATIONS

I.13. On considère un oscillateur harmonique classique à une dimension. Sa position en fonction du temps est donnée par

$$x = A\cos(\omega t + \varphi) ;$$

l'amplitude A est fixée tandis que la phase φ est aléatoire, toutes les valeurs entre 0 et 2π étant également probables. Trouver la probabilité $W(x)\mathrm{d}x$ pour que, à l'instant t, l'oscillateur se trouve en un point d'abscisse comprise entre x et $x+\mathrm{d}x$. Tracer la courbe $W(x)$ en fonction de x.

Mêmes questions si, au lieu des valeurs de la phase φ, ce sont celles de l'abscisse x qui sont également probables (entre $-A$ et $+A$).

I.14. Dans un gaz à l'équilibre macroscopique, la probabilité pour qu'une molécule donnée subisse une collision entre les instants t et $t+\mathrm{d}t$ est indépendante de t et égale à $W\,\mathrm{d}t$. On commence à observer le gaz à l'instant $t=0$ et on note $Q(t)$ la probabilité pour que la molécule considérée n'ait pas encore subi de collision à l'instant t. Écrire la relation différentielle liant $Q(t)$ et W; en déduire $Q(t)$.

Quelle est la probabilité pour que la molécule subisse sa première collision entre les instants t et $t+\mathrm{d}t$? Calculer la durée moyenne séparant deux collisions successives d'une molécule.

Considérant une molécule à un instant donné, on peut calculer l'intervalle de temps moyen au bout duquel se produira la collision suivante, ainsi que l'intervalle de temps moyen qui s'est écoulé depuis la dernière collision qu'elle a subie. Pourquoi la durée moyenne séparant deux collisions n'est-elle pas la somme des deux intervalles de temps précédents?

I.15. Un récipient de volume V contient un gaz de N molécules indépendantes (gaz parfait). On suppose que la probabilité de trouver une molécule est la même en tout point du récipient.

Écrire la probabilité pour qu'une molécule déterminée se trouve dans un petit domaine de volume v.

Calculer la probabilité $P(n)$ pour qu'il y ait exactement n molécules dans ce domaine; en déduire le nombre moyen $\bar{n}$ de molécules contenues dans v.

Montrer que, si $v \ll V$, on peut écrire $P(n)$ sous la forme d'une distribution de Poisson (cf. exercice I.10) :

$$P(n) \simeq \frac{\bar{n}^n}{n!} \mathrm{e}^{-\bar{n}} \quad \text{si} \quad v \ll V .$$

Calculer dans ce cas l'écart quadratique moyen Δn.

Application numérique : calculer $\bar{n}$, Δn et $\Delta n/\bar{n}$ pour un petit domaine de $1\ (\mu\mathrm{m})^3$ situé dans une enceinte contenant un gaz dans les conditions normales de température et de pression.

I.16. On considère, en mécanique classique, une particule se déplaçant librement dans l'espace. A intervalles de temps réguliers, de durée τ, elle subit une collision instantanée qui la renvoie au hasard dans une direction quelconque, avec une probabilité uniformément répartie sur toutes les directions; les collisions ne modifient pas le module v de la vitesse. On choisit comme origine, dans l'espace, l'endroit où se trouve la particule à l'instant $t=0$. A l'instant $t=N\tau$, après N collisions, elles se trouve au point M.

Calculer la valeur moyenne du vecteur $\overrightarrow{OM}$, l'écart quadratique moyen de la projection de $\overrightarrow{OM}$ sur une direction Ox donnée et la valeur moyenne de $\overrightarrow{OM}^2$. Calculer le coefficient de diffusion D, défini par $\overline{\overrightarrow{OM}^2} \widehat{=} 6Dt$.

I.17. Pour réaliser une couche métallique mince, on évapore un métal dans le vide et on recueille les atomes un à un sur une plaque de quartz refroidie. On suppose que la probabilité d'impact de chacun d'eux est répartie de façon uniforme sur la plaque, d'aire $S = 10\ \mathrm{cm}^2$. Lorsqu'un atome se fixe sur la plaque, il y occupe une aire πa^2, où $a \simeq 1$ Å.

1) On choisit un point particulier de la plaque. Évaluer la probabilité P qu'a chacun des atomes qui se condensent de recouvrir ce point.

2) On forme ainsi un film dont l'épaisseur est de 6 couches métalliques en moyenne. Quel pourcentage de l'aire S de la plaque ne comporte aucun atome? Quel pourcentage est recouvert de 6 atomes exactement? de seulement 3 atomes? (on pourra utiliser l'approximation de Poisson, voir exercice I.10).

I.18. On considère un ensemble de N particules (fixées aux nœuds d'un réseau cristallin) de spin 1/2 et de moment magnétique μ, indépendantes les unes des autres. On choisit un axe Oz; chaque particule a 2 états possibles : son moment magnétique peut pointer selon Oz ou dans la direction opposée.

1) On suppose d'abord que ces deux états ont même probabilité. Quelle est la valeur moyenne du moment magnétique de l'ensemble et quel est son écart quadratique moyen?

2) En appliquant un champ magnétique $\vec{B}$ suivant Oz, on modifie les probabilités des deux états, qui deviennent P_+ et $P_- = 1 - P_+$. Calculer à nouveau la valeur moyenne et l'écart quadratique moyen du moment magnétique total.

I.19. Localisation du centre de masse d'un gaz. Considérons un gaz constitué de N particules identiques enfermées dans une boîte. On suppose que leurs positions sont des variables aléatoires *indépendantes,* chacune d'elles étant *distribuée uniformément dans toute la boîte.* On cherche la distribution de probabilité qui régit la *position du centre de masse* du gaz.

Pour simplifier, nous allons raisonner à une dimension : la « boîte » est le segment $[-L/2, +L/2]$ de l'axe Ox.

1) Prenons d'abord une seule particule ($N = 1$). Quelle est la densité de probabilité (normalisée) qui décrit la distribution de son abscisse x?

2) Étudions ensuite le cas où $N = 2$. Les positions des deux particules (discernables) sont distribuées suivant la densité de probabilité précédente. Posons

$$x \widehat{=} \frac{1}{2}(x_1 + x_2), \qquad y \widehat{=} x_1 - x_2 \,. \tag{1}$$

a) Montrer que, si l'on fixe l'abscisse x du centre de masse, la variable relative y varie de $-(L - 2|x|)$ à $+(L - 2|x|)$.

b) En déduire que la densité de probabilité $w(x)$ de l'abscisse x du centre de masse s'écrit

$$w(x) = 2\,\frac{L - 2|x|}{L^2} \,. \tag{2}$$

Vérifier que cette expression est convenablement normalisée.

c) Quelle est la valeur la plus probable de x? Que vaut son écart quadratique moyen? Tracer la courbe représentant $w(x)$.

3) On considère maintenant le cas où $N \gg 1$.

a) En utilisant le théorème de la limite centrale (exercice I.12), montrer que l'abscisse

$$x \widehat{=} \frac{1}{N}(x_1 + x_2 + \ldots + x_N) \tag{3}$$

du centre de masse des N particules a une distribution gaussienne, centrée en $x = 0$ et d'écart quadratique moyen

$$\sigma = \frac{L}{\sqrt{12N}} \,. \tag{4}$$

b) Discuter comment évolue cette distribution lorsque N croît.

Application numérique : avec quelle précision peut-on situer le centre de masse des molécules d'air contenues dans une pièce?

I.20. Un homme « marche au hasard » à une dimension : la probabilité est p pour qu'il fasse un pas vers la droite, $q = 1 - p$ pour qu'il le fasse vers la gauche; les pas, indépendants, ont tous la même longueur ℓ. On choisit comme origine le point de départ, et on oriente l'axe de gauche à droite.

Quelle est la probabilité pour qu'après N pas le marcheur se trouve en un point d'abscisse $n\ell$ (n entier positif, négatif ou nul)?

Que valent alors le nombre moyen $\bar{n}_D$ de pas effectués vers la droite et l'écart quadratique moyen Δn_D correspondant?

I.21. On considère une marche au hasard à une dimension dans laquelle la longueur des pas est variable : la probabilité pour qu'un pas produise un déplacement (compté algébriquement) compris entre s et $s + ds$ est donnée par une loi gaussienne :

$$w(s)\,ds = \frac{1}{\sqrt{2\pi\sigma^2}}\, e^{-(s-\ell)^2/2\sigma^2}\, ds\ .$$

On choisit comme origine, dans l'espace, le point où se trouve le mobile avant le premier pas.

1) Que valent la valeur moyenne de s et son écart quadratique moyen?

2) Soit x la position du mobile après N pas. Calculer la valeur moyenne de x et son écart quadratique moyen. Calculer la probabilité $W_N(x)\,dx$ pour que, après N pas, le mobile se trouve en x à dx près.

Indication : voir complément I.B (§ II.3).

I.22. 1) On étudie un oscillateur harmonique classique, de pulsation ω. Son élongation x est de la forme

$$x = a\,\cos(\omega t + \varphi).$$

Son énergie est proportionnelle au carré de l'amplitude a (positive) et indépendante de la phase à l'origine φ ($0 \leqslant \varphi < 2\pi$).

a) Calculer, pour a et φ données, la *valeur moyenne dans le temps* de x et x^2, puis de $\dot{x}$ et $\dot{x}^2$.

b) On considère maintenant un *ensemble statistique* d'oscillateurs harmoniques de pulsation ω; on suppose que tous les oscillateurs de l'ensemble ont *même énergie* et que la phase à l'origine φ est distribuée de façon équiprobable sur l'intervalle $(0, 2\pi)$. Calculer les valeurs moyennes, dans cet ensemble statistique, de x et x^2, puis de $\dot{x}$ et $\dot{x}^2$.

c) En déduire que l'oscillateur harmonique classique est un système ergodique.

2) On étudie maintenant un système de deux oscillateurs harmoniques classiques indépendants, de même pulsation ω, correspondant par exemple au déplacement d'un point matériel selon deux axes perpendiculaires :

$$x = a\,\cos(\omega t + \varphi)\ ,$$

$$y = b\,\cos(\omega t + \psi)\ .$$

L'énergie du système est ici proportionnelle à la somme $a^2 + b^2$.

a) On considère d'abord un système particulier (a, φ, b et ψ fixés). Calculer la *valeur moyenne dans le temps* de x, x^2, y, y^2, xy, ainsi que celle de $\dot{x}$, $\dot{x}^2$, $\dot{y}$, $\dot{y}^2$, $\dot{x}\dot{y}$.

b) On considère ensuite un *ensemble statistique* de tels systèmes, dans lequel *l'énergie totale est fixée,* c'est-à-dire

$a^2 + b^2 = A^2$ fixé.

On pose

$a = A \cos \alpha$,

$b = A \sin \alpha$,

et on suppose que l'ensemble statistique est caractérisé par une densité de probabilité constante dans le domaine

$$0 \leqslant \varphi < 2\pi, \quad 0 \leqslant \psi < 2\pi, \quad 0 \leqslant \alpha \leqslant \frac{\pi}{2} .$$

Calculer la valeur moyenne d'ensemble des quantités étudiées à la question *a* ci-dessus.

c) Conclusion : un système de deux oscillateurs harmoniques classiques indépendants vérifie-t-il le principe ergodique?

DENSITÉ D'ÉTATS POUR UNE PARTICULE LIBRE OU QUASI-LIBRE

I.23. Considérons un système constitué d'une particule unique enfermée dans une boîte de taille macroscopique. Nous supposons qu'elle n'a pas de spin et que son énergie ε dépend seulement du module de son impulsion $\vec{p}$:

$$\varepsilon = f(p) . \tag{1}$$

La « boîte » qui la contient a pour volume V si elle est à 3 dimensions, mais ce peut être une surface d'aire S ou un segment de longueur L; nous poserons $V_3 \widehat{=} V$, $V_2 \widehat{=} S$ et $V_1 \widehat{=} L$, et raisonnerons de façon générale à d dimensions.

1) Montrer que les vecteurs d'onde $\vec{K} \widehat{=} \vec{p}/\hbar$ permis par les conditions aux limites périodiques sur les parois de la boîte constituent un réseau dont la maille élémentaire a pour volume, dans l'espace des $\vec{K}$, $(2\pi)^d / V_d$.

2) Montrer que la densité d'états $\rho(\varepsilon)$ s'écrit

$$\rho(\varepsilon) = \frac{S_d(K)}{(2\pi)^d / V_d} \cdot \frac{1}{\hbar \, \mathrm{d}f/\mathrm{d}p} , \tag{2}$$

où $S_d(K)$ est l'aire de la « sphère » de « rayon » K dans l'espace des vecteurs d'onde à d dimensions.

3) Montrer que

$$S_1(K) = 2, \quad S_2(K) = 2\pi K, \quad S_3(K) = 4\pi K^2 \tag{3}$$

(pour $d = 1$, la « sphère » de « rayon » K se réduit à l'ensemble des deux points d'abscisse $\pm K$).

4) En déduire les résultats résumés dans le tableau ci-dessous. Si la particule a un spin s et que son énergie est indépendante du spin, la densité d'états doit être multipliée par $(2s + 1)$ dans les première et deuxième colonnes du tableau, par 2 dans la troisième colonne.

5) Supposons maintenant que la particule étudiée (non relativiste, de masse m) ait un spin 1/2 et qu'elle soit soumise à un champ magnétique uniforme $\vec{B}$. Son énergie a pour expression

$$\varepsilon_+ = \frac{p^2}{2m} - \mu B \tag{4.a}$$

si le moment magnétique $\vec{\mu}$ associé au spin est de même sens que $\vec{B}$, et

$$\varepsilon_- = \frac{p^2}{2m} + \mu B \tag{4.b}$$

si $\vec{\mu}$ est de sens opposé à $\vec{B}$ (attention : le «rapport gyromagnétique» entre $\vec{\mu}$ et le spin $\vec{S}$ est parfois négatif; c'est le cas pour l'électron).

Écrire, en utilisant le tableau (ici, $d=3$), les densités d'états $\rho_+(\varepsilon)$ et $\rho_-(\varepsilon)$ correspondant aux deux états de spin possibles. En déduire la densité totale

$$\rho(\varepsilon) = \rho_+(\varepsilon) + \rho_-(\varepsilon) \,. \tag{5}$$

6) Dans certains problèmes (électrons dans les solides, voir complément I.F), on est amené à considérer des particules dont l'énergie est de la forme

$$\varepsilon = \frac{\hbar^2}{2}\left(\frac{K_x^2}{m_1} + \frac{K_y^2}{m_2} + \frac{K_z^2}{m_3}\right), \tag{6}$$

où K_x, K_y, K_z sont les composantes cartésiennes d'un vecteur d'onde $\vec{K}$ permis par les conditions aux limites et m_1, m_2, m_3 trois paramètres positifs. Montrer que la densité d'états s'écrit alors (si le spin est nul)

$$\rho(\varepsilon) = \frac{V}{2\pi^2\hbar^3}(m_1 m_2 m_3)^{1/2}(2\varepsilon)^{1/2} \,. \tag{7}$$

	Particule libre non relativiste de masse m	Particule libre relativiste de masse m	Particule libre relativiste de masse nulle (photon)
$\varepsilon = f(p)$	$\varepsilon = \frac{p^2}{2m} + \varepsilon_0$	$\varepsilon = [p^2c^2 + m^2c^4]^{1/2}$	$\varepsilon = cp$
$d=1$	$\frac{L}{\pi\hbar}\left[\frac{m}{2(\varepsilon-\varepsilon_0)}\right]^{1/2}$	$\frac{L}{\pi\hbar c}\frac{\varepsilon}{(\varepsilon^2 - m^2c^4)^{1/2}}$	$\frac{L}{\pi\hbar c}$
$d=2$	$\frac{Sm}{2\pi\hbar^2}$	$\frac{S\varepsilon}{2\pi\hbar^2c^2}$	$\frac{S\varepsilon}{2\pi\hbar^2c^2}$
$d=3$	$\frac{Vm}{2\pi^2\hbar^3}[2m(\varepsilon-\varepsilon_0)]^{1/2}$	$\frac{V}{2\pi^2\hbar^3c^3}\varepsilon(\varepsilon^2 - m^2c^4)^{1/2}$	$\frac{V\varepsilon^2}{2\pi^2\hbar^3c^3}$
facteur de spin	$2s+1$	$2s+1$	2

ORDRES DE GRANDEUR

I.24. On veut évaluer l'ordre de grandeur des effets liés aux *interactions gravitationnelles* dans un système macroscopique.

1) Montrer, par des arguments dimensionnels, que l'énergie potentielle gravitationnelle E_G d'un corps de masse totale M et de volume V est de l'ordre de

$$E_G \sim -G\frac{M^2}{V^{1/3}}, \tag{1}$$

où G est la constante gravitationnelle. On pourra aussi montrer (à l'aide du théorème de Gauss) que, si la masse M est uniformément répartie dans une sphère de rayon R, l'expression exacte de E_G s'écrit

$$E_G = -\frac{3}{5}G\frac{M^2}{R}. \tag{2}$$

2) Pour avoir une idée de l'importance des effets gravitationnels, on peut comparer $|E_G|$ à NkT, où N est le nombre de particules du système considéré. Estimer $|E_G|/NkT$ pour 1 m^3 d'eau à la température ordinaire, pour 1 kg de cuivre à la même température, puis pour le Soleil. Conclusion?

3) Si, comme au paragraphe II du complément II.A, on multiplie par λ la taille du système, comment varie son énergie gravitationnelle? En déduire que l'énergie totale d'un système n'est une grandeur extensive que si la contribution des interactions gravitationnelles est négligeable.
(Solution : complément VI.B, § II.)

I.25. Pour *une mole* de fluide occupant le volume V sous la pression p à la température T, l'équation d'état de van der Waals s'écrit

$$\left(p+\frac{a}{V^2}\right)(V-b)=RT,$$

où R est la constante des gaz parfaits et a, b deux constantes caractéristiques du fluide considéré. Écrire l'équation de van der Waals pour n moles.

I.26. On considère un plasma contenant deux types de particules, de charge $+e$ et $-e$. On appelle « longueur de Landau » la distance minimum d'approche de deux particules de même charge lorsqu'elles sont animées de vitesses « thermiques », c'est-à-dire lorsque leur énergie cinétique est de l'ordre de kT. Évaluer la longueur de Landau à la température T. Montrer qu'elle doit être grande devant la longueur d'onde thermique de de Broglie pour que le plasma puisse être traité par la mécanique classique; y a-t-il d'autres conditions à prendre en compte?

I.27. Estimer le nombre de collisions par seconde qu'effectue une molécule d'oxygène de l'air.

I.28. Lorsqu'on exerce une force de traction F sur une barre de longueur L et de section s, son allongement ΔL est déterminé par le « module d'Young » Y du matériau constituant la barre :

$$\frac{\Delta L}{L}=\frac{1}{Y}\frac{F}{s}.$$

En admettant que l'interaction entre atomes dans un cristal peut être assimilée à des forces de rappel harmoniques entre plus proches voisins, estimer la constante K caractérisant ces interactions (pour les métaux usuels, les valeurs de Y s'échelonnent entre 2.10^{10} Pa (fer pur) et 21.10^{10} Pa (nickel); pour le diamant, $Y = 12.10^{11}$ Pa). Comparer avec l'estimation (A.15) du complément I.A.

I.29. Dans un liquide, chaque molécule interagit principalement avec ses plus proches voisines.

1) Estimer le nombre de plus proches voisines d'une molécule à l'intérieur du liquide et d'une molécule à la surface.

2) Relier l'énergie de liaison ε_0 d'une molécule de l'intérieur à la chaleur latente L d'évaporation *par unité de volume* de liquide.

3) On appelle «tension de surface» γ l'énergie nécessaire pour augmenter la surface du liquide d'une aire unité. Relier γ et ε_0.

4) En éliminant ε_0 entre les deux expressions précédentes, trouver une relation (en ordre de grandeur) entre L et γ. Comparer le résultat obtenu aux valeurs expérimentales pour quelques liquides usuels (pour l'eau, $L_0 = 22{,}6.10^5$ J/kg, $\gamma = 7{,}6.10^{-2}$ N/m; pour l'éthanol de masse volumique 0,789 kg/litre, $L_0 = 8{,}8.10^5$ J/kg et $\gamma = 2{,}2.10^{-2}$ N/m; pour le mercure de masse volumique 13,5 kg/litre, $L_0 = 2{,}96.10^5$ J/kg et $\gamma = 47.10^{-2}$ N/m).

II

I. LA DISTRIBUTION MICROCANONIQUE

A. Quelques définitions préliminaires
   1. Paramètres extérieurs et variables internes
   2. Grandeurs conservées

B. Le postulat fondamental de la mécanique statistique
   1. Énoncé du postulat
      a. Paramètres extérieurs et états accessibles
      b. Probabilités d'occupation des états accessibles
   2. L'ensemble microcanonique
   3. Discussion du postulat
      a. Est-ce le seul postulat possible?
      b. Le principe du « désordre » maximum

C. L'entropie microcanonique
   1. Entropie microcanonique et nombre d'états accessibles
   2. Entropie d'une réunion de systèmes isolés

II. PROPRIÉTÉS ET COMPORTEMENT D'UN SYSTÈME ISOLÉ

A. Distribution statistique d'une variable interne
   1. Formule générale
   2. Résultat pratique pour un grand système

B. Évolution spontanée d'un système isolé après modification des contraintes extérieures
   1. Détente de Joule
   2. Généralisation

III. ÉQUILIBRE ENTRE SOUS-SYSTÈMES D'UN SYSTÈME ISOLÉ

A. Température et équilibre thermique
   1. Définitions
      a. Température microcanonique
      b. Contact thermique entre deux systèmes
   2. Équilibre thermique entre deux systèmes
      a. Position du problème
      b. L'établissement du contact thermique comme relâchement d'une contrainte
      c. Caractérisation de l'équilibre thermique
      d. Sens de l'échange de chaleur

B. Généralisation : échange de volume et de particules
   1. Définitions : pression et potentiel chimique microcanoniques
   2. Établissement du contact entre deux systèmes
   3. Équilibre entre les deux systèmes
      a. Conditions d'équilibre
      b. Sens des échanges hors d'équilibre

CHAPITRE II

Système isolé à l'équilibre - Distribution microcanonique

Le *postulat fondamental de la mécanique statistique* s'introduit commodément dans le cas simple de systèmes isolés. On appelle traditionnellement *« distribution microcanonique »* la distribution statistique des états microscopiques d'un tel système *à l'équilibre.* Dans la première partie de ce chapitre, nous énonçons le postulat et nous calculons l'entropie statistique correspondante, dite *« entropie microcanonique »*. Nous indiquons dans la deuxième partie comment on en déduit les propriétés d'un système isolé. Enfin, dans la troisième partie, nous introduisons les notions fondamentales de *température, pression et potentiel chimique microcanoniques,* qui permettent d'exprimer simplement les conditions d'équilibre entre deux systèmes en contact.

I. LA DISTRIBUTION MICROCANONIQUE

A. Quelques définitions préliminaires

1. PARAMÈTRES EXTÉRIEURS ET VARIABLES INTERNES

Pour décrire les propriétés et le comportement macroscopiques d'un système, on fait appel à des grandeurs telles que volume, énergie, nombre de particules, pression, température, densité, ...

Certaines d'entre elles sont *fixées* par des contraintes extérieures imposées au système : par exemple, un gaz est enfermé dans un récipient de volume fixé, ou bien c'est sa pression qui est imposée de l'extérieur[1]. De telles grandeurs seront appelées *« paramètres extérieurs »*. Du point de vue statistique, les paramètres extérieurs ont des *valeurs certaines* (et non pas diverses valeurs possibles affectées de probabilités). Remarquons cependant que ces valeurs ne sont pas connues avec une exactitude parfaite, mais sont entachées d'une *incertitude expérimentale* (voir plus loin notes 4 et 5).

Par opposition aux paramètres extérieurs, nous appellerons *« variables internes »* les grandeurs qui sont *libres de fluctuer* au gré de l'agitation microscopique, et auxquelles on doit donc associer une *distribution statistique :* c'est par exemple le cas de la densité d'un gaz au voisinage d'un point, c'est-à-dire du nombre de molécules contenues dans un petit volume fixe entourant ce point.

---

1. Il suffit pour cela que le gaz soit contenu dans un cylindre fermé par un piston étanche mais pouvant se déplacer sans frottement sous l'effet des forces de pression : le piston se fixe à une position d'équilibre telle que la pression du gaz à l'intérieur du cylindre est égale à la pression imposée à l'extérieur (cf. § III.B.3).

Bien entendu une même grandeur peut, suivant la situation considérée, se classer parmi les paramètres extérieurs ou parmi les variables internes : le volume d'un gaz enfermé dans un récipient indéformable est un paramètre extérieur; il est au contraire variable interne si c'est la pression qui est fixée de l'extérieur.

2. GRANDEURS CONSERVÉES

L'évolution d'un système macroscopique est régie par les lois microscopiques d'interaction entre ses constituants, et respecte de ce fait certaines *lois de conservation :* citons par exemple la conservation de l'*énergie,* celle du *nombre de molécules* d'un corps pur donné si le système est inerte chimiquement[2], ...

Les lois de conservation s'expriment le plus simplement dans le cas d'un *système isolé :* si le système est placé dans des conditions telles qu'il ne peut échanger d'énergie avec son environnement, son énergie totale reste constante au cours du temps; si les parois qui l'enferment ne laissent passer aucune matière, le nombre total de molécules (ou d'atomes) qu'il contient reste constant. *Dans de telles situations, les grandeurs conservées se comportent comme des paramètres extérieurs,* puisqu'elles gardent une valeur fixe et certaine, imposée par les conditions de préparation du système.

Mais *si le système n'est pas isolé, certaines de ces grandeurs sont des variables internes :* par exemple, si des échanges d'énergie se produisent entre le système et son environnement, l'énergie du système lui-même ne reste pas fixée; même à l'équilibre macroscopique, elle fluctue et possède donc une distribution statistique. Cependant, les fluctuations d'une telle grandeur doivent respecter la loi de conservation : lorsque l'énergie du système diminue, celle de son environnement augmente exactement d'autant.

Ces considérations appellent une remarque importante : lorsqu'on veut appliquer la mécanique statistique à un problème concret, il est indispensable de commencer par *définir de manière précise le système* que l'on va considérer et de *bien distinguer entre les paramètres extérieurs* fixés *et les variables internes* libres de fluctuer. Ceci est essentiel : beaucoup d'erreurs et d'incompréhensions proviennent d'ambiguïtés concernant la délimitation du système auquel on applique tel ou tel résultat général, et la situation précise dans laquelle il se trouve.

B. Le postulat fondamental de la mécanique statistique

1. ÉNONCÉ DU POSTULAT

a. Paramètres extérieurs et états accessibles

Considérons un *système isolé.* Concrètement, cela signifie que les parois qui l'enferment ne laissent passer aucune particule et ne permettent aucun échange d'énergie avec l'extérieur, ce qui implique en particulier qu'elles sont fixes. Nous supposons ce système à l'*équilibre macroscopique.*

---

2. Si le système considéré est constitué d'un mélange de corps purs donnant lieu à des réactions chimiques, ce sont les nombres d'atomes (ou de radicaux chimiques) qui sont conservés. Bien entendu, la conservation du nombre de particules matérielles n'est valable que dans le domaine non relativiste, auquel nous nous limitons ici.

Son état macroscopique est alors déterminé par les valeurs d'un certain nombre de paramètres extérieurs[3] : pour un gaz, ce sera l'énergie totale E, le volume V, et le nombre total de molécules N (ou, s'il s'agit d'un mélange, les nombres N_1, N_2, ... de molécules des divers corps purs (1), (2), ... qui le constituent). L'ensemble des paramètres extérieurs définissant l'état macroscopique d'un système isolé quelconque comprend toujours l'énergie totale E, un (ou plusieurs) nombre(s) de particules N, et le plus souvent le volume V. Dans les raisonnements généraux, nous noterons cet ensemble de paramètres extérieurs

$$\{E, x_1, x_2, \ldots, x_q\}, \tag{I.1.a}$$

que nous abrégerons en

$$\{E, x\}. \tag{I.1.b}$$

Il est clair qu'*il existe un grand nombre d'états microscopiques vérifiant les contraintes extérieures,* c'est-à-dire compatibles avec les valeurs des paramètres extérieurs : si par exemple le système est constitué de particules indépendantes (cf. I, § III.B.2 et 3), son énergie totale E peut être partagée d'un grand nombre de façons possibles entre les diverses particules. Il existe bien sûr un nombre encore plus grand d'états microscopiques qui sont incompatibles avec les contraintes, et donc exclus : par exemple tous ceux dont l'énergie (ou le nombre de particules) est différente de la valeur fixée, même lorsqu'on tient compte de l'incertitude qui affecte celle-ci. Un état microscopique compatible avec les valeurs macroscopiques des paramètres extérieurs est dit *accessible.*

b. Probabilités d'occupation des états accessibles

On pose alors le postulat suivant :

Pour un système isolé à l'équilibre (macroscopique), tous les états microscopiques accessibles ont même probabilité.

Pour préciser cet énoncé, appelons E la valeur (fixée) de l'énergie du système, et δE celle de l'incertitude (macroscopique) qui lui est associée[4]. Les états

3. La quantité de mouvement totale et le moment cinétique total d'un système vraiment isolé seraient des paramètres extérieurs comme son énergie, puisqu'il leur correspond aussi une loi de conservation. Cependant, s'il est relativement facile d'interdire les échanges d'énergie entre un système et son environnement (en enfermant le système dans une « boîte » à parois rigides et imperméables à la chaleur), il n'en est pas de même pour la quantité de mouvement et le moment cinétique. La « boîte » contenant le système est en effet presque toujours solidaire d'un support (elle est posée sur une table, ou tenue à la main...) dont la masse est beaucoup plus grande que celle du système. La quantité de mouvement et le moment cinétique du système ne sont donc pas fixés, puisqu'ils sont couplés à ceux du support.

Dans la suite, nous dirons qu'un système est isolé si son énergie et son nombre de particules sont fixés; en général, sa quantité de mouvement et son moment cinétique ne le seront pas (cf. complément II.E).

4. Lorsqu'on prépare un système macroscopique, on ne contrôle pas de façon suffisamment précise les conditions extérieures pour connaître la valeur exacte de son énergie. De plus, un tel système n'est jamais parfaitement isolé (cf. I, § IV.A.2), de sorte que son énergie ne reste pas rigoureusement constante. C'est là l'origine de l'incertitude δE.

Il faut tenir compte d'une telle incertitude, car sinon le nombre d'états accessibles présenterait des variations brutales et artificielles, puisqu'il ne serait différent de zéro que pour les valeurs de E strictement égales aux valeurs propres discrètes E_ℓ de l'hamiltonien H. Cependant, la valeur précise de δE est sans importance dans les applications, tellement est grand, de toutes façons, le nombre d'états accessibles (voir paragraphe C.1).

microscopiques (ℓ) accessibles sont ceux qui vérifient toutes les autres contraintes extérieures, et dont l'énergie E_ℓ est compatible avec ces valeurs :

$$E \leqslant E_\ell \leqslant E + \delta E \,. \tag{I.2}$$

Pour un gaz par exemple, on cherchera les états propres (ℓ) de l'hamiltonien H correspondant à un système constitué de N molécules (ou de N_1 molécules de type (1), N_2 molécules de type (2), ...), enfermées dans le volume V fixé; parmi ces états, on sélectionnera ceux dont l'énergie E_ℓ vérifie la condition (I.2)[5]. Si nous notons Ω le *nombre d'états accessibles,* le postulat ci-dessus se traduit par la distribution suivante de probabilités pour les divers états microscopiques (ℓ) :

$$P^*_\ell = \begin{cases} \dfrac{1}{\Omega} & \text{si } E \leqslant E_\ell \leqslant E + \delta E \ ; \\ 0 & \text{pour tous les autres états.} \end{cases} \tag{I.3}$$

Bien entendu, la valeur commune $1/\Omega$ des probabilités des états accessibles est déterminée par la condition de normalisation

$$\sum_{(\ell)} P^*_\ell = 1 \,. \tag{I.4}$$

Les probabilités (I.3) constituent ce que l'on appelle la *distribution microcanonique*[6], valable pour un *système isolé à l'équilibre macroscopique.*

La probabilité d'une propriété quelconque, dans un tel système, s'en déduit très facilement : il suffit de compter le *nombre d'états accessibles qui réalisent cette*

---

5. Il faudrait en principe compter comme accessibles tous les états qui (tout en vérifiant la condition (I.2)) correspondent à un volume compris entre V et $V + \delta V$, un nombre de particules compris entre N et $N + \delta N$, ... Ce n'est pas nécessaire : s'il faut tenir compte de l'existence d'une incertitude (cf. note 4), il suffit de la faire porter sur un seul des paramètres extérieurs, et le choix de l'énergie est le plus commode.

6. Le postulat fondamental s'applique également dans les cas où le système étudié est décrit par la mécanique classique (I, § III.C). A la grandeur physique énergie est associée la *fonction de Hamilton* $\mathcal{H}(q_i, p_i)$. Donc, pour un système isolé à l'équilibre macroscopique, la densité de probabilité w^* dans l'espace des phases est uniforme dans le domaine défini par

$$E \leqslant \mathcal{H}(q_i, p_i) \leqslant E + \delta E \,, \tag{1}$$

et nulle en dehors de ce domaine.

Comme l'ensemble des valeurs possibles pour l'énergie du système est ici continu, et non plus discret comme en mécanique quantique, le rôle de l'incertitude δE devient secondaire (cf. note 4) : on peut faire comme si l'énergie du système avait une valeur E bien déterminée, et prendre w^* uniforme sur l'hypersurface d'équation

$$\mathcal{H}(q_i, p_i) = E \,, \tag{2}$$

et nulle partout ailleurs dans l'espace des phases. Autrement dit, la densité de probabilité microcanonique $w^*(q_i, p_i)$ est proportionnelle à $\delta\left[E - \mathcal{H}(q_i, p_i)\right]$, où δ est la fonction de Dirac; le coefficient de proportionnalité est ici aussi déterminé par la condition de normalisation, ce qui donne

$$w^*(q_i, p_i) = \frac{\delta\left[E - \mathcal{H}(q_i, p_i)\right]}{\int \ldots \mathrm{d}q'_i \ldots \mathrm{d}p'_i \ldots \delta\left[E - \mathcal{H}(q'_i, p'_i)\right]} \,. \tag{3}$$

propriété et de le diviser par le nombre total Ω d'états accessibles, puisqu'ils ont tous même probabilité.

2. L'ENSEMBLE MICROCANONIQUE

La distribution microcanonique que nous venons d'introduire doit être interprétée dans le cadre d'un ensemble statistique constitué d'un grand nombre de systèmes isolés identiques, tous préparés de la même façon du point de vue macroscopique (cf. I, § IV); cet ensemble statistique est appelé *« ensemble microcanonique »*.

Nous supposons comme toujours que chacun des $\mathcal{N}$ systèmes constituant l'ensemble est dans un état microscopique (ℓ). Dans l'ensemble microcanonique, seuls apparaissent les états (ℓ) accessibles, et ils apparaissent tous avec la même fréquence (dans la limite $\mathcal{N} \to \infty$)[7].

3. DISCUSSION DU POSTULAT

a. Est-ce le seul postulat possible?

Soulignons à nouveau que le postulat précédent concerne seulement la *mécanique statistique des états d'équilibre;* il devra donc être complété si l'on veut analyser le comportement des systèmes hors d'équilibre (voir chapitre IV). Mais si l'on se limite aux propriétés de l'équilibre macroscopique, il est tout à fait général : il s'applique à des systèmes de nature très variée, et permet de développer *toute* la mécanique statistique d'équilibre.

Comme tout postulat physique, celui-ci n'a d'intérêt que dans la mesure où il permet de *décrire correctement les phénomènes observés.* Le développement de la mécanique statistique et la confrontation de ses prédictions avec l'expérience ont jusqu'ici confirmé le bien-fondé de ce postulat, comme nous le constaterons tout au long de ce livre.

Cela étant, on voit mal quelle autre distribution de probabilités pourrait être valable pour les états microscopiques d'un système isolé : la conservation de l'énergie impose que seuls les états définis plus haut comme « accessibles » puissent avoir une probabilité non nulle; d'autre part, on n'a pas de raison valable de favoriser certains de ces états par rapport aux autres. Cependant, la réalité aurait pu être différente : il aurait pu se faire que, précisément, certains états soient « favorisés ». Le postulat qu'il aurait alors fallu poser aurait été moins simple, puisqu'il aurait nécessité des informations supplémentaires permettant de distinguer entre eux les divers états accessibles. L'énoncé ci-dessus est le seul possible *si l'on peut se contenter du minimum d'information,* c'est-à-dire s'il est suffisant de savoir, pour chaque état microscopique, s'il est compatible ou non avec les contraintes extérieures.

Finalement, la question cruciale est celle que pose le *principe ergodique* (I, § IV.B.3) : est-ce que l'état macroscopique d'*un* système isolé peut être décrit en termes de probabilités de ses états microscopiques? Il semble en effet clair d'après ce qui précède que, dans un *ensemble statistique* constitué d'un grand nombre de systèmes isolés identiques vérifiant tous les mêmes contraintes extérieures, les états microscopiques accessibles seraient tous également représentés. Reste à savoir si cet ensemble microcanonique permet de décrire correctement les propriétés macroscopiques d'un *système unique* du type considéré, propriétés qui résultent de moyennes sur les

7. Le nombre Ω d'états accessibles étant toujours fantastiquement grand, il est hors de question d'essayer de réaliser concrètement une collection de systèmes identiques qui ressemble tant soit peu à l'ensemble microcanonique (puisque $\mathcal{N}$ doit être au moins égal à Ω). C'est seulement un outil théorique commode.

fluctuations temporelles, au niveau microscopique, de ce système particulier. Si c'est bien le cas[8], le postulat microcanonique est sans conteste le seul possible.

b. Le principe du « désordre » maximum

Il est possible d'exprimer le postulat précédent sous une forme apparemment plus générale, mais qui s'avère en fait équivalente. C'est ce que l'on appelle le *« principe du désordre maximum »*.

Lorsque nous constatons que le postulat ci-dessus est le seul possible si l'on se contente de l'information minimale, cela peut être traduit quantitativement en disant que la distribution microcanonique est celle qui rend l'entropie statistique maximale (cf. I, § II.B.3). Lorsqu'on cherche à généraliser cette constatation à des situations où le système considéré n'est pas nécessairement isolé, on aboutit à l'énoncé suivant du postulat fondamental : dans tous les cas, *la distribution de probabilité des états microscopiques est celle qui maximise l'entropie statistique, compte tenu des contraintes extérieures* imposées au système.

Or on dit parfois que l'entropie statistique mesure le « désordre » du système. Cette façon de parler imagée, qui cherche à utiliser le pouvoir évocateur du mot « désordre » pour fonder une compréhension intuitive de la notion d'entropie, est dangereuse, car le mot est trop ambigu pour ne pas induire souvent de grossières erreurs. Sachons bien que, s'il nous arrive de l'utiliser, il ne signifiera rien d'autre que l'expression « manque d'information » introduite au paragraphe II du chapitre I. C'est dans ce cadre que le postulat fondamental devient le « principe du désordre maximum ».

Nous étudierons cette formulation générale du postulat fondamental dans le complément V.D.

C. L'entropie microcanonique

A la distribution de probabilité (I.3) est associée une entropie statistique (I, § II), que nous appellerons « entropie microcanonique ».

---

8. Supposons d'abord que les systèmes considérés soient décrits dans le cadre de la mécanique classique. L'évolution d'un système isolé donné est en principe fixée par les conditions initiales. Il pourrait alors se faire que, au sortir de sa préparation, un système particulier se trouve dans des conditions initiales qui ne lui permettent d'atteindre, au cours de son mouvement microscopique ultérieur, qu'une classe restreinte d'états accessibles. Un tel système ne vérifierait pas le principe ergodique. En réalité, tout système est soumis à des perturbations aléatoires venant de son environnement; bien que très faibles pour un système « isolé » tel que ceux qui font l'objet de ce chapitre, elles ne sont jamais strictement nulles. Leur influence fait rapidement « oublier » au système considéré les conditions initiales dont il est parti (en modifiant très légèrement mais constamment, de façon aléatoire, son état microscopique). Ce sont ces perturbations extérieures qui assurent la validité du principe ergodique, même pour des types de systèmes qui seraient non ergodiques s'ils pouvaient être parfaitement isolés (voir complément IV.B).

Dans le cadre plus général de la mécanique quantique, imaginons qu'initialement chaque système de l'ensemble microcanonique se trouve dans un état propre (ℓ) particulier de l'hamiltonien H. Si H était l'hamiltonien exact, aucun de ces systèmes n'évoluerait dans le temps, puisque les états (ℓ) seraient véritablement stationnaires. En réalité, les perturbations aléatoires mentionnées ci-dessus induisent des transitions entre les divers états (ℓ) (cf. IV, § I.A), de sorte qu'un système donné passe sans arrêt d'un état microscopique à un autre. Si la composition globale de l'ensemble statistique ne varie pas (la distribution microcanonique est indépendante du temps), c'est qu'il y a en moyenne, pendant un intervalle de temps fixé quelconque, autant de systèmes qui arrivent dans un état (ℓ) particulier que de systèmes qui quittent cet état (ℓ). Ce sont ici aussi ces transitions aléatoires entre états microscopiques qui permettent la description statistique d'un système unique, dans le cadre du principe ergodique.

1. ENTROPIE MICROCANONIQUE ET NOMBRE D'ÉTATS ACCESSIBLES

D'après le postulat fondamental, les divers états microscopiques accessibles sont équiprobables. Par conséquent, si leur nombre est Ω, la valeur correspondante S^* de l'entropie statistique est simplement (I, § II.B.3)

$$\boxed{S^* = k \ln \Omega} \,. \tag{I.5}$$

Le nombre Ω d'états microscopiques accessibles dépend des paramètres extérieurs $\{E, x\}$: dans le cas d'un gaz, par exemple, le nombre d'états propres de l'hamiltonien dont l'énergie est comprise entre E et $E + \delta E$ dépend de E, mais aussi du volume V et du nombre N de molécules[9].

Mais Ω dépend également de l'incertitude δE sur l'énergie. Pourtant, pour un système de taille macroscopique, cette dernière dépendance devient totalement négligeable au niveau de l'entropie. En effet, le nombre $\Omega(E, x; \delta E)$ d'états accessibles est fantastiquement grand, et de la forme

$$\Omega(E, x; \delta E) = \rho(E, x)\, \delta E \,, \tag{I.6}$$

où $\rho(E, x)$ est la *densité d'états* du système considéré (I, § III.B.1). Si l'on accepte de prendre le logarithme de grandeurs ayant des dimensions, on en déduit

$$\ln \Omega = \ln \rho(E, x) + \ln \delta E \,. \tag{I.7}$$

On constate alors que, quelque unité d'énergie que l'on ait choisie, *le second terme est totalement négligeable* devant le premier. En effet, le raisonnement semi-quantitatif du paragraphe III.B.3 du chapitre I (voir formule (III.49)) indique que *le logarithme de la densité d'états* d'un système macroscopique *est de l'ordre du nombre N de particules* du système, c'est-à-dire du nombre d'Avogadro; or jamais δE n'est mesurée par un nombre si petit ou si grand que son logarithme puisse approcher de telles valeurs[10].

En fin de compte, l'entropie microcanonique est une fonction des paramètres extérieurs, et la formule fondamentale (I.5) s'écrit en pratique

$$S^*(E, x) = k \ln \rho(E, x) \,. \tag{I.8}$$

La densité d'états ρ dont on prend le logarithme a des dimensions (celles de l'inverse d'une énergie), mais la modification de S^* provoquée par un changement quelconque d'unité d'énergie est totalement inappréciable[10]. L'entropie microcanonique S^* est

---

9. C'est ce que montre concrètement, par exemple, le calcul du paragraphe III.B.2 du chapitre I : ce calcul n'est valable que dans le cas très particulier de particules *libres* et *discernables*, mais la dépendance du nombre d'états par rapport à E, V et N y apparaît explicitement.

10. Le logarithme d'une quantité ayant des dimensions change de valeur si l'on change d'unités. Mais supposons, pour fixer les idées, que δE vaille par exemple 1/1 000 de joule. Si l'on choisit le joule comme unité d'énergie (c'est une unité adaptée au domaine macroscopique, cf. I, paragraphe I.A.3), $\ln \delta E$ vaudra $-6{,}9$, ce qui est absolument inappréciable devant N. Imaginons maintenant que, malgré son ordre de grandeur, on mesure δE en électron-volt; sa valeur sera multipliée par $\simeq 10^{19}$, mais celle de son logarithme sera seulement augmentée d'environ 44, ce qui reste encore ridiculement faible devant N.

C'est seulement dans des cas limites exceptionnels que, le nombre d'états accessibles étant moins élevé, il faudra se préoccuper de la valeur de l'incertitude δE. Mais l'application du postulat microcanonique à de tels cas pose de toutes façons des problèmes de fond, et c'est l'ensemble de ces problèmes qu'il convient alors d'analyser.

La validité de la formule (I.6) et de sa conséquence (I.8) est analysée dans l'exercice II.17. Cet exercice montre également que l'entropie microcanonique peut aussi s'écrire

$$S^* = k \ln \Phi(E) \,,$$

où $\Phi(E)$ est le nombre d'états microscopiques dont l'énergie est inférieure à E.

donc le produit de la constante de Boltzmann k, qui lui donne ses dimensions, par un *nombre de l'ordre de* N pour N grand[11].

2. ENTROPIE D'UNE RÉUNION DE SYSTÈMES ISOLÉS

Considérons deux systèmes macroscopiques $\mathcal{S}_1$ et $\mathcal{S}_2$ séparément isolés et à l'équilibre. Soit Ω_1 le nombre d'états accessibles du système $\mathcal{S}_1$, Ω_2 celui du système $\mathcal{S}_2$; leurs entropies microcanoniques respectives S_1^* et S_2^* valent, d'après la formule (I.5),

$$S_1^* = k \ln \Omega_1 , \tag{I.9.a}$$

$$S_2^* = k \ln \Omega_2 . \tag{I.9.b}$$

Réunissons maintenant ces deux systèmes en un système unique $\mathcal{S}$ (fig. 1); puisque $\mathcal{S}_1$ et $\mathcal{S}_2$ sont séparément isolés et à l'équilibre, leur réunion $\mathcal{S}$ est isolée et à l'équilibre. Nous voulons calculer son entropie microcanonique S^*.

Tout état accessible de $\mathcal{S}$ est obtenu en prenant $\mathcal{S}_1$ dans l'un de ses états accessibles et $\mathcal{S}_2$ dans l'un des siens. A un état accessible déterminé du système $\mathcal{S}_1$ on peut associer successivement les Ω_2 états accessibles du système $\mathcal{S}_2$ pour former Ω_2 états accessibles distincts du système global $\mathcal{S}$; puisqu'on peut partir pour cela de n'importe lequel des Ω_1 états accessibles de $\mathcal{S}_1$, le nombre Ω d'états accessibles de $\mathcal{S}$ est le produit

$$\Omega = \Omega_1 . \Omega_2 . \tag{I.10}$$

En utilisant à nouveau la formule (I.5), on conclut que

$$S^* = S_1^* + S_2^* . \tag{I..11}$$

Le résultat se généralise sans difficulté à la réunion d'un nombre quelconque de systèmes isolés. Nous l'exprimerons sous la forme suivante : *si un système est constitué de sous-systèmes séparément isolés, son entropie microcanonique est la somme des entropies microcanoniques de ces sous-systèmes.*

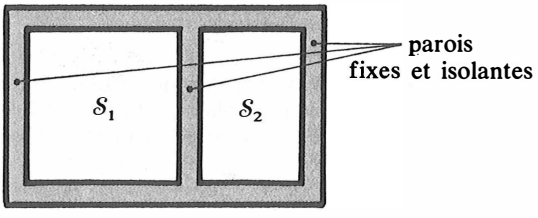


FIGURE 1
Système isolé $\mathcal{S}$ constitué par la réunion de deux sous-systèmes $\mathcal{S}_1$ et $\mathcal{S}_2$ séparément isolés.

11. Comme k est de l'ordre de 10^{-23} J/K et N de l'ordre de 10^{23}, les valeurs de S^* sont des nombres raisonnables lorsqu'on les exprime en *joule par degré Kelvin.* Nous serons d'ailleurs amenés à poser

$$R \mathrel{\widehat{=}} N_A k ,$$

où N_A est le nombre d'Avogadro. La constante R, appelée pour des raisons historiques « constante des gaz parfaits », a pour valeur numérique

$$R = 8{,}31 \text{ J/K} .$$

II. PROPRIÉTÉS ET COMPORTEMENT D'UN SYSTÈME ISOLÉ

Les caractéristiques macroscopiques d'un système isolé à l'équilibre sont entièrement déterminées par la distribution statistique de ses états microscopiques, que nous venons de postuler. Nous allons en donner quelques exemples importants.

A. Distribution statistique d'une variable interne

Considérons un système isolé, dans un état macroscopique d'équilibre caractérisé par les valeurs $\{E, x\}$ des paramètres extérieurs. Nous voulons expliciter la distribution statistique d'une variable interne (§ I.A.1), que nous noterons y.

1. FORMULE GÉNÉRALE

L'ensemble des valeurs possibles pour la variable interne considérée est constitué par les valeurs y_ℓ qu'elle prend dans les divers états (ℓ) accessibles. Choisissons l'une de ces valeurs possibles y. Sa probabilité est égale (« loi d'addition », appendice I, § I.2.a) à la probabilité $1/\Omega$ de chacun des états accessibles multipliée par le *nombre d'états accessibles pour lesquels* $y_\ell = y$.

En pratique, la variable interne étudiée est une grandeur macroscopique; les valeurs qu'elle prend dans un système isolé ne peuvent pas être parfaitement précisées, mais sont entachées d'une incertitude δy, comme pour l'énergie (cf. note 4). On cherche donc la probabilité $w^*(E, x; y)\,\delta y$ pour que la valeur de cette grandeur soit comprise entre y et $y + \delta y$, les paramètres extérieurs ayant les valeurs $\{E, x\}$. Parmi tous les états accessibles, comptons ceux pour lesquels la valeur y_ℓ de la grandeur considérée est telle que

$$y \leqslant y_\ell \leqslant y + \delta y \,. \tag{II.1}$$

Soit $\omega(E, x; y)\,\delta y$ leur nombre[12]. La probabilité cherchée vaut

$$w^*(E, x; y)\,\delta y = \frac{\omega(E, x; y)\,\delta y}{\Omega(E, x)} \,. \tag{II.2}$$

Bien entendu, cette probabilité est nulle si la valeur y choisie n'est réalisée par aucun état accessible. De plus, si l'on somme les nombres $\omega(E, x; y)\,\delta y$ sur toutes les valeurs possibles de y, on retrouve le nombre total d'états accessibles :

$$\int_{-\infty}^{+\infty} \omega(E, x; y)\,\mathrm{d}y = \Omega(E, x) \,, \tag{II.3}$$

de sorte que la distribution (II.2) est correctement normalisée :

$$\int_{-\infty}^{+\infty} w^*(E, x; y)\,\mathrm{d}y = 1 \,. \tag{II.4}$$

---

12. Comme le nombre total d'états accessibles

$$\Omega(E, x; \delta E) = \rho(E, x)\,\delta E \,, \tag{1}$$

le nombre de ceux qui vérifient la condition supplémentaire (II.1) est proportionnel à δE : $\omega(E, x; y)\,\delta y$ est en réalité de la forme

$$\omega(E, x; \delta E; y)\,\delta y = \rho(E, x; y)\,\delta E\,\delta y \,, \tag{2}$$

où $\rho(E, x; y)$ est la densité d'états microscopiques qui vérifient (II.1), de sorte que

$$w^*(E, x; y) = \rho(E, x; y)/\rho(E, x) \,. \tag{3}$$

Le nombre total $\Omega(E, x)$ d'états accessibles est énorme. La formule (II.2) indique par conséquent que seules les valeurs de y pour lesquelles $\omega(E, x; y)\delta y$ est lui aussi énorme ont une probabilité appréciable. Considérons alors

$$\ln w^*(E, x; y) = \ln \omega(E, x; y) - \ln \Omega(E, x) . \tag{II.5}$$

Dans cette expression, le second terme est proportionnel à l'entropie microcanonique :

$$\ln \Omega(E, x) = \frac{1}{k} S^*(E, x) , \tag{II.6}$$

et il est de l'ordre de N pour N grand (cf. § I.C). La probabilité (II.2) n'est appréciable que si le premier terme de l'expression (II.5) est lui aussi de l'ordre de N. Par analogie avec (II.6), nous poserons[13]

$$s^*(E, x; y) \mathrel{\widehat{=}} k \ln \omega(E, x; y) , \tag{II.7}$$

que nous appellerons *entropie microcanonique partielle.* La formule (II.2) s'écrit alors sous la forme

$$w^*(E, x; y) = \exp \frac{1}{k} \left[s^*(E, x; y) - S^*(E, x) \right] . \tag{II.8}$$

Pour des valeurs fixées des paramètres extérieurs $\{E, x\}$, la dépendance en y de la densité de probabilité $w^*(E, x; y)$ est donnée par le premier terme de (II.8)[14] :

$$w^*(E, x; y) \propto \exp \frac{1}{k} s^*(E, x; y) \quad \text{pour } \{E, x\} \text{ fixés} . \tag{II.9}$$

La valeur la plus probable y_m de la variable interne y est donc celle qui rend s^* maximum.

2. RÉSULTAT PRATIQUE POUR UN GRAND SYSTÈME

Le nombre total N de particules du système étudié est toujours très grand. Ceci permet d'obtenir une expression approchée pour la densité de probabilité de la variable interne y (les paramètres extérieurs $\{E, x\}$ étant fixés), en appliquant une méthode usuelle en mécanique statistique, dont les conditions de validité sont précisées au complément I.C à propos de la distribution binomiale (notamment dans la note 6).

La *valeur la plus probable* y_m de la variable étudiée est donnée, d'après (II.9), par l'équation

$$\frac{1}{k} \frac{\partial s^*}{\partial y} (E, x; y_m) = 0 , \quad \{E, x\} \text{ fixés} . \tag{II.10}$$

(nous sous-entendrons désormais la dépendance par rapport à $\{E, x\}$). Nous supposons que l'équation (II.10) admet *une seule solution* y_m, et que celle-ci rend $s^*(y)$ (et donc

13. Nous prenons à nouveau le logarithme d'une quantité qui a des dimensions; la justification est analogue à celle qui a été développée pour l'entropie microcanonique S^* (voir paragraphe I.C, en particulier note 10).

14. *Attention :* la densité de probabilité $w^*(y)$ est *proportionnelle* au nombre d'états accessibles réalisant la valeur y (c'est ce qu'exprime la formule (II.9), compte tenu de (II.7)), alors que la probabilité microcanonique P^*_ξ est *inversement proportionnelle* au nombre total d'états accessibles. Cette différence est souvent source de confusions, alors que son origine est parfaitement claire, et réside dans le fait que tous les états accessibles sont également probables : la probabilité de chacun d'eux est, par normalisation, égale à l'inverse de leur nombre total; mais la probabilité d'un événement que plusieurs d'entre eux réalisent est proportionnel au nombre de ceux-ci, d'après la « loi d'addition » des probabilités.

la densité de probabilité $w^*(y)$) *maximum*[15]. *Développons alors* s^* *autour de* $y = y_m$: à cause de la condition (II.10), le terme du premier ordre est nul, de sorte que

$$\frac{1}{k}\, s^*(y) \simeq \frac{1}{k}\, s^*(y_m) + \frac{1}{2}\,(y - y_m)^2\, \frac{1}{k}\, \frac{\partial^2 s^*}{\partial y^2}\,(y_m)\,. \tag{II.11}$$

Posons

$$\frac{1}{\Delta^2} \widehat{=} -\frac{1}{k}\, \frac{\partial^2 s^*}{\partial y^2}\,(y_m)\,; \tag{II.12}$$

comme y_m rend par hypothèse $s^*(y)$ maximum, Δ^2 est positif; il dépend de $\{E, x\}$. Reportons enfin ce développement limité dans (II.8) :

$$w^*(y) \simeq w^*(y_m)\, \exp\left[-(y - y_m)^2/2\Delta^2\right]. \tag{II.13}$$

La distribution statistique d'une variable interne y *est pratiquement une gaussienne*[16], *centrée sur la valeur la plus probable* y_m.

Par conséquent, pour un grand système, la valeur moyenne $\bar{y}$ d'une variable interne y coïncide avec sa valeur la plus probable :

$$\bar{y} \simeq y_m\,. \tag{II.14}$$

---

15. Si ces hypothèses ne sont pas vérifiées, il faut revenir à la formule générale et examiner chaque cas particulier. Il peut se faire également que la valeur la plus probable y_m se situe à l'une des bornes de l'intervalle de variation permis à y par les conditions extérieures, auquel cas le raisonnement qui suit n'est pas valable non plus.

16. Pour analyser la validité de l'approximation (II.13), nous nous appuyons sur la note 6 du complément I.C.

Supposons que la valeur la plus probable y_m de la variable interne considérée se comporte comme N^α pour N grand[18] :

$$y_m \underset{N \to \infty}{\sim} \gamma N^\alpha\,, \tag{1}$$

où γ et α sont des constantes indépendantes de N. Nous avons souligné au paragraphe 1 que s^*/k se comporte comme N. On en déduit le comportement avec N de ses dérivées successives :

$$\frac{1}{k}\, \frac{\partial^n s^*}{\partial y^n}\,(y_m) \underset{N \to \infty}{\sim} c_n N^{1-n\alpha} \tag{2}$$

(au numérateur, l'accroissement de s^* est d'ordre N; au dénominateur, la puissance n-ième de l'accroissement de y est d'ordre $(N^\alpha)^n$). En particulier, la formule (II.12) indique que

$$\Delta \underset{N \to \infty}{\sim} \delta N^{\alpha - 1/2}\,. \tag{3}$$

Considérons alors le développement de s^*/k :

$$\frac{1}{k}\, s^*(y) = \frac{1}{k}\, s^*(y_m) + \frac{1}{2}\,(y - y_m)^2\, \frac{1}{k}\, \frac{\partial^2 s^*}{\partial y^2}\,(y_m) + \ldots + \frac{1}{n!}\,(y - y_m)^n\, \frac{1}{k}\, \frac{\partial^n s^*}{\partial y^n}\,(y_m) + \ldots\,, \tag{4}$$

et évaluons l'ordre de grandeur de ses divers termes pour $(y - y_m)$ d'ordre Δ :

$$(y - y_m)^n\, \frac{1}{k}\, \frac{\partial^n s^*}{\partial y^n}\,(y_m) \underset{N \to \infty}{\sim} \delta^n N^{n\alpha - n/2} \cdot c_n N^{1-n\alpha} \sim \delta^n c_n N^{1-n/2}\,. \tag{5}$$

Donc, le premier terme du développement (4) est d'ordre N, le terme en $(y - y_m)^2$ d'ordre 1, et les termes suivants d'ordre $1/\sqrt{N}$, puis $1/N$, ... Par conséquent, *les corrections* à la formule (II.11) et à l'expression approchée (II.13) de la densité de probabilité *sont négligeables pour* N *grand*, sauf évidemment cas particuliers (cf. note 15). En pratique, l'approximation (II.13) est impossible à distinguer de la distribution exacte dans un intervalle centré en y_m et dont la largeur est d'ordre Δ; en dehors de cet intervalle, $w^*(y)$ est totalement négligeable.

Son écart quadratique moyen Δ est donné par (II.12). On voit facilement[17] que, pour N grand, Δ / y_m se comporte comme $1/\sqrt{N}$: *les fluctuations sont d'autant plus petites*, en valeur relative, *que N est plus grand*[18]; pour des systèmes suffisamment grands, on peut considérer à une bonne approximation que toute variable interne est pratiquement fixée à sa valeur la plus probable, ses fluctuations devenant inappréciables. On démontre alors (exercice II-19) que l'entropie $S^*(E, x)$ du système est égale à $s^*(E, x ; y_m(E, x))$.

B. Évolution spontanée d'un système isolé après modification des contraintes extérieures

Tout système macroscopique abandonné à lui-même finit par atteindre l'état d'équilibre permis par les conditions extérieures. Ce fait fondamental sera analysé au chapitre IV. Jusque-là, nous le considérons comme un fait d'expérience, qui justifie l'importance accordée à l'étude de l'équilibre macroscopique. Dans ce paragraphe, nous allons nous intéresser aux *changements d'état d'équilibre* pour un système *isolé*[19], changements provoqués par des modifications dans les contraintes extérieures qui lui sont imposées. Commençons par un exemple simple.

1. DÉTENTE DE JOULE

Considérons un gaz isolé à l'équilibre macroscopique; son énergie totale est E (à δE près), le nombre de molécules est N et le volume V. Supposons que, en maintenant le système isolé c'est-à-dire en gardant E et N constants, on modifie le volume V. Le gaz va évoluer spontanément vers un nouvel état d'équilibre, dont nous voulons comparer les caractéristiques à celles de l'état de départ.

Pour pouvoir appliquer les considérations du paragraphe I, il est indispensable que *le gaz reste isolé.* Il est donc exclu par exemple que nous diminuions son volume, car ceci exigerait que les forces extérieures qui maintiennent en place les parois du récipient effectuent un certain travail; il y aurait alors échange d'énergie entre le gaz et l'extérieur; le gaz ne constituerait plus un système isolé. Par contre, il est possible d'*augmenter le volume,* à condition d'opérer de façon telle qu'il n'y ait *pas d'échange de travail avec l'extérieur :* initialement, le gaz qui constitue le système est enfermé dans une enceinte de volume V, séparée par une cloison d'un compartiment vide, de volume ΔV (fig. 2); l'ensemble est isolé de l'extérieur par des parois fixes et totalement imperméables à l'énergie et aux molécules. Le gaz ayant atteint l'état d'équilibre caractérisé par ces conditions (énergie E, nombre de molécules N, volume V), nous ouvrons la cloison séparant les deux compartiments. Ceci peut se faire en ne mettant

17. Les formules (1) et (3) de la note 16 donnent aussitôt

$$\frac{\Delta}{y_m} \underset{N \to \infty}{\sim} \frac{\delta}{\gamma} N^{-1/2} .$$

18. Les grandeurs physiques usuelles se divisent en deux grandes classes (complément II.A, § II) : pour certaines, que l'on qualifie d'*extensives,* la valeur la plus probable (ou la valeur moyenne) se comporte comme N, et l'écart quadratique moyen comme $\sqrt{N}$; pour les autres, dites *intensives,* la valeur moyenne reste finie quand $N \to \infty$, et l'écart quadratique moyen décroît comme $1/\sqrt{N}$.

19. Ces changements doivent donc se faire sans échange d'énergie ni de particules avec l'extérieur.

en jeu qu'un travail négligeable[20]. Nous savons par expérience que le gaz va alors évoluer spontanément vers un *nouvel état d'équilibre,* dans lequel il va occuper *tout le volume* $V+\Delta V$ *disponible,* son énergie E et le nombre de molécules N restant inchangés.

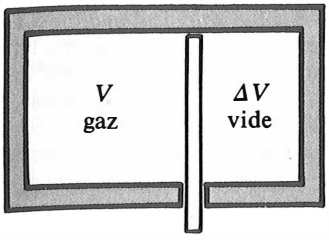


FIGURE 2
Dispositif schématique permettant d'augmenter de ΔV le volume V d'un gaz sans modification de son énergie : les parois grisées sont totalement isolantes, la cloison intermédiaire est amovible.

D'après le postulat fondamental, le nouvel état d'équilibre correspond à une équiprobabilité de tous les microétats accessibles dans la nouvelle situation. Il semble clair intuitivement que l'ensemble de ces états accessibles comprend tous ceux qui étaient déjà accessibles dans la situation initiale[21] : après tout, les molécules auraient pu rester toutes confinées dans le compartiment initial sans violer aucune loi microscopique[22]. Mais l'augmentation du volume offert ouvre des possibilités supplémentaires, rendant accessibles des états qui étaient auparavant interdits (ceux qui donnent aux molécules une probabilité de présence non nulle dans le compartiment de volume ΔV). En bref, *le nombre d'états accessibles a augmenté :*

$$\Omega(E,\ N,\ V+\Delta V)>\Omega(E,\ N,\ V) \quad \text{si} \quad \Delta V>0\ . \tag{II.15}$$

Par conséquent, d'après la formule (I.5), *l'évolution spontanée du gaz (maintenu isolé) s'est accompagnée d'une augmentation de son entropie* microcanonique :

$$\Delta S^* \mathrel{\widehat{=}} S^*(E,\ N,\ V+\Delta V)-S^*(E,\ N,\ V)>0\ . \tag{II.16}$$

20. On peut par exemple faire glisser la cloison parallèlement à elle-même, c'est-à-dire perpendiculairement aux forces de pression qui s'exercent sur elle, et suffisamment lentement pour que les forces de viscosité restent négligeables.

21. Paradoxalement, ceci ne ressort pas de façon évidente de la méthode utilisée au paragraphe III.B.2 du chapitre I pour calculer la densité d'états dans le cas où les molécules du gaz sont libres («gaz parfait») : un état propre de l'hamiltonien individuel correspondant à des conditions aux limites strictes dans le volume V ne le reste pas quand on passe à $V+\Delta V$ (la fonction d'onde associée présente une discontinuité de sa dérivée première sur la limite du volume V, qui se trouve ensuite au beau milieu du nouveau domaine disponible). Il s'agit là encore d'un artefact provoqué par des conditions aux limites peu réalistes. Mais la formule (III.42) du chapitre I indique clairement que le nombre d'états accessibles est une fonction très rapidement croissante du volume, ce qui confirme, pour des molécules libres, la conclusion (II.16).

22. Pourquoi n'y restent-elles donc pas? Parce que le nombre d'états accessibles réalisant cette condition est extrêmement faible par rapport au nombre total d'états accessibles dans la nouvelle situation; la probabilité pour que le compartiment de volume ΔV se retrouve vide est ainsi tellement petite que cette éventualité n'est jamais observée. On peut dire que le volume effectivement occupé par le gaz s'établit, parmi toutes les valeurs possibles, à la valeur maximale $V+\Delta V$ compatible avec les conditions extérieures, parce que c'est celle qui a, de loin, la probabilité la plus forte (cf. note 15 ci-dessus).

2. GÉNÉRALISATION

La conclusion (II.16) est en fait générale, comme nous le démontrerons au chapitre IV (« théorème H » de Boltzmann). Nous nous contenterons ici de préciser son domaine de validité et de souligner son origine.

Considérons un système isolé quelconque, ayant atteint son état d'équilibre macroscopique dans des conditions extérieures données. Supposons que l'on modifie alors ces conditions extérieures, *en maintenant le système isolé.* Comme dans l'exemple du gaz, cette exigence fondamentale s'avère assez restrictive; elle permet pourtant, dans certains cas, de *relâcher certaines contraintes* qui pèsent initialement sur le système sans être essentielles pour assurer son isolement (nous en verrons d'autres exemples au paragraphe III). Si l'on relâche effectivement une ou plusieurs de ces contraintes, le système évolue vers un *nouvel état d'équilibre.* Celui-ci correspond forcément à un *nombre accru d'états accessibles,* car la situation initiale du système n'est pas exclue par les nouvelles conditions extérieures, qui ajoutent par contre de nouvelles possibilités.

> *L'évolution spontanée d'un système isolé s'accompagne d'une augmentation de son entropie* microcanonique; le nouvel *état d'équilibre* atteint par le système après relâchement d'une contrainte correspond au *maximum de son entropie* compatible avec les contraintes restantes.

III. ÉQUILIBRE ENTRE SOUS-SYSTÈMES D'UN SYSTÈME ISOLÉ

L'équilibre d'un système macroscopique peut être analysé selon deux points de vue complémentaires : dans le *point de vue global* que nous avons adopté jusqu'ici, le système est étudié comme un tout, et son équilibre se manifeste par le fait que les diverses distributions statistiques qui lui sont associées (distribution des états microscopiques et des variables internes) sont indépendantes du temps; on peut aussi considérer l'équilibre macroscopique du système comme résultant d'un *équilibre mutuel entre ses diverses parties* (sous-systèmes). C'est ce second point de vue que nous allons maintenant développer.

A. Température et équilibre thermique

1. DÉFINITIONS

a. Température microcanonique

Soit un système isolé à l'équilibre macroscopique. Son entropie microcanonique S^* est une fonction des paramètres extérieurs $\{E, x\}$ qui déterminent son état macroscopique :

$$S^* = S^*(E, x) \,. \tag{III.1}$$

On définit la *température microcanonique* T^* de ce système par la relation

$$\boxed{\frac{1}{T^*} \widehat{=} \frac{\partial S^*}{\partial E}} \,. \tag{III.2}$$

La dérivée partielle indique que les autres paramètres extérieurs $\{x\}$ définissant l'état macroscopique du système *isolé* (tels que volume et nombre de particules) sont maintenus constants.

Comme l'entropie, la température est une *grandeur à caractère statistique :* elle n'a de sens que pour des systèmes macroscopiques. A partir de l'expression (III.1)

de S^*, supposée connue, la définition (III.2) donne T^* comme *fonction des paramètres extérieurs :*

$$T^* = T^*(E, x) . \qquad \text{(III.3)}$$

Nous verrons au paragraphe 2 que les propriétés de la grandeur T^* justifient qu'on l'appelle « température ». L'*unité de température* est ici fixée par la définition (III.2) et par le choix de la constante de Boltzmann k dans la définition de l'entropie statistique (I, § II.A); cette unité est le *degré kelvin*[23], noté K.

Énonçons enfin une propriété fondamentale, qui est démontrée de façon générale dans le complément II.E : quel que soit le système considéré, son entropie microcanonique S^* est une fonction croissante de son énergie E, car le nombre d'états accessibles croît avec l'énergie (les autres paramètres extérieurs $\{x\}$ étant maintenus constants); par conséquent, *la température microcanonique est toujours positive*[24], [25] :

$$T^* > 0 . \qquad \text{(III.4)}$$

b. Contact thermique entre deux systèmes

α. Définition pratique

On sait depuis longtemps que les échanges d'énergie entre deux systèmes ne font pas nécessairement intervenir le travail mécanique des forces macroscopiques

23. Cette unité provient de celle qui avait été définie par Celsius (1742) pour son « échelle centigrade » : c'était la centième partie de la différence de température entre le point d'ébullition de l'eau à la pression atmosphérique (100 °C), et son point de fusion à cette même pression (0 °C). Mais la division de l'intervalle en cent « parties égales » supposait le choix d'un thermomètre particulier. La thermodynamique permit ensuite de s'affranchir de cet arbitraire, en définissant une « température absolue ». Il en résulta l'*« échelle Kelvin »*, que nous utiliserons presque exclusivement dans la suite. L'*« échelle Celsius »*, qu'on lui préfère dans la vie courante, est simplement décalée par rapport à l'échelle Kelvin : si T et t mesurent la même température dans les échelles Kelvin et Celsius, respectivement, ces nombres sont reliés, à une excellente approximation, par

$$t\ (°C) = T\ (K) - 273{,}15.$$

Nous montrerons plus loin (III, § IV.B.3) que le choix de la constante de Boltzmann dans la définition de l'entropie a en fait pour but l'identification de la « température statistique » que nous définissons ici avec la « température des gaz parfaits » et la « température absolue » qu'utilisaient déjà les thermodynamiciens avant l'introduction de la mécanique statistique.

24. On est amené dans certains cas à parler de « températures négatives » (cf. complément II.B). Il ne s'agit pas alors d'un système dans son ensemble, ni même d'un de ses sous-systèmes, mais seulement de certains de ses degrés de liberté (degrés de liberté de spin, par exemple), dans des situations particulières où ils sont provisoirement et approximativement découplés des autres; le système n'est donc pas à l'équilibre.

25. L'échelle des températures ne paraît pas être bornée supérieurement : en laboratoire, on atteint 10^8 K dans de faibles volumes et pendant des temps très courts; des températures du même ordre règnent dans les étoiles; les modèles théoriques actuels concernant l'évolution de l'Univers envisagent une explosion initiale, au cours de laquelle auraient été produites des températures de l'ordre de 10^{32} K, peut-être même supérieures.

Le zéro de l'échelle Kelvin, dit *« zéro absolu »*, est une limite théorique impossible à atteindre : plus la température est basse, plus il est difficile de l'abaisser davantage. On arrive actuellement à faire « descendre » des échantillons de certains matériaux particuliers jusqu'à 10^{-6} K environ, mais les températures les plus basses vraiment utilisables se situent aux alentours du millikelvin.

qu'ils exercent l'un sur l'autre, et l'*on appelle «chaleur» l'énergie échangée qui n'est pas du travail.* Nous nous appuierons sur ce fait d'expérience et adopterons cette définition empirique de la chaleur; nous reviendrons plus en détail, au complément III.L (§ III), sur la réalité microscopique sous-jacente.

Nous dirons alors que deux systèmes sont en *contact thermique* lorsqu'ils sont placés dans des conditions telles qu'*ils peuvent échanger librement de la chaleur.* Concrètement, ces échanges se font le plus souvent à travers une cloison commune aux deux systèmes : de telles cloisons permettant les échanges de chaleur sont dites *diathermes;* on appelle au contraire *adiabatiques* les parois qui interdisent tout échange de chaleur[26]. Si les deux systèmes peuvent échanger seulement de la chaleur, c'est-à-dire s'ils peuvent échanger de l'énergie mais pas du travail, leur contact sera qualifié de *«purement thermique».*

β. Description théorique

Considérons deux systèmes $\mathcal{S}_1$ et $\mathcal{S}_2$, et supposons d'abord qu'ils soient séparément isolés. Nous pouvons les réunir en un système unique $\mathcal{S}$, comme au paragraphe I.C.2. Si H_1 et H_2 sont les hamiltoniens des deux premiers systèmes, celui du système global $\mathcal{S}$ sera simplement

$$H_0 = H_1 + H_2 \,, \tag{III.5}$$

puisqu'il n'y a, par hypothèse, aucune interaction entre $\mathcal{S}_1$ et $\mathcal{S}_2$. Dans une telle situation, les énergies E_1 et E_2 des systèmes isolés $\mathcal{S}_1$ et $\mathcal{S}_2$ sont séparément conservées, ainsi évidemment que leur somme E, énergie du système $\mathcal{S}$.

Supposons maintenant que $\mathcal{S}_1$ et $\mathcal{S}_2$ ne soient plus isolés l'un de l'autre, leur réunion $\mathcal{S}$ restant cependant isolée de l'extérieur. L'hamiltonien H du système global $\mathcal{S}$ comprend alors un terme d'interaction, ou terme de couplage, H_{12} :

$$H = H_1 + H_2 + H_{12} \,. \tag{III.6}$$

A strictement parler, seul le système $\mathcal{S}$, qui reste isolé, possède une énergie bien définie E, valeur propre de l'hamiltonien H : en général, les états propres de H ne sont plus en même temps états propres de H_1 et H_2, à cause de la présence de H_{12}.

Comment pouvons-nous, dans ce cadre théorique, décrire une situation où $\mathcal{S}_1$ et $\mathcal{S}_2$ sont en contact thermique? Dans la définition du paragraphe α, il est question d'échange de chaleur, c'est-à-dire d'énergie, entre les deux systèmes, ce qui suppose que leurs énergies individuelles E_1 et E_2 continuent à exister, bien qu'elles varient au cours des échanges. Ceci est possible si *le terme de couplage est petit* devant chacun des termes individuels H_1 et H_2.

On aboutit ainsi à l'image suivante du contact thermique : *les deux systèmes sont couplés, mais leur couplage est faible.* D'une part, la faiblesse du couplage permet de considérer que, à une excellente approximation, les états propres de H sont également états propres de H_1 et H_2 : *les énergies E_1 et E_2 des deux systèmes continuent à être définies,* l'énergie du système global en étant la somme

$$E = E_1 + E_2 \,; \tag{III.7}$$

de plus, *dans un état microscopique (ℓ) du système $\mathcal{S}$,* on peut considérer que $\mathcal{S}_1$ *et* $\mathcal{S}_2$ *sont eux-mêmes dans des états microscopiques* (ℓ_1) *et* (ℓ_2) séparément déterminés :

$$(\ell) = (\ell_1, \ell_2) \,. \tag{III.8}$$

Mais d'autre part l'existence du couplage, même faible, implique que *les énergies* E_1

26. Aucune paroi réelle n'est adiabatique au sens strict du terme. Mais on la considère comme telle si l'on peut négliger la quantité de chaleur qu'elle laisse passer pendant la durée de l'expérience envisagée.

et E_2 ne sont plus séparément conservées; seule l'est leur somme E[27]. La validité et la signification physique de cette image sont discutées dans le complément II.A.

2. ÉQUILIBRE THERMIQUE ENTRE DEUX SYSTÈMES

Le postulat fondamental du paragraphe I va maintenant nous permettre d'analyser les propriétés de l'équilibre atteint par deux systèmes que l'on a mis en contact thermique.

a. Position du problème

De façon plus précise, nous allons étudier le problème suivant. Deux systèmes macroscopiques $\mathcal{S}_1$ et $\mathcal{S}_2$ sont disposés comme l'indique schématiquement la figure 3. Initialement, chacun d'eux est séparément isolé et à l'équilibre. On les met ensuite en contact thermique en rendant diatherme la cloison qui les sépare (nous supposons ici que la cloison reste fixe, de sorte que le contact est purement thermique). Quelles sont les caractéristiques de l'état d'équilibre (dit « équilibre thermique ») qu'atteignent finalement les deux systèmes?

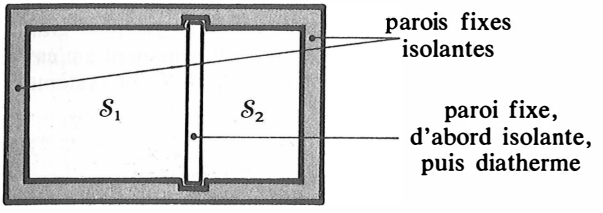


FIGURE 3
Établissement du contact thermique entre deux systèmes $\mathcal{S}_1$ et $\mathcal{S}_2$ initialement isolés.

Le système global $\mathcal{S}$ constitué par la réunion des deux systèmes $\mathcal{S}_1$ et $\mathcal{S}_2$ *reste isolé* tout au long de l'expérience. C'est donc en appliquant à ce système $\mathcal{S}$ le postulat fondamental que nous obtiendrons la clef du problème.

b. L'établissement du contact thermique comme relâchement d'une contrainte

Les considérations du paragraphe 1.b.β indiquent que l'énergie E du système global $\mathcal{S}$ est constamment la somme

$$E = E_1 + E_2 \tag{III.9}$$

des énergies de ses sous-systèmes $\mathcal{S}_1$ et $\mathcal{S}_2$, que ceux-ci soient isolés ou en contact thermique.

Dans la *situation initiale* (cloison adiabatique, $\mathcal{S}_1$ et $\mathcal{S}_2$ séparément isolés), les énergies E_1 et E_2 sont des *paramètres extérieurs*, fixés à des valeurs que nous noterons $E_1^{(i)}$ et $E_2^{(i)}$:

$$\left.\begin{aligned} E_1 &= E_1^{(i)} \\ E_2 &= E_2^{(i)} \end{aligned}\right\} \text{ dans l'état initial.} \tag{III.10}$$

27. Le contact entre les deux systèmes est purement thermique si seules peuvent varier les énergies E_1 et E_2, tous les autres paramètres extérieurs qui caractérisent l'état macroscopique des systèmes $\mathcal{S}_1$ et $\mathcal{S}_2$ étant maintenus constants.

De façon plus générale, l'état macroscopique du système $\mathcal{S}$ est caractérisé dans la situation initiale par la réunion des deux ensembles $\{E_1^{(i)}, x_1\}$ et $\{E_2^{(i)}, x_2\}$ de paramètres extérieurs associés aux deux systèmes $\mathcal{S}_1$ et $\mathcal{S}_2$. Son entropie microcanonique $S^{*(i)}$ est simplement (§ I.C.2) la somme des entropies microcanoniques de ces deux systèmes :

$$S^{*(i)} = S_1^*(E_1^{(i)}, x_1) + S_2^*(E_2^{(i)}, x_2) \,. \tag{III.11}$$

L'établissement du *contact thermique* a pour effet de *relâcher les contraintes* (III.10) : les échanges d'énergie entre les deux systèmes changent en général les valeurs de E_1 et E_2. Lorsque le nouvel état d'équilibre est atteint par le système $\mathcal{S}$, E_1 *et* E_2 *sont devenues des variables internes :* la répartition de l'énergie totale E entre les sous-systèmes $\mathcal{S}_1$ et $\mathcal{S}_2$ n'est plus imposée, et peut donc fluctuer.

C'est en réalité une seule contrainte qui a été relâchée, donnant lieu à la transformation d'un seul paramètre extérieur en variable interne, puisque, l'énergie totale étant toujours fixée, la donnée de E_1 détermine aussitôt E_2 :

$$E_2 = E - E_1 \tag{III.12}$$

(comme il n'y a pas d'ambiguïté, nous notons simplement E l'énergie totale du système $\mathcal{S}$). Nous pouvons appliquer à la variable interne E_1 les arguments généraux du paragraphe II.A : dans le nouvel état d'équilibre, sa distribution sera donnée, à une excellente approximation, par une gaussienne centrée sur sa valeur la plus probable E_{1m}, et d'autant plus étroite (en valeur relative) que le système est plus grand. C'est donc la *valeur la plus probable* E_{1m} qui nous intéresse principalement : c'est elle qui caractérisera l'équilibre thermique entre les sous-systèmes $\mathcal{S}_1$ et $\mathcal{S}_2$ du système isolé $\mathcal{S}$.

c. Caractérisation de l'équilibre thermique

α. Condition d'équilibre

Il nous faut d'abord expliciter la condition qui fixe la valeur la plus probable de la variable interne E_1.

Nous commençons par compter les états accessibles du système global $\mathcal{S}$ pour lesquels l'énergie du sous-système $\mathcal{S}_1$ vaut E_1 (à δE_1 près). La faiblesse du couplage thermique (§ 1.b.β) permet de considérer que tout état microscopique (ℓ) du système global $\mathcal{S}$ est construit à partir d'un état (ℓ_1) du système $\mathcal{S}_1$ et un état (ℓ_2) de $\mathcal{S}_2$, comme lorsque ces deux systèmes étaient encore isolés. Soit alors $\rho_1(E_1)$ la densité d'états du système $\mathcal{S}_1$, $\rho_2(E_2)$ celle de $\mathcal{S}_2$ (nous sous-entendons la dépendance par rapport aux paramètres extérieurs $\{x_1\}$ et $\{x_2\}$, qui restent inchangés). Le nombre d'états microscopiques du *système global* $\mathcal{S}$ vérifiant les conditions

$$E_1 \leqslant E_{\ell_1} \leqslant E_1 + \delta E_1 \,, \tag{III.13.a}$$

$$E_2 \leqslant E_{\ell_2} \leqslant E_2 + \delta E_2 \tag{III.13.b}$$

est donc $\rho_1(E_1)\,\delta E_1 \times \rho_2(E_2)\,\delta E_2$ (cf. formule (I.10)). Mettons ce résultat sous la forme générale introduite au paragraphe II.A, c'est-à-dire écrivons que le nombre $\omega(E, E_1)\delta E_1$ d'états microscopiques du système global $\mathcal{S}$ vérifiant les conditions[28]

$$E \leqslant E_\ell \leqslant E + \delta E \tag{III.14.a}$$

et

$$E_1 \leqslant E_{\ell_1} \leqslant E_1 + \delta E_1 \tag{III.14.b}$$

28. La première est la condition habituelle définissant les états accessibles (les autres paramètres extérieurs étant fixés une fois pour toutes), la seconde est la condition (II.1) écrite pour la variable interne $y \equiv E_1$.

vaut[29]

$$\omega(E, E_1)\,\delta E_1 = \rho_1(E_1)\,\rho_2(E_2 = E - E_1)\,\delta E\,\delta E_1 \,. \tag{III.15}$$

Nous en déduisons *l'entropie microcanonique partielle* associée à ce nombre d'états; en négligeant, dans les mêmes conditions qu'au paragraphe I.C.1, le logarithme des incertitudes δE et δE_1, on obtient

$$s^*(E, E_1) = k \ln \rho_1(E_1) + k \ln \rho_2(E_2 = E - E_1) \,, \tag{III.16}$$

où l'on reconnaît l'expression de l'entropie microcanonique de chacun des deux sous-systèmes (formule (I.8))[30] :

$$s^*(E, E_1) = S_1^*(E_1) + S_2^*(E_2 = E - E_1) \,. \tag{III.17}$$

La *valeur la plus probable* E_{1m} de la variable interne E_1 est celle qui rend l'entropie partielle maximale :

$$\left.\frac{\partial s^*}{\partial E_1}\right|_{E_1 = E_{1m}} = 0 \,. \tag{III.18}$$

D'après (III.17), cette condition s'écrit[31]

$$\left.\frac{\partial S_1^*}{\partial E_1}\right|_{E_1 = E_{1m}} - \left.\frac{\partial S_2^*}{\partial E_2}\right|_{E_2 = E - E_{1m}} = 0 \,. \tag{III.19}$$

Mais les températures microcanoniques T_1^* et T_2^* des systèmes $\mathcal{S}_1$ et $\mathcal{S}_2$ sont par définition telles que

$$\frac{1}{T_1^*} = \frac{\partial S_1^*}{\partial E_1} \,, \quad \frac{1}{T_2^*} = \frac{\partial S_2^*}{\partial E_2} \,. \tag{III.20}$$

---

29. Du point de vue technique, on a effectué un changement de variables

$$(E_1, E_2) \Longrightarrow (E, E_1),$$

défini par

$$E = E_1 + E_2 \,. \tag{1}$$

Le jacobien de la transformation vaut

$$\begin{vmatrix} \dfrac{\partial E}{\partial E_1} & \dfrac{\partial E}{\partial E_2} \\ \dfrac{\partial E_1}{\partial E_1} & \dfrac{\partial E_1}{\partial E_2} \end{vmatrix} = \begin{vmatrix} 1 & 1 \\ 1 & 0 \end{vmatrix} = -1 \,, \tag{2}$$

de sorte que

$$\mathrm{d}E_1\,\mathrm{d}E_2 = \mathrm{d}E\,\mathrm{d}E_1 \,. \tag{3}$$

30. L'entropie microcanonique partielle correspondant à la valeur E_1 de la variable interne dans l'état d'équilibre final se trouve donc être égale à l'entropie microcanonique qu'aurait le système global si les deux sous-systèmes étaient isolés et si l'énergie de $\mathcal{S}_1$ était fixée à la valeur E_1 (comparer les formules (III.17) et (III.11)). Ceci provient du fait que le couplage entre $\mathcal{S}_1$ et $\mathcal{S}_2$ est faible (§ 1.b.β).

31. $$\frac{\partial s^*}{\partial E_1} = \frac{\partial S_1^*}{\partial E_1} + \frac{\partial S_2^*}{\partial E_2}\frac{\mathrm{d}E_2}{\mathrm{d}E_1} = \frac{\partial S_1^*}{\partial E_1} - \frac{\partial S_2^*}{\partial E_2}$$

car $E_2 = E - E_1$.

Par conséquent, l'énergie la plus probable E_{1m} du système $\mathcal{S}_1$ dans l'état d'équilibre final est donnée par la condition[32]

$$T_1^*(E_{1m}) = T_2^*(E_2 = E - E_{1m}) : \tag{III.21}$$

la répartition la plus probable de l'énergie totale E entre les deux sous-systèmes correspond à l'égalité de leurs températures microcanoniques.

Le résultat (III.21) justifie la dénomination de « température » que nous avons adoptée pour la grandeur T^* définie au paragraphe 1.a : intuitivement, on attend en effet d'une température qu'elle prenne des valeurs égales pour deux corps pouvant échanger de la chaleur, lorsque l'équilibre thermique a été atteint; c'est ce qu'exprime la condition (III.21). Bien entendu, E_1 fluctue, et ne reste donc pas constamment fixée à la valeur E_{1m} qui égalise exactement les températures T_1^* et T_2^*; toutefois, comme les systèmes $\mathcal{S}_1$ et $\mathcal{S}_2$ sont grands, ces fluctuations sont de faible amplitude (en valeur relative).

β. Condition de stabilité

La condition (III.18) exprime seulement le fait que la valeur E_{1m} de E_1 rend extrémale l'entropie partielle $s^*(E_1)$. Pour que E_{1m} soit véritablement la valeur *la plus probable,* il faut que cet extremum de s^* soit un *maximum,* c'est-à-dire que

$$\frac{\partial^2 s^*}{\partial E_1^2}(E_{1m}) < 0 . \tag{III.22}$$

Si E_{1m} rend ainsi $s^*(E_1)$ maximum, les fluctuations autour de l'état d'équilibre amènent l'énergie E_1 à des valeurs moins probables que E_{1m}, de sorte que le système revient spontanément vers E_{1m}[33]. Au contraire, une solution de l'équation (III.18) qui correspondrait à un minimum de s^* ne donnerait pas un état d'équilibre stable, car le système évoluerait aussitôt vers des configurations plus problables. C'est pourquoi l'inégalité (III.22) est appelée *« condition de stabilité ».*

En dérivant deux fois l'expression (III.17) de s^*, on explicite cette condition sous la forme

$$\frac{\partial^2 S_1^*}{\partial E_1^2}(E_{1m}) + \frac{\partial^2 S_2^*}{\partial E_2^2}(E_2 = E - E_{1m}) < 0 . \tag{III.23}$$

Nous allons montrer qu'en réalité chacun des deux termes du premier membre est négatif. Ceci résulte du fait que chacun des deux systèmes $\mathcal{S}_1$ et $\mathcal{S}_2$ doit lui-même être stable du point de vue thermique. Ainsi, la stabilité de l'équilibre thermique entre les deux systèmes $\mathcal{S}_1$ et $\mathcal{S}_2$ (condition (III.23)) est assurée par la stabilité thermique de chacun d'eux séparément (inégalité (III.25) ci-dessous).

Analysons donc la condition de stabilité pour un système unique, que nous notons simplement $\mathcal{S}$, et que nous supposons homogène[34]. Imaginons qu'on le divise

---

32. Comme au paragraphe II.A.2, nous supposons qu'il existe une seule valeur de E_1 vérifiant cette condition, et qu'elle correspond à un maximum de l'entropie partielle $s^*(E_1)$. Cette hypothèse peut ici être démontrée à partir d'arguments généraux sur la stabilité de l'équilibre thermique (voir paragraphes c.β et d ci-dessous).

33. Dans le cadre du principe ergodique, nous traduisons en termes d'évolution temporelle les fluctuations statistiques de la variable interne E_1.

34. Si le système considéré n'est pas homogène mais constitué de plusieurs parties elles-mêmes homogènes (système comprenant un liquide et un gaz, ou deux liquides non miscibles), l'argument s'applique à chacune de ces parties, et donc finalement au système lui-même : lorsque par exemple on divise ce système en deux sous-systèmes de même taille, il suffit de s'assurer que chacun des sous-systèmes contient une moitié de chacune des parties homogènes.

en deux sous-systèmes $\mathcal{S}'$ et $\mathcal{S}''$ de même taille (c'est-à-dire de même volume), séparés par une cloison diatherme. A l'équilibre, la répartition la plus probable de l'énergie E du système $\mathcal{S}$ donne par symétrie $E/2$ à chacun des deux sous-systèmes. L'inégalité (III.23), appliquée au système $\mathcal{S}$ considéré comme la réunion des deux sous-systèmes identiques $\mathcal{S}'$ et $\mathcal{S}''$, s'écrit

$$2\times\frac{\partial^2 S'^*}{\partial E'^2}(E/2)<0 \qquad \text{(III.24)}$$

(E' et S'^* étant l'énergie et l'entropie microcanonique de l'une des deux moitiés, $\mathcal{S}'$ par exemple). Cette inégalité est une *condition nécessaire de stabilité* pour le système $\mathcal{S}$, puisqu'elle exprime que la répartition « naturelle » de l'énergie entre les deux moitiés de ce système est effectivement la plus probable. On peut évidemment imaginer que $\mathcal{S}$ est divisé en 3, 4, ... sous-systèmes de même taille. On peut également envisager de mettre $\mathcal{S}$ en contact thermique avec un autre système qui lui soit identique (ou avec 3, 4, ... autres systèmes qui lui soient identiques). On obtient ainsi des inégalités analogues à la précédente, valables pour des systèmes de même nature que $\mathcal{S}$ mais de taille différente. Pour tous ces systèmes, et donc pour $\mathcal{S}$ lui-même, la dérivée seconde de l'entropie microcanonique par rapport à l'énergie est toujours négative[35] :

$$\frac{\partial^2 S^*}{\partial E^2}<0 \quad \text{pour toute E .} \qquad \text{(III.25)}$$

Cette condition de stabilité thermique, valable pour tout système macroscopique, peut être exprimée sous la forme suivante[36] : *la température microcanonique d'un système est une fonction croissante de son énergie* (les autres paramètres extérieurs $\{x\}$ étant maintenus constants)[37] :

$$\frac{\partial T^*}{\partial E}>0 . \qquad \text{(III.26)}$$

Lorsque nous aurons précisé, au paragraphe d, le sens dans lequel s'effectuent les échanges de chaleur, nous pourrons comprendre de façon plus intuitive comment la stabilité thermique de chacun des deux systèmes $\mathcal{S}_1$ et $\mathcal{S}_2$, exprimée sous la forme (III.26), suffit à assurer la stabilité de leur équilibre mutuel lorsqu'ils sont mis en contact thermique, c'est-à-dire la validité de la condition (III.22) (voir note 41).

d. Sens de l'échange de chaleur

La notion intuitive de température implique également que, si l'on met en contact thermique deux corps ayant initialement des températures différentes, la chaleur

---

35. On peut aussi raisonner de la façon suivante. L'entropie microcanonique S^* et l'énergie E sont toutes deux de l'ordre de N, nombre total de particules du système (cf. paragraphe I.C.1 et complément II.A), de sorte que $\partial^2 S^*/\partial E^2$ est d'ordre $1/N$. Si donc on met le système considéré en contact thermique avec un autre système beaucoup plus gros que lui, le terme de (III.23) correspondant à ce dernier sera négligeable, et (III.23) se réduira à (III-25) (*attention :* c'est la contribution du gros système qui est négligeable devant celle du plus petit, et non l'inverse).

36. D'après la définition (III.2) de la température microcanonique,

$$\frac{\partial^2 S^*}{\partial E^2}=\frac{\partial}{\partial E}\left(\frac{1}{T^*}\right)=-\frac{1}{T^{*2}}\frac{\partial T^*}{\partial E} .$$

Le signe de $\partial T^*/\partial E$ est donc l'opposé de celui de $\partial^2 S^*/\partial E^2$.

37. La dérivée $\partial T^*/\partial E$ est l'inverse de la *capacité calorifique à volume constant* (III, § II.B.1.b). La stabilité thermique d'un système exige donc que sa capacité calorifique soit *positive.*

passe de celui qui a la température la plus élevée vers celui qui a la température la plus basse. Voyons si la grandeur T^* définie plus haut possède une telle propriété.

Portons sur un même graphique (fig. 4) la température T_1^* du système $\mathcal{S}_1$ en fonction de l'énergie E_1 de ce système, et la température T_2^* du système $\mathcal{S}_2$, également en fonction de E_1. D'après le résultat (III.26), la première courbe a une pente constamment positive, la seconde une pente constamment négative[38]. On en déduit en premier lieu que l'équation d'équilibre (III.21) admet une solution unique, puisque les deux courbes se coupent en un seul point.

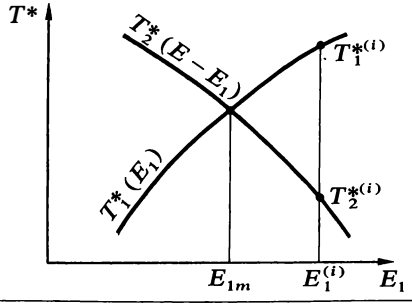


FIGURE 4
Résolution graphique de la condition d'équilibre (III.21) et détermination du sens de l'échange de chaleur entre deux sous-systèmes.

Plaçons ensuite sur le graphique la valeur initiale $E_1^{(i)}$ de l'énergie du système $\mathcal{S}_1$ avant l'établissement du contact thermique. Il est clair que si $E_1^{(i)}$ est supérieure à E_{1m} la température initiale $T_1^{*(i)}$ de $\mathcal{S}_1$ est supérieure à celle $T_2^{*(i)}$ du système $\mathcal{S}_2$, et que c'est le contraire si $E_1^{(i)}$ est inférieure à E_{1m} :

$$E_{1m} - E_1^{(i)} < 0 \iff T_1^{*(i)} > T_2^{*(i)} ;$$
$$E_{1m} - E_1^{(i)} > 0 \iff T_1^{*(i)} < T_2^{*(i)} . \qquad \text{(III.27)}$$

Comme l'énergie échangée entre les deux systèmes l'est seulement sous forme de chaleur, on constate effectivement que *la chaleur est cédée*[39] *par le système dont la température microcanonique est initialement*[40] *la plus élevée à celui dont la température est la plus basse;* on dit couramment que la chaleur passe du corps « le plus chaud » au corps « le plus froid »[41]. La figure 4 montre également que *la température finale* commune aux deux systèmes *est toujours intermédiaire entre leurs températures initiales.*

Attention : c'est la *différence de température* entre les deux systèmes, et *non la quantité d'énergie* contenue dans chacun d'eux, qui fixe le sens de l'échange de

38. T_2^* est fonction croissante de $E_2 = E - E_1$, donc fonction décroissante de E_1. La position de la courbe correspondante, sur le graphique de la figure 4, dépend bien sûr de l'énergie totale E.

39. Il se produit également entre $\mathcal{S}_1$ et $\mathcal{S}_2$ des échanges d'énergie aléatoires, correspondant aux fluctuations de la variable interne E_1 autour de E_{1m}, et dont le sens est lui aussi aléatoire. Ils sont toutefois quantitativement beaucoup moins importants que l'échange considéré ici, auquel ils se superposent (essentiellement comme sur la figure 3 du chapitre I).

40. *Attention :* entre le moment où le contact thermique est établi entre $\mathcal{S}_1$ et $\mathcal{S}_2$ et celui où l'équilibre final a été atteint, les deux systèmes n'ont pas de température définie, puisqu'ils sont hors d'équilibre; sur la figure 4, les points autres que ceux d'abscisses $E_1^{(i)}$ et E_{1m} représentent des états d'équilibre possibles, mais qui ne sont pas réalisés dans le processus considéré.

41. En combinant ce résultat avec la propriété (III.26) que doit vérifier la température de chacun des deux systèmes, on peut comprendre de façon simple et intuitive la stabilité de l'équilibre thermique entre les deux systèmes (§ c.β). En effet, lorsque l'énergie E_1 du système $\mathcal{S}_1$ prend, au cours de ses fluctuations, une valeur légèrement supérieure à E_{1m}, la température T_1^* a tendance

.../...

chaleur; ce n'est pas nécessairement le système qui a l'énergie la plus grande qui cède de l'énergie à l'autre.

B. Généralisation : échange de volume et de particules

1. DÉFINITIONS : PRESSION ET POTENTIEL CHIMIQUE MICROCANONIQUES

La définition (III.2) de la température fait intervenir la dérivée de l'entropie microcanonique S^* du système par rapport à son énergie E. Or S^* est également fonction des autres paramètres extérieurs $\{x\}$ qui déterminent l'état macroscopique du système lorsqu'il est isolé. Pour un fluide pur, par exemple, ce sont le volume V et le nombre de molécules N; pour un système quelconque, les paramètres extérieurs comprennent aussi le plus souvent le volume V et au moins un nombre de particules N :

$$S^* = S^*(E,\ V,\ N,\ \ldots)\ . \tag{III.28}$$

On introduit alors la *pression* p^* et le *potentiel chimique* μ^* microcanoniques par les définitions suivantes[42], [43] :

$$\frac{p^*}{T^*} \widehat{=} \frac{\partial S^*}{\partial V}\ , \tag{III.29.a}$$

$$\frac{\mu^*}{T^*} \widehat{=} -\frac{\partial S^*}{\partial N}\ . \tag{III.29.b}$$

à croître d'après (III.26); corrélativement, l'énergie E_2 du système $\mathcal{S}_2$ a diminué (puisque $E = E_1 + E_2$ reste constante), ce qui fait de même décroître T_2^* :

$$E_1 > E_{1m} \implies T_1^* > T_2^*$$

(cf. figure 4). Dans cette situation, de la chaleur passe de $\mathcal{S}_1$ à $\mathcal{S}_2$, ce qui ramène E_1 et E_2 vers leur valeur la plus probable. Bien entendu, les phénomènes inverses se produisent lorsque E_1 devient inférieure à E_{1m}, mais la conclusion reste la même. C'est ainsi que la propriété (III.26), expression de la stabilité thermique pour chacun des deux systèmes, entraîne la stabilité de leur équilibre thermique mutuel.

42. Le nombre de particules N est traité comme une variable continue, car ses accroissements à l'échelle macroscopique sont toujours très grands devant 1.

Si le système est constitué d'un mélange de plusieurs types de particules, son entropie microcanonique $S^*(E, V, N_1, N_2, \ldots)$ est fonction des nombres N_1, N_2, ... de particules de chaque type et l'on définit un potentiel chimique μ_i^* pour chacun d'eux :

$$\frac{\mu_i^*}{T^*} \widehat{=} -\frac{\partial S^*}{\partial N_i}\ ; \quad i = 1,\ 2,\ \ldots\ .$$

43. Les formules (III.2) et (III.29.b) sont des définitions au plein sens du terme : c'est d'elles que nous déduirons toutes les propriétés de la température et du potentiel chimique (nous avons commencé à le faire pour la température au paragraphe A.2) La situation est moins simple pour la pression p^*. Il faudra en effet montrer que la grandeur définie par (III.29.a) s'identifie à la pression définie par ailleurs en mécanique, c'est-à-dire qu'elle correspond effectivement aux forces exercées par un gaz, par exemple, sur les parois du récipient qui l'enferme. Nous examinerons ce problème au complément II.D (§ I.3).

On comprend alors l'importance physique du fait que le nombre d'états accessibles, et donc l'entropie microcanonique, d'un système tel qu'un gaz dépendent seulement de son volume total V et pas de la forme du récipient : en dérivant S^* par rapport aux paramètres géométriques qui caractérisent l'état macroscopique du système à l'équilibre, on obtient seulement une pression; par conséquent, les forces exercées par le système sur les parois qui le maintiennent sont partout normales à ces parois, et la force par unité de surface est partout la même.

Comme l'entropie et la température, la pression et le potentiel chimique sont des *grandeurs à caractère statistique,* qui n'ont de sens que pour des systèmes macroscopiques. Les dimensions physiques des paramètres p^* et μ^* se déduisent immédiatement de (III.29) et (III.2) : $1/T^*$ étant homogène à une entropie divisée par une énergie, p^* a bien les dimensions d'une énergie divisée par un volume, c'est-à-dire d'une force par unité de surface; quant à μ^*, c'est une énergie divisée par un nombre de particules. L'entropie microcanonique S^* étant dans tous les cas une fonction croissante du volume[44], *la pression est toujours positive. Le potentiel chimique,* quant à lui, *peut suivant les cas avoir une valeur positive ou négative*[45].

2. ÉTABLISSEMENT DU CONTACT ENTRE DEUX SYSTÈMES

Reprenons le système isolé $\mathcal{S}$, divisé en deux sous-systèmes $\mathcal{S}_1$ et $\mathcal{S}_2$, dans l'état initial où $\mathcal{S}_1$ et $\mathcal{S}_2$ sont séparément isolés et à l'équilibre (fig. 3). L'entropie microcanonique $S^{*(i)}$ du système global $\mathcal{S}$ dans cet état initial est simplement la somme des entropies microcanoniques de $\mathcal{S}_1$ et $\mathcal{S}_2$ (§ I.C.2) :

$$S^{*(i)} = S_1^*(E_1^{(i)}, V_1^{(i)}, N_1^{(i)}, \ldots) + S_2^*(E_2^{(i)}, V_2^{(i)}, N_2^{(i)}, \ldots). \qquad \text{(III.30)}$$

Les volumes $V_1^{(i)}$ et $V_2^{(i)}$ de $\mathcal{S}_1$ et $\mathcal{S}_2$, ainsi que leurs nombres de particules $N_1^{(i)}$ et $N_2^{(i)}$, sont des paramètres extérieurs comme leurs énergies $E_1^{(i)}$ et $E_2^{(i)}$.

Supposons que, comme au paragraphe A.2.a pour l'énergie, nous relâchions les contraintes assurant que V_1, V_2, N_1 et N_2 sont fixés, en maintenant cependant le système global $\mathcal{S}$ isolé : pour les volumes V_1 et V_2, il suffit que la cloison intermédiaire de la figure 3 puisse se déplacer ou se déformer; pour les nombres de particules N_1 et N_2, on rend cette cloison perméable aux molécules, qui peuvent donc passer d'un compartiment à l'autre. Cependant, le volume total V et le nombre total N de particules du système $\mathcal{S}$ restent fixés :

$$V_1 + V_2 = V \text{ fixé} ; \qquad \text{(III.31.a)}$$

$$N_1 + N_2 = N \text{ fixé} . \qquad \text{(III.31.b)}$$

Après avoir ainsi relâché les trois contraintes fixant E_1 (et E_2), V_1 (et V_2) et N_1 (et N_2), nous laissons le système $\mathcal{S}$ atteindre un nouvel état d'équilibre, dans lequel E_1,

---

44. Si $\partial S^*/\partial V$ était négative, le système ne serait pas stable. Il pourrait en effet augmenter son entropie en diminuant de volume, ce qu'il ferait donc spontanément (cf. § II.B) : le récipient qui le contient fixerait seulement la valeur maximale permise au volume, et le système se contracterait tant qu'il n'aurait pas atteint un état où sa pression redevienne positive.

45. Le potentiel chimique μ^* est souvent négatif comme le suggère un examen superficiel de la définition (III.29.b). Mais il faut tenir compte du fait que la dérivée $\partial S^*/\partial N$ doit être prise *à énergie constante* (et volume constant). Prenons par exemple un gaz constitué de N molécules. Rajoutons une molécule de même nature dans le récipient de volume V fixe qui le contient. Si cette particule a une énergie cinétique nulle lorsque nous l'ajoutons et si ses interactions avec les autres molécules sont négligeables, l'énergie E du système n'aura pas varié. Par contre, son entropie S^* aura augmenté, lorsque la nouvelle molécule se sera mise en équilibre avec les autres, par augmentation du nombre d'états accessibles (le nombre de manières possibles de partager l'énergie totale entre les $(N+1)$ molécules est plus grand qu'auparavant). Le potentiel chimique μ^* est dans ce cas (gaz parfait) *négatif.* Mais l'argument suppose qu'on puisse rajouter une particule qui n'apporte pas d'énergie. Ceci n'est pas toujours possible. Par exemple, si les *interactions entre particules* sont *répulsives,* la particule supplémentaire augmentera forcément l'énergie du système en même temps que son entropie S^*; il faudra donc retirer de l'énergie au système pour revenir à la valeur initiale E, et ceci diminuera S^*; cette diminution peut être supérieure à l'augmentation provoquée par l'arrivée de la nouvelle particule. Ainsi, lorsque les interactions sont suffisamment répulsives, le potentiel chimique est *positif.*

V_1 et N_1 sont devenus des variables internes. Leur distribution statistique sera comme toujours une gaussienne (ici à trois dimensions) dont la largeur sera d'autant plus faible que le système $\mathcal{S}$ est plus grand. Nous voulons d'abord déterminer les *conditions d'équilibre entre les deux sous-systèmes* $\mathcal{S}_1$ et $\mathcal{S}_2$, c'est-à-dire les équations fixant les valeurs les plus probables E_{1m}, V_{1m} et N_{1m} des trois variables internes dans l'état d'équilibre final. Nous cherchons également à préciser *dans quel sens se font les échanges* lorsque $\mathcal{S}_1$ et $\mathcal{S}_2$ ne sont pas en équilibre.

3. ÉQUILIBRE ENTRE LES DEUX SYSTÈMES

a. Conditions d'équilibre

Comme pour les échanges de chaleur, nous supposons que *le couplage entre les deux sous-systèmes est faible* (cf. complément II.A). Ceci permet de calculer simplement la probabilité pour que, dans l'état d'équilibre final, l'énergie de $\mathcal{S}_1$ soit E_1, son volume V_1 et son nombre de particules N_1 (l'énergie totale, le volume total et le nombre total de particules étant fixés à E, V et N)[46]. Cette probabilité est en effet proportionnelle au nombre d'états accessibles du système global réalisant les conditions énoncées pour le système $\mathcal{S}_1$ (cf. § II.A), et le décompte de ces états se fait facilement (cf. § A.2.c) : leur nombre $\omega(E, V, N; E_1, V_1, N_1)$ est, à cause du couplage faible, égal au produit

$$\omega(E, V, N; E_1, V_1, N_1) = \Omega_1(E_1, V_1, N_1).\Omega_2(E_2 = E - E_1, V_2 = V - V_1, N_2 = N - N_1) . \qquad \text{(III.32)}$$

L'*entropie partielle* associée,

$$s^*(E, V, N; E_1, V_1, N_1) \widehat{=} k \ln \omega(E, V, N; E_1, V_1, N_1) , \qquad \text{(III.33)}$$

vaut donc

$$s^*(E, V, N; E_1, V_1, N_1) = S_1^*(E_1, V_1, N_1) + S_2^*(E_2 = E - E_1, V_2 = V - V_1, N_2 = N - N_1) , \qquad \text{(III.34)}$$

où $S_j^*(E_j, V_j, N_j)$ (avec $j = 1$ ou 2) est l'entropie microcanonique qu'aurait le système $\mathcal{S}_j$ s'il était isolé, son énergie, son volume et son nombre de particules étant fixés aux valeurs E_j, V_j et N_j.

La *situation la plus probable* correspond au maximum de l'entropie partielle précédente, à E, V et N fixés. Les valeurs E_{1m}, V_{1m} et N_{1m} qu'y prennent les trois variables internes sont donc données par les conditions (la dépendance par rapport à E, V et N étant sous-entendue)

$$\frac{\partial s^*}{\partial E_1}(E_{1m}, V_{1m}, N_{1m}) = 0 ,$$

$$\frac{\partial s^*}{\partial V_1}(E_{1m}, V_{1m}, N_{1m}) = 0 ,$$

$$\frac{\partial s^*}{\partial N_1}(E_{1m}, V_{1m}, N_{1m}) = 0 . \qquad \text{(III.35)}$$

46. Il s'agit en réalité de la probabilité pour que l'énergie de $\mathcal{S}_1$ soit comprise entre E_1 et $E_1 + \delta E_1$, son volume entre V_1 et $V_1 + \delta V_1$ et son nombre de molécules entre N_1 et $N_1 + \delta N_1$. Comme les incertitudes δE_1, δV_1 et δN_1 ne vont jouer aucun rôle dans la suite (cf. notes 5 et 10), nous les « oublions » dès le départ.

Compte tenu de l'expression (III.34) de s^* et des définitions (III.2) et (III.29), ces conditions d'équilibre s'écrivent (cf. note 31)

$$\frac{1}{T_1^*}=\frac{1}{T_2^*}\,,\quad \frac{p_1^*}{T_1^*}=\frac{p_2^*}{T_2^*}\,,\quad \frac{\mu_1^*}{T_1^*}=\frac{\mu_2^*}{T_2^*}\,, \tag{III.36}$$

c'est-à-dire

$$T_1^*=T_2^*\,,\quad p_1^*=p_2^*\,,\quad \mu_1^*=\mu_2^*\,. \tag{III.37}$$

Rappelons que T_1^*, p_1^* et μ_1^* sont des fonctions de E_1, V_1, N_1 et T_2^*, p_2^*, μ_2^* des fonctions de $E_2=E-E_1$, $V_2=V-V_1$ et $N_2=N-N_1$; les égalités (III.37) sont ainsi des *équations en* E_1, V_1, N_1 (E, V, N étant fixés), dont la solution[47] donne les valeurs les plus probables E_{1m}, V_{1m}, N_{1m}.

Donc, *la répartition la plus probable de l'énergie, du volume et du nombre de particules* entre les deux sous-systèmes couplés *correspond à l'égalité des températures, des pressions et des potentiels chimiques* microcanoniques des deux sous-systèmes.

b. Sens des échanges hors d'équilibre

Nous supposons que, dans l'état initial où les systèmes $\mathcal{S}_1$ et $\mathcal{S}_2$ sont complètement isolés, les valeurs $E_1^{(i)}$, $V_1^{(i)}$ et $N_1^{(i)}$ de l'énergie, du volume et du nombre de particules de $\mathcal{S}_1$ sont différentes des valeurs E_{1m}, V_{1m}, N_{1m} solutions des conditions d'équilibre (III.37). Comme au paragraphe A.2.d, nous cherchons le signe de $(E_{1m}-E_1^{(i)})$, $(V_{1m}-V_1^{(i)})$ et $(N_{1m}-N_1^{(i)})$.

La généralisation du raisonnement développé au paragraphe A.2.d pose quelques problèmes délicats, dus à la présence simultanée de trois variables[48]. Nous nous contenterons ici d'arguments plus simples, quoique de validité plus restreinte : nous allons choisir un état initial qui diffère suffisamment peu de l'état d'équilibre final pour que les différences puissent être traitées au *premier ordre.*

Dans l'état d'équilibre final, la probabilité $P_f(E_1^{(i)}, V_1^{(i)}, N_1^{(i)})$ pour que l'énergie, le volume et le nombre de particules reprennent les valeurs initiales $E_1^{(i)}$, $V_1^{(i)}$ et $N_1^{(i)}$ est de la forme

$$P_f(E_1^{(i)}, V_1^{(i)}, N_1^{(i)}) \propto \exp\left[\frac{1}{k}\, s^*(E_1^{(i)}, V_1^{(i)}, N_1^{(i)})\right], \tag{III.38}$$

où l'entropie microcanonique partielle s^* est égale à l'entropie microcanonique du système $\mathcal{S}$ dans l'état initial :

$$\begin{aligned} s^*(E_1^{(i)}, V_1^{(i)}, N_1^{(i)}) = {} & S_1^*(E_1^{(i)}, V_1^{(i)}, N_1^{(i)}) \\ & + S_2^*(E_2=E-E_1^{(i)}, V_2=V-V_1^{(i)}, N_2=N-N_1^{(i)})\,. \end{aligned} \tag{III.39}$$

---

47. En réalité, si l'état du système est défini par trois variables seulement (E, V et N), les trois équations (III.37) ne sont pas indépendantes et l'état d'équilibre reste indéterminé (voir exercice II.11). Par contre l'état final est parfaitement déterminé si le système est constitué de deux sortes de particules dont une seule peut traverser la paroi (paroi semi-perméable).

48. Ces trois variables sont indépendantes : pour un système isolé, on peut fixer séparément le volume, le nombre de particules et l'énergie totale. Pourtant, ces trois grandeurs ne jouent pas un rôle symétrique dans les processus considérés ici : si deux échantillons de gaz peuvent facilement échanger de l'énergie (il s'agit alors de chaleur) sans modification de leur volume et de leur nombre de particules, il est plus difficile qu'ils échangent du volume ou des particules sans échanger en même temps de l'énergie (dans une variation de volume, les forces de pression s'exerçant sur la cloison mobile effectuent du travail; d'autre part, lorsque des particules passent d'un système à l'autre, l'énergie cinétique et l'énergie potentielle totales de chacun des deux systèmes en sont modifiées).

Les valeurs les plus probables E_{1m}, V_{1m}, N_{1m} des variables internes sont celles qui rendent l'entropie microcanonique partielle maximale, de sorte que

$$\delta s^* \widehat{=} s^*(E_{1m}, V_{1m}, N_{1m}) - s^*(E_1^{(i)}, V_1^{(i)}, N_1^{(i)}) > 0 . \tag{III.40}$$

Au premier ordre dans les différences

$$\begin{aligned} \delta E_1 &\widehat{=} E_{1m} - E_1^{(i)} , \\ \delta V_1 &\widehat{=} V_{1m} - V_1^{(i)} , \\ \delta N_1 &\widehat{=} N_{1m} - N_1^{(i)} , \end{aligned} \tag{III.41}$$

l'accroissement δs^* de l'entropie partielle à partir de sa valeur (III.39) s'écrit

$$\delta s^* = \left[\frac{\partial S_1^*}{\partial E_1} - \frac{\partial S_2^*}{\partial E_2}\right] \delta E_1 + \left[\frac{\partial S_1^*}{\partial V_1} - \frac{\partial S_2^*}{\partial V_2}\right] \delta V_1 + \left[\frac{\partial S_1^*}{\partial N_1} - \frac{\partial S_2^*}{\partial N_2}\right] \delta N_1 , \tag{III.42}$$

où les dérivées sont à prendre pour les valeurs *initiales* $E_1^{(i)}$, $V_1^{(i)}$, $N_1^{(i)}$ des variables. Compte tenu des définitions (III.2) et (III.29), l'inégalité (III.40) prend la forme

$$\left[\frac{1}{T_1^{*(i)}} - \frac{1}{T_2^{*(i)}}\right] \delta E_1 + \left[\frac{p_1^{*(i)}}{T_1^{*(i)}} - \frac{p_2^{*(i)}}{T_2^{*(i)}}\right] \delta V_1 - \left[\frac{\mu_1^{*(i)}}{T_1^{*(i)}} - \frac{\mu_2^{*(i)}}{T_2^{*(i)}}\right] \delta N_1 > 0 . \tag{III.43}$$

Pour comprendre la signification de ce résultat, analysons deux *cas particuliers simples*. Supposons que l'on ait d'abord établi le contact thermique entre $\mathcal{S}_1$ et $\mathcal{S}_2$, de sorte que l'équilibre thermique soit déjà réalisé dans l'état initial :

$$T_1^{*(i)} = T_2^{*(i)} . \tag{III.44}$$

On permet ensuite aux deux sous-systèmes $\mathcal{S}_1$ et $\mathcal{S}_2$ d'échanger, outre de la chaleur, soit seulement du volume (cloison mobile ou déformable mais imperméable aux particules), soit seulement des particules (cloison fixe mais permettant le passage des molécules). L'inégalité (III.43) se simplifie alors en

$$[p_1^{*(i)} - p_2^{*(i)}] \delta V_1 > 0 \tag{III.45.a}$$

dans le premier cas, ou

$$[\mu_1^{*(i)} - \mu_2^{*(i)}] \delta N_1 < 0 \tag{III.45.b}$$

dans le second. On en déduit que, à l'équilibre thermique,

$$\left.\begin{aligned} p_1^{*(i)} > p_2^{*(i)} &\iff \delta V_1 > 0 \\ p_1^{*(i)} < p_2^{*(i)} &\iff \delta V_1 < 0 \end{aligned}\right\} \text{ à } N_1 \text{ et } N_2 \text{ constants,} \tag{III.46.a}$$

ou bien

$$\left.\begin{aligned} \mu_1^{*(i)} > \mu_2^{*(i)} &\iff \delta N_1 < 0 \\ \mu_1^{*(i)} < \mu_2^{*(i)} &\iff \delta N_1 > 0 \end{aligned}\right\} \text{ à } V_1 \text{ et } V_2 \text{ constants.} \tag{III.46.b}$$

La conclusion (III.46.a) est bien conforme à l'identification entre la grandeur p^* et la pression exercée par le système considéré sur les parois qui le contiennent (cf. note 43) : lorsque la cloison séparant les sous-systèmes $\mathcal{S}_1$ et $\mathcal{S}_2$ devient mobile ou déformable, *c'est le sous-système dont la pression est la plus forte qui augmente de volume.* Dans le cas d'échange de particules, *c'est le sous-système de potentiel chimique le plus élevé qui cède des particules à l'autre.*

Si les systèmes $\mathcal{S}_1$ et $\mathcal{S}_2$ n'ont pas initialement la même température et s'ils échangent à la fois de l'énergie, du volume et des particules (paroi diatherme, mobile et poreuse), ces échanges se font dans le sens qui accroît l'entropie microcanonique partielle (inégalité (III.43)). Mais il n'est pas nécessaire que chacun des termes du premier membre de (III.43) soit séparément positif : il peut se faire par exemple que ce soit le système dont la pression est initialement la plus faible qui augmente de volume, pourvu que l'inégalité (III.43) soit globalement vérifiée.

COMPLÉMENTS DU CHAPITRE II

II.A. Couplage entre systèmes de taille macroscopique	Facile et important du point de vue physique. La notion d'extensivité et d'intensivité sera souvent utilisée dans la suite.
II.B. Le cristal paramagnétique parfait. Températures négatives **II.C. Le gaz parfait classique : entropie de mélange et paradoxe de Gibbs**	Applications directes des notions introduites dans le chapitre à deux systèmes physiques simples.
II.D. Transformations thermodynamiques et variations d'entropie	Introduit la thermodynamique comme conséquence de la mécanique statistique (la démonstration sera poursuivie au complément III.L.). Très important du point de vue théorique mais assez difficile.
II.E. Mouvement macroscopique d'un système totalement isolé en équilibre statistique	Étudie les mouvements macroscopiques compatibles avec l'équilibre statistique et en déduit deux conséquences très importantes sur le plan théorique. Difficile. Nécessite une bonne connaissance de la mécanique classique.
II.F. Distribution et entropie microcanoniques en mécanique classique	Reprend, dans le cadre de la mécanique classique, les notions et résultats que le chapitre présente dans le cadre quantique. Assez facile si l'on sait raisonner dans l'espace des phases.

COMPLÉMENT II.A

Couplage entre systèmes de taille macroscopique

I. Portée des forces et intensité du couplage
1. Analyse d'un exemple simple : gaz de molécules non polaires
2. Discussion et généralisation
II. Grandeurs extensives et grandeurs intensives

Les raisonnements conduisant aux conditions d'équilibre entre deux sous-systèmes $\mathcal{S}_1$ et $\mathcal{S}_2$ d'un système isolé $\mathcal{S}$ (II, § III.A.2.c et § III.B.3.a) sont basés sur l'hypothèse d'un couplage faible entre les sous-systèmes : de façon plus précise, nous avons supposé que l'hamiltonien d'interaction H_{12} entre $\mathcal{S}_1$ et $\mathcal{S}_2$ est petit devant chacun de leurs hamiltoniens individuels H_1 et H_2 (II, § III.A.1.b.β). C'est cette hypothèse physique que nous voulons discuter ici.

I. PORTÉE DES FORCES ET INTENSITÉ DU COUPLAGE

1. ANALYSE D'UN EXEMPLE SIMPLE : GAZ DE MOLÉCULES NON POLAIRES

Imaginons l'expérience suivante. Deux récipients, de volumes V_1 et V_2, contiennent N_1 et N_2 molécules d'un même corps pur gazeux. Initialement, ces deux récipients sont suffisamment éloignés l'un de l'autre pour que toute interaction soit exclue entre les systèmes $\mathcal{S}_1$ et $\mathcal{S}_2$ constitués par les échantillons de gaz qu'ils contiennent; ceux-ci sont alors isolés, leurs énergies respectives valant E_1 et E_2. Il est clair que dans ces conditions l'énergie E du système $\mathcal{S}$ défini comme la réunion de $\mathcal{S}_1$ et $\mathcal{S}_2$ vaut simplement

$$E = E_1 + E_2 \,. \tag{A.1}$$

Rapprochons les deux récipients pour les amener au contact, comme sur la figure 3 du chapitre II. Ceci ne coûte aucun travail car les deux échantillons de gaz n'exercent l'un sur l'autre aucune force appréciable; l'énergie du système global $\mathcal{S}$ est encore donnée par la formule (A.1) lorsque les deux récipients ont une paroi commune.

En réalité, les interactions entre $\mathcal{S}_1$ et $\mathcal{S}_2$ ne sont pas strictement nulles : les deux systèmes exercent l'un sur l'autre des forces dues à la gravitation universelle et des forces d'origine électromagnétique. Les forces gravitationnelles sont tellement faibles, pour les masses de gaz considérées habituellement[1], que l'énergie potentielle correspondante est totalement négligeable devant chacune des énergies E_1 et E_2. Les forces électromagnétiques sont plus intenses, et leur portée est en principe infinie comme celle des forces gravitationnelles (potentiel en $1/r$). Mais l'existence de charges des deux signes change radicalement la situation : les charges des électrons et des noyaux constituant une molécule neutre compensent si bien leurs effets (on dit que ces charges opposées « s'écrantent » les unes les autres) que la molécule n'interagit de façon appréciable qu'avec ses plus proches voisines à l'échelle microscopique; les

1. Les forces de gravitation jouent par contre un rôle primordial en astrophysique. Les questions de couplage et d'équilibre entre sous-systèmes y sont dès lors moins simples que pour les systèmes considérés ici, et justifient une étude spécifique (voir complément VI.B, § II).

interactions résiduelles à des distances macroscopiques comme celles qui nous intéressent ici sont elles aussi négligeables[2]. Ainsi, *les forces intermoléculaires sont à courte portée.*

Les deux récipients sont maintenant au contact, c'est-à-dire qu'ils possèdent une paroi commune. Examinons les phénomènes énergétiques mis en jeu au niveau microscopique lorsqu'on modifie cette cloison de façon à permettre des échanges d'énergie entre $\mathcal{S}_1$ et $\mathcal{S}_2$ (II, § III.A.1.b), accompagnés éventuellement d'échanges de volume et de particules (II, § III.B.2). Si, comme nous venons de le souligner, la portée des forces intermoléculaires est d'ordre microscopique, les molécules qui se trouvent à une distance macroscopique de la cloison « ne s'aperçoivent pas » du changement, car elles n'interagissent qu'avec leur environnement immédiat et celui-ci n'est pas modifié; seules celles qui se situent tout près de la cloison sont affectées.

---

2. Les molécules d'un gaz ordinaire sont neutres. Nous supposons ici qu'elles sont de plus *« non polaires »*, c'est-à-dire qu'elles ne possèdent pas de moment dipolaire électrique permanent (c'est le cas des molécules monoatomiques comme celles des gaz rares, des molécules homopolaires comme O_2, N_2, H_2, ... et de molécules linéaires et symétriques comme CO_2, CH_2, ...). Elles exercent pourtant les unes sur les autres des forces d'origine électromagnétique, dites *« forces de van der Waals »* : le mouvement des électrons crée en réalité dans chaque molécule un moment dipolaire électrique fluctuant; bien que sa valeur moyenne soit nulle (pas de moment permanent), l'interaction entre deux molécules n'est pas nulle par suite des corrélations entre leurs dipôles fluctuants (voir C.D.L., complément C_{XI}). Cet effet produit une énergie potentielle d'interaction qui décroît en $1/r^6$ lorsque la distance r entre les deux molécules croît au-delà de quelques angströms.

De façon un peu plus précise, nous avons indiqué dans le complément I.A (§ III.3) que cette énergie potentielle d'interaction est de la forme

$$u(r) = -u_0 f\left(\frac{r}{r_0}\right), \tag{1}$$

où u_0 est de l'ordre de quelques centièmes d'électron-volt, r_0 de l'ordre de quelques angströms, et f est assez bien représentée par l'expression simple

$$f(x) = \frac{2}{x^6} - \frac{1}{x^{12}}. \tag{2}$$

C'est le terme en $(r_0/r)^6$ qui nous intéresse ici, car il domine dès que r est supérieur à r_0.

Choisissons alors une molécule déterminée, située au beau milieu du récipient, c'est-à-dire à une distance macroscopique de toute paroi. L'énergie d'interaction entre cette molécule et le reste du gaz est donnée approximativement par (complément I.A, formule (A.49))

$$u_1 \simeq -u_0 \int_{r_0}^{+\infty} \frac{N}{V} \cdot 4\pi r^2 \, \mathrm{d}r \, f\left(\frac{r}{r_0}\right), \tag{3}$$

c'est-à-dire

$$u_1 \simeq -8\pi N u_0 \frac{r_0^3}{V} \int_1^{+\infty} \frac{\mathrm{d}x}{x^4} = -\frac{8\pi}{3} N u_0 \frac{r_0^3}{V}. \tag{4}$$

L'intégration ne s'étend pas en toute rigueur jusqu'à l'infini, mais jusqu'aux parois du récipient. Cependant, la décroissance rapide de la fonction à intégrer rend la différence inappréciable. Plus précisément, si l'on impose une borne supérieure $x_M r_0$ à l'intégrale de la formule (3), le résultat (4) est remplacé par

$$u_1(x_M) \simeq -\frac{8\pi}{3} N u_0 \frac{r_0^3}{V}\left[1 - \frac{1}{x_M^3}\right]. \tag{5}$$

On constate que $u_1(x_M)$ atteint pratiquement la valeur totale (4) dès que x_M est de l'ordre de 4 ou 5. Ainsi, une molécule donnée n'interagit de façon appréciable qu'avec ses plus proches voisines, c'est-à-dire celles qui se trouvent à une distance inférieure à $d \simeq 5r_0 \simeq 15$ à 50 Å.

Le nombre de ces dernières étant très petit devant le nombre total de molécules de chacun des deux systèmes[3], la modification de l'énergie du système global est, elle aussi, très faible.

On peut traduire les raisonnements précédents en termes d'hamiltoniens : dans la mesure où les forces gravitationnelles sont négligeables, l'hamiltonien du système global est

$$H = H_1 + H_2 \,, \tag{A.2}$$

où H_j est l'hamiltonien du système $\mathcal{S}_j$, tant que $\mathcal{S}_1$ et $\mathcal{S}_2$ sont séparés par une cloison fixe, adiabatique et imperméable aux molécules; si $\mathcal{S}_1$ et $\mathcal{S}_2$ peuvent échanger de l'énergie, du volume et des molécules, cet hamiltonien devient

$$H = H_1 + H_2 + H_{12} \,, \tag{A.3}$$

mais *le terme de couplage H_{12} est petit devant chacun des deux autres :*

$$H_{12} \ll H_1 \,,\ H_2 \,. \tag{A.4}$$

2. DISCUSSION ET GÉNÉRALISATION

Le résultat précédent se généralise à d'autres systèmes que des échantillons de gaz. Mais les raisonnements du paragraphe 1 montrent clairement que ce résultat n'est valable que si *deux conditions* sont simultanément remplies. Il faut en premier lieu que *les forces d'interaction* entre constituants microscopiques *soient à courte portée;* il importe de noter à ce sujet que *ce n'est pas nécessairement l'intensité de ces forces qui doit être faible* pour que le couplage entre systèmes macroscopiques le soit, mais leur portée (cf. notes 2 et 3). Il faut en outre que les systèmes considérés soient macroscopiques dans toutes leurs dimensions, de façon que l'on puisse *négliger les effets de surface;* il est clair par exemple que les propriétés de films minces dont l'épaisseur est inférieure au dixième de micron[4] seront différentes de celles du même matériau lorsqu'il se présente en cubes ou en boules : dans un film mince, aucune particule n'est suffisamment loin de la surface pour ne pas en ressentir la présence.

Par ailleurs, il ne faudrait pas croire que les forces d'interaction sont toujours à courte portée. Elles ne le sont pas par exemple pour les systèmes de conducteurs macroscopiques chargés, en équilibre électrostatique, car les forces qui s'exercent entre eux ne sont pas en général écrantées. Lorsqu'on approche l'un de l'autre deux tels systèmes $\mathcal{S}_1$ et $\mathcal{S}_2$, l'état de chacun d'eux est profondément perturbé : le champ électrique produit par l'un des systèmes au niveau de l'autre provoque une redistribution des charges à la surface des conducteurs. Corrélativement, l'énergie du système $\mathcal{S}$ réunissant $\mathcal{S}_1$ et $\mathcal{S}_2$ n'est pas simplement la somme de deux termes dont chacun est associé à un seul des deux systèmes, car l'énergie d'interaction est du même ordre de grandeur que les deux énergies initiales. Lors de l'étude thermodynamique des systèmes

---

3. En première approximation, les molécules sont réparties uniformément dans le volume de chaque récipient. La proportion de celles qui se trouvent à une distance des parois inférieure à une longueur donnée d est donc $\simeq Sd/V$, où V est le volume du récipient et S l'aire des parois. Si L caractérise les dimensions linéaires du récipient (rayon s'il est sphérique, arête s'il est cubique, ...), cette proportion est donc de l'ordre de d/L, c'est-à-dire très faible si d est d'ordre microscopique et L d'ordre macroscopique. De façon plus précise, L varie comme $V^{1/3}$, c'est-à-dire comme $N^{1/3}$ pour une densité de molécules N/V fixée; le terme négligé est donc d'ordre $N^{-1/3}$ par rapport à l'énergie potentielle de l'ensemble des molécules.

4. On sait réaliser à l'heure actuelle des films dont l'épaisseur atteint les dimensions des molécules qui les constituent : membranes lipidiques de 70 Å d'épaisseur (ce qui correspond à une double couche de molécules de lipides orientées perpendiculairement à la surface), ou dépôts monomoléculaires de divers corps sur un substrat qui assure leur rigidité.

électrostatiques et magnétiques, on est obligé de tenir compte des énergies d'interaction et du couplage entre sous-systèmes macroscopiques.

On peut alors se demander comment se comportent les systèmes constitués, non pas de molécules neutres, mais (au moins en partie) d'ions chargés, comme les plasmas, les électrolytes ou les solides ioniques. La question se pose également pour les fluides ou les solides polaires, c'est-à-dire dont les molécules sont neutres mais possèdent un moment dipolaire électrique permanent (molécules hétéropolaires comme HCl ou coudées comme H_2O, dans lesquelles le barycentre des charges positives ne coïncide pas avec celui des charges négatives)[5]. La réponse à ces questions ne fait aucun doute du point de vue empirique : deux grains de sel ou deux récipients pleins d'eau n'exercent pas plus de forces macroscopiques l'un sur l'autre que les deux échantillons de gaz du paragraphe 1. Le problème est moins simple du point de vue théorique; il est résolu pour les systèmes ioniques (références 1 et 2)[6], mais pas complètement encore pour les corps polaires[7].

---

5. L'interaction dipôle-dipôle (l'un des dipôles crée au niveau de l'autre un champ électrique qui agit sur le second dipôle) donne naissance à une énergie potentielle qui décroît seulement comme $1/r^3$ lorsque la distance r entre les deux dipôles croît. On ne peut donc pas reprendre dans ce cas l'argument développé dans la deuxième partie de la note 2 : l'intégrale de $4\pi r^2\, dr/r^3$ ne converge pas à l'infini.

6. Le mécanisme de base est toujours l'écrantage : une particule chargée s'entoure préférentiellement de particules de charge opposée à la sienne, qui tendent à compenser ses effets à grande distance.

Pour comprendre un peu mieux comment joue ce mécanisme, prenons l'exemple d'un cristal ionique, le chlorure de sodium (fig. 1); il est constitué d'ions Cl^- et Na^+ alternant régulièrement aux sommets d'un réseau cubique (dans la classification cristallographique, ce réseau est appelé « cubique à faces centrées »).

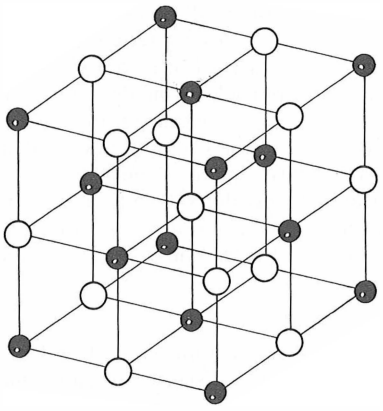


FIGURE 1
Structure du cristal ionique de chlorure de sodium : les centres des ions Cl^- (sphères de rayon $\simeq 1{,}81$ Å) sont repérés par des boules noires, ceux des ions Na^+ (rayon $\simeq 0{,}95$ Å) par des boules blanches. La distance entre ions voisins est $a = 2{,}82$ Å.

Choisissons un ion particulier, de charge q_0, et évaluons son énergie d'interaction avec les autres. Si nous prenons l'origine des coordonnées sur cet ion et les axes selon les arêtes du réseau cubique, les coordonnées d'un ion quelconque sont

$$n_1 a, \quad n_2 a, \quad n_3 a\,, \tag{1}$$

où a est la distance entre deux ions voisins et n_1, n_2, n_3 des entiers (positifs, négatifs ou nuls); la charge de l'ion repéré par (n_1, n_2, n_3) est simplement

$$q_{n_1, n_2, n_3} = q_0(-1)^{n_1+n_2+n_3}\,. \tag{2}$$

Calculons alors l'énergie d'interaction $u_1(n)$ entre l'ion choisi et l'ensemble des ions situés sur

.../...

II. GRANDEURS EXTENSIVES ET GRANDEURS INTENSIVES

Les systèmes que nous étudierons dans la suite possèderont, sauf exceptions que nous signalerons explicitement, la propriété fondamentale de couplage macroscopique faible décrite ci-dessus.

Considérons un tel système macroscopique, constitué d'un très grand nombre N de particules; pour fixer les idées, prenons $N \simeq 10^{23}$. Nous supposons que ce système vérifie les deux conditions énoncées au paragraphe I.2 : forces à faible portée, effets

---

la surface d'un cube dont les faces sont à la distance na de cet ion central, avec n entier positif fixé. Ce cube ayant 8 sommets, 12 arêtes et 6 faces, on trouve facilement

$$u_1(n) = \frac{q_0^2}{4\pi\varepsilon_0}\frac{1}{a}\left\{8\frac{(-1)^{3n}}{\sqrt{3n^2}} + 12\left[\frac{(-1)^{2n}}{\sqrt{2n^2}} + 2\sum_{p=1}^{n-1}\frac{(-1)^{2n+p}}{\sqrt{2n^2+p^2}}\right]\right.$$
$$\left. + 6\left[\frac{(-1)^n}{\sqrt{n^2}} + 4\sum_{p=1}^{n-1}\frac{(-1)^{n+p}}{\sqrt{n^2+p^2}} + 4\sum_{p=1}^{n-1}\sum_{q=1}^{n-1}\frac{(-1)^{n+p+q}}{\sqrt{n^2+p^2+q^2}}\right]\right\}. \quad (3)$$

Cette expression se prête à un calcul numérique relativement simple, dont les résultats sont indiqués dans le tableau I.

L'énergie d'interaction totale d'un ion avec l'ensemble du cristal vaut (référence 3)

$$u_1 = \sum_{n=1}^{\infty} u_1(n) = -1{,}747\,\frac{q_0^2}{4\pi\varepsilon_0}\frac{1}{a}. \quad (4)$$

On constate que cette limite est approchée par excès (en valeur absolue) si l'on tronque la série à une valeur impaire de n, et par défaut si c'est à une valeur paire. La convergence de cette série alternée n'est pas très rapide, mais elle s'effectue quand même sur une distance d'ordre de grandeur microscopique : la somme des 25 premiers termes donne à mieux que 2 % près la somme totale (4), alors qu'elle correspond à des distances de l'ordre de $25a$, c'est-à-dire 70 Å. Ceci explique que le couplage entre cristaux ioniques macroscopiques est faible.

TABLEAU I **Premiers termes et premières sommes partielles de la série (4) de la note 6.**

n	$u_1(n)\Big/\frac{q_0^2}{4\pi\varepsilon_0 a}$	$\sum_{p=1}^{n} u_1(p)\Big/\frac{q_0^2}{4\pi\varepsilon_0 a}$
1	−2,133 5	−2,133 5
2	+0,616 9	−1,516 6
3	−0,395 9	−1,912 5
4	+0,293 2	−1,619 3
5	−0,233 3	−1,852 6
10	+0,115 8	−1,692 6
15	−0,077 1	−1,784 9
20	+0,057 8	−1,719 5
25	−0,046 2	−1,770 3

7. Dans un fluide polaire, par exemple, deux molécules voisines ont tendance à s'orienter de façon à minimiser leur énergie potentielle; cette tendance est contrecarrée par l'agitation thermique et les collisions, qui tendent au contraire à distribuer les dipôles au hasard. Lorsqu'on examine le problème plus en détail, on trouve que la *corrélation* entre deux dipôles, qui mesure l'influence de l'un sur l'orientation de l'autre, décroît en $1/r^3$ comme l'énergie potentielle entre deux dipôles fixes (cf. note 5) et *dépend de la forme macroscopique de l'échantillon :* ces corrélations ne sont pas les mêmes pour un récipient sphérique que pour un récipient allongé. C'est là l'origine des difficultés : le résultat lorsque le volume devient très grand n'est pas le même suivant que la limite est prise à partir d'une sphère dont on fait croître indéfiniment le rayon ou à partir d'un parallélépipède dont les arêtes tendent successivement vers l'infini (voir référence 2).

de surface négligeables. Imaginons que nous le découpions par la pensée en 10^5 sous-systèmes identiques, comportant donc chacun $N' \simeq 10^{18}$ particules. Bien que nombreux et petits par rapport au système initial, *ces sous-systèmes sont encore de taille macroscopique.* Mieux, *les fluctuations y sont négligeables :* pour une variable interne quelconque y' relative à l'un de ces sous-systèmes, les fluctuations sont d'ordre $1/\sqrt{N'} \simeq 10^{-9}$ en valeur relative (II, § II.A.2); tout se passe donc en fait comme si la variable y' avait une valeur unique (sa valeur la plus probable y'_m).

L'énergie d'interaction entre ces sous-systèmes est négligeable devant leur énergie propre[8]. En conséquence, l'énergie totale du système est simplement la *somme des énergies des sous-systèmes.* De même, l'état microscopique d'un sous-système donné n'est influencé que de façon négligeable par les autres : ces sous-systèmes macroscopiques sont *statistiquement indépendants* les uns des autres, de sorte que l'entropie totale du système est la *somme des entropies des sous-systèmes* (voir exercice II.19). Comme les fluctuations sont négligeables, chaque sous-système a une énergie déterminée (son énergie la plus probable) et une entropie déterminée (son entropie microcanonique prise pour la valeur la plus probable de son énergie, de son volume et de son nombre de particules). L'énergie et l'entropie sont ainsi réparties en quelque sorte localement, c'est-à-dire entre les sous-systèmes macroscopiques en lesquels on peut décomposer le système initial.

De façon plus générale, cette propriété d'additivité sera valable pour toute une classe de grandeurs physiques, qui seront dites *extensives.* Il est clair par exemple que le nombre de particules et le volume sont des grandeurs extensives; nous en rencontrerons d'autres dans la suite de ce cours.

Mais nous avons montré dans la partie III du chapitre II que la *température* (§ III.A.2), la *pression* et le *potentiel chimique* (§ III.B.3) se comportent de façon totalement différente des grandeurs précédentes : le couplage entre les sous-systèmes, bien que faible, a pour effet d'*égaliser les valeurs* qu'elles prennent *dans les divers sous-systèmes.* De telles grandeurs sont dites *intensives.* Ici aussi, la valeur de chacune de ces grandeurs est unique au niveau de chaque sous-système, d'après leur définition à partir des dérivées de l'entropie.

La différence entre grandeurs extensives et grandeurs intensives peut être caractérisée de la façon suivante. Supposons que, toutes choses restant égales par ailleurs[9], *la taille du système global soit multipliée par un nombre* positif quelconque[10] λ. Ceci se traduit par le fait que *le nombre de sous-systèmes est multiplié par* λ, chacun d'eux gardant les mêmes propriétés (nous supposons que le système a été divisé en sous-systèmes égaux). On en déduit aussitôt que, dans le système global, *la valeur de chaque grandeur extensive est multipliée par* λ, *celle des grandeurs intensives reste inchangée.* Mais *le rapport de deux grandeurs extensives,* telle la densité de particules N/V ou l'énergie par particule E/N, *se comporte dans cette opération comme les grandeurs intensives.* Aussi, bien qu'ils soient de nature fondamentalement différente, nous classerons les rapports de grandeurs extensives parmi les grandeurs intensives.

---

8. Nous supposons que le découpage n'a pas artificiellement introduit des effets de surface : les sous-systèmes ont grossièrement la forme de boules ou de cubes, pas de lamelles très minces dans une direction.

9. Pour préciser, on peut imaginer que l'on prend 2 (ou 3, ...) systèmes identiques, préparés de la même façon, qu'on les amène au contact de part et d'autre d'une cloison commune, puis qu'on les réunit en un seul système en supprimant la cloison. On aura ainsi multiplié la taille du système par $\lambda = 2$ (ou 3, ...).

10. λ peut être aussi grand que l'on veut. Il ne doit par contre pas être petit au point de remettre en cause la division du système en sous-systèmes *macroscopiques.*

La classification des grandeurs physiques en extensives et intensives joue un rôle important en thermodynamique. Mais elle s'applique seulement aux systèmes qui vérifient les deux conditions du paragraphe I.2 : *couplage faible* entre les sous-systèmes macroscopiques, dû à la faible portée des forces; *effets de surface négligeables.* Elle suppose de plus que *les fluctuations sont trop petites en valeur relative pour être décelables.* Les deux premières conditions portent sur la nature du système, et seront le plus souvent remplies par les systèmes que nous aurons à étudier; elles assurent l'additivité des grandeurs extensives et l'égalisation des grandeurs intensives. La condition supplémentaire concerne la taille des sous-systèmes et donc du système lui-même; elle permet d'attribuer une valeur définie à chaque grandeur dans chaque sous-système.

Références

1. E.H. Lieb and J.L. Lebowitz, The Constitution of Matter : Existence of Thermodynamics for Systems Composed of Electrons and Nuclei. *Advances in Mathematics 9,* 316-398 (1972).

   E.H. Lieb, The Stability of Matter. *Reviews of Modern Physics 48,* 553-569 (1976).

2. S.W. de Leeuw, J.W. Perram and E.R. Smith, Simulation of Electrostatic Systems in Periodic Boundary Conditions. I - Lattice Sums and Dielectric Constants. II - Equivalence of Boundary Conditions. *Proceedings of the Royal Society of London A 373,* 27-56 et 57-66 (1980).

3. E.M. Purcell, *Berkeley, Cours de Physique.* Vol. 2 : *Électricité et Magnétisme,* A. Colin, Paris (1973), paragraphe 1.6.

COMPLÉMENT II.B

Le cristal paramagnétique parfait. Températures négatives

I. Étude d'un cristal paramagnétique dans un champ magnétique extérieur
   1. Le modèle du cristal paramagnétique parfait de spins 1/2
   2. Calcul de l'entropie
   3. Calcul de la température

II. Les températures négatives
   1. Conditions d'existence
   2. Les températures négatives sont « plus chaudes » que les positives

III. Température de spin
   1. Spins et réseau. Temps de relaxation
   2. Comment modifier la température de spin?

Le cristal paramagnétique parfait est probablement le système le plus simple auquel on peut appliquer la mécanique statistique. Nous le traitons ici dans le formalisme microcanonique. Il possède une propriété curieuse : sa température peut devenir négative. Nous préciserons la signification théorique des températures négatives et nous indiquerons dans quelles conditions expérimentales on peut les obtenir.

I. ÉTUDE D'UN CRISTAL PARAMAGNÉTIQUE DANS UN CHAMP MAGNÉTIQUE EXTÉRIEUR

1. LE MODÈLE DU CRISTAL PARAMAGNÉTIQUE PARFAIT DE SPINS 1/2

Considérons un système de N spins 1/2 fixés aux nœuds d'un réseau cristallin, placé dans un champ magnétique uniforme $\vec{B}$. Nous ne nous intéressons qu'à l'ensemble des N spins[1] et négligeons tous les autres degrés de liberté : au spin $\vec{S}_i$ situé sur le i-ième site est associé un moment magnétique

$$\vec{\mu}_i = \gamma \vec{S}_i \,, \tag{B.1}$$

où γ est le « rapport gyromagnétique ». Ce système atteint son équilibre macroscopique grâce aux interactions entre spins, mais nous supposons que les énergies d'interaction sont négligeables devant les énergies de couplage entre les moments magnétiques de spin et le champ extérieur. Dans cette approximation, l'hamiltonien du système s'écrit

$$H = -\sum_{i=1}^{N} \vec{\mu}_i . \vec{B} = -B \sum_{i=1}^{N} \mu_{iz} \,, \tag{B.2}$$

où μ_{iz} est la composante du i-ième moment magnétique dans la direction de $\vec{B}$.

L'état microscopique du système est déterminé lorsqu'on connaît l'état du spin situé en chaque site du réseau. Chaque spin a deux états possibles. Les états propres

1. Un corps est paramagnétique lorsque ses constituants microscopiques (ou tout au moins certains d'entre eux) possèdent un moment magnétique permanent. Un tel moment est toujours associé à un moment cinétique non nul, que nous supposons ici être un spin 1/2.

$|+\rangle$ et $|-\rangle$ de la composante S_z de l'opérateur $\vec{S}$ qui le décrit ont une énergie ε bien déterminée :

$$\left.\begin{aligned} S_z|+\rangle &= +\frac{1}{2}|+\rangle \\ S_z|-\rangle &= -\frac{1}{2}|-\rangle \end{aligned}\right\} \Longrightarrow \begin{cases} \varepsilon_+ = -\dfrac{\gamma}{2}B = -\mu B \\ \varepsilon_- = +\dfrac{\gamma}{2}B = +\mu B \end{cases} \tag{B.3}$$

(nous avons pris pour simplifier $\vec{S}$ sans dimension, le facteur $\hbar$ étant intégré dans le rapport gyromagnétique, et nous supposons pour fixer les idées que γ est positif). Si l'on note n_+ et n_- les nombres de spins dans les états $|+\rangle$ et $|-\rangle$, l'énergie totale du système vaut

$$E = (n_- - n_+)\mu B\,. \tag{B.4}$$

Elle peut varier entre $-N\mu B$ et $+N\mu B$; la première limite correspond au cas où tous les spins sont dans l'état $|+\rangle$, la deuxième au cas où ils sont tous dans l'état $|-\rangle$. Si on fixe l'énergie à une valeur E comprise entre ces deux bornes, les nombres n_+ et n_- sont déterminés de façon unique. En effet

$$\left.\begin{aligned} n_+ + n_- &= N \\ (n_- - n_+)\mu B &= E \end{aligned}\right\} \Longrightarrow \begin{cases} n_+ = \dfrac{1}{2}\left(N - \dfrac{E}{\mu B}\right) \\ n_- = \dfrac{1}{2}\left(N + \dfrac{E}{\mu B}\right). \end{cases} \tag{B.5}$$

Les interactions entre spins peuvent changer leur état, on dit qu'elles les « retournent ». Toutefois, pour un système isolé[2], ces changements doivent conserver l'énergie totale (B.4) : chaque fois qu'un spin passe de l'état $|-\rangle$ à l'état $|+\rangle$, un autre doit simultanément se retourner en sens inverse, de façon que les nombres n_+ et n_- restent inchangés.

2. CALCUL DE L'ENTROPIE

L'entropie d'un système dont l'énergie E est fixée s'écrit

$$S^* = k \ln \Omega\,, \tag{B.6}$$

où Ω est le *nombre d'états accessibles*. C'est ici le nombre d'états qui, pour un nombre de particules N et un champ extérieur $\vec{B}$ donnés, ont l'énergie E. D'après (B.5), tous les états qui ont la même énergie E ont le même nombre n_- de spins dans l'état $|-\rangle$; ils diffèrent les uns des autres par l'emplacement des spins $|-\rangle$ dans le cristal. Le nombre de ces états est donc

$$\Omega = C_N^{n_-} = \frac{N!}{n_-!\,n_+!}\,. \tag{B.7}$$

2. A strictement parler, le système de spins placé dans un champ magnétique peut perdre de l'énergie par *émission spontanée* de rayonnement : un spin qui se trouve dans son état d'énergie la plus élevée (l'état $|-\rangle$) peut se retourner en émettant un photon de pulsation $2\mu B/\hbar$. Mais ce rayonnement se situe dans le domaine des radiofréquences (longueur d'onde de l'ordre du mètre pour un champ magnétique de quelques teslas), où l'émission spontanée est très peu probable (cf. complément VI.E, § II.3). Si nous négligeons ce phénomène, nous pouvons considérer que le système garde une énergie constante et l'étudier dans le cadre microcanonique, bien qu'il ne soit pas véritablement isolé.

On en déduit l'entropie[3]

$$S^* = k \ln \frac{N!}{n_-!n_+!} = k \ln \frac{N!}{\left[\frac{N+E/\mu B}{2}\right]!\left[\frac{N-E/\mu B}{2}\right]!} . \tag{B.8}$$

Sauf dans les cas limites où l'un des nombres n_+ et n_- est voisin de 1, on peut écrire les logarithmes des factorielles dans l'approximation de Stirling :

$$\ln N! \simeq N \ln N - N ; \tag{B.9}$$

on trouve alors :

$$S^* = k\left\{N \ln N - \frac{N+E/\mu B}{2} \ln \frac{N+E/\mu B}{2} - \frac{N-E/\mu B}{2} \ln \frac{N-E/\mu B}{2}\right\} . \tag{B.10}$$

Pour un cristal donné (N déterminé) et un champ magnétique B fixé, l'entropie est une fonction de l'énergie E (fig. 1). Lorsque l'énergie prend sa valeur minimum $-N\mu B$, tous les spins se trouvent dans l'état $|+\rangle$; le système n'a alors qu'un état accessible, et son entropie est nulle. Le même raisonnement est valable lorsque l'énergie atteint sa valeur maximum $+N\mu B$. L'entropie est donc une fonction de l'énergie d'abord croissante puis décroissante. La courbe est symétrique, son maximum est obtenu pour $E = 0$ et vaut[4]

$$S_m = Nk \ln 2 . \tag{B.11}$$

---

3. En pratique, l'énergie n'est pas strictement fixée, mais comprise entre E et $E + \delta E$; le nombre n_- peut varier entre la valeur donnée par (B.5) et une valeur légèrement supérieure $n_- + \delta n_-$, correspondant à $E + \delta E$. Le nombre d'états accessibles est donc égal à

$$\Omega = \sum_{n=n_-}^{n_-+\delta n_-} C_N^n . \tag{1}$$

Lorsque le nombre n_- varie d'une unité, l'énergie du système varie de $2\mu B$; la somme (1) compte donc $\delta E/2\mu B$ termes. On trouve ainsi pour le nombre d'états accessibles

$$\Omega = \frac{\delta E}{2\mu B} C_N^{n_-} , \tag{2}$$

et pour l'entropie

$$S^* = k \ln C_N^{n_-} + k \ln (\delta E/2\mu B)$$

$$= k \ln \frac{N!}{n_-!n_+!} + k \ln (\delta E/2\mu B) . \tag{3}$$

Le deuxième terme de cette expression étant fantastiquement plus petit que le premier, on peut le négliger et on retrouve (B.8).

En toute rigueur, l'expression (2) de Ω devrait faire intervenir un nombre $\bar{n}_-$ compris entre n_- et $n_- + \delta n_-$, et non pas la borne inférieure de l'intervalle. Mais ceci n'affecte pratiquement pas la valeur de l'entropie (voir exercice II-17).

4. Cette valeur, obtenue à partir de (B.8) en utilisant l'approximation de Stirling sous la forme (B.9), peut surprendre car c'est celle que l'on trouverait en calculant le logarithme du nombre total d'états du système, *quelle que soit leur énergie :* il y a N sites qui ont chacun deux états possibles; le nombre total d'états pour le système est donc 2^N.

En fait, le nombre d'états d'énergie nulle est

$$\Omega(E=0) = C_N^{N/2} = \frac{N!}{(N/2)!(N/2)!} . \tag{1}$$

En précisant l'approximation de Stirling :

$$N! \simeq \sqrt{2\pi N} N^N e^{-N} , \tag{2}$$

.../...

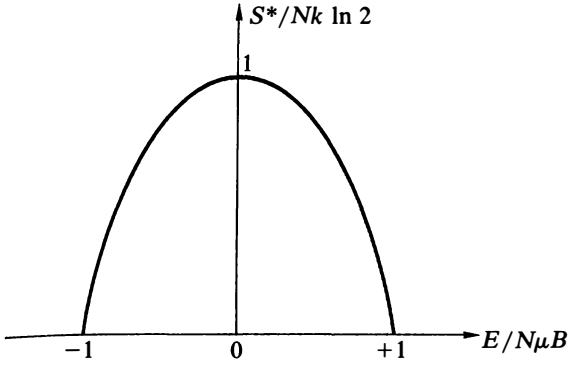


FIGURE 1
Relation entre l'énergie E et l'entropie S^* d'un système de N spins 1/2 indépendants placés dans un champ magnétique $\vec{B}$.

3. CALCUL DE LA TEMPÉRATURE

La température microcanonique T^* du système est définie par la relation

$$\frac{1}{T^*} \widehat{=} \frac{\partial S^*}{\partial E} . \tag{B.12}$$

On trouve facilement

$$\frac{1}{T^*} = \frac{k}{2\mu B} \ln \frac{N - E/\mu B}{N + E/\mu B} = \frac{k}{2\mu B} \ln \frac{n_+}{n_-} . \tag{B.13}$$

A N et B fixés, la température calculée par (B.13) est une fonction de l'énergie E du système[5]. *Lorsque l'énergie est positive,* l'argument du logarithme est inférieur à

---

on trouve

$$\Omega(E=0) = \frac{2^N}{\sqrt{\pi N/2}} ; \tag{3}$$

d'où

$$S^*(E=0) = Nk \ln 2 - \frac{1}{2} k \ln (N\pi/2) . \tag{4}$$

Pour N de l'ordre du nombre d'Avogadro, le deuxième terme est totalement négligeable devant le premier. Le nombre $\Omega(E=0)$ d'états d'énergie nulle est évidemment inférieur au nombre total 2^N d'états possibles, il est même petit devant lui puisque $\sqrt{N}$ est grand devant 1; mais leur différence est beaucoup trop faible, en valeur relative, pour être appréciable dans l'entropie.

5. On peut inverser la relation entre E et T^* :

$$E = N\mu B \frac{1 - e^{2\mu B/kT^*}}{1 + e^{2\mu B/kT^*}} = -N\mu B \operatorname{th} \mu B/kT^* . \tag{1}$$

A chaque température T^* correspond ainsi une énergie E unique et donc, d'après les relations (B.5), un nombre n_+ de spins dans l'état $|+\rangle$ ainsi qu'un nombre n_- de spins dans l'état $|-\rangle$:

$$n_+ = N \frac{e^{\varepsilon/kT^*}}{e^{\varepsilon/kT^*} + e^{-\varepsilon/kT^*}} ,$$

$$n_- = N \frac{e^{-\varepsilon/kT^*}}{e^{\varepsilon/kT^*} + e^{-\varepsilon/kT^*}} , \tag{2}$$

où nous avons posé

$$\varepsilon \widehat{=} \mu B . \tag{3}$$

Les expressions (2) ont une forme qui nous deviendra familière au chapitre III.

1 de sorte que *la température T* est négative.* Ce résultat vient de ce que, pour $E>0$, l'entropie est une fonction décroissante de l'énergie.

La figure 2 représente les variations de $1/T^*$ et de T^* en fonction de l'énergie E : $1/T^*$ est une fonction continue tandis que T^* a une asymptote verticale en $E=0$. Partons de l'état d'énergie minimum $-N\mu B$ où tous les spins sont dans l'état $|+\rangle$ et la température nulle. Lorsque l'énergie augmente (et donc le rapport n_+/n_- diminue), la température est positive et croît. Lorsque l'énergie approche de 0, c'est-à-dire lorsque n_+/n_- tend vers 1 par valeurs supérieures, la température tend vers $+\infty$. Si l'énergie devient positive, la température croît toujours mais de $-\infty$ (pour $E=0_+$) à 0_- (pour $E=+N\mu B$). Les températures négatives correspondent aux cas où n_+/n_- est inférieur à 1; on dit alors que «les populations sont inversées[6]». Lorsque la température est égale à 0_-, tous les spins sont dans l'état $|-\rangle$. Les températures 0_+ et 0_- correspondent donc à des états physiques totalement différents, tandis qu'il y a peu de différence entre les états correspondant aux températures $+\infty$ et $-\infty$. Ceci apparaît clairement sur la courbe représentant $1/T^*$ (fig. 2.a).

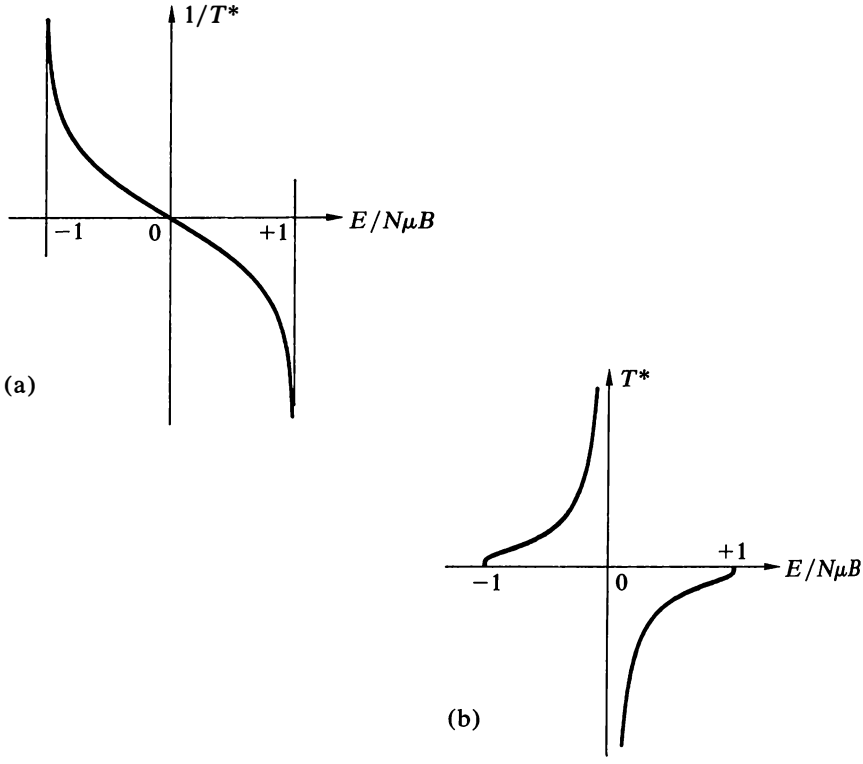


FIGURE 2

Variation de la température T^* en fonction de l'énergie E pour un système de N spins 1/2 indépendants : la courbe (a) représente $1/T^*$, la courbe (b) T^* elle-même.

6. Nous verrons au chapitre III que, lorsqu'un système constitué de N particules indépendantes identiques est en contact avec un thermostat, le nombre de particules qui se trouvent dans un état individuel d'énergie ε est proportionnel à $\exp -\varepsilon/kT$ (où T est la température du thermostat, qui est *positive*) : les états de basse énergie sont donc les plus «peuplés». Quand on dit que les populations sont inversées, c'est par rapport à cette situation de référence.

II. LES TEMPÉRATURES NÉGATIVES

1. CONDITIONS D'EXISTENCE

Les résultats que nous venons d'établir (entropie fonction décroissante de l'énergie et donc température négative dans un certain domaine), ne sont pas limités à l'exemple très simple étudié ici; on les retrouve pour tous les systèmes dont les constituants ont chacun un *nombre fini d'états individuels possibles.* Pour de tels systèmes en effet, l'énergie E est bornée inférieurement ($E \geqslant E_{\min}$), comme toujours, mais aussi supérieurement ($E \leqslant E_{\max}$). Si, pour simplifier le raisonnement, nous négligeons les interactions et supposons que l'état individuel d'énergie la plus basse et celui d'énergie la plus élevée sont uniques[7], l'entropie du système sera nulle aux deux bornes de l'intervalle des énergies permises : $E = E_{\max}$, par exemple, est obtenue pour un seul état microscopique du système, celui où tous ses constituants se trouvent dans leur état d'énergie la plus élevée. L'entropie S^* du système, nulle pour $E = E_{\min}$ et $E = E_{\max}$, est comme précédemment une fonction d'abord croissante de E, ce qui correspond à des températures T^* positives, puis décroissante, ce qui donne lieu à des températures négatives[8],[9]. Au contraire, quand le nombre d'états possibles pour chaque constituant est infini, l'entropie du système est une fonction toujours croissante de l'énergie, et sa température est toujours positive[10].

---

7. Si le niveau fondamental individuel et le niveau individuel d'énergie la plus élevée sont dégénérés g_m et g_M fois respectivement, l'entropie du système vaut, non plus zéro, mais

$$S^*(E_{\min}) = Nk \ln g_m$$

et $$S^*(E_{\max}) = Nk \ln g_M$$

aux bornes de l'intervalle permis pour l'énergie. Mais $S^*(E)$ présente toujours un maximum pour une valeur intermédiaire de E, au-dessus de laquelle les températures T^* sont négatives.

8. Le résultat reste valable en présence d'interactions. En effet, l'espace des états du système (produit tensoriel des N espaces d'états individuels; voir appendice II, § III) est de dimension finie. Le nombre d'états accessibles Ω est donc borné, de sorte qu'il existe au moins une valeur E_0 de l'énergie pour laquelle la fonction $\Omega(E)$, et donc l'entropie, présentent un maximum. Pour les énergies supérieures à E_0 il existe nécessairement un domaine de températures négatives (sauf cas particulier où le maximum se produirait à l'extrémité supérieure de l'intervalle permis pour l'énergie).

9. Nous verrons dans le complément II.E qu'un système de température négative serait instable car il y apparaîtrait spontanément des mouvements macroscopiques qui le détruiraient. Cet argument ne s'applique pas ici : les systèmes que nous étudions n'ont pas d'énergie cinétique et ne peuvent donc pas se mettre en mouvement.

Il existe un autre argument général assurant que la température d'un système doit être positive. Nous établirons en effet au chapitre III la distribution de probabilité des états microscopiques d'un système à l'équilibre thermique avec un thermostat de température fixée. Nous verrons alors que, pour normaliser cette distribution, il faut calculer une somme sur tous les états du système (fonction de partition). Si l'énergie de ces états n'est pas bornée, la série correspondante diverge nécessairement si la température est négative. Mais, dans les systèmes qui nous intéressent ici, le nombre d'états est fini; aucun problème de convergence ne peut donc se poser. Ce second argument ne s'applique pas non plus.

10. Pour nous en persuader, considérons un système constitué de N éléments indépendants ayant chacun un nombre infini de niveaux d'énergie. Examinons comment varie le nombre d'états accessibles du système lorsque l'énergie croît de E à $E' > E$. A partir de chaque état d'énergie E on peut construire un grand nombre d'états d'énergie E' : on peut en effet rajouter l'énergie supplémentaire de nombreuses façons différentes, puisque l'énergie de chacun des éléments du système peut être accrue (ce qui n'est pas le cas lorsque le spectre d'énergie de chaque élément est borné supérieurement). Le nombre d'états accessibles du système est donc nécessairement une fonction croissante de l'énergie.

Pour que la température puisse être négative, il faut donc que chacun des éléments constituant le système ait un nombre fini d'états possibles. *Ceci exclut en particulier que ces éléments aient de l'énergie cinétique,* puisque le spectre de l'énergie cinétique n'est pas borné supérieurement. Or les particules, électrons ou noyaux, qui possèdent un spin possèdent aussi de l'énergie cinétique. Il semble donc qu'aucun système réel ne soit dans les conditions requises pour que sa température puisse être négative. Nous verrons cependant au paragraphe III que, dans certains systèmes hors d'équilibre, les degrés de liberté de spin peuvent être décrits par une température négative pendant suffisamment longtemps pour que l'on puisse observer leurs propriétés.

2. LES TEMPÉRATURES NÉGATIVES SONT « PLUS CHAUDES » QUE LES POSITIVES

L'expression « température négative » pourrait faire croire qu'un système à température négative est plus « froid » qu'un système à température positive. Il n'en est rien et nous allons montrer qu'il est en fait plus « chaud » que tout système à température positive.

Nous avons vu au chapitre II que, lorsqu'on établit le contact thermique entre deux corps dont les températures initiales sont T_1^* et T_2^* respectivement, à l'équilibre les deux corps ont une même température, intermédiaire entre T_1^* et T_2^*. On dit que le corps qui a cédé de l'énergie à l'autre était initialement le plus « chaud ». Nous allons montrer maintenant que *lorsqu'on met en contact thermique un corps à température négative avec un corps à température positive, c'est toujours le corps à température négative qui cède de l'énergie à l'autre.*

Pour le voir, il suffit de nous rappeler que, pour un système donné, les températures négatives (si elles existent) sont toujours obtenues pour des valeurs de l'énergie plus élevées que les températures positives. Envisageons tout d'abord le contact d'un corps à température négative avec un corps dont la température est toujours positive (on qualifie un tel corps de normal). A l'équilibre les deux températures doivent être égales; elles sont donc nécessairement positives. L'énergie du corps initialement à température négative a donc diminué et c'est lui qui a cédé de l'énergie lors du contact thermique. Imaginons maintenant le contact entre deux corps pouvant avoir des températures négatives, l'un ayant initialement une température positive et l'autre une température négative. Supposons que ce dernier reçoive de l'énergie; sa température resterait négative et la température finale serait négative; mais alors l'énergie de l'autre corps aurait elle aussi augmenté. Or les deux corps ne peuvent pas avoir tous deux reçu de l'énergie puisque le système global qu'ils constituent est isolé. Cette situation est donc impossible; le corps dont la température initiale était négative doit nécessairement avoir perdu de l'énergie.

Dans les deux cas, le corps initialement à température négative apparaît comme le plus chaud, puisqu'il cède de l'énergie à celui qui avait au départ une température positive. Si par exemple on met en contact thermique deux corps aux températures $T_1^* = +300$ K et $T_2^* = -300$ K, la température finale intermédiaire entre T_1^* et T_2^* peut être $+1\,000$ K mais certainement pas 0. Si l'on ordonne les températures du « plus froid » au « plus chaud », on obtient la succession

$$0_+, \quad T>0, \quad +\infty, \quad -\infty, \quad T<0, \quad 0_- . \tag{B.14}$$

Ces paradoxes apparents proviennent de ce que la quantité historiquement appelée température en thermodynamique n'est pas celle qui s'introduit naturellement en mécanique statistique. On ne trouverait aucun résultat étonnant si l'on avait choisi de caractériser de « degré de chaud ou de froid » d'un corps par la quantité $1/T^* = \partial S^*/\partial E$ qui est continue sur tout le domaine de variation de l'énergie, ou mieux par $-1/T^*$, qui est d'autant plus grande que le corps est plus chaud.

III. TEMPÉRATURE DE SPIN

Nous avons vu que, s'il est facile de concevoir une situation théorique dans laquelle la température d'un système est négative, la condition à remplir (le nombre d'états du système doit être fini) semble impossible à satisfaire puisqu'elle interdit que le système ait de l'énergie cinétique. Dans ce paragraphe, nous allons introduire le concept de température de spin; c'est elle qui dans certaines conditions peut être négative. Nous examinerons comment on peut, en pratique, faire varier la température de spin d'un système et la rendre négative.

1. SPINS ET RÉSEAU. TEMPS DE RELAXATION

Dans un système comportant des spins placés dans un champ magnétique extérieur, on appelle traditionnellement «le réseau» l'ensemble des degrés de liberté autres que les spins. Dans certains corps[11], les interactions des spins nucléaires avec le réseau sont très faibles, bien inférieures aux interactions entre éléments du réseau ou entre les spins eux-mêmes. Un échange appréciable d'énergie entre les spins et le réseau prend alors beaucoup de temps; bien qu'ils se trouvent dans la même région de l'espace, tout se passe comme si l'ensemble des spins et le réseau étaient séparés par une paroi presque adiabatique. Chacun de ces deux sous-systèmes a donc le temps

11. On distingue deux sortes de matériaux paramagnétiques parmi les cristaux ioniques isolants.

Dans un premier type de cristaux, le paramagnétisme est attaché au cortège électronique de l'un au moins des ions constituant la maille élémentaire : lorsque le moment cinétique total $\vec{J}$ de l'ensemble des électrons est non nul dans l'état fondamental, il lui est associé un moment magnétique

$$\vec{\mu} = -g_J \mu_B \vec{J}, \tag{1}$$

où $$\mu_B = \frac{e\hbar}{2m} \tag{2}$$

désigne le «magnéton de Bohr» ($-e$ est la charge de l'électron et m sa masse), et g_J le «facteur de Landé» de l'état. Les ions Cu^{++}, Mn^{++}, Fe^{3+}, Cr^{3+}, Gd^{3+}, ... sont dans ce cas et leurs sels (notamment les aluns) sont très utilisés comme matériaux paramagnétiques. Dans les corps de cette espèce, seuls les ions paramagnétiques sont sensibles à la présence d'un champ magnétique (en première approximation); mais ces ions sont couplés au reste du cristal, essentiellement à ses modes de vibration, et il n'est pas possible de traiter le sous-système des ions paramagnétiques comme isolé. Nous étudierons ces corps dans le complément III.A.

Mais il existe un deuxième type de cristaux ioniques paramagnétiques, par exemple les halogénures alcalins comme le fluorure de lithium LiF. Dans ce cas, le cortège électronique de chacun des ions a un moment cinétique total et donc un moment magnétique nuls (les ions Li^+ et F^- sont comme les gaz rares constitués de couches électroniques toutes complètes). C'est alors le noyau qui possède un moment cinétique $\vec{I}$ non nul, auquel est encore associé un moment magnétique $\vec{\mu}$ colinéaire à $\vec{I}$ selon une formule analogue à (1). Le magnétisme nucléaire est beaucoup plus faible que le magnétisme électronique : le coefficient de proportionnalité entre $\vec{\mu}$ et $\vec{I}$ (rapport gyromagnétique) fait intervenir le «magnéton de Bohr nucléaire»

$$\mu_N = \frac{e\hbar}{2m_p}, \tag{3}$$

où figure la masse du proton m_p à la place de celle de l'électron (notons que le neutron, bien qu'électriquement neutre, possède comme le proton un moment magnétique non nul). Le moment magnétique du noyau est masqué par celui des électrons lorsque ceux-ci ont un moment cinétique total $\vec{J}$ non nul. Mais lorsque $\vec{J}$ est nul, comme dans Li^+ et F^-, les moments magnétiques nucléaires se manifestent directement. Leur couplage aux autres degrés de liberté du cristal est considérablement moins intense que pour les moments électroniques.

de passer un grand nombre de fois par tous ses états accessibles et d'atteindre ainsi son propre état d'équilibre macroscopique avant que les interactions entre les spins et le réseau ne finissent par établir leur équilibre mutuel. On dit que *« le temps de relaxation spins-réseau »* τ est long devant les « temps de relaxation » τ_S du système de spins et τ_R du réseau[12]. Pendant la période transitoire où les deux sous-systèmes peuvent être considérés comme séparément isolés et à l'équilibre, leurs températures sont en général différentes; celle du réseau est toujours positive, celle du système de spins peut être positive ou négative.

En pratique, les expériences sont réalisées dans des cristaux ioniques tels que le fluorure de lithium LiF[13]. Le temps de relaxation spins-réseau est, à la température ambiante, d'environ 5 minutes pour les ions Li^+ et 2 minutes pour les ions F^-, alors que le temps de relaxation de chacun des deux ensembles de spins est de l'ordre de 10^{-5} s.

2. COMMENT MODIFIER LA TEMPÉRATURE DE SPIN?

Le problème est de faire varier la température de spin indépendamment de celle du réseau. La température d'un système de N spins dans un champ magnétique $\vec{B}$ est donnée par la relation (B.13). On peut envisager *a priori* deux façons de modifier la température de spin d'un système : soit modifier le rapport n_+/n_- en maintenant le champ constant, soit faire varier le champ magnétique $\vec{B}$ en maintenant le rapport n_+/n_- constant.

On peut modifier les populations des états en faisant agir sur le système une onde électromagnétique sinusoïdale de pulsation $\omega_0 = 2\mu B/\hbar$ (radiofréquence). Les coefficients d'absorption et d'émission induite étant égaux (et l'émission spontanée négligeable), tant que l'état $|+\rangle$ est plus peuplé que l'état $|-\rangle$ il y a plus d'absorption que d'émission et le rapport n_-/n_+ augmente. Mais lorsque n_+ et n_- s'égalisent, l'émission et l'absorption se compensent et les populations des niveaux ne varient plus. On peut donc ainsi amener le système dans un état tel que sa température de spin soit très élevée et positive ($T^* \to +\infty$), mais on ne peut pas atteindre les températures négatives.

Il est facile de modifier le champ magnétique. Si on le fait de façon suffisamment lente (mais quand même en un temps court devant le temps de relaxation spins-réseau τ), et en conservant sa direction et son sens, on arrive à changer l'énergie de chacun des spins sans modifier son état : les nombres n_+ et n_- restent inchangés, seul μB varie[14]. On voit aussitôt sur les formules (B.4) et (B.13) que l'énergie totale du système change proportionnellement au champ magnétique :

$$\frac{E_2}{E_1} = \frac{B_2}{B_1}, \tag{B.15}$$

de même que sa température de spin T^* :

$$\frac{T_2^*}{T_1^*} = \frac{B_2}{B_1} \tag{B.16}$$

(la température du réseau n'est bien sûr pas affectée). Si l'on diminue la valeur de

12. On appelle « temps de relaxation » le temps nécessaire à l'établissement de l'équilibre dans un système que l'on a préparé dans une situation hors d'équilibre (cf. IV, § II.B.1.b).

13. Le spin I du noyau de fluor est bien 1/2, mais celui du noyau de lithium 7 est 3/2.

14. On parvient donc à changer l'énergie du système sans modifier son entropie : si n_+ et n_- restent constants, il en est de même pour S^* d'après (B.8). On réalise ainsi une transformation adiabatique réversible (voir complément II.D, § III.3.c).

B, la température de spin diminue d'autant : ce phénomène est connu sous le nom de *«refroidissement par désaimantation adiabatique»*.

Mais ce procédé de variation lente du champ magnétique ne permet pas de faire franchir au système de spins la frontière entre le domaine des températures positives et celui des températures négatives. Pour y parvenir, on agit de la façon suivante. Partant d'un état initial où le champ B a une valeur relativement importante (de l'ordre de 6000 gauss), on effectue d'abord une désaimantation adiabatique jusqu'à une valeur nettement plus faible (100 gauss). On inverse alors, en un temps très court[15], le sens du champ, que l'on ramène ensuite de façon adiabatique à la valeur initiale de 6000 gauss en conservant sa nouvelle orientation. L'inversion du sens du champ peut être réalisée sans que les populations des états individuels aient le temps de changer. Renverser ainsi le champ magnétique revient à intervertir les états $|+\rangle$ et $|-\rangle$. Partant d'un état où n_+/n_- est supérieur à 1 (température de spin positive), on aboutit dans la nouvelle configuration à un rapport n_+/n_- inférieur à 1, c'est-à-dire que la température est devenue négative. Cette température négative se maintient pendant un temps de l'ordre de τ, jusqu'à ce que l'énergie fournie au système de spins par l'inversion du champ magnétique soit transférée au réseau qu'elle échauffe. Purcell et Pound ont obtenu ainsi, pour la première fois en 1950, dans un cristal de fluorure de lithium, une température négative qui s'est maintenue pendant plusieurs minutes (référence 1).

Le point crucial est que puissent coexister dans un même cristal deux températures différentes, celle du réseau et celle du système de spins. Que c'est bien le cas, c'est-à-dire que la température de spin est une véritable température, a été démontré par Proctor et Abragam (référence 2) grâce à l'expérience suivante. Un échantillon de fluorure de lithium est d'abord placé dans un thermostat à la température de 3 K suffisamment longtemps pour que le système de spins et le réseau acquièrent tous deux cette température; la température de spin est mesurée en étudiant le signal de résonance magnétique nucléaire à la pulsation ω_0, dont l'intensité donne le rapport des populations des états de spin individuels. Dans un second temps, le même échantillon est pris à température ambiante (300 K) et on lui fait subir une désaimantation adiabatique calculée pour amener la température de spin à 3 K sans modifier celle du réseau; l'étude par résonance magnétique confirme qu'il s'agit bien de la même température de spin que précédemment. Une fois établie l'existence autonome de la température de spin, les raisonnements du paragraphe II montrent qu'elle est susceptible de devenir négative.

Références

1. E.M. Purcell et R.V. Pound, *Phys. Rev. 81,* 279 (1951).

2. A. Abragam et W.G. Proctor, *Phys. Rev. 109,* 1441 (1958).

15. La durée de cette opération doit être petite devant le temps caractéristique de l'évolution des moments magnétiques individuels. Dans un champ de 100 gauss, la pulsation de Larmor pour un moment magnétique nucléaire typique vaut

$$\omega_L = \mu_N B/\hbar \simeq 4.10^5 \text{ radian} \cdot \text{s}^{-1} ,$$

de sorte qu'il faut inverser le champ en un temps inférieur à 10^{-5} s.

Si l'on voulait inverser directement le champ initial de 6000 gauss, il faudrait réaliser l'opération en un temps 60 fois plus court, alors qu'en pratique plus le champ est grand, plus il est long à établir ou à supprimer.

COMPLÉMENT II.C

Le gaz parfait classique : entropie de mélange et paradoxe de Gibbs

I. Description microcanonique d'un système de particules indépendantes et discernables
   1. Entropie
   2. Température et pression. Équation d'état du gaz parfait
   3. Mélange idéal de deux gaz parfaits
      a. Entropie, température et pression
      b. Entropie de mélange
   4. Le paradoxe de Gibbs
      a. Énoncé
      b. Origine du paradoxe

II. Système de particules indépendantes et indiscernables. Approximation de Maxwell-Boltzmann
   1. Évaluation de la densité d'états
   2. Expression des propriétés physiques des gaz parfaits
      a. Entropie
      b. Équation d'état. Entropie de mélange
      c. Potentiel chimique
   3. Domaine de validité de l'approximation de Maxwell-Boltzmann

Nous nous proposons d'étudier les propriétés d'un système de N particules sans structure interne, libres, indépendantes et enfermées dans un récipient de volume V. Ces particules constituent un *gaz parfait monoatomique*[1].

Dans un premier temps, nous supposerons les N particules discernables et marquerons d'un indice supérieur ^{D} les grandeurs correspondantes. Lorsqu'on l'applique à un gaz, ce modèle conduit à une difficulté fondamentale connue sous le nom de *paradoxe de Gibbs*. Dans la deuxième partie, nous montrerons comment on peut lever le paradoxe en tenant compte de l'indiscernabilité des particules que nous traiterons dans l'approximation de Maxwell-Boltzmann.

I. DESCRIPTION MICROCANONIQUE D'UN SYSTÈME DE PARTICULES INDÉPENDANTES ET DISCERNABLES

1. ENTROPIE

Au chapitre II (formule (I.8)), nous avons montré que l'entropie $S^*(E, V, N)$ d'un système $\mathcal{S}$ isolé, d'énergie E, contenant N particules dans un volume V, s'exprime simplement à l'aide de la densité d'états $\rho(E, V, N)$:

$$S^*(E, V, N) = k \ln \rho(E, V, N) . \tag{C.1}$$

Pour un système de N particules de masse m, *discernables*, libres et indépendantes, cette densité d'états a pour expression (complément I.E, formule (E.22)) :

$$\rho^D(E, V, N) = \frac{3N}{2} \left(\frac{m}{2\pi\hbar^2}\right)^{3N/2} \frac{V^N}{\Gamma\left(\frac{3N}{2}+1\right)} E^{\frac{3N}{2}-1} . \tag{C.2}$$

1. Les propriétés du gaz parfait seront étudiées en détail au chapitre III (§ IV.B et complément III.B). Ici les atomes sont assimilés à des points matériels (sans spin) dont l'énergie se réduit à l'énergie cinétique de translation.

Si le système est macroscopique, N est très grand, ce qui permet d'approximer $\Gamma(3N/2+1)$ par la formule de Stirling :

$$\ln \Gamma\left(\frac{3N}{2}+1\right) \simeq \frac{3N}{2} \ln \frac{3N}{2} - \frac{3N}{2} . \tag{C.3}$$

On trouve ainsi[2]

$$S^{*D}(E, V, N) = Nk\left[\ln V + \frac{3}{2} \ln\left(\frac{m}{3\pi\hbar^2} \frac{E}{N}\right) + \frac{3}{2}\right] . \tag{C.4}$$

2. TEMPÉRATURE ET PRESSION. ÉQUATION D'ÉTAT DU GAZ PARFAIT

La température et la pression se déduisent immédiatement de l'entropie (II, § III.A.1.a et § III.B.1) :

$$\frac{1}{T^*} \widetilde{=} \frac{\partial S^*}{\partial E} = \frac{3}{2} \frac{Nk}{E} \tag{C.5}$$

et

$$\frac{p^*}{T^*} \widetilde{=} \frac{\partial S^*}{\partial V} = \frac{Nk}{V} . \tag{C.6}$$

Ces deux relations resteront valables pour des particules indiscernables. La première relie l'énergie du gaz (parfait et monoatomique) à sa température; nous la retiendrons sous la forme

$$E = \frac{3}{2} NkT^* . \tag{C.7}$$

La seconde n'est autre que *l'équation d'état des gaz parfaits :*

$$p^*V = NkT^* . \tag{C.8}$$

3. MÉLANGE IDÉAL DE DEUX GAZ PARFAITS

a. Entropie, température et pression

Considérons un gaz parfait constitué de deux corps purs différents : dans un récipient de volume V sont mélangés N_1 atomes de type (1), de masse m_1, et N_2 atomes de types (2), de masse m_2, *sans interactions.* C'est un *mélange idéal* des gaz parfaits (1) et (2).

Si le système est isolé, l'énergie totale

$$E = E_1 + E_2 \tag{C.9}$$

est fixée (à δE près) alors que les énergies E_1 et E_2 des sous-systèmes (1) et (2) (constitués respectivement par les N_1 particules (1) et les N_2 particules (2)) sont des variables internes libres de fluctuer.

Lorsque l'énergie du sous-système (1) est fixée à E_1 (à δE_1 près), le nombre d'états microscopiques du système global vaut (II, formule (III.15))

$$\omega^D(E, E_1)\delta E_1 = \rho_1^D(E_1, V, N_1)\rho_2^D(E_2 = E - E_1, V, N_2)\delta E \delta E_1 , \tag{C.10}$$

où $\rho_1^D(E_1, V, N_1)$ et $\rho_2^D(E_2, V, N_2)$ sont les densités d'états des sous-systèmes (1) et

2. Pour obtenir (C.4), nous gardons les termes d'ordre N et négligeons ceux d'ordre $\ln N$ ou inférieur.

(2). On en déduit la densité d'états microscopiques du système global en sommant sur toutes les valeurs possibles de la variable interne E_1 :

$$\Omega^D(E, V, N_1, N_2) \widehat{=} \rho^D(E, V, N_1, N_2)\,\delta E = \delta E \int_0^E \rho_1^D(E_1, V, N_1)\rho_2^D(E - E_1, V, N_2)\,\mathrm{d}E_1 . \qquad \text{(C.11)}$$

Si l'on remplace ρ_1^D et ρ_2^D par leur expression tirée de (C.2), l'intégrale se calcule aisément[3] et on obtient

$$\rho^D(E, V, N_1, N_2) = \left(\frac{m_1}{2\pi\hbar^2}\right)^{3N_1/2} \left(\frac{m_2}{2\pi\hbar^2}\right)^{3N_2/2} \frac{V^{N_1+N_2} E^{3(N_1+N_2)/2-1}}{\Gamma\left(\frac{3}{2}(N_1+N_2)\right)} . \qquad \text{(C.12)}$$

Pour $N_1 \gg 1$ et $N_2 \gg 1$, l'entropie du système formé d'un mélange idéal de gaz parfaits s'écrit donc

$$S^{*D}(E, V, N_1, N_2) = N_1 k \left[\frac{3}{2} \ln \frac{m_1}{3\pi\hbar^2} + \ln V + \frac{3}{2}\right] + N_2 k \left[\frac{3}{2} \ln \frac{m_2}{3\pi\hbar^2} + \ln V + \frac{3}{2}\right] + \frac{3}{2}(N_1 + N_2) k \ln \frac{E}{N_1 + N_2} . \qquad \text{(C.13)}$$

On en déduit la température T^* :

$$\frac{1}{T^*} = \frac{3}{2} k \frac{N_1 + N_2}{E} , \qquad \text{(C.14)}$$

et la pression p^* :

$$p^* = \frac{N_1 + N_2}{V} kT^* . \qquad \text{(C.15)}$$

On obtient donc les mêmes expressions qu'en (C.7) et (C.8), N étant remplacé par le nombre total d'atomes $(N_1 + N_2)$, indépendamment des différences d'espèces

---

3. Il vient en effet

$$\rho^D(E, V, N_1, N_2) = \left(\frac{m_1}{2\pi\hbar^2}\right)^{3N_1/2} \left(\frac{m_2}{2\pi\hbar^2}\right)^{3N_2/2} \frac{V^{N_1+N_2}}{\Gamma\left(\frac{3N_1}{2}\right)\Gamma\left(\frac{3N_2}{2}\right)} \times I , \qquad (1)$$

avec $$I = \int_0^E E_1^{3N_1/2-1} (E - E_1)^{3N_2/2-1} \,\mathrm{d}E_1 . \qquad (2)$$

Le changement de variable

$$x \widehat{=} \frac{E_1}{E} \qquad (3)$$

donne

$$I = E^{3(N_1+N_2)/2-1} \int_0^1 x^{3N_1/2-1} (1 - x)^{3N_2/2-1} \,\mathrm{d}x , \qquad (4)$$

d'où (L. Schwartz, *Méthodes mathématiques pour les sciences physiques*, Hermann, Paris (1961), p. 348)

$$I = E^{3(N_1+N_2)/2-1} \times \frac{\Gamma\left(\frac{3N_1}{2}\right)\Gamma\left(\frac{3N_2}{2}\right)}{\Gamma\left(\frac{3}{2}(N_1+N_2)\right)} . \qquad (5)$$

chimiques. Il est d'usage de définir la *pression partielle* d'un constituant dans le mélange comme la pression qu'aurait ce constituant s'il occupait seul le même volume à la même température : ici, la pression partielle p_α^* du constituant (α) s'écrit évidemment

$$p_\alpha^* = \frac{N_\alpha}{V} kT^* , \quad \alpha = 1, 2 . \tag{C.16}$$

La relation (C.15) exprime alors que la pression totale du mélange de gaz est la somme des pressions partielles de ses constituants *(loi de Dalton).* Soulignons que cette loi n'est valable que pour des mélanges idéaux (où les interactions entre les gaz constituants sont négligeables).

b. Entropie de mélange

Supposons que, pour obtenir le mélange de gaz parfaits décrit au paragraphe précédent, nous ayons procédé comme suit. Dans un premier temps, le récipient de volume V, isolé dans son ensemble, est séparé en deux compartiments par une cloison diatherme et coulissant sans frottement, mais imperméable aux particules (fig. 1). Dans la partie gauche sont enfermés les N_1 atomes de type (1) et dans la partie droite les N_2 atomes de type (2). Comme ces deux systèmes échangent de la chaleur et du volume, leurs énergies E_1 et E_2, ainsi que leurs volumes V_1 et V_2, sont des variables internes fluctuantes. On retire ensuite la cloison; le système évolue alors vers un nouvel état d'équilibre où les deux types d'atomes sont répartis dans l'ensemble du volume V.

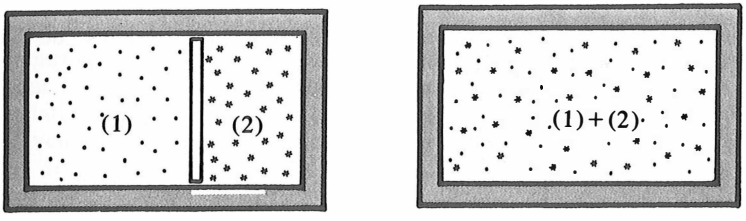


FIGURE 1
Dans l'état initial, les deux gaz (1) et (2) sont séparés par une cloison diatherme et mobile. Lorsqu'on retire la cloison, les atomes (1) et (2) se répartissent dans l'ensemble du volume V.

Calculons d'abord l'entropie du système *dans l'état initial,* lorsque les deux gaz sont séparés. A l'équilibre thermique, la répartition la plus probable de l'énergie entre les deux sous-systèmes correspond à l'*égalité de leurs températures microcanoniques* (II, § III.A.2.c). Ici la relation entre énergie et température est connue (formule (C.7)); les valeurs les plus probables E_{1m} et E_{2m} des énergies des deux sous-systèmes sont donc données par

$$\frac{E_{1m}}{E_{2m}} = \frac{N_1}{N_2} , \tag{C.17}$$

avec $$E_{1m} + E_{2m} = E . \tag{C.18}$$

De même, les valeurs les plus probables V_{1m} et V_{2m} des volumes des deux compartiments, qui vérifient évidemment

$$V_{1m} + V_{2m} = V , \tag{C.19}$$

sont déterminées par l'*égalité des pressions* de part et d'autre de la cloison (II, § III.B.3.a), qui implique d'après (C.8)

$$\frac{V_{1m}}{V_{2m}} = \frac{N_1}{N_2} . \tag{C.20}$$

Si les sous-systèmes sont suffisamment grands, on peut négliger les fluctuations des variables internes E_1, E_2, V_1 et V_2 autour de leur valeur la plus probable et l'entropie du système s'écrit (voir exercice II.19)

$$S_i^{*D} = S_1^{*D}(E_{1m}, V_{1m}, N_1) + S_2^{*D}(E_{2m}, V_{2m}, N_2)\,, \tag{C.21}$$

c'est-à-dire d'après (C.4)

$$S_i^{*D} = N_1 k\left[\ln V_{1m} + \frac{3}{2}\ln\left(\frac{m_1}{3\pi\hbar^2}\frac{E_{1m}}{N_1}\right) + \frac{3}{2}\right] + N_2 k\left[\ln V_{2m} + \frac{3}{2}\ln\left(\frac{m_2}{3\pi\hbar^2}\frac{E_{2m}}{N_2}\right) + \frac{3}{2}\right]. \tag{C.22}$$

Comparons cette expression à l'entropie S_f^{*D} que possède le système total *après le mélange* des deux gaz (1) et (2). Lorsqu'on retire la cloison, le système dans son ensemble reste isolé et son énergie totale E n'est pas modifiée. L'entropie S_f^{*D} cherchée est donc simplement

$$S_f^{*D} = S^{*D}(E, V, N_1, N_2) \tag{C.23}$$

donnée par (C.13).

Le mélange des deux gaz, processus spontané qui se produit après relâchement d'une contrainte dans le système global isolé, doit s'accompagner d'une *augmentation d'entropie*

$$\Delta S^{*D} \widehat{=} S_f^{*D} - S_i^{*D} > 0 \tag{C.24}$$

que l'on appelle *« entropie de mélange »*. On trouve facilement, à partir de (C.13) et (C.22)[4],

$$\begin{aligned}\Delta S^{*D} &= N_1 k \ln\frac{N_1+N_2}{N_1} + N_2 k \ln\frac{N_1+N_2}{N_2} \\ &= k\left[(N_1+N_2)\ln(N_1+N_2) - N_1\ln N_1 - N_2\ln N_2\right].\end{aligned} \tag{C.25}$$

On constate effectivement que l'entropie de mélange est *toujours positive*. Notons que *sa valeur ne dépend pas de la nature des deux gaz*, c'est-à-dire des masses m_1 et m_2. Cette augmentation d'entropie est liée à une *perte d'information* sur le système : on ne sait plus maintenant où se trouvent les atomes (1) et (2). Pour revenir à l'état initial, il ne suffit pas de replacer la cloison, il faut en plus *trier* les atomes pour remettre chaque espèce dans son compartiment d'origine.

Il est intéressant de remarquer que, contrairement à l'entropie, *la température du système et sa pression ne sont pas modifiées par le mélange des deux gaz parfaits.*

---

4. Il suffit de remarquer que

$$\frac{E_{1m}}{N_1} = \frac{E_{2m}}{N_2} = \frac{E}{N_1+N_2} \tag{1}$$

pour réécrire (C.22) sous la forme

$$S_i^{*D} = N_1 k\left[\frac{3}{2}\ln\frac{m_1}{3\pi\hbar^2} + \ln V_{1m} + \frac{3}{2}\right] + N_2 k\left[\frac{3}{2}\ln\frac{m_2}{3\pi\hbar^2} + \ln V_{2m} + \frac{3}{2}\right] + \frac{3}{2}(N_1+N_2)\,k\ln\frac{E}{N_1+N_2}\,. \tag{2}$$

Le résultat (C.25) découle alors aussitôt des relations

$$\frac{V_{1m}}{N_1} = \frac{V_{2m}}{N_2} = \frac{V}{N_1+N_2}\,. \tag{3}$$

En effet, dans l'état initial, les valeurs les plus probables E_{1m} et E_{2m} des énergies des deux sous-systèmes sont fixées par la condition

$$T_i^* = \frac{2}{3}\frac{E_{1m}}{N_1 k} = \frac{2}{3}\frac{E_{2m}}{N_2 k} . \tag{C.26}$$

Quant à la température T_f^* dans l'état final, elle est donnée par (C.14) :

$$T_f^* = \frac{2E}{3(N_1+N_2)k} . \tag{C.27}$$

Comme

$$\frac{E_{1m}}{N_1} = \frac{E_{2m}}{N_2} = \frac{E}{N_1+N_2} , \tag{C.28}$$

on voit que

$$T_f^* = T_i^* . \tag{C.29}$$

Un raisonnement analogue, sur les volumes cette fois, conduit à l'égalité des pressions initiale et finale.

4. LE PARADOXE DE GIBBS

a. Énoncé

Supposons maintenant que les atomes (1) et (2) soient identiques, et reprenons l'expérience. Lorsqu'on retire la cloison, les particules (de même masse m) diffusent d'un compartiment dans l'autre. Il suffit de remplacer m_1 et m_2 par m dans (C.22) et (C.13) pour obtenir les expressions de l'entropie dans l'état initial et dans l'état final. L'augmentation d'entropie ΔS^{*D} est toujours donnée par (C.25) : *l'entropie de mélange* que nous venons de calculer *ne s'annule pas lorsque les atomes sont de même espèce.* Pourtant, si l'on replace la cloison, *rien ne permet de distinguer à l'échelle macroscopique ce nouvel état de l'état initial;* l'expérience n'a provoqué aucune perte d'information, car il est impossible de reconnaître les atomes pour constater que ceux qui se trouvent maintenant dans un compartiment ne sont pas vraiment ceux qui s'y trouvaient dans l'état initial.

Exprimons la même idée d'une autre manière. Comme les atomes n'interagissent pas, on peut considérer que le système est constitué dans l'état final de deux sous-systèmes indépendants, comprenant l'un les N_1 atomes situés le plus à gauche et l'autre les N_2 atomes situés le plus à droite (même si la cloison n'a pas été remise en place). L'énergie et le volume de ces deux sous-systèmes sont évidemment des variables internes dont les valeurs les plus probables sont les mêmes que plus haut. Si on raisonne ainsi, on trouve que l'entropie du système vaut

$$S^{*D} = S^{*D}(E_{1m}, V_{1m}, N_1) + S^{*D}(E_{2m}, V_{2m}, N_2) \tag{C.30}$$

puisque l'entropie est une grandeur additive pour des systèmes faiblement couplés (complément II.A, § II). Si l'on explicite le second membre selon la formule (C.4), cette expression s'écrit

$$S^{*D} \overset{?}{=} N_1 k\left[\ln V_{1m} + \frac{3}{2}\ln\left(\frac{m}{3\pi\hbar^2}\frac{E_{1m}}{N_1}\right) + \frac{3}{2}\right] + N_2 k\left[\ln V_{2m} + \frac{3}{2}\ln\left(\frac{m}{3\pi\hbar^2}\frac{E_{2m}}{N_2}\right) + \frac{3}{2}\right] . \tag{C.31}$$

Mais si on calcule directement S^{*D} en appliquant (C.4) à l'ensemble du système, on trouve

$$S^{*D} \overset{?}{=} (N_1+N_2)k\left[\ln V + \frac{3}{2}\ln\left(\frac{m}{3\pi\hbar^2}\frac{E}{N_1+N_2}\right) + \frac{3}{2}\right] . \tag{C.32}$$

Les résultats (C.31) et (C.32) sont *contradictoires :* ils diffèrent par l'entropie de mélange ΔS^{*D}. Cette contradiction constitue le *paradoxe de Gibbs :* suivant que l'on considère le système comme un ensemble de $(N_1 + N_2)$ particules enfermées dans le volume V ou comme la juxtaposition de deux sous-systèmes indépendants (N_1, V_{1m}) et (N_2, V_{2m}), on n'obtient pas la même valeur pour son entropie.

b. Origine du paradoxe

Il est clair d'après ce qui précède que le paradoxe apparaît *lorsque les atomes contenus dans les deux compartiments sont identiques.* En effet, si les particules sont différentes, l'état macroscopique du système n'est pas le même lorsque les atomes (1) et (2) sont séparés par une cloison ou lorsqu'ils sont mélangés; il n'y a pas alors de paradoxe. En revanche, si les molécules sont identiques, on peut enlever puis remettre la cloison sans changer l'état macroscopique du gaz, mais le calcul de son entropie donne des résultats différents si l'on considère que la cloison est en place ou au contraire qu'elle a été retirée. *C'est donc la façon dont nous avons traité jusqu'ici les systèmes de particules identiques qui est à revoir.*

En mécanique quantique, l'identité des particules est prise en compte par le *postulat de symétrisation, qui change radicalement le dénombrement des états et donc l'entropie du système.* La solution exacte de ce problème sera donnée au chapitre V, paragraphe IV; nous nous contenterons ici d'un résultat approché (l'approximation de Maxwell-Boltzmann), valable lorsqu'on peut négliger la probabilité pour que deux particules se trouvent dans le même état individuel.

II. SYSTÈME DE PARTICULES INDÉPENDANTES ET INDISCERNABLES. APPROXIMATION DE MAXWELL-BOLTZMANN

1. ÉVALUATION DE LA DENSITÉ D'ÉTATS

Il s'agit de compter, pour un gaz parfait, les états microscopiques dont l'énergie est inférieure à une valeur donnée E.

Considérons un ensemble de N vecteurs d'onde permis $\vec{k}_1, \vec{k}_2, \ldots, \vec{k}_N$. Si, comme nous l'avons supposé au paragraphe I, les particules étaient discernables, elles pourraient être numérotées de 1 à N, et cette numérotation permettrait de les distinguer. Dans ce cas, l'état du système où la particule (1) a l'impulsion $\hbar\vec{k}_1$ et la particule (2) l'impulsion $\hbar\vec{k}_2$ serait distinct de celui où c'est la particule (2) qui possède $\hbar\vec{k}_1$ et la particule (1) $\hbar\vec{k}_2$. Au contraire, lorsqu'on applique le postulat de symétrisation (appendice II, § IV.2), ces deux états ne sont pas physiquement distincts : *un état physique du système est caractérisé par un ensemble non ordonné de N vecteurs d'onde permis.*

Comment dès lors rectifier le dénombrement effectué au paragraphe I? Lorsque les N vecteurs d'onde $\vec{k}_i$ sont tous différents, la réponse est simple : à cet ensemble étaient associés $N!$ états distincts dans le cas de particules discernables; il n'y en a plus qu'un seul maintenant. Les choses se compliquent cependant pour les ensembles comprenant plusieurs vecteurs d'onde égaux[5].

5. Considérons par exemple un ensemble de N vecteurs d'onde permis dans lequel deux d'entre eux sont égaux, tous les autres étant différents. A cet ensemble (non ordonné) correspondraient dans le cas discernable $N!/2!$ états distincts (nombre de permutations distinctes de l'ensemble des N vecteurs considérés), et non plus $N!$ comme précédemment. Le problème de la symétrisation se présente de façon différente suivant que les particules sont des fermions ou des bosons (cf. III, § IV.A.3); mais la difficulté provient dans les deux cas du fait que la façon dont il faut corriger le nombre d'états discernables varie d'un ensemble $(\vec{k}_1, \vec{k}_2, \ldots, \vec{k}_N)$ à l'autre suivant le nombre de vecteurs égaux figurant dans chacun de ces ensembles.

Mais si nous supposons que le nombre d'états individuels accessibles est très grand devant N, la probabilité pour que deux particules se trouvent dans le même état individuel est négligeable; nous n'avons alors à considérer pratiquement que le cas où les N vecteurs d'onde $\vec{k}_i$ caractérisant un état du système sont diffférents les uns des autres. Cette approximation, dont nous préciserons plus loin les conditions de validité, est appelée *« approximation de Maxwell-Boltzmann »*.

Le nombre d'états $\Phi(E)$ dont l'énergie est inférieure à E est dans ce cas facile à obtenir : il suffit de diviser le nombre d'états calculé pour des particules discernables par le nombre de permutations des N vecteurs d'onde[6] :

$$\Phi(E,\ V,\ N)=\frac{1}{N!}\,\Phi^{D}(E,\ V,\ N)\ . \tag{C.33}$$

La densité d'états s'écrit de même

$$\rho(E,\ V,\ N)=\frac{1}{N!}\,\rho^{D}(E,\ V,\ N)\ . \tag{C.34}$$

2. EXPRESSION DES PROPRIÉTÉS PHYSIQUES DES GAZ PARFAITS

a. Entropie

La densité d'états ayant été modifiée, l'entropie l'est également. La formule (C.34) donne immédiatement

$$S^{*}(E,\ V,\ N)=S^{*D}(E,\ V,\ N)-k\ \ln N!\ \tag{C.35}$$

ou encore, si on utilise (C.4) et la formule de Stirling,

$$S^{*}(E,\ V,\ N)=Nk\left[\ln\frac{V}{N}+\frac{3}{2}\ln\left(\frac{m}{3\pi\hbar^{2}}\,\frac{E}{N}\right)+\frac{5}{2}\right]\ . \tag{C.36}$$

Cette expression de l'entropie, valable pour des *particules indiscernables,* est la *formule de Sackur-Tetrode.* L'entropie ainsi calculée est extensive[7] (complément II.A, § II). On vérifie facilement qu'elle ne souffre plus du défaut que nous avons mis en évidence plus haut : si on utilise (C.36) plutôt que (C.4) pour calculer les expressions analogues

---

6. Si les particules sont des fermions, la densité d'états exacte est inférieure à $\rho^{D}(E,\ V,\ N)/N!$ puisque un certain nombre d'états sont interdits par le principe de Pauli. Pour les bosons, au contraire, l'approximation de Maxwell-Boltzmann surcorrige la contribution des états dans lesquels deux ou plusieurs vecteurs d'onde sont égaux (cf. V, note 27).

7. La formule de Sackur-Tetrode est de la forme

$$S^{*}(E,\ V,\ N)=N\,f\left(\frac{E}{N},\ \frac{V}{N}\right)\ ,$$

où l'on voit clairement que l'entropie est extensive : si on multiplie par λ la taille du système (c'est-à-dire si on multiplie par un même facteur λ les trois quantités E, V, N), l'entropie est aussi multipliée par λ.

Notons que, en revanche, l'expression (C.4) obtenue en traitant les particules comme discernables n'est pas extensive (ce qui est une autre façon d'exprimer le paradoxe de Gibbs).

à (C.31) et (C.32), celles-ci sont maintenant compatibles[8] : les deux façons de calculer l'entropie des $(N_1 + N_2)$ particules donnent alors le même résultat. *Le paradoxe de Gibbs est résolu*[9].

b. Équation d'état. Entropie de mélange

La nouvelle expression (C.36) de l'entropie n'affecte pas l'équation d'état du gaz (§ I.2), puisque le terme supplémentaire ne dépend pas de V. La relation $p^*V = NkT^*$ reste donc valable.

L'entropie de mélange calculée au paragraphe I.3.b pour deux espèces différentes garde, elle aussi, la même expression[10] :

$$\Delta S^* = N_1 k \ln \frac{N_1 + N_2}{N_1} + N_2 k \ln \frac{N_1 + N_2}{N_2} . \tag{C.37}$$

Notons que ce calcul n'est valable que pour des particules *d'espèces différentes :* il n'est pas possible de considérer un gaz parfait de N particules identiques occupant le volume V comme la réunion de deux sous-systèmes de N_1 et N_2 particules occupant le même volume V.

c. Potentiel chimique

Le potentiel chimique μ^* est défini par l'expression (II, § III.B.1)

$$\frac{\mu^*}{T^*} \widehat{=} -\frac{\partial S^*}{\partial N} , \tag{C.38}$$

---

8. La formule de la note précédente implique que

$$S^*(E_{1m} + E_{2m},\ V_{1m} + V_{2m},\ N_1 + N_2) = (N_1 + N_2)\ f\left(\frac{E_{1m} + E_{2m}}{N_1 + N_2},\ \frac{V_{1m} + V_{2m}}{N_1 + N_2}\right) \tag{1}$$

est bien égale, compte tenu des relations

$$\frac{E_{1m} + E_{2m}}{N_1 + N_2} = \frac{E_{1m}}{N_1} = \frac{E_{2m}}{N_2} \tag{2}$$

et

$$\frac{V_{1m} + V_{2m}}{N_1 + N_2} = \frac{V_{1m}}{N_1} = \frac{V_{2m}}{N_2} , \tag{3}$$

à la somme

$$N_1\ f\left(\frac{E_{1m}}{N_1}, \frac{V_{1m}}{N_1}\right) + N_2\ f\left(\frac{E_{2m}}{N_2}, \frac{V_{2m}}{N_2}\right) = S^*(E_{1m},\ V_{1m},\ N_1) + S^*(E_{2m},\ V_{2m},\ N_2) . \tag{4}$$

9. Dans un cadre purement *classique,* l'état d'une particule est décrit par un point de l'espace des phases (I, § III.C.1). Les coordonnées de ce point variant de façon continue, la probabilité pour que deux particules occupent le même état est alors rigoureusement nulle et le facteur $1/N!$ devient exact. C'est ainsi que Gibbs avait résolu le paradoxe qui porte son nom pour des particules classiques indiscernables.

10. Reprenons en effet le calcul du paragraphe I.3.b sur la base de la formule de Sackur-Tetrode. L'expression de S_f^* tirée de (C.13) est remplacée par

$$S_f^* = N_1 k \left[\ln \frac{V}{N_1} + \frac{3}{2} \ln \frac{m_1}{3\pi\hbar^2} + \frac{5}{2}\right] + N_2 k \left[\ln \frac{V}{N_2} + \frac{3}{2} \ln \frac{m_2}{3\pi\hbar^2} + \frac{5}{2}\right] + \frac{3}{2}(N_1 + N_2) k \ln \frac{E}{N_1 + N_2} \tag{1}$$

et la formule (C.22) par

$$S_i^* = N_1 k \left[\ln \frac{V_{1m}}{N_1} + \frac{3}{2} \ln \left(\frac{m_1}{3\pi\hbar^2} \frac{E_{1m}}{N_1}\right) + \frac{5}{2}\right] + N_2 k \left[\ln \frac{V_{2m}}{N_2} + \frac{3}{2} \ln \left(\frac{m_2}{3\pi\hbar^2} \frac{E_{2m}}{N_2}\right) + \frac{5}{2}\right] , \tag{2}$$

mais leur différence donne encore (C.25).

qui donne ici

$$\mu^* = -kT^* \ln\left[\left(\frac{m}{3\pi\hbar^2}\frac{E}{N}\right)^{3/2}\frac{V}{N}\right]$$

$$= -kT^* \ln\left[\left(\frac{mkT^*}{2\pi\hbar^2}\right)^{3/2}\frac{V}{N}\right]. \qquad \text{(C.39)}$$

Le résultat aurait évidemment été différent si nous étions partis de l'entropie S^{*D} obtenue pour des particules discernables, puisque la dépendance en N de l'entropie a été modifiée.

3. DOMAINE DE VALIDITÉ DE L'APPROXIMATION DE MAXWELL-BOLTZMANN

L'expression (C.34) de la densité d'états d'un gaz parfait de N particules indiscernables est seulement approchée : elle suppose que deux particules du gaz n'occupent pratiquement jamais le même état individuel. Pour délimiter le domaine de validité de cette approximation[11], nous allons écrire que la probabilité d'occupation de chacun des états individuels est très faible; la probabilité de double occupation sera alors négligeable[12]. Nous montrerons ainsi que l'approximation de Maxwell-Boltzmann est applicable si l'énergie par particule est suffisamment grande ou la densité de particules suffisamment petite.

Considérons donc un état individuel particulier, caractérisé par le vecteur d'onde $\vec{k}$; l'énergie associée à cet état est

$$\varepsilon = \frac{\hbar^2 \vec{k}^2}{2m}. \qquad \text{(C.40)}$$

La probabilité $P_1(\vec{k})$ pour que cet état individuel soit occupé par une particule (au moins) s'obtient en comptant combien, parmi les $\Omega(E, V, N)$ états microscopiques accessibles au système, comportent au moins une particule ayant le vecteur d'onde $\vec{k}$. Or leur nombre est évidemment égal à $\Omega(E-\varepsilon, V, N-1)$, nombre d'états accessibles aux $(N-1)$ autres particules lorsque l'état $(\vec{k})$ (d'énergie ε) de l'une d'entre elles est fixé. On a donc

$$P_1(\vec{k}) = \frac{\Omega(E-\varepsilon, V, N-1)}{\Omega(E, V, N)}, \qquad \text{(C.41)}$$

c'est-à-dire, d'après (C.34) et (C.2),

$$P_1(\vec{k}) = \frac{N!}{(N-1)!}\left(\frac{2\pi\hbar^2}{m}\right)^{3/2}\frac{\Gamma(3N/2)}{\Gamma\left(3(N-1)/2\right)}\frac{1}{V}\frac{(E-\varepsilon)^{\frac{3(N-1)}{2}-1}}{E^{\frac{3N}{2}-1}}. \qquad \text{(C.42)}$$

---

11. Nous cherchons uniquement ici à justifier *a posteriori* l'approximation de Maxwell-Boltzmann, en montrant dans quelles conditions les résultats qu'elle fournit sont cohérents. Pour démontrer véritablement la validité de cette approximation, il nous faudra attendre de connaître le résultat exact du décompte des états (V, § IV.C.1).

12. On peut d'ailleurs calculer, par la même méthode que ci-dessous pour $P_1(\vec{k})$, la probabilité $P_2(\vec{k})$ pour que deux particules au moins se trouvent dans le même état individuel $(\vec{k})$. On s'assure ainsi que les conditions

$$\frac{P_2(\vec{k})}{P_1(\vec{k})} \ll 1$$

et $\quad P_1(\vec{k}) \ll 1$

sont équivalentes.

N étant grand, on peut utiliser la formule de Stirling pour calculer le rapport des deux fonctions Γ, ce qui donne[13]

$$P_1(\vec{k}) = \left(\frac{3\pi\hbar^2}{m}\right)^{3/2} \left(\frac{N}{E}\right)^{3/2} \frac{N}{V} \left(1 - \frac{\varepsilon}{E}\right)^{\frac{3N-5}{2}} . \tag{C.43}$$

La valeur maximum de $P_1(\vec{k})$ étant obtenue pour $\varepsilon = 0$ (c'est-à-dire $\vec{k} = \vec{0}$), la condition de validité de l'approximation de Maxwell-Boltzmann s'écrit

$$\left(\frac{3\pi\hbar^2}{m} \frac{N}{E}\right)^{3/2} \frac{N}{V} \ll 1 . \tag{C.44}$$

Il faut donc que E/N soit suffisamment grand et/ou N/V suffisamment petit. Si l'on exprime E/N en fonction de la température T^* selon (C.7), on transforme (C.44) en

$$\left(\frac{2\pi\hbar^2}{mkT^*}\right)^{3/2} \frac{N}{V} \ll 1 . \tag{C.45}$$

Nous retrouverons cette condition au chapitre III (§ IV.B.2.b).

---

13. On sait que, pour $n \gg 1$,

$$\Gamma(n+1) \simeq n^n e^{-n} \sqrt{2\pi n} . \tag{1}$$

On en déduit

$$\frac{\Gamma(3N/2)}{\Gamma(3(N-1)/2)} \simeq \left(\frac{3N-2}{3N-5}\right)^{3N/2-5/2} \left(\frac{3}{2} N - 1\right)^{3/2} e^{-3/2} \sqrt{\frac{3N-2}{3N-5}} . \tag{2}$$

Or
$$\frac{3N-2}{3N-5} \simeq 1 , \tag{3}$$

$$\left(\frac{3}{2} N - 1\right)^{3/2} \simeq \left(\frac{3N}{2}\right)^{3/2} . \tag{4}$$

On montre enfin en prenant le logarithme que

$$\left(\frac{3N-2}{3N-5}\right)^{3N/2-5/2} \simeq e^{3/2} , \tag{5}$$

de sorte que

$$\frac{\Gamma(3N/2)}{\Gamma(3(N-1)/2)} \simeq \left(\frac{3}{2} N\right)^{3/2} . \tag{6}$$

COMPLÉMENT II.D

Transformations thermodynamiques et variations d'entropie

I. Paramètres extérieurs et entropie
  1. Force conjuguée d'un paramètre extérieur
  2. Variation de l'entropie microcanonique par modification d'un paramètre extérieur
    a. Cas où la force conjuguée est la même dans tous les états accessibles
    b. Cas général
    c. Cas d'un système macroscopique
    d. Cas d'un système purement mécanique
  3. Exemple : la pression, force moyenne conjuguée du volume
  4. Équilibre entre systèmes pouvant échanger un paramètre extérieur
    a. Condition générale d'équilibre
    b. Équilibre avec un système purement mécanique

II. Transformations d'un système
  1. Définition générale
  2. Énergie échangée au cours d'une transformation
    a. Travail reçu par le système
    b. Énergie « chimique » reçue par le système
    c. Chaleur reçue par le système
    d. Le « premier principe » de la thermodynamique
  3. Transformations particulières
    a. Transformation adiabatique
    b. Transformation quasistatique
    c. Transformation réversible

III. Variation d'entropie dans certaines classes de transformations
  1. Expression générale de l'accroissement d'entropie dans une transformation infinitésimale
  2. Variation d'entropie dans une transformation réversible
  3. Variation d'entropie dans une transformation adiabatique
    a. Transformation adiabatique quelconque
    b. Application à une transformation adiabatique infinitésimale mettant en jeu un seul paramètre extérieur
    c. Cas d'une transformation adiabatique réversible

Après la découverte de la « force motrice du feu », l'analyse du fonctionnement des *moteurs thermiques,* destinés à transformer de la chaleur en travail mécanique, a conduit au cours du dix-neuvième siècle à la mise au point d'une théorie physique extrêmement élaborée, la *thermodynamique.* Celle-ci, dans sa forme finale, est fondée sur un petit nombre de *principes* (essentiellement deux), d'où découlent des résultats d'une diversité et d'une portée étonnantes.

Pourtant la *signification physique des principes de la thermodynamique,* particulièrement celle du « second principe », et la véritable nature de la chaleur restaient mystérieuses. C'est la *mécanique statistique,* développée à partir de 1870, qui allait fournir la compréhension physique des bases de la thermodynamique, en reliant le comportement des systèmes macroscopiques à leurs propriétés mécaniques à l'échelle microscopique (cf. chapitre I, § I).

Nous nous proposons, dans ce complément puis dans le complément III.L qui lui fait suite, de montrer que *les « principes » de la thermodynamique apparaissent comme des conséquences de la mécanique statistique.* Nous commençons ici par introduire certaines notions indispensables, notamment celle de *« transformations thermodynamiques »* d'un système, et par étudier les *variations d'entropie* qu'elles induisent. Nous en tirerons un résultat essentiel : *l'entropie d'un système macroscopique ne peut qu'augmenter au cours d'une transformation adiabatique.* Ce résultat sera ensuite exploité dans le complément III.L : la notion de thermostat (III, § I.A.1) nous permettra d'en déduire l'expression générale du « second principe » et de préciser la nature physique de la chaleur.

I. PARAMÈTRES EXTÉRIEURS ET ENTROPIE

Étudions en premier lieu comment le nombre d'états accessibles à un système isolé, et par conséquent son entropie microcanonique, dépendent des valeurs fixées pour les paramètres extérieurs.

1. FORCE CONJUGUÉE D'UN PARAMÈTRE EXTÉRIEUR

L'état macroscopique d'un *système isolé* est défini par la donnée de son énergie E, du nombre N de particules qu'il contient, et d'un certain nombre d'autres paramètres extérieurs que nous noterons $x_1, x_2, \ldots, x_n$. La connaissance des x_i permet en principe de résoudre l'équation aux valeurs propres de l'hamiltonien H du système, c'est-à-dire de déterminer ses états microscopiques possibles (ℓ) et les énergies E_ℓ correspondantes. *Les états (ℓ) et les énergies associées E_ℓ dépendent des x_i;* nous admettrons que ce sont des *fonctions continues*[1], [2] de ces paramètres, et écrirons en particulier

$$E_\ell = E_\ell(x_1, x_2, \ldots, x_n) \equiv E_\ell(x) \,. \tag{D.1}$$

La *force conjuguée* X_i du paramètre extérieur x_i est alors une *variable interne* définie par la donnée de ses valeurs possibles

$$X_{i,\ell} \mathrel{\widehat{=}} -\frac{\partial E_\ell}{\partial x_i} \tag{D.2}$$

dans chacun des états (ℓ) accessibles au système isolé que nous étudions. La distribution statistique de X_i s'obtient par la méthode exposée au chapitre II, paragraphe II.A, valable pour toute variable interne : la probabilité $\mathrm{d}P(X_i)$ pour que la force conjuguée de x_i prenne une valeur comprise entre X_i et $X_i + \mathrm{d}X_i$ s'écrit

$$\mathrm{d}P(X_i) = \frac{\omega(E, x; X_i)\,\mathrm{d}X_i}{\Omega(E, x)} \,, \tag{D.3}$$

où $\omega(E, x; X_i)\,\mathrm{d}X_i$ est le nombre d'états (ℓ) accessibles qui vérifient la condition

$$X_i \leqslant X_{i,\ell} \leqslant X_i + \mathrm{d}X_i \,, \tag{D.4}$$

---

1. Par exemple, l'état d'une particule libre enfermée dans une boîte parallélépipédique de côtés L_x, L_y, L_z est repéré par trois entiers relatifs ℓ_x, ℓ_y, ℓ_z (appendice II, § II.4). L'énergie correspondante

$$E_{\ell_x, \ell_y, \ell_z} = \frac{2\pi^2\hbar^2}{m}\left(\frac{\ell_x^2}{L_x^2} + \frac{\ell_y^2}{L_y^2} + \frac{\ell_z^2}{L_z^2}\right)$$

est bien une fonction continue des paramètres extérieurs L_x, L_y, L_z. Il en est de même pour les énergies d'un système de N particules libres indépendantes enfermées dans une boîte (I, § III.B.2.b).

2. Les états d'un système ne sont pas des fonctions continues du nombre N de particules (voir l'exemple simple de la note 1). N joue donc ici un rôle particulier et c'est pourquoi nous ne l'incluons pas dans l'ensemble des x_i.

et $\Omega(E, x)$ le nombre total d'états accessibles :

$$\Omega(E, x) = \int_{-\infty}^{+\infty} \omega(E, x; X_i)\, \mathrm{d}X_i \,. \tag{D.5}$$

On en déduit la *force moyenne conjuguée* $\overline{X_i}$ du paramètre extérieur x_i :

$$\overline{X_i} = \frac{1}{\Omega(E, x)} \int_{-\infty}^{+\infty} X_i\, \omega(E, x; X_i)\, \mathrm{d}X_i \,. \tag{D.6}$$

La force moyenne $\overline{X_i}$ conjuguée de x_i est elle-même fonction de l'ensemble des paramètres extérieurs (y compris l'énergie et le nombre de particules).

La force conjuguée est une *« force généralisée »* : c'est seulement si x_i a les dimensions d'une longueur que X_i est la composante d'une force au sens de la mécanique; mais le produit $x_i X_i$ d'un paramètre extérieur par sa force conjuguée a toujours les dimensions d'une énergie.

2. VARIATION DE L'ENTROPIE MICROCANONIQUE PAR MODIFICATION D'UN PARAMÈTRE EXTÉRIEUR

Nous cherchons à préciser comment le nombre total $\Omega(E, x)$ d'états accessibles dépend des paramètres extérieurs x_i. Pour la simplicité de l'exposé, nous considérons un seul paramètre extérieur x, et nous notons X sa force conjuguée.

a. Cas où la force conjuguée est la même dans tous les états accessibles

Pour bien comprendre le raisonnement, commençons par un cas simple; supposons que tous les X_ℓ soient égaux[3] :

$$X_\ell = X \text{ quel que soit } (\ell) \,. \tag{D.7}$$

Alors, si le paramètre extérieur x est augmenté de $\mathrm{d}x$, au premier ordre en $\mathrm{d}x$ toutes les énergies E_ℓ sont déplacées de la même quantité : d'après (D.2),

$$E_\ell(x + \mathrm{d}x) = E_\ell(x) - X\, \mathrm{d}x \,. \tag{D.8}$$

Par conséquent, à chaque état (ℓ) correspondant à la valeur $x + \mathrm{d}x$ du paramètre extérieur et tel que

$$E \leqslant E_\ell(x + \mathrm{d}x) \leqslant E + \delta E \,, \tag{D.9}$$

est associé *un état et un seul* (repéré par le même indice ℓ) correspondant à la valeur x du paramètre et tel que

$$E + X\, \mathrm{d}x \leqslant E_\ell(x) \leqslant E + X\, \mathrm{d}x + \delta E \,. \tag{D.10}$$

Le nombre $\Omega(E, x + \mathrm{d}x)$ d'états microscopiques correspondant à la valeur $x + \mathrm{d}x$ du paramètre et dont l'énergie est comprise entre E et $E + \delta E$ est donc égal au nombre

---

3. C'est le cas par exemple pour un système de N spins indépendants (cristal paramagnétique parfait) placés dans un champ magnétique uniforme $\vec{B}$. L'énergie du système est

$$E = -M \,.\, B,$$

où M est la somme algébrique des projections des N moments magnétiques individuels le long de $\vec{B}$. Pour tous les états correspondant à cette énergie (il en existe en général un grand nombre, voir complément II.B), la dérivée de l'énergie par rapport au paramètre extérieur B est la même puisque M est la même.

d'états microscopiques correspondant à la valeur x mais dont l'énergie est cette fois comprise entre $E + X\,dx$ et $E + X\,dx + \delta E$ (fig. 1) :

$$\Omega(E,\, x + dx) = \Omega(E + X\,dx,\, x)\,. \tag{D.11}$$

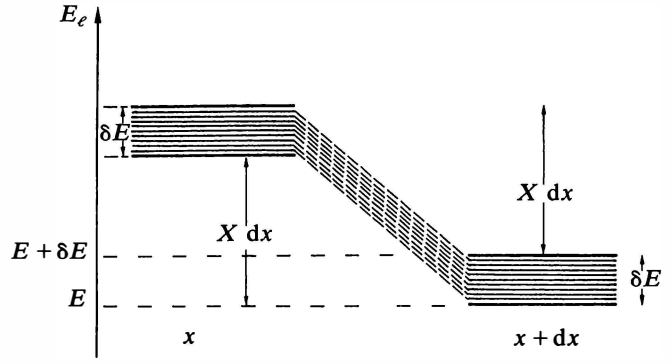


FIGURE 1

Représentation schématique de l'évolution des états (ℓ) et des énergies E_ℓ associées lorsqu'un paramètre extérieur varie de x à $x + dx$ (on a supposé pour simplifier que la force conjuguée X a même valeur dans tous les états accessibles).

Au premier ordre en dx, cette égalité implique

$$\frac{\partial \Omega}{\partial x} = X \frac{\partial \Omega}{\partial E}\,. \tag{D.12}$$

b. Cas général

Le plus souvent, la valeur X_ℓ de la force conjuguée varie d'un état (ℓ) à l'autre. Mais, parmi tous les états dont l'énergie est comprise entre E et $E + \delta E$ (pour la valeur x du paramètre extérieur), regroupons ceux qui donnent la même valeur X à la force conjuguée, à dX près. Leur nombre est $\omega(E, x; X)\,dX$. Le raisonnement du paragraphe a reste valable si on l'applique à ce sous-ensemble d'états accessibles : la relation (D.11) s'écrit ici

$$\omega(E,\, x + dx;\, X)\,dX = \omega(E + X\,dx,\, x;\, X)\,dX\,, \tag{D.13}$$

ou encore, au premier ordre en dx,

$$\frac{\partial \omega}{\partial x}\,dX = X \frac{\partial \omega}{\partial E}\,dX\,. \tag{D.14}$$

La relation équivalente à (D.12) s'obtient en sommant cette égalité sur toutes les valeurs possibles de X :

$$\int_{-\infty}^{+\infty} dX \frac{\partial}{\partial x}\,\omega(E, x; X) = \int_{-\infty}^{+\infty} dX \cdot X \frac{\partial}{\partial E}\,\omega(E, x; X)\,, \tag{D.15}$$

et en intervertissant, dans les deux membres, la sommation et la dérivation; il suffit enfin d'utiliser les formules (D.5) et (D.6) pour trouver

$$\frac{\partial}{\partial x}\,\Omega(E, x) = \frac{\partial}{\partial E}\left[\overline{X}(E, x) \cdot \Omega(E, x)\right], \tag{D.16}$$

où $\overline{X}(E, x)$ est la force moyenne conjuguée du paramètre x.

c. Cas d'un système macroscopique

Les systèmes isolés que nous étudions en mécanique statistique sont de taille macroscopique. Pour de tels systèmes, la relation (D.16) se simplifie.

Explicitons en effet le second membre de cette relation; elle devient

$$\frac{\partial \Omega}{\partial x} = \overline{X} \frac{\partial \Omega}{\partial E} + \Omega \frac{\partial \overline{X}}{\partial E}, \tag{D.17}$$

ou encore, si l'on divise les deux membres par Ω,

$$\frac{\partial}{\partial x} \ln \Omega = \overline{X} \frac{\partial}{\partial E} \ln \Omega + \frac{\partial \overline{X}}{\partial E}. \tag{D.18}$$

Or nous savons (I, § III.B.3) que, pour un système macroscopique, Ω est une fonction très rapidement croissante de l'énergie, de la forme $E^{N'}$, où N' est un nombre fantastiquement grand, de l'ordre du nombre d'Avogadro. Dans le second membre de (D.18), le premier terme est donc de l'ordre de

$$\overline{X} \frac{\partial}{\partial E} \ln \Omega \sim \overline{X} \frac{N'}{E}. \tag{D.19}$$

Au contraire, $\overline{X}$ est une fonction de E à variation lente, dont la dérivée peut être évaluée grossièrement comme

$$\frac{\partial \overline{X}}{\partial E} \sim \frac{\overline{X}}{E}. \tag{D.20}$$

Le second terme de (D.18) est donc totalement négligeable devant le premier : pour un système macroscopique,

$$\frac{\partial}{\partial x} \ln \Omega = \overline{X} \frac{\partial}{\partial E} \ln \Omega. \tag{D.21}$$

Introduisons alors l'entropie microcanonique S^*, en remplaçant $\ln \Omega$ par S^*/k. Si l'on se rappelle que

$$\frac{\partial S^*}{\partial E} \widehat{=} \frac{1}{T^*}, \tag{D.22}$$

on obtient finalement[4]

$$\boxed{\frac{\partial}{\partial x} S^*(E, x) = \frac{\overline{X}}{T^*}}. \tag{D.23}$$

Lorsque l'entropie microcanonique est fonction de plusieurs paramètres extérieurs x_i, la relation précédente est évidemment valable pour chacun d'eux[5].

---

4. Les paramètres extérieurs x_i caractérisant l'état macroscopique d'un système *isolé* sont en général *extensifs* (complément II.A, § II), comme l'énergie, le nombre de particules et le volume. L'entropie étant également extensive, la formule (D.23) montre que *les forces conjuguées moyennes* $\overline{X}_i$ *sont alors intensives* comme la température, le potentiel chimique et la pression.

5. Comme nous l'avons déjà remarqué (note 2), le nombre N de particules joue un rôle à part. Le potentiel chimique microcanonique μ^* est *directement défini* (II, § III.B.1) par une égalité analogue à (D.23) :

$$\frac{\mu^*}{T^*} \widehat{=} -\frac{\partial S^*}{\partial N},$$

ce qui permet de le considérer comme la « force » conjuguée de N (au signe près).

d. Cas d'un système purement mécanique

En général, pour un système macroscopique, on peut fixer de façon indépendante la valeur de l'énergie et celle des autres paramètres extérieurs (nombre de particules, volume, ...) : les paramètres extérieurs autres que l'énergie imposent des conditions sur les états microscopiques possibles, mais l'équation aux valeurs propres de l'hamiltonien admet en général une infinité de solutions compatibles avec ces conditions, et donne donc une série infinie d'énergies possibles $E_\epsilon(x)$; on peut par conséquent choisir l'énergie E du système indépendamment des valeurs des autres paramètres extérieurs.

Par opposition au cas général, on appelle *« système purement mécanique »* un système macroscopique tel que *la donnée des paramètres extérieurs x_i détermine de façon unique la valeur de son énergie.* Dans un système purement mécanique, une fois les paramètres extérieurs x_i fixés, les états microscopiques possibles ont donc tous la même énergie :

$$E_\epsilon(x) = E(x) \ . \tag{D.24}$$

Un exemple simple de système purement mécanique est le ressort idéal (sans masse) dont l'énergie

$$E = \frac{1}{2} K x^2 \tag{D.25}$$

est fixée par la valeur de son allongement x, si l'on considère, comme c'est l'usage en mécanique, que sa raideur K est une constante[6].

Pour un système purement mécanique, le nombre d'états accessibles ne peut être fonction que des paramètres extérieurs x_i autres que l'énergie, puisque celle-ci n'est pas indépendante des x_i :

$$\Omega\left[E(x),\, x\right] \equiv \Omega(x) \ . \tag{D.26}$$

L'application à un tel système du raisonnement du paragraphe 2.a (ici, la force conjuguée est la même dans tous les états accessibles, cf. (D.24)) montre que *le nombre d'états accessibles est en réalité constant*[7] :

$$\Omega(x + \mathrm{d}x) = \Omega(x) \ . \tag{D.27}$$

---

6. Pour un ressort réel, la raideur K dépend de la température, de sorte que E et x peuvent en fait être choisis indépendamment (à l'intérieur de certaines limites). La mécanique néglige cet aspect des choses et ne considère que des ressorts idéaux. Il arrive que l'on tienne compte des effets anharmoniques, c'est-à-dire de la dépendance en x de K, mais le ressort reste encore dans ce cas un système purement mécanique : on a toujours

$$E = E(x),$$

même si la fonction $E(x)$ est alors moins simple qu'en (D.25).

7. Lorsque les valeurs des paramètres x_i sont fixées, le nombre d'états accessibles est nul si l'énergie E que l'on se donne en outre (à δE près) diffère de $E(x)$. On peut donc écrire de façon plus précise l'égalité (D.26) sous la forme

$$\Omega(E,\, x) = \Omega(x)\, \bar{\delta}\left[E(x) - E\right], \tag{1}$$

où par définition

$$\begin{aligned} \bar{\delta}(u) &= 1 \quad \text{si} \quad 0 \leqslant u \leqslant \delta E \ , \\ &= 0 \quad \text{ailleurs} \ . \end{aligned} \tag{2}$$

Avec ces notations, la relation (D.11) donne

$$\Omega(x + \mathrm{d}x)\, \bar{\delta}\left[E(x + \mathrm{d}x) - E\right] = \Omega(x)\, \bar{\delta}\left[E(x) - (E + X\, \mathrm{d}x)\right] . \tag{3}$$

.../...

Par conséquent, *l'entropie d'un système purement mécanique est invariable :* elle ne dépend pas des paramètres extérieurs x_i, ni donc de l'énergie. C'est là la caractéristique essentielle d'un système purement mécanique. Ses propriétés peuvent être décrites dans le cadre de la mécanique classique sans faire appel à aucune notion de type statistique telle qu'entropie et température. En particulier, *un système purement mécanique ne peut pas échanger de chaleur* avec d'autres systèmes : son énergie ne peut être modifiée que par variation des paramètres x_i, c'est-à-dire par échange de travail avec l'environnement.

3. EXEMPLE : LA PRESSION, FORCE MOYENNE CONJUGUÉE DU VOLUME

Considérons un système *macroscopique* isolé, d'énergie E, occupant un volume V délimité par une surface s. Soit $\vec{r}$ le rayon-vecteur joignant l'origine des coordonnées à un point M de cette surface. Nous prenons autour du point M un élément de surface, caractérisé par le vecteur $\delta\vec{s}$ normal en M à la surface s et dirigé vers l'extérieur (fig. 2), et nous cherchons la force (au sens habituel) qu'exerce le système sur cet élément de surface.

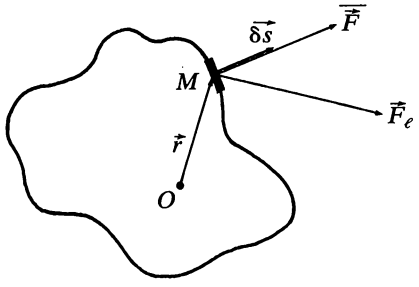


FIGURE 2
Lorsque le système est dans l'état (ℓ), il exerce sur l'élément de surface $\delta\vec{s}$ la force $\vec{F}_\ell$. A l'équilibre, la force moyenne $\vec{F}$ est normale à la surface et proportionnelle à δs.

Le rayon-vecteur $\vec{r}$ qui repère la position de l'élément de surface joue le rôle d'un paramètre extérieur[8] : à une variation $d\vec{r}$ de $\vec{r}$ correspond un déplacement de l'élément de surface; or l'énergie E_ℓ de l'état microscopique (ℓ) du système dépend de l'espace qu'il occupe, et donc en particulier de la position $\vec{r}$ de l'élément de surface considéré. La force $\vec{F}_\ell$ qu'exerce le système, lorsqu'il se trouve dans l'état (ℓ), sur l'élément de surface $\delta\vec{s}$ s'écrit comme en (D.2)[9] :

$$\vec{F}_\ell(\vec{r}) = -\overrightarrow{\text{grad}}\, E_\ell(\vec{r})\,. \tag{D.28}$$

Comme par définition (cf. (D.2) et (D.24))

$$X = -\frac{\partial}{\partial x} E(x)\,, \tag{4}$$

l'argument de la fonction $\bar{\delta}$ est le même dans les deux membres de l'égalité (3). On en déduit

$$\Omega(x + dx) = \Omega(x)\,. \tag{5}$$

8. Il faudrait aussi caractériser l'orientation du vecteur $\delta\vec{s}$. Mais un changement d'orientation de $\delta\vec{s}$ ne change pas, au premier ordre, le volume V du système (voir note 11). Il en résulte (voir plus loin la formule (D.30)) que la « force conjuguée » associée aux paramètres d'orientation est constamment nulle en moyenne.

9. La formule (D.28) permet de faire le lien avec la mécanique. Elle reproduit en effet la *relation entre force et énergie en mécanique :* lors d'un déplacement (virtuel) $d\vec{r}$ de l'élément de surface, la force $\vec{F}_\ell$ effectue le travail $\vec{F}_\ell . d\vec{r}$, qui est pris sur l'énergie E_ℓ du système :

$$\vec{F}_\ell . d\vec{r} = -dE_\ell\,.$$

D'après (D.23), la force moyenne correspondante vaut donc

$$\overrightarrow{\bar{F}}(\vec{r}) = T^* \overrightarrow{\text{grad}}\, S^*(\vec{r}) \,. \tag{D.29}$$

Mais l'entropie est une grandeur additive : si l'on divise (par la pensée) le système considéré en un grand nombre de sous-systèmes petits mais encore macroscopiques (complément II.A, § II), son entropie microcanonique est la somme des entropies de ces divers sous-systèmes. Il en résulte que *S^* ne dépend que du volume V, pas de sa forme*[10] : en effet, un changement de forme du système sans changement de son volume total peut toujours être considéré comme un réarrangement des sous-systèmes, et n'affecte donc pas les grandeurs additives telles que son énergie et son entropie. *L'entropie S^* ne dépend donc de $\vec{r}$ que par l'intermédiaire de V,* de sorte que

$$\overrightarrow{\bar{F}} = T^* \frac{\partial}{\partial V} S^*(E, V) \overrightarrow{\text{grad}}\, V(\vec{r}) \,. \tag{D.30}$$

Or[11]

$$\overrightarrow{\text{grad}}\, V(\vec{r}) = \delta\vec{s} \;; \tag{D.31}$$

10. Ceci n'est vrai que si les interactions entre les sous-systèmes macroscopiques sont à courte portée (complément II.A, § I) et si l'on peut négliger la contribution des phénomènes de surface à l'énergie et à l'entropie.

11. Lorsqu'on déplace (virtuellement) l'élément de surface $\delta\vec{s}$ de $\mathrm{d}\vec{r}$ (fig. 3.a), le volume du système augmente de

$$\mathrm{d}V = \delta\vec{s} \,.\, \mathrm{d}\vec{r} \,. \tag{1}$$

En identifiant avec la formule générale

$$\mathrm{d}V = \overrightarrow{\text{grad}}\, V \,.\, \mathrm{d}\vec{r} \,, \tag{2}$$

on obtient la relation (D.31).

Un changement d'orientation de $\delta\vec{s}$ sans changement de $\vec{r}$ n'affecte le volume V qu'au second ordre (fig. 3.b).

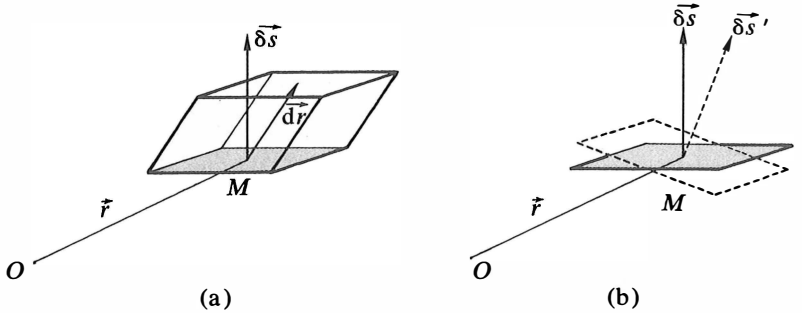


FIGURE 3

L'élément de surface $\delta\vec{s}$, repéré par le rayon-vecteur $\vec{r}$, est déplacé de $\mathrm{d}\vec{r}$ parallèlement à lui-même (a), ou changé d'orientation sans déplacement du point M (b).

on obtient finalement

$$\overrightarrow{\delta F} = T^* \frac{\partial S^*}{\partial V} \overrightarrow{\delta s} . \tag{D.32}$$

La force moyenne agissant sur l'élément de surface $\overrightarrow{\delta s}$ est donc *normale à la surface* au point M, et son intensité est *proportionnelle à* δs[12]. Le système exerce donc sur les parois du récipient qui le contient une *pression*

$$p^* = T^* \frac{\partial S^*}{\partial V} . \tag{D.33}$$

Ce résultat justifie la définition que nous avons donnée (II, formule (III.29.a)) de la pression microcanonique : la « force » moyenne conjuguée du volume V apparaît comme une véritable pression au sens mécanique du terme, c'est-à-dire comme une force (ordinaire) par unité de surface, s'exerçant perpendiculairement à la paroi. La formule finale (D.33) montre clairement que *la pression est la même en tous les points de la surface s* limitant le volume occupé par le système[13].

4. ÉQUILIBRE ENTRE SYSTÈMES POUVANT ÉCHANGER UN PARAMÈTRE EXTÉRIEUR

a. Condition générale d'équilibre

Les conditions d'équilibre entre sous-systèmes d'un système isolé, que nous avons établies au chapitre II, paragraphe III.B pour les échanges de volume et de particules (égalité des pressions et des potentiels chimiques), se généralisent pour n'importe quel paramètre extérieur à l'aide de la notion de force conjuguée.

Considérons en effet deux sous-systèmes A et B d'un système isolé $\mathcal{S}$. A et B sont en contact thermique et peuvent échanger, outre de l'énergie, du paramètre extérieur x. Le fait que $\mathcal{S}$ soit isolé se traduit en particulier par une relation entre les valeurs x^A et x^B que prend le paramètre extérieur considéré dans chacun des sous-systèmes. Cette « relation de conservation » s'écrit sous la forme

$$x^A + \varepsilon x^B = \text{c}^{\text{te}}, \qquad \varepsilon = +1 \text{ ou } -1 . \tag{D.34}$$

Si x est le volume ou le nombre de particules, grandeurs à valeurs toujours positives, ε vaut $+1$ et la constante du second membre est liée à la valeur de x pour le système global $\mathcal{S}$; pour un paramètre x à valeurs algébriques (abscisse d'un piston, par exemple[14]), on pourra avoir $\varepsilon = -1$; à chaque paramètre extérieur x_i est en tout cas attaché un ε_i déterminé.

A partir de la « relation de conservation » (D.34) et de l'expression (D.23) de la force moyenne conjuguée $\overline{X}$ du paramètre extérieur x, on peut reprendre point par point la démonstration du chapitre II, paragraphe III.B.3 : la valeur la plus probable

12. La force (D.32) est *dirigée vers l'extérieur,* car $\partial S^*/\partial V$ est (comme T^*) toujours positif : s'il en était autrement, le système pourrait accroître son entropie microcanonique en diminuant de volume et il s'effondrerait sur lui-même (puisqu'il chercherait à maximiser son entropie S^*).

13. Si l'on divise le système considéré en sous-systèmes macroscopiques, ceux-ci peuvent échanger du volume; à l'équilibre, leur pression microcanonique doit donc être la même (II, § III.B.3). Par conséquent, la pression est également définie à l'intérieur du récipient contenant le système, et elle y est égale à celle qui s'exerce sur les parois.

de x^A (et celle de x^B qui en découle) est déterminée par la *condition d'équilibre*[14]

$$\overline{X}^A = \varepsilon \overline{X}^B \,. \tag{D.35}$$

b. Équilibre avec un système purement mécanique

Le raisonnement qui vient d'être esquissé n'est pas applicable aux cas où l'un au moins des deux systèmes en contact est purement mécanique (§ 2.d) : on ne peut pas utiliser son entropie, qui reste invariable. Que devient dans ce cas la condition d'équilibre? La réponse est évidente (équilibre des forces) si l'on se place dans le cadre de la mécanique pure. Il est facile également de retrouver ce résultat en mécanique statistique.

Considérons en effet un système macroscopique $\mathcal{S}$ en contact avec un système purement mécanique $\mathcal{M}$, par exemple un gaz enfermé dans un récipient dont le volume est contrôlé par un ressort idéal (fig. 5). Soit x le paramètre que les deux systèmes peuvent échanger, $x_{\mathcal{S}}$ et $x_{\mathcal{M}}$ les valeurs de ce paramètre pour chacun des systèmes. Plaçons-nous dans le cas (correspondant à celui de la figure 5) où la « relation de conservation » s'écrit

$$x_{\mathcal{S}} - x_{\mathcal{M}} = \mathrm{c^{te}} \,. \tag{D.36}$$

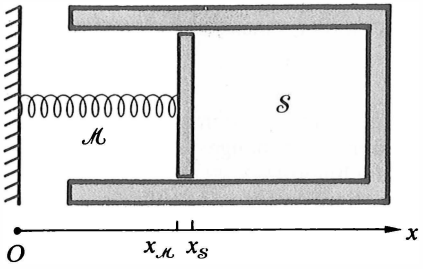


FIGURE 5
Le volume du gaz $\mathcal{S}$ est contrôlé par le ressort purement mécanique $\mathcal{M}$ qui agit sur le piston (adiabatique, comme les parois du cylindre). Les deux systèmes $\mathcal{S}$ et $\mathcal{M}$ échangent du paramètre x.

14. Dans le cas d'un piston mobile séparant deux gaz (fig. 4), on choisit généralement comme paramètre extérieur le volume. La relation

$$V_A + V_B = V = \mathrm{c^{te}} \tag{1}$$

conduit alors à l'égalité des pressions :

$$p_A^* = p_B^* \,. \tag{2}$$

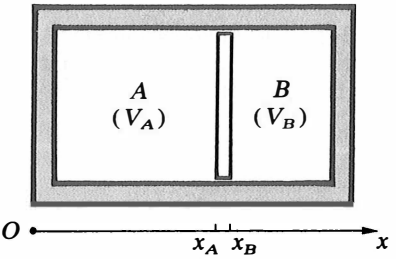


FIGURE 4
L'état (macroscopique) d'un système isolé constitué par deux gaz A et B séparés par un piston mobile (et diatherme) peut être caractérisé, soit par la valeur V_A du volume de A (avec $V_A + V_B = \mathrm{c^{te}}$), soit par l'abscisse x_A de la paroi mobile qui limite l'espace accessible à A (auquel cas $x_A - x_B = \mathrm{c^{te}} = -e$, si e est l'épaisseur du piston).

Mais on peut également prendre comme paramètre extérieur l'abscisse x de la paroi mobile qui, pour chacun des deux gaz, limite l'espace qu'il peut occuper. Dans ce cas, la « relation de conservation » s'écrit

$$x_A - x_B = \mathrm{c^{te}} \tag{3}$$

.../...

L'entropie du système $\mathcal{M}$ étant invariable (indépendante de $x_{\mathcal{M}}$), la valeur la plus probable de x_S est celle qui rend l'entropie microcanonique $S_S^*(E_S, x_S)$ du système S maximum, sachant que son énergie est donnée par

$$E_S = E_{\text{tot}} - E_{\mathcal{M}}(x_{\mathcal{M}}) , \tag{D.37}$$

où E_{tot} est l'énergie totale, fixée, du système isolé $S \cup \mathcal{M}$. La condition d'équilibre s'écrit donc

$$\frac{\partial S_S^*}{\partial E_S}\frac{\partial E_S}{\partial x_S} + \frac{\partial S_S^*}{\partial x_S} = 0 , \tag{D.38}$$

ou encore, si l'on tient compte de (D.36), (D.37) et (D.23),

$$-\frac{1}{T_S^*}\frac{\partial E_{\mathcal{M}}}{\partial x_{\mathcal{M}}} + \frac{1}{T_S^*}\overline{X}_S = 0 , \tag{D.39}$$

$\overline{X}_S$ étant la force moyenne conjuguée du paramètre x dans le système S. Finalement, la valeur la plus probable de x_S est déterminée par la condition

$$\overline{X}_S - \frac{\partial E_{\mathcal{M}}}{\partial x_{\mathcal{M}}} = 0 . \tag{D.40}$$

Comme on devait s'y attendre, cette condition exprime que *la somme des forces* s'exerçant sur la paroi commune aux deux systèmes (le piston sur la figure 5) *est nulle à l'équilibre :* la force exercée par le système purement mécanique $\mathcal{M}$ a l'expression habituelle en mécanique $-\partial E_{\mathcal{M}}/\partial x_{\mathcal{M}}$ (cf. l'exemple (D.25) du ressort idéal), celle qu'exerce le système S est la force moyenne conjuguée du paramètre extérieur x considéré.

II. TRANSFORMATIONS D'UN SYSTÈME

1. DÉFINITION GÉNÉRALE

Considérons un système macroscopique isolé, dans l'état d'équilibre où les paramètres extérieurs caractérisant son état macroscopique ont les valeurs $E^{(1)}$, $N^{(1)}$ et $x_i^{(1)}$. Modifions alors ces valeurs, en rompant l'isolement du système; cette modification dure un certain temps, et amène les paramètres extérieurs aux valeurs finales $E^{(2)}$, $N^{(2)}$ et $x_i^{(2)}$. Isolons à nouveau le système et attendons qu'il atteigne l'état d'équilibre macroscopique correspondant aux nouvelles conditions. On dit qu'on a ainsi fait subir au système une *transformation,* qui l'a fait passer de *l'état initial* $\{E^{(1)}, N^{(1)}, x_i^{(1)}\}$ à *l'état final* $\{E^{(2)}, N^{(2)}, x_i^{(2)}\}$, par un certain *« chemin »* caractérisé par la façon dont, de l'extérieur, nous avons procédé pour faire varier les paramètres de leurs valeurs initiales à leurs valeurs finales. Pour que la transformation soit bien définie, il faut qu'elle parte d'un état d'équilibre initial et aboutisse à un état d'équilibre

---

(la constante étant égale à l'épaisseur du piston). Elle conduit à

$$\overline{X}_A + \overline{X}_B = 0 : \tag{4}$$

la force que le gaz A exerce sur le piston est, en mesure algébrique sur l'axe Ox, opposée à celle qu'exerce sur ce même piston le gaz B.

Les deux choix possibles (volume ou abscisse de la paroi mobile) décrivent ici la même réalité physique et aboutissent à deux expressions équivalentes de l'équilibre mécanique du piston.

final[15]; cependant, entre des états initial et final donnés, il existe en général une infinité de « chemins » possibles, caractérisant chacun une transformation différente.

Si les valeurs finales des paramètres extérieurs s'écartent infiniment peu (à l'échelle macroscopique) de leurs valeurs initiales, la transformation est qualifiée d'*infinitésimale.*

Nous appellerons *« milieu extérieur »,* ou simplement *« l'extérieur »,* l'ensemble des autres systèmes avec lesquels le système considéré est en contact au cours d'une transformation. Le comportement détaillé du « milieu extérieur » ne nous intéresse pas; seuls nous importent les *actions* (forces généralisées) qu'il exerce à chaque instant sur notre système et les *échanges* (d'énergie et de particules) qu'il a avec lui. Différents milieux extérieurs exerçant les mêmes actions sur le système et opérant avec lui les mêmes échanges sont équivalents pour ce qui concerne la transformation du système.

2. ÉNERGIE ÉCHANGÉE AU COURS D'UNE TRANSFORMATION

a. Travail reçu par le système

Nous supposons, pour chaque paramètre extérieur x_i autre que l'énergie et le nombre de particules, que le milieu extérieur exerce sur le système une force généralisée X_i^{ex} bien définie tout au long de la transformation. Dans les états initial (1) et final (2), cette force maintient le paramètre considéré à la valeur fixée, $x_i^{(1)}$ ou $x_i^{(2)}$. Au cours de la transformation, les divers paramètres x_i varient, et les forces correspondantes effectuent un certain travail $W_{1\to 2}$, que l'on peut calculer en suivant les variations des X_i^{ex} et des paramètres x_i^{ex}, analogues des x_i mais *se rapportant au milieu extérieur :*

$$W_{1\to 2} = \sum_i \int_{(1)}^{(2)} X_i^{ex}\, \mathrm{d}x_i^{ex} , \tag{D.41}$$

où l'intégrale est prise *le long du chemin suivi par la transformation.* Suivant la nature physique des paramètres extérieurs qui varient dans la transformation, le travail $W_{1\to 2}$ peut être d'origine mécanique, électrique, magnétique, ...

De façon générale, si $W_{1\to 2}$ est positif (travail moteur), l'extérieur fournit du travail au système, dont l'énergie s'accroît d'autant (cf. § d ci-dessous); si au contraire $W_{1\to 2}$ est négatif (travail résistant), c'est le système qui fournit du travail à l'extérieur, en prenant sur son énergie. En grandeur et en signe, $W_{1\to 2}$ est donc le *travail reçu par le système* au cours de la transformation.

Insistons sur le fait que $W_{1\to 2}$ est défini à partir des grandeurs X_i^{ex} et $\mathrm{d}x_i^{ex}$ se référant au *milieu extérieur.* Une « relation de conservation » de la forme (cf. § I.4.a)

$$x_i + \varepsilon_i x_i^{ex} = \mathrm{c^{te}} \tag{D.42}$$

permet le plus souvent de remplacer $\mathrm{d}x_i^{ex}$ par $\mathrm{d}x_i$:

$$\mathrm{d}x_i^{ex} = -\varepsilon_i\, \mathrm{d}x_i . \tag{D.43}$$

Mais, comme le système n'est en général pas en équilibre pendant le déroulement de la transformation, les forces conjuguées moyennes $\overline{X_i}$ des paramètres du système ne

15. L'exigence que le système soit *isolé* dans l'état initial puis dans l'état final pourra être abandonnée : nous montrerons en effet que, si la taille du système est suffisante pour que les fluctuations des diverses variables internes y soient négligeables (système dit « thermodynamique »), il est équivalent, à l'équilibre, qu'il soit isolé, ou maintenu à température constante par un thermostat (voir III, § III.C), ou à pression constante par un réservoir de volume (complément III.D), ou à potentiel chimique constant par un réservoir de particules (V, § III.B). Il est en revanche essentiel que le système soit à l'équilibre macroscopique dans l'état initial puis dans l'état final.

sont bien définies que dans les états initial et final, et pas le long du chemin imposé de l'extérieur par la transformation[16].

Par exemple, si x_i est le volume,

$$W_{1\to 2}^{(V)} = \int_{(1)}^{(2)} p^{ex}\, dV^{ex}\,, \tag{D.44}$$

où p^{ex} est la *pression exercée par l'extérieur sur le système* à chaque instant de la transformation, et dV^{ex} l'accroissement de volume du milieu extérieur[17]; sauf transformations très particulières[18], on a dans ce cas

$$dV^{ex} = -dV\,, \tag{D.45}$$

dV étant l'accroissement de volume du système, mais la pression p du système lui-même n'est en général bien définie que dans les états (1) et (2), pas dans les étapes intermédiaires de la transformation.

Le travail reçu par le système dans une *transformation infinitésimale* est évidemment infinitésimal. Mais, comme dans le cas général $W_{1\to 2}$ dépend du chemin suivi entre les états (1) et (2), le travail infinitésimal n'est pas une différentielle totale, c'est-à-dire qu'il n'existe pas de fonction des paramètres extérieurs dont il soit la différentielle; c'est pourquoi nous le noterons đW :

$$đW = \sum_i X_i^{ex}\, dx_i^{ex} = -\sum_i \varepsilon_i X_i^{ex}\, dx_i \tag{D.46}$$

(la seconde expression étant valable si la relation (D.43) l'est).

---

16. Pour pouvoir mesurer ou calculer le travail reçu par le système considéré, il faut contrôler les propriétés du milieu extérieur tout au long de la transformation. Ceci n'est souvent pas facile. Par exemple, si l'on met en contact, sans précautions, deux systèmes semblables de taille voisine, chacun d'eux fait pour l'autre partie du « milieu extérieur », et il est donc impossible de connaître à chaque instant de la transformation les caractéristiques de ce « milieu extérieur ». Il n'en reste pas moins que, si l'on choisit de s'intéresser à l'un de ces deux systèmes, il subit une transformation bien déterminée (à condition qu'il passe d'un état d'équilibre initial à un état d'équilibre final), même s'il est impossible de préciser quel travail il a reçu dans cette transformation.

17. En un point M de la surface séparant le système du milieu extérieur, prenons un élément de surface caractérisé par un vecteur $\delta\vec{s}$: $\delta\vec{s}$, normal à la surface en M et de module égal à l'aire de l'élément de surface, est choisi ici *dirigé de l'extérieur vers le système.* Si p^{ex} est la pression exercée par le milieu extérieur sur le système au point M, la force correspondante $\vec{F}^{ex}$ (au sens mécanique) qui agit sur l'élément de surface a pour expression

$$\vec{F}^{ex} = p^{ex}\delta\vec{s}\,. \tag{1}$$

Lorsque, au cours d'une transformation, le point M est déplacé de $d\vec{r}$, la force $\vec{F}^{ex}$ effectue le travail

$$\vec{F}^{ex}.d\vec{r} = p^{ex}\delta\vec{s}\,.d\vec{r}\,. \tag{2}$$

Il est facile de voir (cf. note 11 et fig. 3.a) que

$$\delta\vec{s}\,.d\vec{r} = dV^{ex} \tag{3}$$

est ici l'*accroissement de volume du milieu extérieur,* accompagnant le déplacement $d\vec{r}$.

18. Au cours d'une *« détente de Joule »* (décrite au II, § II.B.1), le volume du gaz s'accroît ($\Delta V > 0$) sans que celui du milieu extérieur varie ($\Delta V^{ex} = 0$); le travail reçu par le gaz est donc *nul.*

De façon générale, c'est toujours à la formule (D.41) qu'il faut revenir en cas d'ambiguïté.

b. Énergie « chimique » reçue par le système

On peut envisager des transformations qui modifient le nombre de particules du système ($N^{(2)} \neq N^{(1)}$) : on peut par exemple ménager des passages poreux dans les parois qui limitent le système, lui permettant ainsi d'échanger des particules avec l'extérieur. Un changement dans le nombre de particules s'accompagne nécessairement d'un changement d'énergie : l'énergie cinétique totale du système est la somme de celles de ses N constituants, et son énergie potentielle totale dépend aussi de N.

Nous supposons ici aussi que le *potentiel chimique* μ^{ex} *du milieu extérieur*[19] est bien défini (mais pas nécessairement constant) pendant toute la durée de l'échange de particules. Par analogie avec le travail, on *définit* l'énergie « chimique » $\Phi_{1\to 2}$ reçue par le système par suite de la variation du nombre de particules qui le constituent, comme

$$\Phi_{1\to 2} \mathrel{\widehat{=}} -\int_{(1)}^{(2)} \mu^{ex}\, dN^{ex}\,, \tag{D.47}$$

où N^{ex} est le nombre de particules du milieu extérieur et l'intégrale est à prendre le long de la transformation envisagée. La conservation du nombre total de particules implique

$$dN^{ex} + dN = 0\,, \tag{D.48}$$

de sorte que

$$\Phi_{1\to 2} = +\int_{(1)}^{(2)} \mu^{ex}\, dN\,. \tag{D.49}$$

Comme le travail, l'énergie « chimique » $\Phi_{1\to 2}$ est comptée *algébriquement*. Dans une transformation infinitésimale, l'énergie « chimique » reçue par le système est infiniment petite; on l'écrit

$$đ\Phi \mathrel{\widehat{=}} -\mu^{ex}\, dN^{ex} = +\mu^{ex}\, dN\,. \tag{D.50}$$

c. Chaleur reçue par le système

Comme nous l'avons indiqué au chapitre II, paragraphe III.A.1.b, le système et le milieu extérieur peuvent échanger de l'énergie même si tous les autres paramètres extérieurs restent constants, à condition qu'une paroi diatherme établisse entre eux le *contact thermique*. Cette énergie échangée, qui n'est ni du travail ni de l'énergie « chimique », est appelée *chaleur*. Dans une transformation, le système et le milieu extérieur échangent en général à la fois du travail (par variation des paramètres extérieurs x_i), de l'énergie « chimique » (si N varie) et de la chaleur.

Nous compterons ici aussi positivement la chaleur reçue par le système du milieu extérieur, négativement celle que fournit le système à l'extérieur, et noterons $Q_{1\to 2}$ cette quantité de chaleur algébrique. L'accroissement de l'énergie du système au cours d'une transformation s'écrit de façon générale

$$E^{(2)} - E^{(1)} = W_{1\to 2} + \Phi_{1\to 2} + Q_{1\to 2}\,. \tag{D.51}$$

Les contributions $W_{1\to 2}$ et $\Phi_{1\to 2}$ ayant été définies par ailleurs (formules (D.41) et (D.47)), l'égalité précédente constitue la *définition de la quantité de chaleur* $Q_{1\to 2}$.

Comme $W_{1\to 2}$ et $\Phi_{1\to 2}$, la chaleur $Q_{1\to 2}$ reçue par le système dépend des états initial (1) et final (2) mais aussi du chemin suivi par la transformation. Comme eux,

19. Si le système contient des particules de plusieurs espèces différentes, il faut distinguer les nombres N_1, N_2, ... de particules de chaque espèce et les potentiels chimiques μ_1^{ex}, μ_2^{ex}, ... correspondants, et sommer les termes (D.47) ou (D.49) associés aux divers types de particules.

elle peut être déterminée en analysant les *modifications du milieu extérieur* au cours de la transformation[20].

La chaleur đQ reçue par le système dans une transformation infinitésimale est infiniment petite; pas plus que đW ou đΦ, la chaleur đQ n'est une différentielle totale. Pour une transformation infinitésimale, l'égalité (D.51) s'écrit

$$dE = đW + đ\Phi + đQ \,. \qquad \text{(D.52)}$$

Soulignons que $W_{1\to 2}$, $\Phi_{1\to 2}$ et $Q_{1\to 2}$ (ou đW, đΦ et đQ) correspondent à des *manières différentes d'échanger de l'énergie.* Mais, si elle peut être modifiée de diverses façons, *l'énergie d'un système est une grandeur unique :* des expressions comme « la chaleur du système » ou « le travail du système » sont totalement dépourvues de sens.

d. Le « premier principe » de la thermodynamique

Les relations (D.51) et (D.52) expriment ce que l'on appelle le « premier principe de la thermodynamique ».

Sa formulation initiale (Joule-1842) est celle d'un *« principe d'équivalence »,* affirmant que la chaleur, dont la nature était jusque là totalement inconnue, est comme le travail une forme d'énergie échangée entre des systèmes. Pour nous, maintenant, cette affirmation sert à *définir la chaleur* (cf. § c), c'est-à-dire que le « premier principe » apparaît comme une simple conséquence de la loi générale de *conservation de l'énergie.*

Dans la pratique, lorsqu'on analyse une transformation, il est commode d'expliciter le contenu du « premier principe » de la manière suivante : $W_{1\to 2}$, $\Phi_{1\to 2}$ et $Q_{1\to 2}$ dépendent chacun du chemin suivi par la transformation, mais leur somme ne dépend d'après (D.51) que des états initial et final; ou bien, pour une transformation infinitésimale, đW, đΦ et đQ ne sont pas des différentielles totales, mais leur somme en est une d'après (D.52).

3. TRANSFORMATIONS PARTICULIÈRES

a. Transformation adiabatique

Une transformation est dite *« adiabatique »* si elle ne donne lieu à *aucun échange de particules* ni *aucun échange de chaleur* entre le système considéré et l'extérieur[21]. Dans une transformation adiabatique, la variation d'énergie du système est seulement due au travail (mécanique, électrique, magnétique, ...) $W_{1\to 2}$ qu'il reçoit de l'extérieur :

$$[E^{(2)} - E^{(1)}]^{\text{ad}} = W^{\text{ad}}_{1\to 2} \,. \qquad \text{(D.53)}$$

b. Transformation quasistatique

En général, les états intermédiaires par lesquels passe un système au cours d'une transformation ne sont pas des états d'équilibre. Par exemple, lorsqu'on déplace

---

20. Par exemple, si la chaleur est échangée par le système avec une masse connue d'eau, ou de tout autre corps dont on connaît la capacité calorifique (c'est-à-dire la façon dont son énergie dépend de la température, cf. III, § II.B.1.b), il suffit de mesurer la variation de la température de ce corps de référence au cours de la transformation pour en déduire la chaleur reçue par le système.

21. Pour une transformation finie, cette définition impose non seulement que l'énergie « chimique » et la chaleur reçues par le système soient globalement nulles :

$$\Phi_{1\to 2} = 0 \,, \quad Q_{1\to 2} = 0 \,,$$

mais en outre qu'il n'y ait à aucun stade intermédiaire de la transformation des échanges de particules ou de chaleur entre le système et l'extérieur.

le piston qui fixe le volume V offert à un gaz (comme sur la figure 5), celui-ci passe par des états transitoires où sa densité est plus faible ou plus forte au voisinage du piston qu'au fond du cylindre; il faut un certain temps, après que le piston se soit à nouveau immobilisé, pour que ces densités s'égalisent[22] et que le gaz revienne à l'équilibre.

Nous qualifierons de *« quasistatique »* une transformation qui amène le système de l'état initial à l'état final par une *succession continue d'états d'équilibre :* au cours d'une transformation quasistatique, le système se trouve à chaque instant dans l'état d'équilibre correspondant aux valeurs que prennent à cet instant les paramètres extérieurs E, N, x_i. Pour que ceci soit possible, il faut que la variation des paramètres extérieurs soit opérée assez lentement pour que le système ait le temps d'atteindre son nouvel état d'équilibre avant qu'ils n'aient encore changé de façon appréciable; autrement dit, le processus (imposé) de variation des paramètres doit être lent devant le processus (spontané) d'établissement de l'équilibre dans le système considéré[23].

Notons que toute transformation infinitésimale est évidemment quasistatique. Inversement, toute transformation quasistatique est une suite de transformations infinitésimales (ce qui n'est pas vrai pour une transformation dans laquelle les états intermédiaires du système ne sont pas des états d'équilibre).

Au cours d'une transformation quasistatique, *les forces moyennes $\overline{X}_i$ conjuguées des paramètres x_i sont constamment définies dans le système, de même que sa température et son potentiel chimique,* bien que ces diverses grandeurs varient d'un instant à l'autre.

c. Transformation réversible

Une transformation sera dite *« réversible »* si elle est *quasistatique* et si en outre, à chaque instant de la transformation, il y a *équilibre entre le système considéré et les systèmes extérieurs* avec lesquels il est en contact. Pour préciser, prenons un paramètre x_i particulier dont la valeur varie au cours de la transformation. Celle-ci étant quasistatique, la force moyenne $\overline{X}_i$ conjuguée de x_i est bien définie à chaque instant dans le système. Pour que la transformation soit réversible, il faut en outre que $\overline{X}_i$ et X_i^{ex} vérifient à chaque instant la condition d'équilibre (D.35). Il faut également que la température du système soit constamment égale à celle du milieu extérieur s'il échange avec lui de la chaleur, et que son potentiel chimique soit égal à celui du milieu extérieur s'il y a échange de particules[24] :

$$\text{transformation réversible} : \begin{cases} \overline{X}_i = \varepsilon_i X_i^{ex} & \text{si } x_i \text{ varie,} \\ T^* = T^{ex} & \text{si échange de chaleur,} \\ \mu^* = \mu^{ex} & \text{si échange de particules.} \end{cases} \qquad \text{(D.54)}$$

En réalité, s'il y avait équilibre parfait entre le système et le milieu extérieur, il ne se produirait pas de transformation : pour que celle-ci puisse se développer, il faut un *très léger déséquilibre,* dans un sens ou dans l'autre, entre les forces généralisées, les températures et les potentiels chimiques du système et de l'extérieur, ou tout au moins certaines de ces grandeurs. Mais ce déséquilibre est assez faible pour être

22. Les inhomogénéités de densité produites par le déplacement du piston provoquent des ondes sonores (en fait, souvent, infra-sonores) qui se propagent dans le gaz et se réfléchissent sur les parois, perdant peu à peu de leur amplitude.

23. Dans le cas d'un gaz dont on varie le volume à l'aide d'un piston, la transformation sera quasistatique si la vitesse de déplacement du piston est très inférieure à la vitesse du son dans le gaz (cf. note 22).

24. Une transformation infinitésimale, si elle est toujours quasistatique, n'est en général pas réversible.

négligeable. En outre, *il suffit de modifier de façon presque imperceptible les conditions extérieures* (forces, température, potentiel chimique) *pour que le sens de la transformation s'inverse;* c'est pourquoi une telle transformation est dite « réversible ».

Par opposition, toute autre transformation sera dite *« irréversible »*. Il est clair que les transformations réelles sont en règle générale irréversibles, les transformations réversibles constituant un cas limite idéal.

Prenons à nouveau l'exemple d'un gaz contenu dans un cylindre fermé par un piston mobile. Pour comprimer le gaz, il faudra exercer sur le piston une pression p^{ex} supérieure à la pression p du gaz. Si les frottements du piston le long des parois du cylindre sont négligeables, la compression pourra s'effectuer de façon réversible : on prendra pour cela p^{ex} à peine supérieure à p; il suffira alors de réduire très légèrement p^{ex} pour qu'elle devienne inférieure à p et que le piston se déplace en sens inverse (détente du gaz, et non plus compression). Mais si les *frottements solides* ne sont pas négligeables, le piston ne se mettra en mouvement que si la différence de pression $|p^{ex} - p|$ est supérieure à un certain seuil non-nul; même si l'on arrive alors à opérer de façon quasistatique (mouvement lent du piston, de façon que p soit constamment définie), la transformation ne pourra pas être réversible : pour inverser son sens, il faudra changer p^{ex} d'une quantité finie[25].

III. VARIATION D'ENTROPIE DANS CERTAINES CLASSES DE TRANSFORMATIONS

1. EXPRESSION GÉNÉRALE DE L'ACCROISSEMENT D'ENTROPIE DANS UNE TRANSFORMATION INFINITÉSIMALE

Considérons un système subissant une transformation infinitésimale quelconque : son énergie s'accroît de dE, son nombre de particules de dN, et les autres

25. Un autre exemple intéressant est fourni par les échanges de chaleur entre deux systèmes $\mathcal{S}_1$ et $\mathcal{S}_2$ dont les températures sont initialement différentes (II, § III.A.2). Choisissons $\mathcal{S}_1$ comme notre système; $\mathcal{S}_2$ est alors le milieu extérieur. Si l'on établit le contact thermique sans précautions, le système n'est à l'équilibre que dans les états initial et final; pendant la transformation, il n'est donc pas possible de définir sa température (pas plus d'ailleurs celle de $\mathcal{S}_2$). Pour régulariser et ralentir la transformation, on peut envisager de ne mettre les deux systèmes en contact qu'à travers une mince barre dont une extrémité est plongée dans $\mathcal{S}_1$, l'autre dans $\mathcal{S}_2$ (fig. 6). On aura alors entre les deux systèmes un flux régulier de chaleur, et l'on peut choisir la barre de façon à obtenir un flux suffisamment lent pour que la transformation soit quasistatique : le système $\mathcal{S}_1$ est à chaque instant dans un état d'équilibre, où sa température, notamment, est parfaitement définie. Mais cette transformation n'est pas pour autant réversible, si la différence des températures initiales n'est pas infiniment petite : tout au long de la transformation, la température de $\mathcal{S}_1$ est différente de celle du « milieu extérieur » $\mathcal{S}_2$; pour inverser le flux de chaleur, il faudrait modifier de façon non-négligeable la température de $\mathcal{S}_2$.

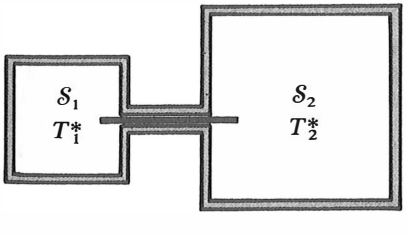


FIGURE 6

Deux systèmes $\mathcal{S}_1$ et $\mathcal{S}_2$, isolés de l'extérieur, sont mis en contact thermique à travers une mince barre. On peut ainsi réaliser un échange de chaleur quasistatique, au cours duquel les températures microcanoniques T_1^* et T_2^* des deux systèmes sont constamment définies, bien qu'elles évoluent dans le temps. Cependant, la transformation ainsi imposée au système $\mathcal{S}_1$, par exemple, n'est pas réversible si les températures sont initialement différentes.

paramètres extérieurs de dx_i (accroissements algébriques). L'accroissement dS^* de son entropie microcanonique s'écrit évidemment

$$dS^* = \frac{\partial S^*}{\partial E}\,dE + \frac{\partial S^*}{\partial N}\,dN + \sum_i \frac{\partial S^*}{\partial x_i}\,dx_i\,, \tag{D.55}$$

c'est-à-dire, d'après la définition de la température et du potentiel chimique microcanoniques (II, formules (III.2) et (III.29.b)), et d'après la relation (D.23) ci-dessus,

$$dS^* = \frac{1}{T^*}\,dE - \frac{\mu^*}{T^*}\,dN + \sum_i \frac{\overline{X}_i}{T^*}\,dx_i\,. \tag{D.56}$$

Si l'on utilise le « premier principe » (formule (D.52)) qui relie l'accroissement d'énergie dE au travail $đW$, à la chaleur $đQ$ et à l'énergie « chimique » $đ\Phi$ reçus par le système, on met l'expression précédente sous la forme

$$dS^* = \frac{1}{T^*}\,[đW + đ\Phi + đQ] - \frac{\mu^*}{T^*}\,dN + \sum_i \frac{\overline{X}_i}{T^*}\,dx_i\,. \tag{D.57}$$

On peut enfin expliciter $đW$ et $đ\Phi$ selon les formules (D.46) et (D.50), ce qui donne

$$dS^* = \frac{1}{T^*}\,đQ - \frac{1}{T^*}\,(\mu^* - \mu^{ex})\,dN + \sum_i \frac{1}{T^*}\,(\overline{X}_i - \varepsilon_i X_i^{ex})\,dx_i\,. \tag{D.58}$$

2. VARIATION D'ENTROPIE DANS UNE TRANSFORMATION RÉVERSIBLE

Une transformation *réversible* est toujours constituée d'une *suite continue de transformations réversibles infinitésimales,* puisque les conditions de réversibilité de la transformation globale portent sur la situation à chaque stade intermédiaire (§ II.3.c).

Or, si l'on tient compte des égalités (D.54) qui définissent une transformation réversible, l'expression (D.58) de l'accroissement infinitésimal d'entropie se simplifie considérablement[26] :

$$dS^{*\text{rév}} = \frac{1}{T^*}\,đQ^{\text{rév}} = \frac{1}{T^{ex}}\,đQ^{\text{rév}}\,. \tag{D.59}$$

Il suffit d'intégrer cette expression le long d'une transformation réversible finie pour obtenir

$$\Delta S^{*\text{rév}} \,\widehat{=}\, [S^{*(2)} - S^{*(1)}]^{\text{rév}} = \int_{(1)}^{(2)} \frac{1}{T^*}\,đQ^{\text{rév}}\,. \tag{D.60}$$

Deux remarques importantes peuvent être formulées à propos de ce résultat. En premier lieu, les égalités (D.54) que nous avons utilisées permettent de *calculer le travail reçu par le système au cours d'une transformation réversible, ainsi que l'énergie « chimique » que lui apporte l'échange réversible de particules, à partir des propriétés du système lui-même :* ces égalités impliquent en effet (cf. (D.46) et (D.49))

$$W_{1\to 2}^{\text{rév}} = -\sum_i \int \overline{X}_i\,dx_i\,, \tag{D.61.a}$$

$$\Phi_{1\to 2}^{\text{rév}} = \int_{(1)}^{(2)} \mu^*\,dN \tag{D.61.b}$$

(noter que l'expression (D.61.a) fait intervenir le même signe pour toutes les contributions, quel que soit le ε_i correspondant). D'autre part, la formule (D.60) fournit un moyen pratique pour *calculer la différence d'entropie entre deux états d'équilibre*

26. Si, dans une transformation réversible infinitésimale, $đW^{\text{rév}}$ et $đ\Phi^{\text{rév}}$ ne sont pas tous deux nuls, $đQ^{\text{rév}}$ n'est pas une différentielle totale. La formule (D.59) montre par contre que $đQ^{\text{rév}}/T^*$ en est une, celle de l'entropie microcanonique.

du système. Cette différence d'entropie est en effet définie de façon unique par la donnée des deux états d'équilibre choisis; or si l'on imagine un *chemin réversible* pouvant les relier, on saura évaluer par (D.60) la variation d'entropie le long de cette transformation réversible; on obtiendra donc ainsi *la* différence d'entropie cherchée.

3. VARIATION D'ENTROPIE DANS UNE TRANSFORMATION ADIABATIQUE

a. Transformation adiabatique quelconque

Considérons un système subissant une transformation *adiabatique.* Nous cherchons à caractériser la variation de son entropie microcanonique

$$\Delta S^{*\mathrm{ad}} \widehat{=} [S^{*(2)} - S^{*(1)}]^{\mathrm{ad}} \tag{D.62}$$

provoquée par la transformation.

Le système n'échange avec l'extérieur que du travail. On peut dès lors imaginer que ces échanges se font exclusivement avec des *systèmes purement mécaniques* (§ I.2.d) : il suffit, pour chaque force généralisée X_i^{ex} et à chaque instant, de remplacer le système extérieur qui produit X_i^{ex} par un système purement mécanique donnant exactement cette même force[27].

Considérons alors l'ensemble réunissant le système étudié et les systèmes purement mécaniques avec lesquels il échange du travail. Cet ensemble constitue un système qui reste *constamment isolé* au cours de la transformation que nous avons imaginée. Par conséquent, son entropie microcanonique ne peut que croître[28]. Comme celle des systèmes purement mécaniques reste invariable, *l'entropie du système étudié ne peut que croître :*

$$\boxed{\Delta S^{*\mathrm{ad}} \geqslant 0} \ . \tag{D.63}$$

b. Application à une transformation adiabatique infinitésimale mettant en jeu un seul paramètre extérieur

Reprenons la formule générale (D.58), et appliquons-la à une transformation infinitésimale *adiabatique* faisant varier *un seul paramètre extérieur*, noté x :

$$\mathrm{d}S^{*\mathrm{ad}} = \frac{1}{T^*} \left(\overline{X} - \varepsilon X^{\mathrm{ex}} \right) \mathrm{d}x^{\mathrm{ad}} \ . \tag{D.64}$$

---

27. Comme le travail $W_{1\to 2}^{\mathrm{ad}}$ reçu par le système est le même pour tous les chemins *adiabatiques* conduisant de l'état (1) à l'état (2) (cf. § II.3.a), il n'est même pas nécessaire d'obliger les systèmes purement mécaniques à « suivre » pas à pas la transformation qu'on étudie. On peut envisager une *autre transformation adiabatique* partant du *même état initial* (1) et aboutissant au *même état final* (2), dans laquelle les systèmes purement mécaniques fournissent au système le *même travail* $W_{1\to 2}^{\mathrm{ad}}$ que la transformation considérée au départ. Prenons par exemple la situation de la figure 5. On veut amener le gaz du volume $V^{(1)}$ au volume $V^{(2)}$ en lui fournissant une quantité de travail $W_{1\to 2}^{\mathrm{ad}}$ prescrite à l'avance. Il suffit d'ajuster la raideur K du ressort pour que

$$\frac{1}{2} K [x_R^{(1)}]^2 - \frac{1}{2} K [x_R^{(2)}]^2 = W_{1\to 2}^{\mathrm{ad}} \ ,$$

où $x_R^{(1)}$ et $x_R^{(2)}$ sont les élongations du ressort correspondant à $V^{(1)}$ et $V^{(2)}$ respectivement.

28. Du point de vue du système global isolé, la transformation consiste à relâcher des contraintes intérieures : dans la transformation « simplifiée » de la note 27, par exemple, les systèmes purement mécaniques sont bloqués initialement dans des positions correspondant aux valeurs $x_i^{(1)}$ des paramètres; on les libère ensuite, et ils ont été ajustés de façon que l'équilibre atteint ainsi par le système global implique, pour le système étudié, l'état d'équilibre final (2).

La conclusion du paragraphe précédent implique

$$(\overline{X} - \varepsilon X^{ex})\, dx^{ad} \geqslant 0 . \tag{D.65}$$

Ce résultat peut être explicité sous la forme suivante :

$$\begin{aligned} \overline{X} > \varepsilon X^{ex} &\implies dx^{ad} > 0 , \\ \overline{X} < \varepsilon X^{ex} &\implies dx^{ad} < 0 . \end{aligned} \tag{D.66}$$

Si par exemple x est le volume V (alors $\varepsilon = +1$), on trouve ainsi que

$$\begin{aligned} p > p^{ex} &\implies dV^{ad} > 0 , \\ p < p^{ex} &\implies dV^{ad} < 0 : \end{aligned} \tag{D.67}$$

dans une transformation *adiabatique,* le système augmente de volume si sa pression est supérieure à la pression extérieure, et diminue de volume dans le cas contraire. Les implications (D.66) s'interprètent de manière analogue si l'on choisit pour x un paramètre extérieur autre que le volume : dans une transformation adiabatique infinitésimale, le sens de variation de x est celui que prédisent les lois de la mécanique.

Notons que, *si l'on permet à deux paramètres extérieurs de varier simultanément, seule la somme des deux termes* figurant alors dans dS^{*ad} *est nécessairement positive;* dans certaines situations, l'un des termes peut être négatif, pourvu que l'autre soit suffisamment positif pour rétablir le sens de l'inégalité.

Ces résultats sont analogues à ceux que nous avons établis au chapitre II, paragraphe III.B.3.b. Toutefois, nous avions alors supposé que les deux systèmes échangeant par exemple du volume étaient initialement en équilibre thermique l'un avec l'autre (égalité des températures microcanoniques); ici, nous supposons au contraire qu'ils n'échangent pas de chaleur. Il est donc remarquable que la conclusion soit la même dans ces deux cas fort différents.

c. Cas d'une transformation adiabatique réversible

Considérons une transformation *réversible.* C'est une suite continue de transformations infinitésimales réversibles, auxquelles s'applique le résultat (D.59). Si la transformation est en outre *adiabatique,* l'égalité (D.59) se réduit à

$$dS^{*rév,\,ad} = 0 , \tag{D.68}$$

de sorte que, pour la transformation globale,

$$\Delta S^{*rév,\,ad} \widehat{=} [S^{*(2)} - S^{*(1)}]^{rév,\,ad} = 0 : \tag{D.69}$$

une transformation réversible et adiabatique ne change pas l'entropie du système; on dit d'une telle transformation qu'elle est *« isentropique »*[29].

On peut également partir de l'inégalité (D.63). Le résultat que nous venons d'établir montre que cette inégalité devient une égalité dans le cas où la transformation adiabatique envisagée est réversible. Inversement, *une transformation adiabatique qui ne change pas l'entropie est nécessairement réversible :* dans le raisonnement qui a conduit à (D.63), le système global isolé ne peut garder une valeur constante pour son entropie microcanonique que s'il est déjà à l'équilibre au début de la transformation (entropie maximale) et qu'il reste à l'équilibre tout au long de cette transformation; ceci correspond exactement à la définition d'une transformation réversible (§ II.3.c).

Les résultats des paragraphes 2 et 3 sont l'expression du *« second principe de la thermodynamique »* dans des cas particuliers : celui des transformations réversibles (formule (D.60)) et celui des transformations adiabatiques (formules (D.63) et (D.69)). Pour compléter la démonstration du « second principe », nous aurons besoin de la notion de thermostat, qui sera introduite au chapitre III. Les raisonnements de ce complément seront poursuivis dans le complément III.L.

---

29. La caractéristique essentielle d'un système purement mécanique étant d'avoir une entropie invariable, on voit que tout système isolé thermiquement se comporte comme un système purement mécanique tant que les transformations qu'il subit peuvent être considérées comme réversibles.

COMPLÉMENT II.E

Mouvement macroscopique d'un système totalement isolé en équilibre statistique

I. Étude statistique d'un système totalement isolé
  1. Système macroscopique infinitésimal
    a. Définition. Paramètres mécaniques
    b. Énergie interne
    c. Entropie
  2. Équilibre statistique d'un système isolé en mouvement macroscopique
    a. Paramètres caractérisant l'état macroscopique du système à un instant donné
    b. Conditions d'équilibre statistique à un instant donné
    c. Mouvement macroscopique compatible avec l'équilibre statistique
    d. Répartition des masses à l'équilibre statistique
  3. Énergie interne d'un système en mouvement macroscopique
  4. Entropie et paramètres extérieurs d'un système totalement isolé

II. Conséquences importantes
  1. La température est une grandeur positive
  2. Étude des systèmes partiellement isolés (à énergie fixée)
    a. Position du problème. Condition d'équilibre
    b. Distribution statistique de la quantité de mouvement et du moment cinétique

Les particules constitutives d'un système macroscopique (atomes, molécules, ions, électrons...) sont animées d'un mouvement incessant à l'échelle microscopique : c'est ce qu'on appelle « l'agitation thermique ». Par opposition, on dit qu'un corps est en « mouvement macroscopique » si un (ou plusieurs) de ses sous-systèmes de taille macroscopique est animé d'un mouvement d'ensemble.

Le but de ce complément est de trouver *les mouvements macroscopiques compatibles avec l'équilibre statistique pour un système totalement isolé.* Nous en déduirons deux conséquences importantes : d'une part le signe de la température microcanonique, d'autre part la distribution statistique de la quantité de mouvement et du moment cinétique dans les systèmes « isolés » que nous avons étudiés au chapitre II (dont l'énergie est fixée, mais pas la quantité de mouvement ni le moment cinétique).

I. ÉTUDE STATISTIQUE D'UN SYSTÈME TOTALEMENT ISOLÉ

Dans l'ensemble de ce complément, nous n'examinerons que *des systèmes ou des sous-systèmes macroscopiques.* Leur mouvement pourra donc être analysé dans le cadre de la *mécanique classique.*

1. SYSTÈME MACROSCOPIQUE INFINITÉSIMAL

a. Définition. Paramètres mécaniques

Nous dirons qu'un système macroscopique est *infinitésimal* s'il est *de taille infinitésimale à l'échelle macroscopique,* c'est-à-dire si, à cette échelle, il peut être considéré comme un « point matériel », bien qu'il soit constitué d'un très grand nombre N de particules. C'est la valeur gigantesque du nombre d'Avogadro N_A qui permet l'existence de tels systèmes : lorsque N_A particules occupent un volume dont les

dimensions linéaires sont de l'ordre du mètre, on peut diviser ce volume en éléments de taille négligeable mais contenant un nombre N de particules tel que[1]

$$1 \ll N \ll N_A \,. \tag{E.1}$$

Cependant, on veut aussi pouvoir négliger les interactions entre systèmes macroscopiques infinitésimaux; il faut donc que leurs dimensions soient grandes devant la portée des interactions entre particules (cf. complément II.A, § I). Comme celle-ci est en général microscopique, de l'ordre d'une dizaine d'angströms, nous supposerons cette condition satisfaite[2].

Du point de vue mécanique, un système macroscopique infinitésimal est caractérisé par sa position $\vec{R}$, sa masse M, sa quantité de mouvement $\vec{P} = M\vec{V}$, et son énergie E; du fait de sa toute petite taille, il n'a pas de moment cinétique propre[3].

b. Énergie interne

Les paramètres mécaniques d'un système infinitésimal sont reliés simplement à ceux des particules qui le constituent :

$$M = \sum_{i=1}^{N} m_i \,, \tag{E.2.a}$$

$$\vec{R} = \frac{1}{M} \sum_{i=1}^{N} m_i \vec{r}_i \,, \tag{E.2.b}$$

$$\vec{P} = \sum_{i=1}^{N} m_i \vec{v}_i \,, \tag{E.2.c}$$

où m_i est la masse de la i-ème particule, $\vec{r}_i$ sa position et $\vec{v}_i$ sa vitesse.

Mais il existe une différence essentielle entre un système macroscopique infinitésimal et un simple point matériel : son énergie n'est pas seulement l'énergie cinétique $P^2/2M$ associée à son mouvement d'ensemble, mais la somme des énergies cinétiques et de l'énergie potentielle d'interaction U de ses N particules :

$$E = \sum_{i=1}^{N} \frac{1}{2} m_i \vec{v}_i^{\,2} + U(\vec{r}_1, \vec{r}_2, \ldots, \vec{r}_N) \,. \tag{E.3}$$

Cette énergie est donc non nulle, même si le système est macroscopiquement au repos ($\vec{P} = \vec{0}$). *On appelle énergie interne E_{int} d'un système macroscopique infinitésimal* (à un instant donné), *son énergie mesurée dans le référentiel galiléen par rapport auquel il est au repos* (à cet instant). Dans ce référentiel, les vitesses $\vec{v}_i^{\,0}$ des particules vérifient

$$\sum_{i=1}^{N} m_i \vec{v}_i^{\,0} = \vec{0} \,, \tag{E.4}$$

et l'énergie interne s'écrit

$$E_{\text{int}} = \sum_{i=1}^{N} \frac{1}{2} m_i (\vec{v}_i^{\,0})^2 + U \,. \tag{E.5}$$

---

1. Par exemple il y a environ 3.10^{10} molécules dans 1 μm^3 d'eau. Le même volume d'un gaz dans les conditions normales en contient encore 3.10^7.

2. Ceci exclut de notre étude les systèmes dans lesquels les forces gravitationnelles jouent un rôle essentiel.

3. Un système macroscopique infinitésimal est toujours un sous-système d'un système plus vaste. Ses dimensions sont très petites devant celles du système dont il fait partie, de sorte que son moment cinétique par rapport à son propre centre de masse est négligeable devant son moment cinétique par rapport au centre de masse du système global.

Si l'on utilise un autre référentiel galiléen, dans lequel le système est animé d'une vitesse d'ensemble $\vec{V}$, son énergie E est la somme de son énergie interne et de son énergie cinétique macroscopique[4] :

$$E = E_{\text{int}} + \frac{1}{2} M\vec{V}^2 = E_{\text{int}} + \frac{\vec{P}^2}{2M} . \tag{E.6}$$

c. Entropie

Considérons un système macroscopique infinitésimal *isolé et à l'équilibre statistique.* Calculons son entropie microcanonique S^* dans le référentiel où il est au repos; elle est fonction de son énergie, c'est-à-dire en l'occurrence de son énergie interne E_{int}. Mais le nombre d'états quantiques accessibles au système est évidemment indépendant du référentiel; il en est de même de l'entropie microcanonique, proportionnelle au logarithme de ce nombre. Si donc on étudie le système dans un référentiel galiléen par rapport auquel son centre de masse est animé de la vitesse $\vec{V} = \vec{P}/M$, son entropie s'écrit

$$S^* = S^*(E_{\text{int}}) = S^*\left(E - \vec{P}^2/2M\right) . \tag{E.7}$$

L'entropie est invariante par changement de référentiel galiléen; il en est de même de la température[5] :

$$\frac{1}{T^*} \widehat{=} \frac{\partial S^*}{\partial E} = \frac{\partial S^*}{\partial E_{\text{int}}} . \tag{E.8}$$

---

4. Soit $\vec{v}_i$ la vitesse de la i-ième particule dans le référentiel où le centre de masse est animé de la vitesse $\vec{V}$:

$$\vec{v}_i = \vec{v}_i^{\,0} + \vec{V} . \tag{1}$$

L'énergie dans ce référentiel est d'après (E.3)

$$\begin{aligned} E &= \sum_{i=1}^{N} \frac{1}{2} m_i (\vec{v}_i^{\,0} + \vec{V})^2 + U \\ &= \sum_{i=1}^{N} \frac{1}{2} m_i (\vec{v}_i^{\,0})^2 + U + \frac{1}{2} \vec{V}^2 \sum_{i=1}^{N} m_i + \vec{V} \cdot \sum_{i=1}^{N} m_i \vec{v}_i^{\,0} . \end{aligned} \tag{2}$$

La relation (E.6) découle alors de (E.2.a), (E.4) et (E.5).

5. La conclusion est différente dans le domaine *relativiste.* En effet le quadrivecteur $(E, \vec{P})$ est alors relié au quadrivecteur $(E_{\text{int}}, \vec{0})$ par la transformation de Lorentz; en particulier la relation entre E et E_{int} s'écrit

$$E = \frac{E_{\text{int}}}{\sqrt{1 - V^2/c^2}} . \tag{1}$$

La température T^* dans le référentiel où la vitesse du centre de masse est $\vec{V}$ est dans ce cas donnée par

$$\frac{1}{T^*} \widehat{=} \frac{\partial S^*}{\partial E} = \frac{\partial S^*}{\partial E_{\text{int}}} \cdot \frac{\partial E_{\text{int}}}{\partial E} ; \tag{2}$$

elle est donc reliée à la température T_0^* mesurée dans le référentiel du centre de masse par

$$T^* = \frac{T_0^*}{\sqrt{1 - V^2/c^2}} . \tag{3}$$

2. ÉQUILIBRE STATISTIQUE D'UN SYSTÈME ISOLÉ EN MOUVEMENT MACROSCOPIQUE

Nous considérons un système $\mathcal{S}$ *totalement isolé,* c'est-à-dire un système dont *l'énergie, la quantité de mouvement et le moment cinétique restent constants au cours du temps.* Par contre *la forme et le volume du système ne sont pas fixés.* Autrement dit, l'énergie, la quantité de mouvement et le moment cinétique sont des paramètres extérieurs fixés une fois pour toutes, alors que la forme et le volume sont caractérisés par des variables internes qui s'ajustent de façon à satisfaire les conditions d'équilibre statistique[6]. Nous nous proposons ici de déterminer ces conditions d'équilibre et d'en déduire les mouvements macroscopiques du système qui sont compatibles avec son équilibre statistique.

Pour simplifier les calculs, nous nous plaçons d'emblée dans le référentiel galiléen où *le centre de masse du système $\mathcal{S}$ est immobile;* on s'affranchit ainsi du mouvement rectiligne uniforme du système dans son ensemble, qui ne présente pas d'intérêt[7]. Nous supposons aussi $\mathcal{S}$ constitué d'un seul type de particules[8], de masse m (l'extension à des systèmes plus complexes se fait sans difficulté).

a. Paramètres caractérisant l'état macroscopique du système à un instant donné

Considérons le système $\mathcal{S}$ à un *instant donné.* Son état macroscopique, à cet instant, est défini par l'état de chacun de ses « points matériels ». Pour préciser, divisons l'espace en petits volumes de taille infinitésimale (à l'échelle macroscopique), caractérisés par leur position $\vec{R}_\alpha$ ($\alpha = 1, 2, 3, \ldots$). Chacun de ces petits volumes, de dimensions et de position fixées, constitue un sous-système macroscopique infinitésimal $\mathcal{S}_\alpha$ du système $\mathcal{S}$. L'état macroscopique de $\mathcal{S}$ est déterminé par la donnée du nombre de particules N_α, de l'énergie E_α et de la quantité de mouvement $\vec{P}_\alpha$ de chacun des sous-systèmes $\mathcal{S}_\alpha$.

L'ensemble des N_α, E_α, $\vec{P}_\alpha$ sont des variables internes du système $\mathcal{S}$, libres de fluctuer, à condition de satisfaire aux relations de conservation que doit vérifier un système totalement isolé[9] :

(i) Conservation du nombre total de particules N :

$$\sum_\alpha N_\alpha = N \; ; \tag{E.9.a}$$

(ii) Conservation de l'énergie totale E :

$$\sum_\alpha E_\alpha = E \; ; \tag{E.9.b}$$

---

6. S'il s'agit d'un liquide ou d'un solide, nous négligeons les phénomènes d'évaporation ou de sublimation. Dans le cas d'un solide, les modifications nécessaires pour parvenir à l'état d'équilibre statistique peuvent prendre un temps extrêmement long. Mais notre propos est ici seulement de déterminer le *véritable état d'équilibre* du système; la façon dont cet état peut être atteint est un problème beaucoup plus compliqué.

7. D'après le principe de relativité, les conclusions physiques sont les mêmes dans tous les référentiels galiléens.

8. Cette hypothèse n'est pas réaliste dans le cas d'un gaz : il faut l'enfermer dans une enveloppe (de nature différente) déformable mais assez rigide pour l'empêcher de s'épandre dans tout l'espace. Dans un tel cas, il faudrait appliquer les raisonnements qui suivent à chacun des deux sous-systèmes (gaz d'une part, enveloppe d'autre part) puis imposer les conditions d'équilibre dues à la présence de l'interface.

9. Il se peut que certains des petits volumes qui « pavent » l'espace soient vides de particules; pour un tel sous-système $\mathcal{S}_\alpha$, les valeurs de N_α, E_α et $\vec{P}_\alpha$ sont nulles ainsi que sa contribution aux relations de conservation.

(iii) Conservation de la quantité de mouvement totale (qui est nulle dans le référentiel du centre de masse) :

$$\sum_\alpha \vec{P}_\alpha = \vec{0} \; ; \tag{E.9.c}$$

(iv) Conservation du moment cinétique total $\vec{\mathcal{M}}$ [10] :

$$\sum_\alpha \vec{R}_\alpha \wedge \vec{P}_\alpha = \vec{\mathcal{M}} \; . \tag{E.9.d}$$

Ces quatre quantités, nombre de particules, énergie, quantité de mouvement et moment cinétique, sont des *constantes du mouvement,* indépendantes les unes des autres, dont les valeurs caractérisent l'état du système totalement isolé $\mathcal{S}$.

b. Conditions d'équilibre statistique à un instant donné

Quelles sont, à l'instant considéré, les valeurs d'équilibre (c'est-à-dire les valeurs les plus probables) des variables internes N_α, E_α, $\vec{P}_\alpha$? Ce sont celles qui rendent l'entropie microcanonique partielle s^* maximum (II, § II.A). Cette entropie partielle est la somme des entropies microcanoniques S^*_α des sous-systèmes $\mathcal{S}_\alpha$, prises pour les valeurs N_α, E_α, $\vec{P}_\alpha$. Or (formule (E.7)) S^*_α ne dépend pas séparément de E_α et $\vec{P}_\alpha$ mais seulement de la combinaison $(E_\alpha - \vec{P}_\alpha{}^2/2mN_\alpha)$, énergie interne du système $\mathcal{S}_\alpha$; on a donc

$$s^*(\{N_\alpha\}, \{E_\alpha\}, \{\vec{P}_\alpha\}) = \sum_\alpha S^*_\alpha \left(E_\alpha - \frac{\vec{P}_\alpha{}^2}{2mN_\alpha}, N_\alpha \right) . \tag{E.10}$$

Nous avons à rechercher les valeurs des N_α, E_α, $\vec{P}_\alpha$, soumises aux conditions (E.9), *qui rendent s^* maximum; ce sont ces valeurs qui caractérisent l'équilibre statistique du système totalement isolé $\mathcal{S}$.* Les $\mathcal{S}_\alpha$ étant macroscopiques, tout écart appréciable à ces valeurs est hautement improbable.

La méthode la plus élégante pour trouver le maximum d'une fonction de plusieurs variables liées est la *méthode des multiplicateurs de Lagrange* (appendice IV). Elle consiste à former, à partir de la fonction (E.10) et des relations (E.9), la fonction auxiliaire

$$\begin{aligned} &Y(\{N_\alpha\}, \{E_\alpha\}, \{\vec{P}_\alpha\}) \\ &\quad \widehat{=} \sum_\alpha \left\{ S^*_\alpha \left(E_\alpha - \frac{\vec{P}_\alpha{}^2}{2mN_\alpha}, N_\alpha \right) + aN_\alpha + bE_\alpha + \vec{c} \, . \vec{P}_\alpha + \vec{d} \, . (\vec{R}_\alpha \wedge \vec{P}_\alpha) \right\} , \end{aligned} \tag{E.11}$$

où a, b et les composantes des vecteurs $\vec{c}$ et $\vec{d}$ sont huit paramètres, appelés multiplicateurs de Lagrange. Les valeurs cherchées des variables N_α, E_α, $\vec{P}_\alpha$ sont celles qui rendent Y maximum *comme si ces variables étaient indépendantes.* Les multiplicateurs sont déterminés (si nécessaire) à la fin du calcul, en reportant les valeurs qui rendent Y extrémale dans les relations (E.9).

La dérivée de Y par rapport à chacune des variables doit donc s'annuler pour les valeurs d'équilibre de ces variables. A chaque type de dérivée correspond une condition d'équilibre statistique du système.

---

10. Dans le référentiel du centre de masse, $\vec{\mathcal{M}}$ ne dépend pas du choix de l'origine des coordonnées. En effet, si l'on change d'origine, chaque $\vec{R}_\alpha$ est remplacé par

$$\vec{R}'_\alpha = \vec{R}_\alpha + \vec{R}_0 \; . \tag{1}$$

Le nouveau moment cinétique s'écrit

$$\vec{\mathcal{M}}' \widehat{=} \sum_\alpha \vec{R}'_\alpha \wedge \vec{P}_\alpha = \sum_\alpha \vec{R}_\alpha \wedge \vec{P}_\alpha + \vec{R}_0 \wedge \sum_\alpha \vec{P}_\alpha \; , \tag{2}$$

c'est-à-dire, compte tenu de (E.9.c),

$$\vec{\mathcal{M}}' = \vec{\mathcal{M}} \; . \tag{3}$$

Les dérivées par rapport aux E_α donnent

$$\frac{\partial Y}{\partial E_\alpha}=\frac{\partial S_\alpha^*}{\partial E_\alpha}+b=\frac{1}{T_\alpha^*}+b=0\ ,\qquad \alpha=1,\,2,\,3,\,\ldots \tag{E.12}$$

T_α^* est donc une constante $T^*=-1/b$, indépendante de α. Ce résultat n'est pas nouveau (II, § III.A) : *à l'équilibre, tous les sous-systèmes ont même température.*

L'annulation des dérivées de Y par rapport aux composantes des $\vec{P}_\alpha$ entraîne[11], compte tenu du résultat précédent ($T_\alpha^*=T^*$),

$$-\frac{\vec{P}_\alpha}{mN_\alpha T^*}+\vec{c}+\vec{d}\wedge\vec{R}_\alpha=\vec{0}\ ,\qquad \alpha=1,\,2,\,3,\,\ldots \tag{E.13}$$

Enfin les dérivées de Y par rapport aux N_α conduisent aux relations

$$\frac{\partial Y}{\partial N_\alpha}=\frac{\partial S_\alpha^*}{\partial N_\alpha}+\frac{\partial S_\alpha^*}{\partial E_\alpha}\frac{\vec{P}_\alpha{}^2}{2mN_\alpha^2}+a=0\ ,\qquad \alpha=1,\,2,\,3,\,\ldots \tag{E.14}$$

Le premier terme est la dérivée partielle par rapport à N_α de la fonction $S_\alpha^*(E_\alpha^{\text{int}}, N_\alpha)$; c'est donc, au facteur $-1/T_\alpha^*$ près, le potentiel chimique de $\mathcal{S}_\alpha$ calculé dans le référentiel galiléen où ce sous-système est immobile. Par analogie avec l'énergie interne, nous l'appellerons « potentiel chimique interne » :

$$\mu_\alpha^{\text{int}}\mathrel{\widehat{=}}-T_\alpha^*\frac{\partial S_\alpha^*(E_\alpha^{\text{int}},N_\alpha)}{\partial N_\alpha}\ . \tag{E.15}$$

Les conditions d'équilibre par rapport aux échanges de particules (E.14) s'écrivent donc (puisque $T_\alpha^*=T^*$)

$$-\mu_\alpha^{\text{int}}+\frac{\vec{P}_\alpha{}^2}{2mN_\alpha^2}+aT^*=0\ ,\qquad \alpha=1,\,2,\,3,\,\ldots \tag{E.16}$$

c. Mouvement macroscopique compatible avec l'équilibre statistique

Considérons le système $\mathcal{S}$ à l'équilibre à l'instant $t=0$. Les valeurs des N_α, E_α et $\vec{P}_\alpha$ sont alors déterminées par les relations (E.9), (E.12), (E.13) et (E.16). Supposons maintenant que chacun des sous-systèmes $\mathcal{S}_\alpha$ se déplace aux instants ultérieurs en suivant le mouvement d'ensemble des particules qu'il contient, c'est-à-dire de façon que

$$\frac{\mathrm{d}\vec{R}_\alpha}{\mathrm{d}t}=\frac{\vec{P}_\alpha}{mN_\alpha}\mathrel{\widehat{=}}\vec{V}_\alpha\ ,\qquad \alpha=1,\,2,\,3,\,\ldots\ , \tag{E.17}$$

où $\vec{P}_\alpha$ et N_α sont fixés aux valeurs d'équilibre. Les relations qui assurent l'équilibre statistique à $t=0$ resteront valables à tout instant t ultérieur pour les nouvelles positions $\vec{R}_\alpha$, à condition que l'ensemble des volumes occupés par les divers $\mathcal{S}_\alpha$ constitue un pavage complet de l'espace à l'instant t comme à l'instant 0. Or c'est justement ce qu'indiquent les relations (E.13). En effet, prenons la différence de deux d'entre elles :

$$\vec{V}_\alpha-\vec{V}_{\alpha'}-T^*\vec{d}\wedge(\vec{R}_\alpha-\vec{R}_{\alpha'})=\vec{0}\ ,\qquad \alpha,\,\alpha'=1,\,2,\,3,\,\ldots \tag{E.18}$$

En multipliant scalairement (E.18) par $(\vec{R}_\alpha-\vec{R}_{\alpha'})$, on obtient

$$(\vec{V}_\alpha-\vec{V}_{\alpha'})\cdot(\vec{R}_\alpha-\vec{R}_{\alpha'})=\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}(\vec{R}_\alpha-\vec{R}_{\alpha'})^2=0\ , \tag{E.19}$$

dont l'interprétation est très simple : les distances relatives des différents sous-systèmes

---

11. Dérivons Y par rapport à la composante $P_{\alpha x}$ de $\vec{P}_\alpha$. Il vient

$$\frac{\partial Y}{\partial P_{\alpha x}}=\frac{\partial S_\alpha^*}{\partial E_\alpha}\left(-\frac{P_{\alpha x}}{mN_\alpha}\right)+c_x+d_yR_{\alpha z}-d_zR_{\alpha y}=-\frac{1}{T_\alpha^*}\frac{P_{\alpha x}}{mN_\alpha}+c_x+(\vec{d}\wedge\vec{R}_\alpha)_x\ .$$

On trouve des expressions semblables pour $\partial Y/\partial P_{\alpha y}$ et $\partial Y/\partial P_{\alpha z}$, d'où la relation vectorielle (E.13).

$\mathcal{S}_\alpha$ sont des constantes, indépendantes du temps; tout mouvement relatif d'un sous-système par rapport aux autres est exclu. Les conditions d'équilibre (E.12), (E.13) et (E.16), établies à un instant particulier, restent valables à tout instant et *à l'équilibre statistique, le système totalement isolé se comporte comme un solide indéformable.*

Si l'on choisit pour $\mathcal{S}_\alpha$, le sous-système dont la position coïncide avec le centre de masse (immobile) O du système $\mathcal{S}$, et qu'*on prend le point O pour origine,* les relations (E.18) s'écrivent, comme pour un solide indéformable[12],

$$\vec{V}_\alpha = \vec{\omega} \wedge \vec{R}_\alpha \,, \tag{E.20}$$

où
$$\vec{\omega} \mathrel{\widehat{=}} T^* \vec{d} \tag{E.21}$$

est le vecteur vitesse angulaire de rotation du système.

Le résultat que nous venons d'obtenir était prévisible : il exprime simplement que l'équilibre statistique ne peut s'accommoder que d'une évolution stationnaire, sans mouvement macroscopique relatif. Mais on peut dire plus. Dans cet état stationnaire, avec la répartition de masse (c'est-à-dire des N_α) qu'il implique, le système $\mathcal{S}$ possède les propriétés cinématiques d'un solide indéformable[13]. En particulier, la relation entre son moment cinétique $\vec{\mathcal{M}}$ et le vecteur rotation $\vec{\omega}$ fait intervenir trois *axes principaux d'inertie,* constituant un trièdre trirectangle lié à $\mathcal{S}$ et centré en O : si I_1, I_2, I_3 sont les moments d'inertie de $\mathcal{S}$ par rapport à ces axes, de vecteurs unitaires $\vec{e}_1$, $\vec{e}_2$, $\vec{e}_3$, on peut écrire

$$\vec{\mathcal{M}} = I_1(\vec{\omega} \,.\vec{e}_1)\vec{e}_1 + I_2(\vec{\omega} \,.\vec{e}_2)\vec{e}_2 + I_3(\vec{\omega} \,.\vec{e}_3)\vec{e}_3 \,. \tag{E.22}$$

Or le moment cinétique $\vec{\mathcal{M}}$ du système isolé $\mathcal{S}$ est indépendant du temps. Si, à un instant donné, $\vec{\omega}$ n'est colinéaire à aucun des trois axes principaux d'inertie, ceux-ci ne sont pas fixes (ils tournent autour de $\vec{\omega}$); $\vec{\omega}$ lui-même doit alors dépendre du temps pour que $\vec{\mathcal{M}}$ reste constant : $\vec{\omega}$ précesse autour de $\vec{\mathcal{M}}$. Mais dans cette rotation caractérisée par une vitesse angulaire $\vec{\omega}$ dépendant du temps, les diverses parties du système sont soumises à des «forces centrifuges» *variables au cours du temps;* la répartition des masses dans le système ne peut pas correspondre alors à un équilibre statistique (voir paragraphe d ci-dessous), et elles vont donc se déplacer les unes par rapport aux autres[14]. Lorsque l'équilibre est atteint, $\vec{\omega}$ *doit donc être dirigé suivant un axe principal d'inertie du système,* auquel cas on a simplement

$$\vec{\mathcal{M}} = I\vec{\omega} \,, \tag{E.23}$$

---

12. Dans un solide indéformable, la vitesse $\vec{V}(P)$ d'un point quelconque P a pour expression générale

$$\vec{V}(P) = \vec{V}(O) + \vec{\omega} \wedge \overrightarrow{OP} \,, \tag{1}$$

où $\vec{V}(O)$ est la vitesse (quelconque) d'un point particulier O fixe par rapport au solide et $\vec{\omega}$ le «vecteur vitesse angulaire instantané de rotation» du solide. On voit que les relations (E.13) donnent des vitesses $\vec{V}_\alpha$ de cette forme. Si l'on choisit pour O le centre de masse du système, immobile dans le référentiel où l'on s'est placé,

$$\vec{V}(O) = T^*\vec{c} = \vec{0} \,; \tag{2}$$

le multiplicateur de Lagrange $\vec{c}$ est nul.

13. Voir, par exemple, L. Landau et E. Lifchitz, *Mécanique*, Éditions de Moscou (1960), chapitre VI.

14. Un *solide indéformable* isolé peut, en principe, effectuer un mouvement de précession autour de son moment cinétique pendant un temps infini. Mais aucun système, même solide, n'est complètement indéformable. Sous l'action des forces d'inertie, les sous-systèmes se déplacent les uns par rapport aux autres jusqu'à ce que l'équilibre soit atteint, c'est-à-dire jusqu'à ce que l'axe de rotation devienne parallèle à $\vec{\mathcal{M}}$. Dans le cas d'un solide, le temps pour que l'équilibre soit atteint peut être très long, et la question est alors de savoir si ce solide peut être considéré comme véritablement isolé sur cette échelle de temps.

où I est le moment d'inertie par rapport à cet axe; alors, $\vec{\omega}$ est indépendant du temps, comme $\vec{\mathcal{M}}$ et I. Notons cependant que la répartition des masses (§ d), et donc I et $\vec{\omega}$, dépendent de la valeur du moment cinétique $\vec{\mathcal{M}}$.

d. Répartition des masses à l'équilibre statistique

Ce sont les relations (E.16) qui fixent la répartition des particules entre les sous-systèmes $\mathcal{S}_\alpha$. En remplaçant $\vec{P}_\alpha$ par son expression tirée de (E.20), on peut les réécrire sous la forme

$$\mu_\alpha^{\text{int}} - \frac{1}{2} m \vec{V}_\alpha^{\,2} = \mu_\alpha^{\text{int}} - \frac{1}{2} m (\vec{\omega} \wedge \vec{R}_\alpha)^2 = aT^* , \quad \alpha = 1, 2, 3, \ldots \tag{E.24}$$

Le terme $-m(\vec{\omega} \wedge \vec{R}_\alpha)^2/2$ s'interprète facilement si l'on se place dans le référentiel (non-galiléen) où le système $\mathcal{S}$ est immobile : dans ce référentiel qui tourne avec $\mathcal{S}$, chaque particule de $\mathcal{S}_\alpha$ est soumise à une « force centrifuge » qui dérive du potentiel[15] $-m(\vec{\omega} \wedge \vec{R}_\alpha)^2/2$. Les particules se répartissent donc de telle façon que la somme du potentiel chimique interne et du potentiel centrifuge auquel les soumet le mouvement de rotation d'ensemble soit la même en tout point du système global $\mathcal{S}$. En l'absence de mouvement, on retrouve le résultat habituel : μ_α^{int} est le même dans tout le système à l'équilibre. Lorsqu'il y a rotation, le système est modifié par la présence de forces centrifuges, mais c'est toujours la quantité

$$\mu_\alpha^* \widehat{=} - T_\alpha^* \frac{\partial S_\alpha^*}{\partial N_\alpha} \left(E_\alpha - \frac{\vec{P}_\alpha^{\,2}}{2mN_\alpha}, N_\alpha \right) = \mu_\alpha^{\text{int}} - \frac{m}{2} (\vec{\omega} \wedge \vec{R}_\alpha)^2 ,$$

$$\alpha = 1, 2, 3 \ldots \tag{E.25}$$

(où $\vec{R}_\alpha$ est repéré à partir du centre de masse O de $\mathcal{S}$), qui est uniforme dans $\mathcal{S}$ à l'équilibre statistique[16] :

$$\mu_\alpha^* = \mu^*, \quad \alpha = 1, 2, 3, \ldots \tag{E.26}$$

C'est cette quantité μ^* qu'on appellera le potentiel chimique de $\mathcal{S}$ (voir § 4, ci-dessous).

En résumé, *les conditions d'équilibre statistique d'un système totalement isolé animé d'un mouvement macroscopique sont :*

— *température uniforme,*

— *potentiel chimique uniforme* (le terme de potentiel centrifuge étant pris en compte),

— *mouvement d'ensemble comme un solide indéformable autour d'un axe principal d'inertie.* Dans un tel mouvement, les vecteurs-positions $\vec{R}_\alpha$ des divers points matériels du système varient au cours du temps, mais de façon simple : dans le référentiel lié au centre de masse O ils tournent à vitesse constante autour de l'axe parallèle à $\vec{\mathcal{M}}$ passant par O.

---

15. Si $R_{\alpha\perp}$ est la distance de $\mathcal{S}_\alpha$ à l'axe de rotation (axe parallèle à $\vec{\omega}$ passant par O), la « force centrifuge » ressentie, dans le référentiel tournant, par une particule de $\mathcal{S}_\alpha$, est radiale (par rapport à cet axe) et a pour module

$$f_\alpha^{(c)} = m\omega^2 R_{\alpha\perp} . \tag{1}$$

On vérifie facilement que cette force dérive du potentiel

$$-\frac{1}{2} m\omega^2 R_{\alpha\perp}^2 = -\frac{1}{2} m (\vec{\omega} \wedge \vec{R}_\alpha)^2 . \tag{2}$$

Remarquons que, puisque la rotation se fait autour d'un axe fixe ($\vec{\omega}$ indépendant du temps), sans déformation ($R_{\alpha\perp}$ indépendant du temps), le potentiel centrifuge est constant dans le temps (cf. paragraphe c).

16. On comparera le résultat (E.25)-(E.26) avec celui obtenu au complément V.B pour un système placé dans un champ de forces extérieur.

3. ÉNERGIE INTERNE D'UN SYSTÈME EN MOUVEMENT MACROSCOPIQUE

L'énergie interne E_{int} d'un système doit être, comme son énergie totale, une *grandeur additive.* Ce sera donc, par *définition,* la somme des énergies internes de chacune de ses parties (formule (E.6)) :

$$E_{\text{int}} \widehat{=} \sum_{\alpha} E_{\alpha}^{\text{int}} = \sum_{\alpha} \left(E_{\alpha} - \frac{1}{2} m N_{\alpha} \vec{V}_{\alpha}^{\,2} \right) . \tag{E.27}$$

Elle est donc égale à l'énergie totale $\sum_{\alpha} E_{\alpha}$ du système (mesurée dans un référentiel galiléen[17]) diminuée de son énergie cinétique macroscopique $\sum_{\alpha} m N_{\alpha} \vec{V}_{\alpha}^{\,2}/2$. Pour un *système isolé en équilibre statistique,* et dans le référentiel galiléen lié au centre de masse O, on obtient aisément[18], [19]

$$E_{\text{int}} = E - \frac{1}{2} I \vec{\omega}^{\,2} = E - \frac{1}{2} \frac{\vec{\mathcal{M}}^{\,2}}{I} , \tag{E.28}$$

où I est le moment d'inertie de $\mathcal{S}$ par rapport à l'axe de rotation et $\vec{\mathcal{M}}^{\,2}/2I$ l'énergie cinétique macroscopique de rotation du système global (il ne faut pas oublier que la distribution des masses dans le système est modifiée lorsqu'il est en rotation, et que I est lui-même une fonction de $\vec{\mathcal{M}}$).

---

17. *Attention :* l'énergie interne d'un corps en rotation n'est pas son énergie mesurée dans le référentiel où il est immobile. En effet, dans ce *référentiel non galiléen,* chaque particule est soumise à une force centrifuge qui dérive de l'énergie potentielle (note 15) $-m(\vec{\omega} \wedge \vec{r}_i)^2/2$. Pour obtenir l'énergie E' dans le référentiel tournant, il faut donc ajouter à l'énergie interne, somme des énergies cinétiques et d'interaction des diverses particules (formule (E.5)), la somme de leurs énergies potentielles centrifuges

$$E^{(c)} = -\frac{1}{2} \sum_{i=1}^{N} m(\vec{\omega} \wedge \vec{r}_i)^2 = -\frac{1}{2} I\omega^2 \tag{1}$$

(cf. note 18 ci-dessous). On trouve ainsi

$$E' = E_{\text{int}} - \frac{1}{2} I\omega^2 . \tag{2}$$

18. La somme des énergies cinétiques macroscopiques peut alors se calculer en utilisant (E.20) :

$$E_{\text{cin}} \widehat{=} \sum_{\alpha} \frac{1}{2} m N_{\alpha} \vec{V}_{\alpha}^{\,2} = \frac{1}{2} \sum_{\alpha} m N_{\alpha} (\vec{\omega} \wedge \vec{R}_{\alpha})^2 . \tag{1}$$

Si on choisit par exemple l'axe Oz parallèle à $\vec{\omega}$, il vient

$$\sum_{\alpha} m N_{\alpha} (\vec{\omega} \wedge \vec{R}_{\alpha})^2 = \omega^2 \sum_{\alpha} m N_{\alpha} (X_{\alpha}^2 + Y_{\alpha}^2) = I\omega^2 , \tag{2}$$

où I est le moment d'inertie par rapport à l'axe de rotation Oz. On a donc

$$E_{\text{cin}} = \frac{1}{2} I\omega^2 . \tag{3}$$

19. Si le référentiel galiléen que l'on considère n'est plus lié au centre de masse O du système, celui-ci est animé d'une vitesse constante $\vec{V}_0$ et la relation (E.20) devient (cf. note 12)

$$\vec{V}_{\alpha} = \vec{V}_0 + \vec{\omega} \wedge (\vec{R}_{\alpha} - \vec{R}_0) . \tag{1}$$

On en déduit aisément, en calculant la somme des énergies cinétiques des sous-systèmes $\mathcal{S}_{\alpha}$, que la formule (E.28) est alors remplacée par

$$E_{\text{int}} = E - \frac{1}{2} M \vec{V}_0^{\,2} - \frac{1}{2} I\omega^2 = E - \frac{\vec{P}^{\,2}}{2M} - \frac{\vec{\mathcal{M}}^{\,2}}{2I} . \tag{2}$$

4. ENTROPIE ET PARAMÈTRES EXTÉRIEURS D'UN SYSTÈME TOTALEMENT ISOLÉ

Comme nous l'avons vu au paragraphe 2.a, les *paramètres extérieurs* qui fixent l'état macroscopique d'un système totalement isolé sont l'énergie E, la quantité de mouvement $\vec{P}$, le moment cinétique $\vec{\mathcal{M}}$ et le nombre de particules N. L'entropie étant invariante par changement de référentiel galiléen (§ 1.c), nous nous plaçons dans le *référentiel galiléen lié au centre de masse du système,* où $\vec{P} = \vec{0}$ et S^* n'est fonction que de E, $\vec{\mathcal{M}}$ et N[20] :

$$S^* = S^*\left(E, \vec{\mathcal{M}}, N\right). \tag{E.29}$$

S^* est la somme des entropies des sous-systèmes $\mathcal{S}_\alpha$ prises pour les valeurs d'équilibre E^0_α, $\vec{P}^{\,0}_\alpha$, N^0_α :

$$S^*\left(E, \vec{\mathcal{M}}, N\right) = \sum_\alpha S^*_\alpha\left(E^0_\alpha - \frac{(\vec{P}_\alpha{}^0)^2}{2mN^0_\alpha}, N^0_\alpha\right). \tag{E.30}$$

A l'équilibre statistique S^* est évidemment indépendant du temps (les $\vec{P}_\alpha{}^0$ varient, mais pas le carré de leur module $(\vec{P}_\alpha{}^0)^2$).

Si l'on impose des accroissements $\mathrm{d}E$, $\mathrm{d}\vec{\mathcal{M}}$ et $\mathrm{d}N$ des paramètres extérieurs, les valeurs d'équilibre E^0_α, $\vec{P}^{\,0}_\alpha$, N^0_α ($\alpha = 1, 2, \ldots$) s'accroissent de $\mathrm{d}E^0_\alpha$, $\mathrm{d}\vec{P}^{\,0}_\alpha$ et $\mathrm{d}N^0_\alpha$, et chacun des S^*_α de

$$\mathrm{d}S^*_\alpha = \frac{1}{T^*}\left[\mathrm{d}E^0_\alpha - \frac{1}{mN^0_\alpha}\vec{P}_\alpha{}^0.\mathrm{d}\vec{P}_\alpha{}^0 - \mu^*\,\mathrm{d}N^0_\alpha\right]; \tag{E.31}$$

la variation $\mathrm{d}S^*$ de l'entropie du système total s'écrit donc

$$\mathrm{d}S^* = \frac{1}{T^*}\left[\sum_\alpha \mathrm{d}E^0_\alpha - \sum_\alpha \vec{V}_\alpha{}^0.\mathrm{d}\vec{P}_\alpha{}^0 - \mu^*\sum_\alpha \mathrm{d}N^0_\alpha\right]. \tag{E.32}$$

Mais les accroissements $\mathrm{d}E^0_\alpha$, $\mathrm{d}\vec{P}^{\,0}_\alpha$ et $\mathrm{d}N^0_\alpha$ doivent être compatibles avec les lois de conservation (E.9) :

$$\sum_\alpha \mathrm{d}N^0_\alpha = \mathrm{d}N, \tag{E.33.a}$$

$$\sum_\alpha \mathrm{d}E^0_\alpha = \mathrm{d}E, \tag{E.33.b}$$

$$\sum_\alpha \mathrm{d}\vec{P}_\alpha{}^0 = \vec{0}, \tag{E.33.c}$$

$$\sum_\alpha \vec{R}_\alpha \wedge \mathrm{d}\vec{P}_\alpha{}^0 = \mathrm{d}\vec{\mathcal{M}}. \tag{E.33.d}$$

On en déduit[21], en remplaçant $\vec{V}_\alpha{}^0$ par sa valeur d'équilibre tirée de (E.20),

$$\mathrm{d}S^* = \frac{1}{T^*}\left[\mathrm{d}E - \vec{\omega}.\mathrm{d}\vec{\mathcal{M}} - \mu^*\,\mathrm{d}N\right]. \tag{E.34}$$

---

20. L'espace étant isotrope, S^* ne dépend en fait que du module de $\vec{\mathcal{M}}$ et pas de sa direction :

$$S^* = S^*(E, \mathcal{M}, N). \tag{1}$$

D'autre part, si on observe le système dans un référentiel galiléen dans lequel son centre de masse est animé d'une vitesse constante $\vec{V}_0$, on a (cf. (E.7))

$$S^* = S^*\left(E - \frac{1}{2}M\vec{V}_0^{\,2}, \mathcal{M}, N\right). \tag{2}$$

21. On passe de (E.32) à (E.34) en utilisant les relations (E.33) et en remarquant que

$$\sum_\alpha \vec{V}_\alpha{}^0.\mathrm{d}\vec{P}_\alpha{}^0 = \sum_\alpha (\vec{\omega} \wedge \vec{R}_\alpha)\cdot\mathrm{d}\vec{P}_\alpha{}^0 = \vec{\omega}.\sum_\alpha \vec{R}_\alpha \wedge \mathrm{d}\vec{P}_\alpha{}^0 = \vec{\omega}.\mathrm{d}\vec{\mathcal{M}}.$$

On retrouve donc les relations usuelles (II, formules (III.2) et (III.29.b)) :

$$\frac{1}{T^*} = \frac{\partial S^*(E, \vec{\mathcal{M}}, N)}{\partial E}, \tag{E.35}$$

$$\frac{\mu^*}{T^*} = -\frac{\partial S^*(E, \vec{\mathcal{M}}, N)}{\partial N}, \tag{E.36}$$

mais on obtient en outre

$$\frac{\omega_i}{T^*} = -\frac{\partial S^*(E, \vec{\mathcal{M}}, N)}{\partial \mathcal{M}_i}, \quad i = x, y, z : \tag{E.37}$$

les composantes du vecteur vitesse angulaire sont, au facteur $(-T^*)$ *près, les dérivées de l'entropie par rapport aux composantes correspondantes du moment cinétique.*

Par ailleurs $\vec{\omega}$ et $\vec{\mathcal{M}}$ vérifient la relation[22] $\vec{\mathcal{M}} = I\vec{\omega}$. Comme nous l'avons déjà remarqué, le moment d'inertie dépend en général de $\vec{\mathcal{M}}$. Cependant, dans le cas d'un *corps solide* et *si la rotation n'est pas trop rapide,* la répartition des masses dans le corps est pratiquement fixe et *I est une constante.* Dans ce cas la relation (E.34) devient

$$\mathrm{d}S^*_{\text{solide}} = \frac{1}{T^*}\left[\mathrm{d}\left(E - \frac{\mathcal{M}^2}{2I}\right) - \mu^* \,\mathrm{d}N\right]. \tag{E.38}$$

$(E - \mathcal{M}^2/2I)$ n'est autre (cf. (E.28)) que l'énergie interne E_{int}, de sorte que

$$\mathrm{d}S^*_{\text{solide}} = \frac{1}{T^*}\left[\mathrm{d}E_{\text{int}} - \mu^* \,\mathrm{d}N\right]; \tag{E.39}$$

l'entropie du solide est fonction seulement de son énergie interne et du nombre de particules qu'il contient :

$$S^*_{\text{solide}} = S^*_{\text{solide}}\left(E - \frac{\mathcal{M}^2}{2I}, N\right) = S^*_{\text{solide}}(E_{\text{int}}, N). \tag{E.40}$$

Si le corps est totalement isolé, son énergie E et son moment cinétique $\vec{\mathcal{M}}$ sont fixés; l'entropie, quant à elle, doit avoir la valeur maximale compatible avec ces valeurs de E et de $\vec{\mathcal{M}}$. L'entropie étant une fonction croissante de l'énergie (ce résultat sera démontré ci-dessous au paragraphe II.1), l'entropie sera maximale, à E et $\vec{\mathcal{M}}$ fixés, lorsque $E - \mathcal{M}^2/2I = E_{\text{int}}$ sera maximale. A l'équilibre statistique, *la rotation uniforme* du solide isolé s'effectue *autour de celui des trois axes principaux d'inertie qui correspond au plus grand moment d'inertie.*

Ce résultat est en réalité *valable pour un système quelconque totalement isolé* (dont les moments d'inertie dépendent en général de $\vec{\mathcal{M}}$). Supposons en effet qu'on ait atteint la situation où le corps tourne autour d'un axe principal d'inertie (§ 2.c). Si le moment d'inertie associé à cet axe n'est pas le plus grand des trois moments principaux, l'entropie du système n'est pas maximale : il suffit de « geler » la répartition des masses et d'amener l'axe principal de plus grand moment d'inertie selon le moment cinétique $\vec{\mathcal{M}}$ pour réaliser un état macroscopique où l'entropie du système est supérieure à ce qu'elle était dans la situation précédente. Dans le véritable état d'équilibre, l'axe

---

22. Que $\vec{\mathcal{M}}$ et $\vec{\omega}$ soient colinéaires est en fait une conséquence de la relation (E.37). En effet S^* ne dépend que du module $\mathcal{M}$ de $\vec{\mathcal{M}}$ (note 20), de sorte que

$$\frac{\omega_i}{T^*} = -\frac{\partial S^*(E, \mathcal{M}, N)}{\partial \mathcal{M}} \cdot \frac{\partial \mathcal{M}}{\partial \mathcal{M}_i} = -\frac{1}{\mathcal{M}} \frac{\partial S^*(E, \mathcal{M}, N)}{\partial \mathcal{M}} \mathcal{M}_i ; \tag{1}$$

le moment d'inertie *I à l'équilibre* est donc relié à l'entropie par

$$\frac{1}{I} = -\frac{T^*}{\mathcal{M}} \frac{\partial S^*(E, \mathcal{M}, N)}{\partial \mathcal{M}}. \tag{2}$$

de rotation coïncide donc avec l'axe principal de plus grand moment d'inertie correspondant à la distribution des masses qui s'est établie. Cette condition complète donc les trois conditions d'équilibre statistique d'un système totalement isolé que nous avons données à la fin du paragraphe I.2.

II. CONSÉQUENCES IMPORTANTES

1. LA TEMPÉRATURE EST UNE GRANDEUR POSITIVE

Les conditions d'équilibre d'un système totalement isolé ont été établies au paragraphe I en écrivant que l'entropie microcanonique partielle est *stationnaire* par rapport aux accroissements dE_α, $d\vec{P}_\alpha$ et dN_α (compatibles avec les relations de conservation); nous n'avons pas précisé qu'il s'agissait d'un maximum. Les conditions obtenues ne sont donc pas suffisantes pour assurer la stabilité de l'équilibre statistique du système. Or le fait que *l'entropie doit être maximum* conduit à une conséquence fort importante : *la température d'un système à l'équilibre statistique ne peut être que positive.*

Il n'est pas très commode d'écrire de façon générale que s^*, donnée par (E.10), doit être maximum pour les valeurs E_α^0, $\vec{P}_\alpha^{\,0}$ et N_α^0 que nous avons trouvées, d'autant plus que celles-ci ne sont pas indépendantes mais liées par les relations de conservation (E.9). Il est plus simple de raisonner directement à partir de la définition de s^*.

Considérons un système immobile à l'équilibre $(\vec{P}=\vec{0}, \vec{\mathcal{M}}=\vec{0})$; d'après les conditions nécessaires trouvées au paragraphe I, tous les sous-systèmes $\mathcal{S}_\alpha$ sont alors immobiles $(\vec{P}_\alpha^{\,0}=\vec{0})$. Si T^* était négative, ces conditions correspondraient à un minimum de s^*. En effet, *si T^* est négative,* les S_α^* sont des fonctions décroissantes des énergies internes $(E_\alpha - \vec{P}_\alpha^{\,2}/2mN_\alpha)$; elles croissent donc quand les $\vec{P}_\alpha^{\,2}$ augmentent à E_α et N_α fixés. Or il est facile de donner des valeurs non nulles aux $\vec{P}_\alpha$ tout en vérifiant les équations de conservation (E.9) avec $\vec{\mathcal{M}}=\vec{0}$. Il suffit que chaque sous-système se déplace sur une droite passant par le centre de masse de $\mathcal{S}$, de telle sorte que les quantités de mouvement vérifient

$$\sum_\alpha \vec{P}_\alpha = \vec{0}\,. \tag{E.41}$$

Autrement dit, pour $T^*<0$, on obtient une entropie microcanonique partielle s^* plus grande (qu'avec les conditions établies au paragraphe I) si le corps explose, ou implose, spontanément. *L'équilibre d'un corps quelconque est donc impossible si la température est négative*[23]; *l'entropie d'un système à l'équilibre est toujours une fonction croissante de son énergie.*

2. ÉTUDE DES SYSTÈMES PARTIELLEMENT ISOLÉS (À ÉNERGIE FIXÉE)

a. Position du problème. Condition d'équilibre

Jusqu'à présent nous avons étudié dans ce complément des systèmes totalement isolés. En revanche, les systèmes qui ont fait l'objet du chapitre II ne sont que *partiellement isolés : leur énergie est fixée,* mais ils sont *solidaires d'un support* avec lequel ils peuvent *échanger de la quantité de mouvement et du moment cinétique.* Généralement, le support est un corps solide dont les dimensions et la masse sont bien supérieures à celles du système considéré.

23. On parle parfois de systèmes à température négative (complément II.B). Il s'agit toujours d'équilibre incomplet concernant seulement un sous-ensemble des degrés de liberté du système, qui ne contribuent pas à l'énergie cinétique.

Pour ces systèmes partiellement isolés, nous avons admis le principe microcanonique : tous les états accessibles sont également probables. Nous avons compté comme accessibles tous les états compatibles avec les conditions imposées au système, c'est-à-dire tous les états dont l'énergie est égale à la valeur E (à δE près) imposée au système, *quelles que soient les valeurs prises dans ces états par la quantité de mouvement* $\vec{P}$ *et le moment cinétique* $\vec{\mathcal{M}}$; autrement dit, nous avons admis comme également probables deux états de même énergie mais de $\vec{P}$ et $\vec{\mathcal{M}}$ différents. C'est cette propriété que nous voulons justifier ici à l'aide des résultats du paragraphe I.

Nous considérons un système $\mathcal{S}$ posé sur un *support* $\mathcal{A}$ *solide et beaucoup plus gros que* $\mathcal{S}$, avec lequel il échange de la quantité de mouvement et du moment cinétique, mais pas d'énergie. *L'ensemble* $\mathcal{A} \cup \mathcal{S}$ *est totalement isolé et immobile.* Les énergies E et $E_{\mathcal{A}}$ du système et du support sont fixées. Par contre leurs quantités de mouvement $\vec{P}$ et $\vec{P}_{\mathcal{A}}$, et leurs moments cinétiques[24] $\vec{\mathcal{M}}$ et $\vec{\mathcal{M}}_{\mathcal{A}}$ sont libres de fluctuer à condition de vérifier

$$\vec{P} + \vec{P}_{\mathcal{A}} = \vec{0}\,, \tag{E.42}$$

$$\vec{\mathcal{M}} + \vec{\mathcal{M}}_{\mathcal{A}} = \vec{0}\,. \tag{E.43}$$

Il est facile de montrer[25] que *l'équilibre se produit pour* $\vec{P}$ *et* $\vec{\mathcal{M}}$ *nuls* (et donc aussi $\vec{P}_{\mathcal{A}}$ et $\vec{\mathcal{M}}_{\mathcal{A}}$), c'est-à-dire que les valeurs les plus probables de $\vec{P}$ et $\vec{\mathcal{M}}$ correspondent à l'immobilité macroscopique. Ce que nous allons chercher, c'est la probabilité d'un état qui s'écarte de ces valeurs.

b. Distribution statistique de la quantité de mouvement et du moment cinétique

Soit (ℓ) un état de $\mathcal{S}$, d'énergie E, de quantité de mouvement $\vec{P}_\ell$ et de moment cinétique $\vec{\mathcal{M}}_\ell$. Soit (a) un état accessible au support $\mathcal{A}$, quand $\mathcal{S}$ est dans l'état (ℓ). Dans cet état l'énergie de $\mathcal{A}$ est $E_{\mathcal{A}}$ (fixée), sa quantité de mouvement

$$\vec{P}_{\mathcal{A}} = -\vec{P}_\ell\,, \tag{E.44}$$

---

24. $\vec{\mathcal{M}}$ et $\vec{\mathcal{M}}_{\mathcal{A}}$ sont ici définis par rapport au même point fixe, par exemple le centre de masse du système global $\mathcal{A} \cup \mathcal{S}$.

25. La technique est toujours la même. L'entropie partielle du système global $\mathcal{A} \cup \mathcal{S}$, pour la quantité de mouvement $\vec{P}$ et le moment cinétique $\vec{\mathcal{M}}$ de $\mathcal{S}$, s'écrit, compte tenu de (E.42) et (E.43),

$$s^*(E, E_{\mathcal{A}}; \vec{P}, \vec{\mathcal{M}}) = S^*\left(E - \frac{\vec{P}^2}{2M}, \vec{\mathcal{M}}\right) + S^*_{\mathcal{A}}\left(E_{\mathcal{A}} - \frac{\vec{P}^2}{2M_{\mathcal{A}}}, -\vec{\mathcal{M}}\right) \tag{1}$$

(cf. formule (2) de la note 20), où M et $M_{\mathcal{A}}$ sont les masses respectives du système $\mathcal{S}$ et du support $\mathcal{A}$. Les valeurs les plus probables des variables internes $\vec{P}$ et $\vec{\mathcal{M}}$ sont celles qui rendent s^* maximum. La valeur d'équilibre $\vec{P}^{(0)}$ de la quantité de mouvement est donnée par

$$\frac{\partial s^*}{\partial P_i} = -P_i\left(\frac{1}{T^* M} + \frac{1}{T^*_{\mathcal{A}} M_{\mathcal{A}}}\right) = 0\,, \quad i = x, y, z; \tag{2}$$

on en déduit

$$\vec{P}^{(0)} = \vec{0}\,. \tag{3}$$

Celle du moment cinétique $\vec{\mathcal{M}}^{(0)}$ s'obtient à l'aide de (E.37) et (E.23) :

$$\frac{\partial s^*}{\partial \mathcal{M}} = -\frac{\omega}{T^*} + \frac{\omega_{\mathcal{A}}}{T^*_{\mathcal{A}}} = -\mathcal{M}\left[\frac{1}{IT^*} + \frac{1}{I_{\mathcal{A}} T^*_{\mathcal{A}}}\right] = 0\,, \tag{4}$$

où I et $I_{\mathcal{A}}$ sont les moments d'inertie respectifs des deux systèmes par rapport à leur axe de rotation éventuel (noter que, du fait de la condition (3), les moments cinétiques sont indépendants des points par rapport auxquels ils sont calculés (cf. note 10)). On a donc

$$\vec{\mathcal{M}}^{(0)} = \vec{0}\,. \tag{5}$$

et son moment cinétique

$$\vec{\mathcal{M}}_{\mathcal{A}} = -\vec{\mathcal{M}}_{\ell} \,. \tag{E.45}$$

Selon le postulat microcanonique, tous les états accessibles du système totalement isolé ($\mathcal{A} \cup \mathcal{S}$) sont également probables. La probabilité d'observer $\mathcal{S}$ *dans un état* (ℓ), *quel que soit l'état du support* $\mathcal{A}$, est donc proportionnelle au nombre $\Omega_{\mathcal{A}}$ d'états de $\mathcal{A}$ compatibles avec l'état (ℓ) de $\mathcal{S}$:

$$P_{\ell}(E, \vec{P}_{\ell}, \vec{\mathcal{M}}_{\ell}) \propto \Omega_{\mathcal{A}}(E_{\mathcal{A}}, -\vec{P}_{\ell}, -\vec{\mathcal{M}}_{\ell}) \,. \tag{E.46}$$

$\Omega_{\mathcal{A}}$ est lié à l'entropie microcanonique du support par (II, formule (I.5))

$$\Omega_{\mathcal{A}}(E_{\mathcal{A}}, -\vec{P}_{\ell}, -\vec{\mathcal{M}}_{\ell}) = \exp \frac{1}{k} S^*_{\mathcal{A}}(E_{\mathcal{A}}, -\vec{P}_{\ell}, -\vec{\mathcal{M}}_{\ell}) \,. \tag{E.47}$$

Le support étant un corps solide[26], son entropie $S^*_{\mathcal{A}}$ ne dépend que de son énergie interne (§ I.4, formule (E.40)). Celle-ci est égale à son énergie $E_{\mathcal{A}}$ diminuée de l'énergie cinétique macroscopique; si on mesure le moment cinétique $\vec{\mathcal{M}}_{\ell}$ par rapport au centre d'inertie du support (qui est pratiquement confondu avec le centre d'inertie du système global $\mathcal{A} \cup \mathcal{S}$), on obtient

$$E^{\text{int}(\ell)}_{\mathcal{A}} = E_{\mathcal{A}} - \frac{\vec{P}_{\ell}^{\,2}}{2M_{\mathcal{A}}} - \sum_{i=x,\,y,\,z} \frac{\mathcal{M}^2_{\ell,\,i}}{2I_{\mathcal{A},\,i}} \,, \tag{E.48}$$

où $M_{\mathcal{A}}$ est la masse du support, les $I_{\mathcal{A},\,i}$ ses trois moments d'inertie principaux et les $\mathcal{M}_{\ell,\,i}$ les projections de $\vec{\mathcal{M}}_{\ell}$ sur les axes correspondants.

Dans la limite où le support est beaucoup plus gros que le système ($M_{\mathcal{A}} \to \infty$, $I_{\mathcal{A},\,i} \to \infty$), la correction apportée à $E_{\mathcal{A}}$ par l'énergie cinétique macroscopique (à $\vec{P}_{\ell}$ et $\vec{\mathcal{M}}_{\ell}$ donnés) est négligeable; $S^*_{\mathcal{A}}$ et donc $\Omega_{\mathcal{A}}$ ne dépendent pratiquement pas[27] de $\vec{P}_{\ell}$ et $\vec{\mathcal{M}}_{\ell}$:

$$\Omega_{\mathcal{A}}(E_{\mathcal{A}}, -\vec{P}_{\ell}, -\vec{\mathcal{M}}_{\ell}) \simeq \Omega_{\mathcal{A}}(E_{\mathcal{A}}) \,. \tag{E.49}$$

La probabilité P_{ℓ} *est donc la même pour tous les états accessibles au système* $\mathcal{S}$ *:*

$$P_{\ell}(E, \vec{P}_{\ell}, \vec{\mathcal{M}}_{\ell}) \equiv P_{\ell}(E) = \frac{1}{\Omega(E)} \,, \tag{E.50}$$

où $\Omega(E)$ est le nombre d'états accessibles au système $\mathcal{S}$, quand son énergie est E,

---

26. Cette hypothèse n'est pas indispensable; les relations (E.37) et (E.23), qui s'appliquent à un corps quelconque, conduisent en effet directement au développement limité (1) de la note 27.

27. Un développement limité donne

$$S^*_{\mathcal{A}}(E^{\text{int}(\ell)}_{\mathcal{A}}) \simeq S^*_{\mathcal{A}}(E_{\mathcal{A}}) - \frac{1}{T^*_{\mathcal{A}}} \left[\frac{\vec{P}_{\ell}^{\,2}}{2M_{\mathcal{A}}} + \sum_i \frac{\mathcal{M}^2_{\ell,\,i}}{2I_{\mathcal{A},\,i}} \right] . \tag{1}$$

En reportant dans (E.47), on obtient

$$\Omega_{\mathcal{A}}(E^{\text{int}(\ell)}_{\mathcal{A}}) \simeq \Omega_{\mathcal{A}}(E_{\mathcal{A}}) \exp \left\{ -\frac{1}{kT^*_{\mathcal{A}}} \left[\frac{\vec{P}_{\ell}^{\,2}}{2M_{\mathcal{A}}} + \sum_i \frac{\mathcal{M}^2_{\ell,\,i}}{2I_{\mathcal{A},\,i}} \right] \right\} . \tag{2}$$

Pour qu'on puisse confondre $\Omega_{\mathcal{A}}(E^{\text{int}(\ell)}_{\mathcal{A}})$ avec $\Omega_{\mathcal{A}}(E_{\mathcal{A}})$, il faut que l'argument de l'exponentielle soit petit devant 1. Il y a évidemment toujours des valeurs de $\vec{P}_{\ell}$ et $\vec{\mathcal{M}}_{\ell}$ pour lesquelles ceci n'est pas vrai; mais si $M_{\mathcal{A}}$ et les $I_{\mathcal{A},\,i}$ sont suffisamment grands, ces valeurs ne sont pas accessibles à $\mathcal{S}$ puisque son énergie E est fixée.

quels que soient sa quantité de mouvement et son moment cinétique (compatibles avec les contraintes extérieures imposées à $\mathcal{S}$).

Ainsi le support joue le rôle de *réservoir de quantité de mouvement et de moment cinétique :* il est tellement gros qu'il peut fournir à $\mathcal{S}$ pratiquement n'importe quelle quantité de mouvement et n'importe quel moment cinétique sans changer de propriétés macroscopiques[28]. Les échanges de quantité de mouvement et de moment cinétique avec ce réservoir n'affectent pas la probabilité des états de $\mathcal{S}$, comme nous l'avons admis implicitement au chapitre II[29].

28. De même qu'un thermostat impose sa température au système avec lequel il est en contact thermique (voir III, § I.A), de même ici le support impose sa vitesse et sa vitesse angulaire, qui sont nulles (cf. exercice III.29).

29. Soulignons que ce résultat ne serait pas valable si la quantité de mouvement et le moment cinétique totaux n'étaient pas nuls. Il est en effet facile de voir qu'une quantité de mouvement totale $\vec{P}$ non nulle (gardons pour simplifier $\vec{\mathcal{M}}_{\mathcal{A}\cup\mathcal{S}} = \vec{0}$) ferait apparaître un facteur proportionnel à $\exp(-\vec{P}_\ell . \vec{P}/kT^* M_{\mathcal{A}})$ dans la probabilité d'un état (ℓ) de $\mathcal{S}$.

Ce résultat est en fait illusoire : la probabilité d'un état ne peut pas dépendre du référentiel dans lequel on observe le système. On aboutit donc à un paradoxe. Il provient du fait que si deux systèmes échangent de la quantité de mouvement sans échanger d'énergie dans un référentiel donné, cet échange de quantité de mouvement s'accompagne nécessairement d'un échange d'énergie dans un deuxième référentiel animé d'une vitesse non nulle par rapport au premier. En effet, soit E l'énergie d'un système et $\vec{P}$ sa quantité de mouvement dans un référentiel $\mathcal{R}$, on a alors

$$E = E_{\text{int}} + \frac{\vec{P}^2}{2M} . \tag{1}$$

Dans un référentiel $\mathcal{R}'$, animé d'une vitesse constante $\vec{V}_0$ par rapport à $\mathcal{R}$, l'énergie devient

$$E' = E_{\text{int}} + \frac{1}{2M}(\vec{P} + M\vec{V}_0)^2 = E + \frac{1}{2} M\vec{V}_0^2 + \vec{P} . \vec{V}_0 . \tag{2}$$

Si donc E est maintenue constante lorsque $\vec{P}$ varie, il ne peut en être de même de E'.

La question est donc de savoir dans quel référentiel une paroi dite adiabatique interdit les échanges d'énergie tout en permettant les échanges de quantité de mouvement et de moment cinétique; le seul choix raisonnable est le référentiel dans lequel le système global est immobile.

COMPLÉMENT II.F

Distribution et entropie microcanoniques en mécanique classique

I. Densité de probabilité microcanonique dans l'espace des phases

II. Expression classique de l'entropie microcanonique
1. Méthode de calcul
2. Détermination du volume élémentaire dans l'espace des phases
3. Conclusion : expression générale de l'entropie microcanonique en mécanique classique

L'état microscopique d'un système macroscopique dont les propriétés peuvent être décrites par la mécanique classique (cf. appendice III et complément I.A, § I.4) est défini par la donnée de n coordonnées généralisées $q_i (i = 1, 2, \ldots, n)$ et de leurs moments conjugués p_i, que nous noterons simplement $\{q, p\}$. Dans cette notation simplifiée, la fonction de Hamilton, associée à l'énergie totale du système, s'écrit $\mathcal{H}(q, p)$, et l'élément de volume de l'espace des phases $dq.dp$. La mécanique statistique caractérise l'état macroscopique d'un tel système par une densité de probabilité $w(q, p)$ dans l'espace des phases.

I. DENSITÉ DE PROBABILITÉ MICROCANONIQUE DANS L'ESPACE DES PHASES

Lorsque le système est *isolé et à l'équilibre,* la densité de probabilité caractérisant son état macroscopique est la *distribution microcanonique* $w^*(q, p)$, qui découle du postulat fondamental (II, § I.B) de la même manière que dans le cas quantique.

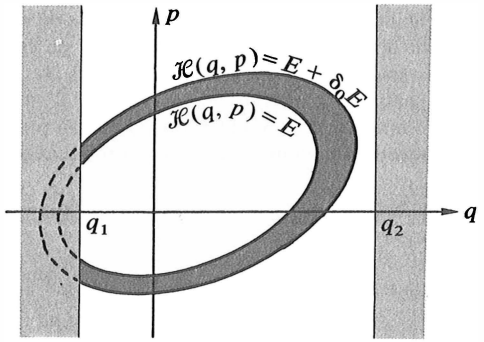


FIGURE 1

Représentation schématique, à deux dimensions, de l'espace des phases et du domaine accessible $\mathcal{D}^*(E, V, \ldots)$ (domaine grisé) : fixer le volume V du système revient ici à borner les valeurs possibles pour q (entre q_1 et q_2); en outre, si l'énergie vaut E à $\delta_0 E$ près, seuls sont accessibles les états dont le point représentatif est situé entre les deux hypersurfaces (ici des courbes) d'équations $\mathcal{H}(q, p) = E$ et $\mathcal{H}(q, p) = E + \delta_0 E$.

Les contraintes extérieures qui s'exercent sur le système fixent certains paramètres macroscopiques comme le volume V, le nombre de particules N, ... et l'énergie E connue avec une incertitude $\delta_0 E$. Le nombre de particules N est pris en compte dans le choix de l'espace des phases : sa dimension est $2n = 6N$ si les particules n'ont pas de structure interne (chaque particule étant repérée par 3 coordonnées, il en faut au total $n = 3N$). Les autres conditions délimitent le *domaine accessible* $\mathcal{D}^*(E, V...)$ *de l'espace des phases :* c'est l'ensemble des points (q, p) de cet espace (fig. 1) tels que

(i) l'état correspondant du système est compatible avec les valeurs fixées pour le volume V et les autres paramètres extérieurs;

(ii) la valeur prise par la fonction de Hamilton est égale à E à $\delta_0 E$ près :

$$E < \mathcal{H}(q, p) < E + \delta_0 E \,. \tag{F.1}$$

Le postulat fondamental affirme alors que *tous les points du domaine accessible de l'espace des phases sont également probables.* La distribution microcanonique $w^*(q, p)$ s'écrit donc

$$w^*(q, p) = \frac{1}{A} \quad \text{si} \quad (q, p) \in \mathcal{D}^*(E, V, \ldots)$$
$$= 0 \quad \text{sinon.} \tag{F.2}$$

La constante A est déterminée par normalisation :

$$\int \mathrm{d}q \; \mathrm{d}p \; w^*(q, p) = 1 \; ; \tag{F.3}$$

elle est égale au volume du domaine accessible de l'espace des phases :

$$A = \int_{\mathcal{D}^*(E, V, \ldots)} \mathrm{d}q \; \mathrm{d}p \,. \tag{F.4}$$

On peut mettre la distribution microcanonique $w^*(q, p)$ sous une forme équivalente, incorporant de façon compacte les contraintes extérieures. Définissons la fonction $\theta(q, p; V, \ldots)$, égale à 1 si (q, p) est compatible avec ces contraintes et à 0 dans le cas contraire. Comme les valeurs possibles pour l'énergie ne sont pas ici quantifiées, on peut le plus souvent « oublier » l'existence d'une incertitude $\delta_0 E$ et utiliser la « fonction » $\delta(x)$ de Dirac[1] pour exprimer le fait que l'énergie est fixée. On écrira alors

$$w^*(q, p) = \frac{1}{B} \, \theta(q, p; V, \ldots) \, \delta\big[E - \mathcal{H}(q, p)\big] \,. \tag{F.5}$$

La constante B, proportionnelle à A :

$$B = \frac{A}{\delta_0 E} \,, \tag{F.6}$$

est donnée par la condition de normalisation :

$$B = \int \mathrm{d}q \; \mathrm{d}p \, \theta(q, p; V, \ldots) \, \delta\big[E - \mathcal{H}(q, p)\big] \,. \tag{F.7}$$

Ces expressions sont plus commodes à manipuler formellement, mais il faut revenir à (F.2) chaque fois que le comportement très singulier de la « fonction » δ pourrait créer des difficultés.

1. La « fonction » $\delta(x)$ peut en effet être considérée comme la limite, quand ε tend vers zéro par valeurs positives, d'une fonction constante sur l'intervalle $[-\varepsilon, +\varepsilon]$, où elle vaut $1/2\varepsilon$ pour que son intégrale reste égale à 1, et nulle en dehors de cet intervalle. Ici, 2ε s'identifie à $\delta_0 E$; d'où la relation (F.6).

II. EXPRESSION CLASSIQUE DE L'ENTROPIE MICROCANONIQUE

L'entropie microcanonique $S^*(E, V, \ldots)$ est l'entropie statistique associée à la distribution de probabilité $w^*(q, p)$. Comme il s'agit d'une distribution continue, la définition de l'entropie statistique correspondante pose problème. Pour éviter les divergences, nous allons discrétiser la distribution (c'est-à-dire que nous utilisons la seconde méthode décrite au paragraphe III.2.b du complément I.G).

1. MÉTHODE DE CALCUL

Il est naturel de choisir les variables canoniques classiques $\{q, p\}$ comme « variables pertinentes » pour effectuer la discrétisation (complément I.G, § III.3)[2]. Nous allons donc *découper l'espace des phases en cellules élémentaires,* dont les dimensions seront suffisamment réduites pour qu'aucune grandeur ne varie de façon appréciable à l'intérieur d'une telle cellule.

Le découpage lui-même s'impose de façon naturelle si l'on se souvient que les variables conjuguées q_i et p_i doivent vérifier en mécanique quantique une « relation d'incertitude » de Heisenberg (C.D.L., p. 286) :

$$\Delta q_i \,.\, \Delta p_i \geqslant \frac{\hbar}{2} \,. \tag{F.8}$$

Si donc on veut rester dans le cadre de la mécanique classique, il faut se cantonner à des situations où les incertitudes avec lesquelles sont connues les valeurs d'une coordonnée q_i et de son moment conjugué p_i ont un produit très supérieur à $\hbar$, de façon que la borne (F.8) ne fasse pas sentir ses effets. Autrement dit, pour chaque couple de variables conjuguées q_i et p_i, le découpage du plan (q_i, p_i) doit être tel que l'aire $\delta q_i \,.\, \delta p_i$ d'un « pavé » élémentaire ne soit pas inférieure à $\hbar$. Notons provisoirement τ^n le volume de la cellule élémentaire dans l'espace des phases global à $2n$ dimensions; nous préciserons plus loin (§ 2) la valeur de τ, mais nous savons d'ores et déjà qu'elle est de l'ordre de $\hbar$.

A chacune des cellules élémentaires est associée une probabilité, égale à l'intégrale de $w^*(q, p)$ dans cette cellule. D'après le postulat microcanonique, qu'exprime ici la forme (F.2) de w^*, cette probabilité est nulle pour les cellules situées en dehors du domaine accessible $\mathcal{D}^*$, et elle est la même pour toutes les cellules situées à l'intérieur de ce domaine. L'extension du domaine accessible est déterminée par l'incertitude $\delta_0 E$ sur l'énergie du système; cette incertitude porte sur une valeur E macroscopique, alors que le volume d'une cellule est d'ordre microscopique ($\sim \hbar^n$). Le domaine accessible de l'espace des phases contient donc un très grand nombre de cellules élémentaires[3], de sorte qu'il n'y a pas à se préoccuper de celles qui se trouvent « à cheval » sur la frontière du domaine.

---

2. Ce choix est rendu possible par le *théorème de Liouville* (appendice III, § II.3.b) : bien que chacun de ses points se déplace au cours du temps, le domaine accessible de l'espace des phases reste constamment le même, et l'on peut indifféremment considérer que son « pavage » en cellules élémentaires reste fixe, ou « suit » les points dans leur mouvement (puisque ce mouvement conserve les volumes dans l'espace des phases).

3. Le raisonnement classique que nous développons ici doit donner les mêmes résultats que le traitement général, fondé sur la description quantique du système au niveau microscopique. Concrètement, le nombre de cellules élémentaires contenues dans le domaine accessible de l'espace des phases va s'identifier au nombre d'états microscopiques accessibles du traitement quantique général (voir formule (F.9)).

Soit $\Omega_c(E, V, \ldots)$ le *nombre de cellules élémentaires contenues dans le domaine accessible* $\mathcal{D}^*(E, V, \ldots)$ *de l'espace des phases*[4]. Ces cellules ayant toutes même probabilité, celle-ci vaut $1/\Omega_c(E, V, \ldots)$, de sorte que l'entropie statistique associée s'écrit

$$S^*(E, V, \ldots) = k \ln \Omega_c(E, V, \ldots) . \tag{F.9}$$

Pour calculer le nombre $\Omega_c(E, V, \ldots)$, il suffit de diviser le volume du domaine accessible $\mathcal{D}^*$ par le volume τ^n de chaque cellule élémentaire :

$$\Omega_c(E, V, \ldots) = \frac{1}{\tau^n} \int_{\mathcal{D}^*(E, V, \ldots)} \mathrm{d}q \, \mathrm{d}p . \tag{F.10}$$

Comme au paragraphe I (cf. note 1), on peut écrire cette expression sous une forme plus commode en faisant intervenir la « fonction » de Dirac[5] :

$$\Omega_c(E, V, \ldots) = \frac{1}{\tau^n} \delta_0 E \int \theta(q, p; V, \ldots) \delta\left[E - \mathcal{H}(q, p)\right] \mathrm{d}q \, \mathrm{d}p . \tag{F.11}$$

(*Attention :* l'argument de la « fonction » δ n'est pas directement l'une des variables d'intégration; c'est une fonction de ces $2n$ variables q_i, p_i; voir l'exemple du gaz parfait ci-dessous.)

2. DÉTERMINATION DU VOLUME ÉLÉMENTAIRE DANS L'ESPACE DES PHASES

Reste à préciser le volume τ^n d'une cellule élémentaire. Pour déterminer le facteur de proportionnalité entre τ et $\hbar$, nous allons calculer l'entropie microcanonique d'un gaz parfait classique par les formules (F.9)-(F.11) et l'identifier avec l'expression trouvée à partir de la description quantique : en effet, les gaz ordinaires sont des systèmes que la mécanique classique décrit correctement à l'échelle microscopique (complément I.A, § III.2); dans l'approximation des gaz parfaits, la description classique et la description quantique sont suffisamment simples pour qu'on puisse dans les deux cas mener les calculs jusqu'au bout. Nous admettons que la constante τ est universelle, c'est-à-dire qu'elle est la même pour tous les systèmes physiques que l'on peut décrire par la mécanique classique.

Considérons donc un gaz parfait, c'est-à-dire un ensemble de N particules ponctuelles de masse m, sans interactions et sans structure interne (gaz parfait monoatomique). Ces particules sont repérées par N coordonnées vectorielles $\vec{r}_j$ ($j = 1, 2, \ldots, N$), dont les moments conjugués sont les impulsions $\vec{p}_j$. La fonction de Hamilton du gaz parfait s'écrit simplement

$$\mathcal{H}(q, p) = \sum_{j=1}^{N} \frac{\vec{p}_j^{\,2}}{2m} . \tag{F.12}$$

---

4. Comme dans le cas quantique le nombre d'états accessibles Ω, ce nombre Ω_c dépend aussi de l'incertitude $\delta_0 E$ sur l'énergie, mais cette dépendance n'est pas essentielle (cf. II, § I.C.1).

5. Il est facile de s'assurer que la formule (F.11) donne bien un nombre Ω_c sans dimension. En effet, la « fonction » $\delta(x)$ a les dimensions de l'inverse de son argument x, puisque

$$\int \delta(x) \mathrm{d}x = 1 . \tag{1}$$

Donc le produit de l'incertitude $\delta_0 E$ par la « fonction » δ portant sur l'énergie E est sans dimension. La fonction θ l'est également. Enfin, si l'on explicite les variables d'intégration :

$$\mathrm{d}q \, \mathrm{d}p = \mathrm{d}q_1 \mathrm{d}p_1 \mathrm{d}q_2 \mathrm{d}p_2 \ldots \mathrm{d}q_n \mathrm{d}p_n , \tag{2}$$

on constate qu'à chacun des produits $\mathrm{d}q_i \mathrm{d}p_i$, qui a les dimensions d'une action, est associé au dénominateur un facteur τ ayant les mêmes dimensions (celles de $\hbar$).

Nous devons calculer

$$\Omega_c(E, V, N)=\frac{1}{\tau^{3N}}.\delta_0 E \int\delta\left[E-\sum_{j=1}^{N}\frac{\vec{p}_j^{\,2}}{2m}\right]\prod_{j=1}^{N}\theta(\vec{r}_j; V)\,d^3r_j.d^3p_j\,. \tag{F.13}$$

La fonction θ exprime simplement que l'extrémité de chacun des N vecteurs de position $\vec{r}_j$ doit être située dans le récipient de volume V. Comme la fonction à intégrer est, à part cela, indépendante des variables $\vec{r}_j$, l'intégration sur chacune de celles-ci donne immédiatement le volume V. Reste une intégrale à $3N$ dimensions sur les impulsions, dans laquelle la fonction à intégrer dépend seulement de la somme des carrés des $3N$ variables. On se ramène (voir complément I.E, § I) à une intégrale portant sur une seule variable, la distance P à l'origine dans l'espace à $3N$ dimensions des impulsions :

$$\Omega_c(E, V, N)=\frac{1}{\tau^{3N}}.\delta_0 E.V^N\int_0^{+\infty}\delta\left(E-\frac{P^2}{2m}\right).C_{3N-1}P^{3N-1}\,dP\,, \tag{F.14}$$

où $C_{3N-1}P^{3N-1}$ est l'« aire » de l'hypersphère à $(3N-1)$ dimensions dans l'espace à $3N$ dimensions (complément I.E, formules (E.4) et (E.12)) :

$$C_{3N-1}=3N\frac{\pi^{3N/2}}{\Gamma\left(\frac{3N}{2}+1\right)}\,. \tag{F.15}$$

On trouve finalement[6]

$$\Omega_c(E, V, N)=\frac{1}{\tau^{3N}}.\delta_0 E.V^N.C_{3N-1}.m(2mE)^{(3N-2)/2}\,. \tag{F.16}$$

Il suffit de comparer ce résultat avec l'expression de la densité d'états calculée au chapitre I (§ III.B.2) pour constater que

$$\tau=2\pi\hbar=h\,. \tag{F.17}$$

Le facteur de proportionnalité entre τ et $\hbar$ est donc simplement égal à 2π. Il est curieux de constater que c'est la constante de Planck h qui s'introduit naturellement en mécanique statistique, plutôt que $\hbar$ qui est partout ailleurs la plus commode à utiliser.

3. CONCLUSION : EXPRESSION GÉNÉRALE DE L'ENTROPIE MICROCANONIQUE EN MÉCANIQUE CLASSIQUE

En définitive, lorsque le comportement microscopique du système étudié est correctement décrit par la mécanique classique, son entropie microcanonique S^* est donnée par

$$S^*(E, V, \ldots)=k\,\ln\,\Omega_c(E, V, \ldots)\,, \tag{F.18}$$

---

6. Si $f(x)$ est une fonction possédant un seul zéro en $x=x_0$, on démontre facilement (C.D.L, appendice II, § 1.c) que

$$\delta[f(x)]=\frac{1}{|f'(x_0)|}\,\delta(x-x_0)\,, \tag{1}$$

où $f'(x)$ est la dérivée de $f(x)$. Ici,

$$\delta\left[E-\frac{P^2}{2m}\right]=\frac{m}{P}\,\delta\left[P-\sqrt{2mE}\right]. \tag{2}$$

L'intégrale se calcule ensuite directement :

$$\int_0^{+\infty}\delta\left[E-\frac{P^2}{2m}\right]P^{3N-1}\,dP=m\,(2mE)^{(3N-2)/2}\,. \tag{3}$$

Le résultat (F.16) en découle aussitôt.

où le nombre Ω_c de cellules accessibles a pour expression

$$\Omega_c(E, V, \ldots)=\frac{1}{(2\pi\hbar)^n}.\delta_0 E \int \theta(q, p; V, \ldots)\,\delta\left[E-\mathcal{H}(q, p)\right] \mathrm{d}q\ \mathrm{d}p \quad . \qquad \text{(F.19)}$$

Dans le cas d'un système de N *particules identiques et indiscernables* (par exemple un gaz), on est amené à *ne pas considérer comme distincts deux états du système qui ne diffèrent que par une permutation des N particules.* Ce principe, comme d'ailleurs la notion d'indiscernabilité elle-même, provient de la mécanique quantique (appendice I, § IV). En mécanique classique, pour un système de 2 particules par exemple, l'état dans lequel la particule numéro 1 se trouve au point $\vec{r}$ avec l'impulsion $\vec{p}$ et la particule numéro 2 en $\vec{r}'$ avec $\vec{p}\,'$ est normalement distinct de celui où c'est la particule (2) qui a $\vec{r}$ et $\vec{p}$ et la particule (1) $\vec{r}'$ et $\vec{p}\,'$ (il suffit de penser au jeu de billard : une situation où la position des trois boules est donnée peut être favorable ou défavorable à un joueur suivant que c'est l'une ou l'autre des deux boules blanches qui lui est affectée). Pourtant, pour des particules *indiscernables,* ces deux configurations doivent, comme en mécanique quantique, être associées à un seul et même état physique (voir complément II.C et chapitre III, § V.A.2.b). Dans ce cas, lorsqu'on intègre sur tout l'espace des phases, c'est-à-dire, pour 2 particules, sur toutes les valeurs permises de $\vec{r}_1$ et $\vec{p}_1$, puis sur toutes celles de $\vec{r}_2$ et $\vec{p}_2$, on compte 2 fois l'état du système où l'une des particules se trouve au point $\vec{r}$ avec l'impulsion $\vec{p}$ et l'autre en $\vec{r}'$ avec $\vec{p}\,'$. Dans le cas général, pour un système de N particules identiques et indiscernables, l'intégrale sur l'espace des phases doit être divisée par un *«facteur d'indiscernabilité»,* égal au nombre de permutations $N!$ de ces particules[7]. On a donc

$$\Omega_c(E, V, N, \ldots)=\frac{1}{(2\pi\hbar)^{3N}}\frac{\delta_0 E}{N!}\int \mathrm{d}^3r_1\,\mathrm{d}^3r_2 \ldots \mathrm{d}^3r_N\,\mathrm{d}^3p_1\,\mathrm{d}^3p_2 \ldots \mathrm{d}^3p_N$$
$$\times\,\theta(\vec{r}, \vec{p}\,; V, \ldots)\,\delta\left[E-\mathcal{H}(\vec{r}_1, \vec{r}_2, \ldots, \vec{r}_N; \vec{p}_1, \vec{p}_2, \ldots, \vec{p}_N)\right] \quad . \qquad \text{(F.20)}$$

7. Les configurations dans lesquelles deux ou plusieurs particules se trouveraient exactement au même point $\vec{r}$ et posséderaient exactement la même impulsion $\vec{p}$ sont infiniment peu probables : dans l'espace des phases, l'ensemble des points correspondants constitue un sous-espace de dimension inférieure, dont la contribution à l'intégrale (F.20) est par conséquent négligeable (on dit en mathématiques que ce sous-espace est «de mesure nulle»). Ainsi, *dans le cadre de la mécanique classique,* si l'on admet la nécessité de prendre en compte l'indiscernabilité des particules, *la division par N! donne une formule exacte,* et non pas approchée comme dans le cas quantique.

Exercices II

ENTROPIE ET TEMPÉRATURE MICROCANONIQUES

II.1. « Ortho » et « para »-hydrogène. Dans son état électronique fondamental, la molécule d'hydrogène H_2 peut exister sous deux formes : l'« ortho-hydrogène », où les spins des deux noyaux sont parallèles, et le « para-hydrogène », où ils sont antiparallèles; la forme « para » possède donc un seul état de spin, dont on prendra l'énergie comme origine; la forme « ortho » présente trois états distincts, de même énergie $\varepsilon > 0$.

On considère un échantillon d'hydrogène solide, constitué de N molécules fixes et discernables, faiblement couplées. On ne s'intéresse qu'aux états de spin « ortho » et « para ». Le système est isolé, son énergie valant $E (E \gg \varepsilon)$.

1) Calculer l'entropie S^* et la température microcanonique T^* du système.

2) Exprimer les nombres de molécules de para-hydrogène et d'ortho-hydrogène en fonction de T^*.

II.2. Défauts de Frenkel. Dans un cristal parfait, les N atomes qui le constituent sont régulièrement disposés sur les « sites » du réseau cristallin. Lorsqu'un atome se place dans un interstice de ce réseau, en laissant un site vacant, on dit qu'il y a un « défaut de Frenkel » dans le cristal. L'énergie nécessaire pour créer un défaut de Frenkel est $\varepsilon > 0$, et on appelle N' le nombre de positions interstitielles (N et N' sont du même ordre de grandeur).

On considère un cristal présentant n défauts de Frenkel (on suppose $n \ll N$). Calculer le nombre d'états accessibles au système et en déduire sa température microcanonique T^*. Montrer que la relation entre T^* et n peut s'écrire

$$n \simeq \sqrt{NN'}\, e^{-\varepsilon/2kT^*} .$$

II.3. Défauts de Schottky. On dit qu'un cristal de N atomes présente n « défauts de Schottky » lorsque n de ses sites sont vacants; le cristal occupe donc un volume légèrement plus grand que s'il était parfait. Soit $\varepsilon (> 0)$ l'énergie nécessaire pour créer un défaut de Schottky (à N constant).

On considère un cristal présentant n défauts de Schottky ($n \ll N$). Calculer le nombre d'états accessibles au système et montrer que l'on peut écrire la relation entre n et la température microcanonique T^* sous la forme

$$n \simeq N e^{-\varepsilon/kT^*} .$$

N.B. Pour établir ce résultat, on considérera que le cristal comporte $N + n$ sites dont n sont inoccupés. Ce faisant on inclut des situations où certaines des n vacances se produisent en surface, auquel cas elles ne constituent évidemment pas des défauts

de Schottky. Mais, pour un cristal de taille macroscopique, le nombre de sites superficiels est très petit devant le nombre total de sites, de sorte que la probabilité pour que l'une des n vacances se trouve sur un site superficiel est négligeable (évaluer cette probabilité pour un cristal cubique d'arête ~ 1 mm).

II.4. On considère un système isolé, constitué de $N (N \gg 1)$ oscillateurs harmoniques indépendants; les niveaux d'énergie de chaque oscillateur sont non-dégénérés et de la forme

$$\varepsilon_n = \left(n + \frac{1}{2}\right)\hbar\omega ,$$

où n est un nombre entier positif ou nul.

1) Calculer le nombre d'états accessibles au système lorsque son énergie vaut E (utiliser le résultat de l'exercice I.2). En déduire sa température microcanonique.

2) L'énergie du système étant toujours fixée à E, calculer la probabilité pour qu'un oscillateur déterminé ait l'énergie ε $\left(\varepsilon = \left(n_0 + \frac{1}{2}\right)\hbar\omega < E\right)$. Exprimer cette probabilité en fonction de la température.

II.5. On considère un système constitué par un oscillateur harmonique classique à une dimension, d'énergie fixée, comprise entre E et $E + \delta E$.

1) Dessiner dans l'espace des phases la surface où se situent les états accessibles à ce système.

2) Calculer la probabilité pour que l'oscillateur se trouve en un point d'abscisse comprise entre x et $x + \mathrm{d}x$. Comparer avec le résultat de l'exercice I.13.

II.6. On considère un système constitué de N oscillateurs harmoniques à une dimension indépendants, dont les propriétés peuvent être décrites par la mécanique classique. Chaque oscillateur est caractérisé par une masse m_i et une pulsation $\omega_i (i = 1, 2, \ldots, N)$. L'état de l'oscillateur (i) est défini par la donnée de sa coordonnée de position q_i et de son moment conjugué p_i, de sorte que la fonction de Hamilton du système s'écrit

$$\mathcal{H} = \sum_{i=1}^{N} \left[\frac{p_i^2}{2m_i} + \frac{1}{2} m_i \omega_i^2 q_i^2\right] .$$

1) Soit $\phi(E)$ le volume occupé dans l'espace des phases (à $2N$ dimensions) par les états du système dont l'énergie totale est inférieure à E. Donner l'expression intégrale de $\phi(E)$ et montrer que, par un changement de variables, le calcul de $\phi(E)$ se ramène à celui du volume de la sphère de rayon $\sqrt{E}$ dans un espace à $2N$ dimensions (complément I.E, formules (E.3) et (E.13)). En déduire $\phi(E)$.

2) Le système étant isolé, avec l'énergie E (à δE près), calculer son entropie et sa température microcanoniques.

3) Comparer avec ce qu'on obtient pour des oscillateurs quantiques (exercice II.4). Pour quelles valeurs de l'énergie, ou quel domaine de température, les résultats classique et quantique sont-ils équivalents?

II.7. On considère un système $\mathcal{S}$ constitué de N_0 particules *indiscernables*, indépendantes, enfermées dans un récipient de volume V. Le système est isolé avec l'énergie E_0. Soit $\{(\lambda)\}$ l'ensemble des états à une particule du système et (λ_0) l'un quelconque de ces états, d'énergie ε_0. Soit $\mathcal{R}$ le sous-système de $\mathcal{S}$ constitué par les particules qui se trouvent dans les états autres que (λ_0); on appelle $\Omega_{\mathcal{R}}(E, N, V)$ le nombre d'états accessibles au système $\mathcal{R}$ quand il comporte N particules d'énergie totale E, dans le volume V.

1) Rappeler les expressions de la température T^* et du potentiel chimique μ^* microcanoniques du système $\mathcal{R}$ en fonction de $\Omega_{\mathcal{R}}$.

2) Les N_0 particules sont des fermions, de sorte que chaque état individuel ne peut accueillir que zéro ou une particule (principe de Pauli).

a) Montrer que le nombre total d'états accessibles au système $\mathcal{S}$ s'écrit dans ce cas

$$\Omega_{\mathcal{S}}(E_0, N_0, V) = \Omega_{\mathcal{R}}(E_0, N_0, V) + \Omega_{\mathcal{R}}(E_0 - \varepsilon_0, N_0 - 1, V).$$

b) En déduire les probabilités P_0 et P_1 pour que l'état (λ_0) soit occupé respectivement par zéro et une particule. En supposant $\varepsilon_0 \ll E_0$ et $N_0 \gg 1$, calculer $\ln P_0/P_1$ puis (à l'aide de la condition de normalisation des probabilités) P_0 et P_1, en fonction de ε_0, μ^* et T^*.

c) En déduire le nombre moyen $\bar{n}_0^F$ de particules dans l'état (λ_0) en fonction des mêmes variables.

3) On considère maintenant le cas où les particules sont des bosons; le nombre de particules qui peuvent occuper un état individuel donné n'est alors pas limité. Adapter le raisonnement de la question 2 à ce cas pour calculer le nombre moyen $\bar{n}_0^B$ de particules dans l'état (λ_0) en fonction de ε_0, T^* et μ^*.

ORDRES DE GRANDEUR

II.8. Un système est constitué de 10^{21} atomes d'hydrogène, indépendants, contenus dans une boîte d'un litre de volume. L'énergie de ce système est fixée, égale à 10 joules à 1 μJ près. Calculer le nombre d'états accessibles au système.

Comparer numériquement le logarithme de ce nombre d'états accessibles et celui du nombre d'états dont l'énergie est inférieure à 10 J (voir note 10 du chapitre II).

II.9. On considère 1 litre d'hélium dans les conditions normales de température et de pression. C'est à une bonne approximation un gaz parfait monoatomique.

1) Évaluer numériquement son entropie (complément II.C, § II.2).

2) Évaluer numériquement l'accroissement de cette entropie dans une détente de Joule (II, § II.B.1) qui augmente de 10 % le volume du gaz.

II.10. La température microcanonique d'un système macroscopique isolé est égale à 300 K. Ce système absorbe un photon de lumière bleue, ce qui augmente son énergie de ΔE.

1) En supposant les accroissements assez petits pour que l'on puisse écrire

$$\frac{\Delta\Omega}{\Delta E} \simeq \frac{d\Omega}{dE},$$

évaluer l'augmentation relative $\Delta\Omega/\Omega$ du nombre Ω d'états accessibles au système. L'approximation suggérée est-elle valable?

2) On suppose que le système est constitué par une mole de gaz parfait monoatomique (hélium par exemple) à la pression atmosphérique. Calculer dans ce cas $\Delta\Omega/\Omega$; vérifier que la température du système ne varie cependant pas lorsqu'il absorbe le photon.

VARIABLES INTERNES ET ÉQUILIBRE D'UN SYSTÈME ISOLÉ

II.11. L'entropie d'un gaz parfait monoatomique constitué de N particules (de masse m), enfermé dans un volume V, et d'énergie totale E, s'écrit (complément II.C, § II.2.a)

$$S^*(E, V, N) = Nk\left[\ln\frac{V}{N} + \frac{3}{2}\ln\left(\frac{m}{3\pi\hbar^2}\frac{E}{N}\right) + \frac{5}{2}\right].$$

1) Calculer la température T^*, la pression p^* et le potentiel chimique μ^* de ce gaz parfait.

2) On considère deux échantillons différents d'un même gaz parfait. Les récipients qui les contiennent ont une paroi commune, comme sur la figure 3 du chapitre II.

a) On suppose d'abord que les deux gaz peuvent échanger seulement de la chaleur. Calculer les valeurs les plus probables E_{1m} et E_{2m} de leurs énergies en fonction de l'énergie totale E du système global et des nombres de particules N_1 et N_2 des deux gaz.

b) La cloison séparant les deux échantillons est maintenant diatherme et mobile, permettant des échanges d'énergie et de volume. Calculer, en fonction de l'énergie totale E et du volume total V, et des nombres de particules N_1 et N_2, les énergies les plus probables et les volumes les plus probables des deux échantillons de gaz.

c) Que se passe-t-il si la cloison devient en outre perméable aux particules? Analyser physiquement la situation et montrer, en utilisant les résultats de l'exercice II.18, qu'elle n'est pas particulière aux gaz parfaits.

II.12. On considère un système de N particules ($N \gg 1$) de spin 1 plongées dans un champ magnétique uniforme $\vec{B}$. Chacune d'elles a 3 états possibles, correspondant aux 3 valeurs possibles ($m_s = +1, 0, -1$) de la projection du spin sur l'axe défini par $\vec{B}$. On notera n_+, n_0 et n_- le nombre de particules du système qui se trouvent dans chacun de ces 3 états individuels, dont l'énergie vaut $\varepsilon_+ = -\mu B$, $\varepsilon_0 = 0$, $\varepsilon_- = +\mu B$.

1) Montrer que l'énergie totale du système est bornée inférieurement et supérieurement. Si l'on choisit une valeur E de cet intervalle, N et B étant déjà fixés, ceci ne suffit pas à déterminer les 3 nombres n_+, n_0 et n_-. Nous considérerons n_0 comme une variable interne.

2) Si N, B et E sont fixés, sur quel intervalle peut fluctuer n_0? Calculer le nombre d'états accessibles correspondant à une valeur donnée de n_0 et en déduire l'entropie microcanonique partielle $s^*(E, N, B; n_0)$.

3) Déterminer la valeur la plus probable n_{0m} de n_0. Que vaut-elle lorsque E prend sa valeur maximale ou sa valeur minimale, et lorsque $E = 0$?

4) On suppose N suffisamment grand pour que l'entropie du système soit pratiquement égale à la valeur de s^* pour $n_0 = n_{0m}$ (cf. exercice II.19). Calculer la température microcanonique du système et discuter son signe (cf. complément II.B).

II.13. A) On considère deux systèmes séparément isolés et à l'équilibre : le système (G), d'énergie E_G, est constitué de N_G atomes libres enfermés dans une boîte de volume V; l'autre, (S), d'énergie E_S, est formé de N_S spins 1/2 indépendants (portés par des particules immobiles) placés dans un champ magnétique $\vec{B}$. On note μ le moment magnétique associé à chaque spin.

1) Calculer la température microcanonique de chacun de ces systèmes.

2) On établit le contact thermique entre ces deux systèmes, et l'on attend que l'équilibre soit atteint.

a) Déterminer graphiquement la température finale T_f en fonction des paramètres caractérisant l'état initial des deux systèmes; cette température peut-elle être négative (complément II.B)?

b) Calculer l'énergie moyenne de chacun des deux systèmes.

c) Calculer la quantité d'énergie ΔE échangée par les deux systèmes lors du contact et étudier le sens de ce transfert d'énergie en fonction des températures initiales (on étudiera en particulier le cas où la température initiale de (S) est négative).

B) Mêmes questions pour le contact thermique entre deux systèmes de spins 1/2 placés dans un champ magnétique $\vec{B}$: (S), constitué de N_S spins de moment μ, et (S'), constitué de N'_S spins de moment μ'. Étudier les différents cas : les températures initiales sont toutes deux positives, ou toutes deux négatives, ou l'une est positive et l'autre négative.

Quelle est la température finale lorsqu'on met en contact deux systèmes de spins identiques ($\mu' = \mu$, $N'_S = N_S$), l'un ayant initialement une température de 300 K et l'autre de -300 K?

II.14. Un gaz parfait constitué de N ($N \gg 1$) atomes de spin 1/2 est enfermé dans un récipient de volume V. Le gaz est soumis à un champ magnétique uniforme $\vec{B}$ et son énergie est fixée à une valeur E. On appelle n_+ le nombre d'atomes dont le moment magnétique $\vec{\mu}$ (associé au spin 1/2) pointe dans la direction du champ.

1) Exprimer sous forme d'une somme sur les valeurs de n_+ le nombre $\Omega(E)$ d'états accessibles au système.

2) Soit n_+^0 la valeur la plus probable de n_+. En développant jusqu'au second ordre, autour de n_+^0, l'entropie microcanonique partielle $s^*(E, n_+)$, calculer $\Omega(E)$ en fonction de n_+^0.

3) En déduire que l'entropie $S^*(E)$ du système est pratiquement égale à $s^*(E, n_+^0)$.

4) Montrer que la température microcanonique T^* du système vérifie

$$E = \frac{3}{2} NkT^* + (N - 2n_+^0)\mu B .$$

5) Calculer n_+^0 en fonction de T^* et du champ magnétique B. Comment varie avec la température la susceptibilité magnétique

$$\chi \widehat{=} \lim_{B \to 0} \frac{M(T^*)}{B},$$

où M est l'aimantation du gaz (moment magnétique moyen par unité de volume)?

II.15. On se propose de déterminer la distribution d'équilibre de N ($\gg 1$) particules d'un gaz parfait monoatomique contenu dans un récipient cylindrique de section s et de hauteur h placé dans le champ de pesanteur terrestre g, son énergie étant fixée à la valeur E. Pour cela on découpe le récipient en tranches horizontales d'épaisseur δz ($\ll h$, mais macroscopique); chaque tranche contient $n(z)\,\delta z$ particules, dont l'énergie totale (cinétique + potentielle de pesanteur) est $e(z)\,\delta z$.

1) Si l'on suppose $e(z)\,\delta z$ et $n(z)\,\delta z$ fixés, quelle est l'entropie microcanonique de la tranche de gaz comprise entre z et $z + \delta z$? En déduire l'entropie microcanonique partielle de l'ensemble du gaz pour la distribution $\{e(z), n(z)\}$?

2) En fait $e(z)$ et $n(z)$ sont des fonctions libres de s'ajuster, soumises cependant aux conditions de conservation de l'énergie totale E et du nombre total N de particules. En utilisant la méthode des multiplicateurs de Lagrange (appendice IV), déterminer la distribution la plus probable de la densité de particules $n(z)$ et de la densité d'énergie $e(z)$.

3) Comment les résultats précédents sont-ils modifiés si l'on tient compte de la variation de g avec l'altitude?

II.16. On considère un système isolé constitué de deux gaz G_1 et G_2 enfermés dans des récipients possédant une paroi commune (comme sur la figure 3 du chapitre II). On cherche la condition d'équilibre de ce système dans le cas où cette paroi est *mobile mais adiabatique* et imperméable aux particules.

Les volumes V_1 et V_2 des deux gaz sont des variables internes, liées par

$$V_1 + V_2 = V = \text{c}^{\text{te}}. \tag{1}$$

De même, l'ensemble des deux gaz étant isolé,

$$E_1 + E_2 = E = \text{cte}. \tag{2}$$

1) Nous allons tout d'abord montrer que, à l'équilibre, les variations de E_1 et E_2 sont déterminées par celles de V_1 et V_2 et en déduire la condition d'équilibre qui relie les valeurs les plus probables de ces grandeurs.

a) Montrer que l'équilibre correspond non seulement au maximum de l'entropie totale S des deux gaz mais aussi au maximum de chacune des entropies S_1 et S_2.

b) A l'équilibre, on a donc

$$\mathrm{d}S_1 = 0, \quad \mathrm{d}S_2 = 0, \quad \mathrm{d}S = 0. \tag{3}$$

Exprimer $\mathrm{d}S_1$ et $\mathrm{d}S_2$ en fonction de $\mathrm{d}E_1$ et $\mathrm{d}V_1$ pour la première et de $\mathrm{d}E_2$ et $\mathrm{d}V_2$ pour la seconde.

c) Montrer que les relations de conservation (1) et (2) et les conditions d'équilibre (3) imposent, comme on pouvait s'y attendre, l'égalité des pressions de part et d'autre de la cloison mobile :

$$p_1^* = p_2^* , \tag{4}$$

les températures T_1^* et T_2^* restant en général différentes.

c) Vérifier que cette égalité des pressions correspond bien à un maximum de l'entropie S^* du système, l'énergie totale $E_1 + E_2$ des deux gaz restant constante au cours des fluctuations de volume.

2) Analysons ensuite la façon dont le système, préparé dans un état (macroscopique) initial où les pressions des deux gaz sont différentes, peut *atteindre une situation d'équilibre.*

Nous supposons pour simplifier que les pressions p_1^* et p_2^* restent définies à chaque instant au cours de l'évolution du système (c'est-à-dire plus précisément que chacun des deux gaz subit une transformation quasistatique) (voir complément II.D, § II.3.b).

a) Montrer que, pendant cette évolution, on a nécessairement $\mathrm{d}E_1 + \mathrm{d}E_2 \neq 0$, si E_1 et E_2 sont les énergies respectives des gaz G_1 et G_2.

b) Montrer que la conservation de l'énergie totale du système s'écrit

$$E_1 + E_2 + E_c = E = \mathrm{c^{te}} , \tag{5}$$

où E_c est l'énergie cinétique du piston séparant les deux gaz (si l'on suppose l'énergie interne du piston, ou plus exactement sa capacité calorifique, négligeable).

c) Décrire qualitativement l'évolution du système à partir de l'état initial. Quels phénomènes lui permettent d'atteindre finalement un état d'équilibre?

3) Pour préciser un peu le lien entre l'état initial hors d'équilibre et l'état d'équilibre final, supposons que les deux gaz G_1 et G_2 sont parfaits.

Dans l'état initial, les paramètres macroscopiques prennent les valeurs $p_1^{(i)}$, $T_1^{(i)}$, $V_1^{(i)}$ pour G_1 et $p_2^{(i)}$, $T_2^{(i)}$, $V_2^{(i)}$ pour G_2, les nombres de particules N_1 et N_2 et l'énergie totale E étant fixés une fois pour toutes. Écrire les 3 relations qui existent entre ces paramètres (cf. complément II.C, § II.2).

Dans l'état final, $p_1^{(f)} = p_2^{(f)} = p$. Écrire à nouveau les 3 relations entre les paramètres. Montrer que le rapport $V_1^{(f)}/V_2^{(f)}$ des volumes à l'équilibre dépend des températures finales $T_1^{(f)}$ et $T_2^{(f)}$, qui ne sont pas complètement déterminées par les équations précédentes (pour une discussion plus détaillée, voir W.G. Hoover et B. Moran, Am. J. Phys. *47*, p. 851 (1979)).

DISCUSSION DES FORMULES FONDAMENTALES

II.17. Pour calculer l'entropie microcanonique d'un système macroscopique isolé et à l'équilibre (II, § I.C.1), nous avons écrit le nombre Ω d'états accessibles sous la forme

$$\Omega(E) \simeq \rho(E)\,\delta E , \tag{1}$$

où $\rho(E)$ est la densité d'états du système et δE l'incertitude sur son énergie E. Nous voulons discuter ici la validité de cette formule.

Lorsqu'on connaît δE, l'expression exacte de Ω est

$$\Omega(E) = \Phi(E + \delta E) - \Phi(E) , \tag{2}$$

où $\Phi(E)$ est le nombre d'états microscopiques d'énergie inférieure à E. Or la dépendance en E de $\Phi(E)$ est de la forme (cf. I, § III.B.2 et § III.B.3)

$$\Phi(E) \simeq A \left(\frac{E}{N\varepsilon_0}\right)^{N\alpha}, \tag{3}$$

où ε_0 est une énergie d'ordre microscopique, α un nombre de l'ordre de 1, et A est indépendant de E.

1) En examinant le comportement avec N des premières dérivées de $\Phi(E)$, montrer que, dans le développement de l'expression (2) en puissances de $\delta E/\varepsilon_0$, les termes d'ordre supérieur à 1 ne sont pas nécessairement négligeables devant le premier. La formule (1) ne peut donc pas être justifiée à partir de ce développement.

2) Mais le « théorème des accroissements finis » permet d'affirmer qu'il existe une valeur $\widetilde{E}$ de l'énergie, vérifiant

$$E < \widetilde{E} < E + \delta E, \tag{4}$$

telle que

$$\Phi(E + \delta E) - \Phi(E) = \rho(\widetilde{E})\delta E. \tag{5}$$

En déduire que l'erreur commise sur l'entropie du système lorsqu'on utilise la formule (1) est de l'ordre de $N\alpha k\delta E/E$. Quel est l'ordre de grandeur de l'erreur relative correspondante?

3) Montrer que, pour N de l'ordre du nombre d'Avogadro, l'entropie du système peut aussi être calculée, à une bonne approximation, par la formule

$$S^* \simeq k \ln \Phi. \tag{6}$$

II.18. On considère un système macroscopique isolé à l'équilibre. Son nombre de particules N et son énergie E sont des grandeurs extensives (complément II.A, § II); nous supposons que les autres paramètres extérieurs, que nous symboliserons par x, le sont également, de même que l'entropie microcanonique $S^*(E, N, x)$.

1) Montrer que, dans la limite où N devient très grand, l'entropie S^* est nécessairement de la forme

$$S^*(E, N, x) \underset{N \to \infty}{\sim} Nf\left(\frac{E}{N}, \frac{x}{N}\right), \tag{1}$$

où $f(u, v)$ est une fonction de 2 variables seulement (en fait $m + 1$, si m est le nombre de paramètres x). Pour fixer les idées, nous écrirons

$$S^*(E, N, x) = Nf_0\left(\frac{E}{N}, \frac{x}{N}\right) + N^a f_1\left(\frac{E}{N}, \frac{x}{N}\right) \tag{2}$$

avec $a < 1$ (le deuxième terme pourrait en réalité faire intervenir $N^a \ln^b N$).

2) En déduire que la température T^*, la pression p^* et le potentiel chimique μ^* sont de la forme

$$g_0\left(\frac{E}{N}, \frac{x}{N}\right) + N^{a-1} g_1\left(\frac{E}{N}, \frac{x}{N}\right). \tag{3}$$

Vérifier que ce résultat est compatible avec le caractère intensif de ces grandeurs.

3) Prenons maintenant une variable interne y extensive. En écrivant l'entropie

microcanonique partielle $s^*(E, N, x; y)$ sous une forme analogue à (2), montrer que ses dérivées successives par rapport à y sont telles que

$$\frac{\partial^n s^*}{\partial y^n} \underset{N \to \infty}{\sim} N^{1-n} h_0 \left(\frac{E}{N}, \frac{x}{N}; \frac{y}{N}\right), \tag{4}$$

et retrouver ainsi les formules de la note 16 du chapitre II.

II.19. On étudie une variable interne y dans un système macroscopique isolé et à l'équilibre, les paramètres extérieurs microcanoniques étant (E, x). On notera $\omega(E, x; y)\,\delta y$ le nombre d'états accessibles dans lesquels la variable interne a une valeur comprise entre y et $y+\delta y$, $s^*(E, x; y)$ l'entropie microcanonique partielle correspondante, et $w^*(E, x; y)$ la densité de probabilité caractérisant la distribution statistique de la variable y.

1) Écrire les relations donnant $s^*(E, x; y)$ et $w^*(E, x; y)$ à partir de $\omega(E, x; y)$.

2) La densité de probabilité w^* est pratiquement une gaussienne :

$$w^*(E, x; y) \simeq w^*(E, x; y_m)\, e^{-(y-y_m)^2/2(\Delta y)^2}. \tag{1}$$

Comment $w^*(E, x; y_m)$ et Δy se déduisent-il de l'entropie microcanonique partielle $s^*(E, x; y)$?

3) Exprimer l'entropie $S^*(E, x)$ du système à partir de $s^*(E, x; y_m)$ et Δy.

4) Supposons que y_m croisse comme le nombre N de particules du système (variable extensive); montrer que, si l'on néglige $\ln N$ devant N,

$$S^*(E, x) \simeq s^*(E, x; y_m). \tag{2}$$

5) La valeur la plus probable y_m de la variable interne dépend des paramètres extérieurs (E, x), de sorte que, à l'approximation précédente,

$$S^*(E, x) = s^*\big(E, x; y_m(E, x)\big). \tag{3}$$

Montrer que, malgré cela, les dérivées premières de S^* ne font pas intervenir celles de y_m : par exemple,

$$\frac{1}{T^*} \widehat{=} \frac{\partial S^*}{\partial E} = \left(\frac{\partial s^*}{\partial E}\right)_{x, y} (E, x; y_m). \tag{4}$$

C'est seulement dans les dérivées d'ordre supérieur de S^* qu'apparaissent les dérivées de y_m par rapport aux paramètres extérieurs.

II.20. La condition d'énergie minimale. On considère un système macroscopique isolé à l'équilibre, d'énergie E. On note x les autres paramètres extérieurs, et y une variable interne dont on cherche la valeur la plus probable. L'entropie microcanonique partielle correspondant à une valeur y de la variable interne est égale à l'entropie qu'aurait le système si cette variable était un paramètre extérieur fixé à la valeur y :

$$s^*(E, x; y) = S^*(E, x; y). \tag{1}$$

On pose

$$\left(\frac{\partial S^*}{\partial y}\right)_{E, x} \widehat{=} \eta(E, x; y). \tag{2}$$

La valeur la plus probable y_m de la variable interne est déterminée par l'équation

$$\eta(E, x, y_m) = 0, \tag{3}$$

à condition que l'entropie soit *maximum* en ce point :

$$\left(\frac{\partial \eta}{\partial y}\right)_{E,x}(y_m)<0\ . \tag{4}$$

1) Montrer que la positivité de la température permet de déterminer de façon unique l'énergie E en fonction de S^*, x, y :

$$E=E(S^*,\ x,\ y)\ . \tag{5}$$

2) Dans ce point de vue, la *condition d'entropie maximale* à E et x constants (formules (3) et (4)) va se traduire par une *condition d'énergie minimale* à S^* et x constants.

a) Montrer en premier lieu que

$$\left(\frac{\partial E}{\partial y}\right)_{S^*,x}=-T^*\eta\ . \tag{6}$$

En déduire que la fonction $E(S^*,\ x,\ y)$ est stationnaire en $y=y_m$, à S^* et x fixés.

b) Calculer la dérivée seconde $(\partial^2 E/\partial y^2)_{S^*,x}$ au point $y=y_m$. En déduire que ce point est un minimum de la fonction E, à S^* et x constants.

3) Montrer que, si le système étudié est *purement mécanique* (complément II.D, § I.2.d), la condition d'entropie maximale est inutilisable, de sorte que la valeur la plus probable d'une variable interne doit être calculée à partir de la condition d'énergie minimale. Comparer cette condition avec celle qui détermine l'équilibre du système en mécanique.

III

I. SITUATION ET DISTRIBUTION CANONIQUES

A. Situation canonique : définitions
   1. Notion de thermostat
   2. Température canonique d'un système
   3. Comparaison avec la situation microcanonique

B. Distribution canonique des états microscopiques
   1. Probabilité d'un état microscopique
   2. Discussion
      a. L'ensemble canonique
      b. L'énergie caractéristique kT

C. Distribution statistique des variables internes
   1. Distribution statistique de l'énergie
      a. Formule générale
      b. Formules approchées
   2. Distribution statistique d'une variable interne quelconque

II. FONCTION DE PARTITION ET ÉNERGIE LIBRE

A. Définition
   1. La fonction de partition
   2. L'énergie libre

B. Rôle fondamental : calcul des propriétés du système
   1. Caractéristiques de la distribution statistique de l'énergie
      a. Énergie moyenne et écart quadratique moyen
      b. Capacité calorifique et chaleur spécifique
   2. Distribution d'une variable interne
   3. Calcul de l'entropie canonique
   4. Pression et potentiel chimique canoniques

C. Système constitué de sous-systèmes macroscopiques indépendants

D. Évolution spontanée du système après relâchement d'une contrainte
   1. Calcul de l'entropie du système global
   2. Évolution du système $\mathcal{S}$ en contact avec un thermostat
   3. Application : échanges isothermes de volume et de particules
      a. Conditions d'équilibre
      b. Sens des échanges et stabilité de l'équilibre

III. LE FORMALISME CANONIQUE A LA LIMITE THERMODYNAMIQUE

A. La notion de limite thermodynamique

B. Limite thermodynamique de la description canonique
   1. Description canonique d'un grand système
   2. Passage à la limite thermodynamique

C. Équivalence avec la description microcanonique dans la limite thermodynamique

IV. DESCRIPTION CANONIQUE DES SYSTÈMES DE PARTICULES INDÉPENDANTES

A. Formules générales pour la fonction de partition et l'énergie libre
   1. Hypothèses et notations
   2. Particules identiques mais discernables
   3. Particules indiscernables
      a. Caractérisation et décompte des états microscopiques du système
      b. Tentative de calcul exact de la fonction de partition
      c. L'approximation de Maxwell-Boltzmann

B. Exemple : les gaz parfaits monoatomiques
   1. Définition d'un gaz parfait
   2. Fonction de partition et énergie libre d'un gaz parfait monoatomique
      a. Fonction de partition d'une particule du gaz
      b. Fonction de partition et énergie libre du gaz
   3. Propriétés du gaz parfait monoatomique
      a. Énergie et chaleur spécifique
      b. Entropie
      c. Pression et potentiel chimique

V. LA DISTRIBUTION CANONIQUE EN MÉCANIQUE CLASSIQUE

A. Formules fondamentales
   1. Densité canonique dans l'espace des phases
   2. Fonction de partition classique
      a. Systèmes classiques ordinaires (éléments tous discernables)
      b. Systèmes de particules indiscernables

B. Domaine de validité du traitement classique

C. Le théorème d'« équipartition de l'énergie »
   1. Énoncé et démonstration
   2. Discussion physique

CHAPITRE III

Système en équilibre avec un thermostat. Description canonique

Le postulat fondamental que nous avons présenté et discuté au chapitre précédent s'applique aux systèmes macroscopiques isolés et à l'équilibre. Lorsque le système auquel on s'intéresse interagit avec l'extérieur, c'est-à-dire avec d'autres systèmes, ceux-ci influent sur ses propriétés et son comportement, qui ne peuvent dès lors être analysés qu'en étudiant l'ensemble des divers systèmes en interaction. Dans certains cas simples cependant, un petit nombre de paramètres suffisent à caractériser les influences extérieures sur le système considéré, de sorte que l'on peut à nouveau restreindre son attention au système lui-même.

Dans la situation qualifiée traditionnellement de « canonique », le système à étudier est en contact thermique avec un « thermostat » qui lui impose, de l'extérieur, sa température; la donnée de ce seul paramètre suffit à caractériser l'influence du thermostat et permet de déterminer les propriétés du système.

I. SITUATION ET DISTRIBUTION CANONIQUES

A. Situation canonique : définitions

Considérons un système $\mathcal{S}$, en *contact thermique* (II, § III.A.1.b) avec un autre système $\mathcal{C}$ beaucoup plus gros que lui. Le système global formé par la réunion de $\mathcal{S}$ et $\mathcal{C}$ est isolé et à l'équilibre[1]; son énergie totale E_{tot}, fixée, est simplement la somme de l'énergie E du système $\mathcal{S}$ que nous voulons étudier et de celle $E_{\mathcal{C}}$ du système $\mathcal{C}$:

$$E + E_{\mathcal{C}} = E_{\text{tot}} \text{ fixée} . \tag{I.1}$$

1. NOTION DE THERMOSTAT

Le système $\mathcal{C}$ étant beaucoup plus gros que $\mathcal{S}$, l'énergie E est toujours beaucoup plus petite que $E_{\mathcal{C}}$:

$$E \ll E_{\mathcal{C}} . \tag{I.2}$$

Ainsi, les échanges de chaleur qui conduisent à l'équilibre thermique entre $\mathcal{S}$ et $\mathcal{C}$, ou qui accompagnent les fluctuations autour de cet équilibre, peuvent changer de façon appréciable l'énergie E du système le plus petit tout en restant négligeables devant $E_{\mathcal{C}}$. De même, ces échanges peuvent modifier l'état du système $\mathcal{S}$ sans que l'état macroscopique de $\mathcal{C}$ soit pratiquement affecté. $\mathcal{C}$ joue alors vis-à-vis de $\mathcal{S}$ le rôle de *« réservoir d'énergie »* ou de *« thermostat »;* à l'équilibre thermique, on dit que le système $\mathcal{S}$ se trouve dans la *« situation canonique »*.

---

1. Il n'est pas nécessaire que le système $\mathcal{S}$ soit macroscopique, pourvu que $\mathcal{C}$ le soit, de façon que le système global $\mathcal{S} \cup \mathcal{C}$, auquel nous allons appliquer le postulat fondamental de la mécanique statistique, le soit aussi.

Précisons. Les échanges de chaleur permis par le contact thermique font varier les énergies E et $E_{\mathcal{C}}$ des deux systèmes; pour chaque valeur de E, l'énergie du gros système vaut $E_{\mathcal{C}} = E_{\text{tot}} - E$ et sa température microcanonique $T^*_{\mathcal{C}}$ est liée à son entropie microcanonique $S^*_{\mathcal{C}}(E_{\mathcal{C}})$ par (définition II (III.2))

$$\frac{1}{T^*_{\mathcal{C}}} = \frac{\partial S^*_{\mathcal{C}}}{\partial E_{\mathcal{C}}} (E_{\mathcal{C}} = E_{\text{tot}} - E) \,. \tag{I.3}$$

La température $T^*_{\mathcal{C}}$ varie donc *a priori* avec E.

Par *définition*, $\mathcal{C}$ constitue un *thermostat* pour le système $\mathcal{S}$ si *sa température microcanonique $T^*_{\mathcal{C}}$ est pratiquement indépendante de l'énergie E du système $\mathcal{S}$*[2]. E restant toujours faible devant E_{tot}, nous pouvons développer le second membre de (I.3) autour de $E_{\mathcal{C}} = E_{\text{tot}}$:

$$\frac{\partial S^*_{\mathcal{C}}}{\partial E_{\mathcal{C}}} (E_{\mathcal{C}} = E_{\text{tot}} - E) = \frac{\partial S^*_{\mathcal{C}}}{\partial E_{\mathcal{C}}} (E_{\mathcal{C}} = E_{\text{tot}}) - E \frac{\partial^2 S^*_{\mathcal{C}}}{\partial E^2_{\mathcal{C}}} (E_{\mathcal{C}} = E_{\text{tot}}) + \dots \,. \tag{I.4}$$

$\mathcal{C}$ constitue donc un thermostat pour $\mathcal{S}$ si ce développement se réduit à son premier terme, c'est-à-dire si

$$E \frac{\partial^2 S^*_{\mathcal{C}}}{\partial E^2_{\mathcal{C}}} (E_{\mathcal{C}} = E_{\text{tot}}) \ll \frac{\partial S^*_{\mathcal{C}}}{\partial E_{\mathcal{C}}} (E_{\mathcal{C}} = E_{\text{tot}}) \,, \tag{I.5}$$

avec des conditions analogues pour les termes d'ordre supérieur. En pratique, ces conditions seront réalisées si le système $\mathcal{C}$ est suffisamment grand par rapport au système $\mathcal{S}$ que l'on désire étudier[3].

Notons que cette définition d'un thermostat est *relative,* puisqu'elle dépend de la taille du système $\mathcal{S}$ que l'on met en contact thermique avec lui : un récipient contenant 5 ou 6 litres d'eau pourra constituer un thermostat pour un échantillon de gaz de quelques cm^3, mais pas pour l'air remplissant une pièce de plusieurs m^3. De plus, la définition repose sur une *approximation :* c'est la précision exigée dans la détermination de la température qui fixe le niveau quantitatif auquel doit être vérifiée la condition (I.5); en pratique, on aura un bon thermostat, dans une expérience donnée, si le thermomètre qui mesure sa température (avec la précision nécessaire) indique une valeur constante pendant toute la durée de l'expérience.

---

2. *Attention :* nous ne demandons pas que l'entropie microcanonique $S^*_{\mathcal{C}}$ du thermostat soit indépendante de l'énergie E du système $\mathcal{S}$; c'est sur la *température* microcanonique $T^*_{\mathcal{C}}$ que porte la condition. En effet, les variations de $S^*_{\mathcal{C}}$ induites par des variations de E sont bien négligeables tant qu'on les compare aux valeurs de $S^*_{\mathcal{C}}$ elle-même, mais elles ne le sont plus si c'est l'entropie du système $\mathcal{S}$ qui sert de référence.

3. L'entropie microcanonique $S^*_{\mathcal{C}}$ du thermostat $\mathcal{C}$ croît comme le nombre de particules $N_{\mathcal{C}}$ de ce système (II, § I.C.1); il en est de même de son énergie $E_{\mathcal{C}}$, de sorte que $\partial^2 S^*_{\mathcal{C}} / \partial E^2_{\mathcal{C}}$ est d'ordre $N_{\mathcal{C}} / N^2_{\mathcal{C}} = 1/N_{\mathcal{C}}$. Quant à E, elle est proportionnelle au nombre N de particules constituant le système $\mathcal{S}$. Le membre de gauche de la relation (I.5) est donc de l'ordre de $N/N_{\mathcal{C}}$; si $N \ll N_{\mathcal{C}}$, il est négligeable devant le membre de droite, qui est d'ordre 1. Les termes d'ordre supérieur du développement (I.4) sont de plus en plus petits : un raisonnement analogue montre qu'ils se comportent comme $N^2/N^2_{\mathcal{C}}$, $N^3/N^3_{\mathcal{C}}$, ... (bien entendu, cette conclusion reste valable si $\mathcal{S}$ est de taille microscopique et $\mathcal{C}$ macroscopique).

On pourrait objecter à ce raisonnement que rien n'interdit *a priori* à E d'atteindre la valeur E_{tot} de l'énergie globale, auquel cas l'inégalité (I.2) ne serait bien sûr pas vérifiée, et pas davantage (I.5). Mais une telle éventualité est tellement peu probable qu'elle est pratiquement impossible : la distribution de $E_{\mathcal{C}}$ autour de sa valeur la plus probable est extrêmement étroite en valeur relative ($\Delta E_{\mathcal{C}} / E_{\mathcal{C}} \sim 1/\sqrt{N_{\mathcal{C}}}$). Il sera donc inutile, dans la suite, d'imposer explicitement une borne supérieure à $E/E_{\mathcal{C}}$; nous raisonnerons en fait comme si E pouvait devenir aussi grande que l'on veut, sachant bien que ses valeurs trop différentes de la valeur moyenne auront de toutes façons des probabilités quasiment nulles.

2. TEMPÉRATURE CANONIQUE D'UN SYSTÈME

Comme nous l'avons indiqué, c'est le système $\mathcal{S}$ qui va nous intéresser dans la suite. S'il est en équilibre thermique avec un thermostat $\mathcal{C}$ dont la température microcanonique $T^*_{\mathcal{C}}$ reste invariable (§ 1), on définit la *température canonique du système* $\mathcal{S}$, que nous noterons simplement T, comme

$$\boxed{T \widehat{=} T^*_{\mathcal{C}}} \tag{I.6}$$

Le système $\mathcal{S}$ est donc *maintenu à la température* T par le thermostat; ce paramètre, T, caractérise à lui seul l'influence du thermostat sur le système étudié. Noter que, pour un système $\mathcal{S}$ macroscopique, la définition (I.6) découle de façon naturelle de la condition d'équilibre thermique entre $\mathcal{S}$ et $\mathcal{C}$ (II, § III.A.2.c.α); nous préciserons ce point plus loin (formule (I.29)).

3. COMPARAISON AVEC LA SITUATION MICROCANONIQUE

Il importe de bien comprendre la différence entre la situation canonique que nous venons de définir et la situation microcanonique étudiée au chapitre II.

La situation *microcanonique* est celle d'un système macroscopique *isolé à l'équilibre,* dont *l'énergie est* par conséquent *fixée* et ne peut donc pas fluctuer (c'est-à-dire que sa valeur est certaine au sens probabiliste). Dans cette situation, la température du système est définie par la formule II (III.2); elle est fixée elle aussi, puisqu'elle se déduit de la dérivée $\partial S^*/\partial E$ de l'entropie microcanonique par rapport à l'énergie, prise pour *la* valeur de l'énergie que renferme le système[4].

Dans la situation *canonique,* au contraire, le système n'est *pas isolé :* il échange de la chaleur avec un thermostat. Par conséquent, *son énergie fluctue* (elle peut prendre différentes valeurs ayant chacune une certaine probabilité); l'énergie du système est donc ici une variable interne, et non plus un paramètre extérieur. C'est la *température canonique, fixée* par le thermostat, qui remplace l'énergie comme paramètre extérieur.

Notons que la température microcanonique se calcule à partir des caractéristiques du système lui-même, alors qu'ici sa température canonique T lui est imposée de l'extérieur, par le thermostat. Bien entendu, température canonique et température microcanonique se réfèrent à la *même grandeur physique,* comme le montre la définition (I.6); c'est la situation du système considéré qui change d'un cas à l'autre.

B. Distribution canonique des états microscopiques

Nous allons maintenant établir, à partir du postulat fondamental énoncé au chapitre II, la distribution de probabilité $\{P^c_\ell\}$ des divers états microscopiques possibles (ℓ) du système $\mathcal{S}$, lorsque celui-ci est en équilibre avec un thermostat de température T.

1. PROBABILITÉ D'UN ÉTAT MICROSCOPIQUE

Nous désignons comme toujours par (ℓ) un état microscopique, d'énergie E_ℓ, du système $\mathcal{S}$ étudié; nous noterons de même (L) un état microscopique, d'énergie E_L, du thermostat $\mathcal{C}$. Le thermostat et le système $\mathcal{S}$ étant en contact thermique (II, § III.A.1.b), un état microscopique quelconque du système global $\mathcal{S} \cup \mathcal{C}$ constitué

4. Au chapitre II (§ III.A.2), nous avons parlé de température microcanonique pour des systèmes qui n'étaient plus isolés, puisqu'ils étaient en contact thermique l'un avec l'autre. C'est que les caractéristiques du contact thermique permettent encore, pour toute répartition de l'énergie totale entre les deux systèmes, de définir pour chacun d'eux un nombre d'états accessibles, et par conséquent une entropie et une température microcanoniques (cf. II, note 30).

par la réunion de $\mathcal{S}$ et $\mathcal{T}$ peut être caractérisé par la donnée de l'état microscopique de $\mathcal{S}$ et de celui de $\mathcal{T}$, c'est-à-dire par le couple (ℓ, L); l'énergie d'un tel état est simplement

$$E_{\ell, L} = E_\ell + E_L \ . \tag{I.7}$$

Le système global $\mathcal{S} \cup \mathcal{T}$ est isolé; on peut donc lui appliquer le postulat fondamental de la mécanique statistique : tous les états accessibles de ce système global sont également probables à l'équilibre. Nous cherchons la probabilité P^c_ℓ pour que le système $\mathcal{S}$ se trouve dans un état (ℓ) particulier, quel que soit par ailleurs l'état du thermostat. Il suffit, comme nous avons appris à le faire au chapitre II (§ II.A.1), de *compter les états accessibles du système global $\mathcal{S} \cup \mathcal{T}$ qui possèdent cette caractéristique;* nous diviserons ensuite leur nombre par le nombre total d'états accessibles.

Nous voulons donc dénombrer les états (ℓ, L) qui vérifient les conditions[5]

$$\begin{cases} \mathcal{S} \text{ dans l'état } (\ell) \text{ choisi} , & \text{(I.8.a)} \\ E_{\text{tot}} \leqslant E_{\ell, L} \leqslant E_{\text{tot}} + \delta E \ . & \text{(I.8.b)} \end{cases}$$

Fixer l'état (ℓ) du système $\mathcal{S}$ revient en particulier à fixer son énergie E_ℓ. Compte tenu de (I.7), les conditions précédentes peuvent donc être exprimées sous la forme

$$\begin{cases} \mathcal{S} \text{ dans l'état } (\ell) \text{ choisi} , & \text{(I.9.a)} \\ E_{\text{tot}} - E_\ell \leqslant E_L \leqslant E_{\text{tot}} - E_\ell + \delta E \ . & \text{(I.9.b)} \end{cases}$$

Puisque l'état du système $\mathcal{S}$ est fixé, le nombre d'états (ℓ, L) à évaluer est égal au nombre des états (L) du thermostat qui satisfont à (I.9.b). Si $\Omega_{\mathcal{T}}(E_{\mathcal{T}})$ est le *nombre d'états microscopiques du thermostat* $\mathcal{T}$ qui ont l'énergie $E_{\mathcal{T}}$ (à δE près), la probabilité cherchée s'écrit

$$P^c_\ell = \frac{\Omega_{\mathcal{T}}(E_{\mathcal{T}} = E_{\text{tot}} - E_\ell)}{\Omega_{\mathcal{S} \cup \mathcal{T}}(E_{\text{tot}})} , \tag{I.10}$$

où $\Omega_{\mathcal{S} \cup \mathcal{T}}(E_{\text{tot}})$ est le nombre total d'états accessibles du système global $\mathcal{S} \cup \mathcal{T}$.

Comme le dénominateur de la formule (I.10) est indépendant de (ℓ), il suffit d'évaluer le numérateur :

$$P^c_\ell = C \Omega_{\mathcal{T}}(E_{\mathcal{T}} = E_{\text{tot}} - E_\ell) \ ; \tag{I.11}$$

la constante C sera ensuite déterminée par normalisation de la distribution :

$$\sum_{(\ell)} P^c_\ell = 1 \ . \tag{I.12}$$

Mais $\Omega_{\mathcal{T}}$ est relié simplement à *l'entropie microcanonique du thermostat* $S^*_{\mathcal{T}}(E_{\mathcal{T}})$ (II, formule (I.5)), de sorte que

$$P^c_\ell = C \exp \left[\frac{1}{k} S^*_{\mathcal{T}}(E_{\mathcal{T}} = E_{\text{tot}} - E_\ell) \right] . \tag{I.13}$$

La probabilité canonique P^c_ℓ d'un état microscopique (ℓ) du système $\mathcal{S}$ ne dépend donc que de l'énergie E_ℓ de cet état. La définition du thermostat et de la température

5. Les paramètres extérieurs autres que l'énergie (volume, nombre de particules, ... de chacun des systèmes $\mathcal{S}$ et $\mathcal{T}$) sont fixés une fois pour toutes et n'interviennent pas directement dans le raisonnement.

canonique T (§ A) permet d'expliciter cette dépendance. En effet, le développement de $S^*_{\mathcal{C}}(E_{\mathcal{C}} = E_{\text{tot}} - E_\ell)$ autour de $E_{\mathcal{C}} = E_{\text{tot}}$ peut être tronqué[6] après le terme du premier ordre :

$$S^*_{\mathcal{C}}(E_{\text{tot}} - E_\ell) = S^*_{\mathcal{C}}(E_{\text{tot}}) - E_\ell \frac{\partial S^*_{\mathcal{C}}}{\partial E_{\mathcal{C}}}(E_{\text{tot}}) , \tag{I.14}$$

puisque le terme suivant est négligeable d'après la condition (I.5). Le coefficient de $-E_\ell$ n'est autre que $1/T^*_{\mathcal{C}}$, c'est-à-dire $1/T$ d'après la définition (I.6). La probabilité P^c_ℓ s'écrit donc finalement[7]

$$\boxed{P^c_\ell = \frac{1}{Z} \mathrm{e}^{-E_\ell/kT}} , \tag{I.15}$$

---

6. Ce développement est destiné à être reporté dans une exponentielle (formule (I.13)). Il *faut* donc conserver tout terme qui est d'ordre 1 ou supérieur (même s'il est petit devant les précédents), mais on peut négliger ceux qui sont petits devant 1 (cf. note 6 du complément I.C). Or $S^*_{\mathcal{C}}(E_{\text{tot}})$ est de l'ordre du nombre $N_{\mathcal{C}}$ de particules du thermostat, $\partial S^*_{\mathcal{C}}/\partial E_{\mathcal{C}}$ est d'ordre 1 et $\partial^2 S^*_{\mathcal{C}}/\partial E^2_{\mathcal{C}}$ d'ordre $1/N_{\mathcal{C}}$. De même, E_ℓ est de l'ordre du nombre N de particules du système $\mathcal{S}$. Ainsi, le terme que nous négligeons,

$$E^2_\ell \times \frac{1}{k} \frac{\partial^2 S^*_{\mathcal{C}}}{\partial E^2_{\mathcal{C}}}(E_{\text{tot}}) \sim \frac{N^2}{N_{\mathcal{C}}} , \tag{1}$$

est petit devant

$$E_\ell \times \frac{1}{k} \frac{\partial S^*_{\mathcal{C}}}{\partial E_{\mathcal{C}}}(E_{\text{tot}}) \sim N , \tag{2}$$

mais il n'est vraiment petit devant 1 que si

$$\frac{N^2}{N_{\mathcal{C}}} \ll 1 . \tag{3}$$

Cette condition est évidemment satisfaite si $\mathcal{S}$ est microscopique. Elle est extrêmement restrictive si $\mathcal{S}$ est macroscopique, mais n'est heureusement pas indispensable pour la validité des raisonnements qui suivent (voir exercice III.46) : comme c'est le logarithme de la fonction de partition (c'est-à-dire l'énergie libre; voir paragraphe II) qui va jouer le rôle central, il suffira en fait que

$$\frac{N}{N_{\mathcal{C}}} \ll 1 . \tag{4}$$

7. Compte tenu de la définition (I.6), la formule (I.14) devient

$$S^*_{\mathcal{C}}(E_{\text{tot}} - E_\ell) = S^*_{\mathcal{C}}(E_{\text{tot}}) - \frac{1}{T} E_\ell , \tag{1}$$

que l'on reporte dans l'expression (I.13) de la probabilité :

$$P^c_\ell = C \exp\left[\frac{1}{k} S^*_{\mathcal{C}}(E_{\text{tot}}) - \frac{E_\ell}{kT}\right] . \tag{2}$$

Il suffit alors de regrouper les facteurs indépendants de (ℓ) en posant

$$\frac{1}{Z} \widehat{=} C \exp\left[\frac{1}{k} S^*_{\mathcal{C}}(E_{\text{tot}})\right] \tag{3}$$

pour obtenir le résultat (I.15).

où la constante multiplicative, traditionnellement notée $1/Z$, est déterminée par normalisation (condition (I.12)) :

$$\boxed{Z=\sum_{(\ell)} e^{-E_\ell/kT}} \quad . \tag{I.16}$$

La formule (I.15) caractérise la *distribution canonique*[8] : elle donne la probabilité des divers états microscopiques (ℓ) d'un système en équilibre thermique avec un thermostat qui le maintient à la température T. Z est appelée *«fonction de partition»* du système. La distribution canonique est applicable à des *systèmes de taille quelconque,* microscopiques[9] ou macroscopiques (cf. note 1).

2. DISCUSSION

a. L'ensemble canonique

L'ensemble statistique associé à la distribution canonique (I.15) est appelé « ensemble canonique ». Il s'agit comme toujours (I, § IV.B.2) d'un ensemble imaginaire constitué d'un très grand nombre $\mathcal{N}$ de systèmes qui sont des répliques du système étudié, et qui sont ici tous maintenus par un thermostat à la même température T. La proportion $\mathcal{N}_\ell/\mathcal{N}$ de ces systèmes qui se trouvent dans un état microscopique (ℓ) donné tend vers P^c_ℓ quand $\mathcal{N}$ tend vers l'infini.

Dans l'ensemble microcanonique, tous les systèmes avaient la même énergie E, à δE près. Ici au contraire, *tous les états microscopiques* (ℓ) *sont a priori représentés,* quelle que soit leur énergie[10]. Ceci est dû à ce que le thermostat peut par définition

---

8. Notons que l'expression trouvée pour P^c_ℓ est invariante dans un changement de l'origine des énergies. Si l'on pose en effet

$$E_\ell = E'_\ell - E_0 \, , \tag{1}$$

le facteur $\exp(E_0/kT)$, indépendant de (ℓ), peut être incorporé dans la constante de normalisation, de sorte que P^c_ℓ s'écrit aussi bien

$$P^c_\ell = \frac{1}{Z'} e^{-E'_\ell/kT} \, , \tag{2}$$

avec

$$Z' = \sum_{(\ell)} e^{-E'_\ell/kT} \, . \tag{3}$$

Il est intéressant de souligner que cette invariance est assurée par la propriété de factorisation

$$\exp\left[-\frac{E_\ell}{kT}\right] = \exp\left[-\frac{E'_\ell}{kT}\right] . \exp\frac{E_0}{kT} \, , \tag{4}$$

qui est caractéristique de la fonction exponentielle. En ce sens, il était nécessaire que la dépendance de P_ℓ par rapport à E_ℓ soit exponentielle (ou constante, comme dans la distribution microcanonique).

9. Un système microscopique se trouve en général placé dans la situation canonique, le thermostat étant constitué par son environnement macroscopique, avec lequel il est en contact thermique (molécule dans un gaz, atome ou ion dans un réseau cristallin, ...). Il faut cependant, pour que le raisonnement précédent s'applique, que le système microscopique soit *faiblement couplé* à son environnement.

10. Ces états doivent cependant être compatibles avec les valeurs fixées par ailleurs pour les paramètres extérieurs tels que volume, nombre de particules, ...

toujours « mettre le complément » pour assurer la conservation de l'énergie totale du système global isolé. Cependant, *la probabilité des divers états dépend fortement de leur énergie,* puisque le rapport des probabilités de deux états (ℓ) et (ℓ') est une fonction exponentielle de leur différence d'énergie :

$$\frac{P^c_\ell}{P^c_{\ell'}} = e^{-(E_\ell - E_{\ell'})/kT}, \tag{I.17}$$

l'état de plus basse énergie étant le plus probable des deux.

Attention : la formule (I.15) donne la probabilité des divers *états* microscopiques (ℓ) du système, pas des énergies E_ℓ correspondantes; il n'est pas vrai en général que les énergies les plus basses soient les plus probables (voir § C.1), car il faut tenir compte du nombre d'états différents associés à chaque énergie.

b. L'énergie caractéristique *kT*

Dans la distribution canonique (I.15) intervient le produit kT de la constante de Boltzmann par la température canonique imposée au système. Ce produit a les dimensions d'une énergie. Ainsi, à chaque température T est associée une *énergie caractéristique* kT. Celle-ci est pratiquement toujours[11] d'un *ordre de grandeur microscopique* (voir complément I.A) : il est commode de garder présent à l'esprit que

$$kT \simeq 1 \text{ eV} \iff T \simeq 12\,000 \text{ K}. \tag{I.18}$$

Notons que la valeur de cette énergie caractéristique est *fixée par le thermostat,* c'est-à-dire imposée *de l'extérieur* au système, dont elle influence fortement les propriétés. Prenons par exemple deux états microscopiques (ℓ) et (ℓ') déterminés, d'énergies E_ℓ et $E_{\ell'}$ différentes. D'après (I.17), leurs probabilités sont pratiquement égales si kT est grand devant la différence d'énergie $|E_\ell - E_{\ell'}|$; elles sont au contraire très différentes si kT est inférieur à $|E_\ell - E_{\ell'}|$ ou du même ordre de grandeur. Le comportement d'un même système peut donc changer de façon radicale suivant la valeur de l'énergie kT caractéristique du thermostat[12].

Pour un *système macroscopique* cependant, cette décroissance exponentielle, en fonction de l'énergie, de la probabilité individuelle des états microscopiques (ℓ) est contrebalancée par une croissance extrêmement rapide du nombre d'états ayant une énergie donnée; cette compétition entre deux effets contraires aboutit (voir paragraphe C.1.b.β) à éliminer complètement les états de basse énergie, malgré leur forte probabilité individuelle, au profit d'états dont l'énergie est grande devant kT, et dont la probabilité individuelle est par conséquent infime (mais le nombre fantastiquement grand).

---

11. Seule peut-être l'explosion initiale (« big-bang ») qui a, pense-t-on actuellement, donné naissance à notre Univers a pu créer des températures si fantastiquement élevées que cette affirmation doive être révisée.

12. Ainsi par exemple, les atomes constituant un gaz comme l'hélium ou le néon (gaz monoatomique) ont à la température ordinaire ($T \simeq 300$ K) une probabilité quasiment nulle de se trouver dans un état excité : leur énergie d'excitation $E_1 - E_0$, différence entre leur énergie la plus basse possible E_0 (niveau fondamental) et l'énergie E_1 immédiatement supérieure dans le spectre (premier niveau excité), est de l'ordre de l'eV, alors que kT vaut seulement 1/40 d'eV à la température ordinaire. On dit que les « degrés de liberté électroniques » de l'atome sont alors « *gelés* ». Au contraire, les états excités, y compris ceux dans lesquels l'atome est dissocié, jouent un rôle important lorsque la température atteint ou dépasse quelques milliers de degrés, comme par exemple dans les étoiles.

C. Distribution statistique des variables internes

1. DISTRIBUTION STATISTIQUE DE L'ÉNERGIE

Lorsqu'un système se trouve dans la situation canonique, son énergie fluctue par suite des échanges de chaleur avec le thermostat; elle est donc devenue une *variable interne* (II, § I.A.1), dont la distribution statistique se déduit de la distribution canonique (I.15).

a. Formule générale

Pour évaluer la probabilité d'*une valeur possible E_ℓ de l'énergie,* il faut tenir compte du fait que E_ℓ est en général une valeur propre dégénérée de l'hamiltonien H, c'est-à-dire qu'il lui correspond *plusieurs états propres* (ℓ) *distincts;* il faut alors sommer les probabilités de ces divers états (« loi d'addition »). Or ces probabilités sont égales entre elles, puisque P^c_ℓ ne dépend que de l'énergie de l'état (ℓ). La probabilité $P^c(E_\ell)$ pour que l'énergie du système soit E_ℓ s'écrit donc

$$P^c(E_\ell) = g(E_\ell)\,\frac{1}{Z}\,\mathrm{e}^{-E_\ell/kT}\,, \tag{I.19}$$

où $g(E_\ell)$ est le degré de dégénérescence du niveau E_ℓ, c'est-à-dire le *nombre d'états distincts*[13] ayant l'énergie E_ℓ. La distribution statistique (I.19) est directement normalisée :

$$\sum_{E_\ell} P^c(E_\ell) = 1\,, \tag{I.20}$$

car l'expression (I.16) de la constante Z peut être transformée en

$$Z = \sum_{E_\ell} g(E_\ell)\,\mathrm{e}^{-E_\ell/kT}\,. \tag{I.21}$$

Attention : les sommes qui apparaissent dans les formules (I.12) et (I.16) comportent *un terme par état,* celles que nous venons d'écrire en (I.20) et (I.21) *un terme par niveau d'énergie.*

Bien entendu, la valeur moyenne $\overline{E}^c$ de l'énergie dans la situation canonique s'obtient par les formules habituelles (appendice I, § I.4) :

$$\overline{E}^c = \sum_{(\ell)} E_\ell P^c_\ell = \frac{1}{Z}\sum_{(\ell)} E_\ell\,\mathrm{e}^{-E_\ell/kT} \tag{I.22.a}$$

$$= \sum_{E_\ell} E_\ell P^c(E_\ell) = \frac{1}{Z}\sum_{E_\ell} E_\ell\, g(E_\ell)\,\mathrm{e}^{-E_\ell/kT}\,. \tag{I.22.b}$$

On peut évaluer de façon analogue l'écart quadratique moyen ΔE^c.

b. Formules approchées

α. *L'approximation continue*

Dans certains cas, les sommes sur les niveaux d'énergie discrets E_ℓ que nous avons écrites au paragraphe a peuvent être approximées par des intégrales, c'est-à-dire que l'on peut *traiter l'énergie du système comme une variable continue.*

---

13. Pour un système microscopique, $g(E_\ell)$ est la dimension du sous-espace de l'espace des états constitué par l'ensemble des vecteurs propres de l'hamiltonien H associés à la valeur propre E_ℓ considérée.

Pour un système macroscopique, nous identifierons ce facteur comme étant le nombre d'états accessibles $\Omega(E)$ correspondant à la valeur de E choisie (§ b.α).

Pour qu'une telle approximation soit valable, il faut que deux conditions soient réunies. D'une part, le spectre d'énergie du système doit être suffisamment serré pour qu'un intervalle «infinitésimal» δE (inférieur à l'incertitude sur une mesure de l'énergie) contienne un grand nombre de niveaux E_ℓ distincts. Il faut d'autre part que la fonction $f(E_\ell)$ à sommer varie très peu d'une énergie possible E_ℓ à la suivante, de façon que l'on puisse définir $f(E)$ comme la valeur de $f(E_\ell)$ pour $E_\ell \simeq E$, et que $f(E+\delta E)$ soit «infiniment» voisin de $f(E)$. La première de ces conditions est celle qui permet de définir une *densité d'états* $\rho(E)$ (I, § III.B.1); elle est vérifiée pour *tout système macroscopique;* elle l'est également pour un système constitué d'*une seule particule* lorsque celle-ci peut se déplacer dans un *volume de taille macroscopique* (cf. I, § III.B.2.a). La seconde condition suppose que l'énergie caractéristique kT fixée par le thermostat est grande devant l'intervalle entre deux niveaux E_ℓ consécutifs, ce qui est pratiquement toujours vérifié dans les deux cas que nous venons de mentionner[14].

Si ces conditions sont remplies, la distribution statistique de l'énergie est caractérisée par une *densité de probabilité* $w^c(E)$ donnée par[15]

$$w^c(E)=\frac{1}{Z}\,\rho(E)\,\mathrm{e}^{-E/kT}\,. \tag{I.23}$$

Cette distribution est ici aussi directement normalisée, car l'expression (I.21) de Z s'écrit, dans le cas où l'énergie est une variable continue,

$$Z=\int_{E_0}^{+\infty}\mathrm{d}E\rho(E)\,\mathrm{e}^{-E/kT} \tag{I.24}$$

(E_0 désigne le niveau fondamental du système, c'est-à-dire son énergie la plus basse). De même, la formule (I.22.b) pour la valeur moyenne de l'énergie devient ici

$$\overline{E}^c=\int_{E_0}^{+\infty}E\,w^c(E)\mathrm{d}E=\frac{1}{Z}\int_{E_0}^{+\infty}E\rho(E)\,\mathrm{e}^{-E/kT}\,\mathrm{d}E\,. \tag{I.25}$$

---

14. Pour que cette affirmation soit mise en défaut, dans le cas d'une particule unique de masse m enfermée dans un volume $V=L^3$, il faudrait (cf. I, note 18) que

$$kT\lesssim\frac{\pi^2\hbar^2}{mL^2}\,.$$

L'évaluation reste la même pour un système macroscopique constitué d'un grand nombre de telles particules, si celles-ci sont indépendantes : l'augmentation minimale possible pour l'énergie du système correspond à l'augmentation minimale pour une particule. Or, même pour un électron ($m\simeq 10^{-30}$ kg) et pour $L=1$ cm, on trouve que T devrait être inférieure à 10^{-10} K. De telles températures sont loin de celles que l'on peut atteindre actuellement ($\simeq 10^{-6}$ K). De toutes façons, l'approximation des particules indépendantes cesse d'être valable bien avant : lorsque kT devient de l'ordre de l'énergie d'interaction entre deux particules, celle-ci ne peut plus être négligée.

Mais il est des situations (chaleurs spécifiques des solides à basse température – voir complément III.E) où l'énergie caractéristique kT n'est plus grande devant la séparation des niveaux du système (de l'ordre de $\hbar w_E$ dans cet exemple). Les considérations qui suivent ne sont alors pas applicables telles quelles.

15. Les $\rho(E)\,\mathrm{d}E$ états du système dont l'énergie est comprise entre E et $E+\mathrm{d}E$ ont pratiquement tous la même probabilité, donnée par la distribution canonique :

$$\frac{1}{Z}\,\mathrm{e}^{-E/kT}\,. \tag{1}$$

La probabilité $w^c(E)\mathrm{d}E$ pour que l'énergie du système soit E à $\mathrm{d}E$ près est donc simplement le produit de la probabilité (1) par le nombre $\rho(E)\mathrm{d}E$ d'états correspondants.

β. L'approximation gaussienne

Pour un *système macroscopique*[16], la densité d'états $\rho(E)$ est une fonction très rapidement croissante de E (cf. I, § III.B.3); dans la formule (I.23), cette fonction est multipliée par une exponentielle décroissante. Ceci donne à la courbe représentative de $w^c(E)$ l'allure d'un *pic très prononcé et très étroit,* centré sur la valeur la plus probable E_m de E (fig. 1).

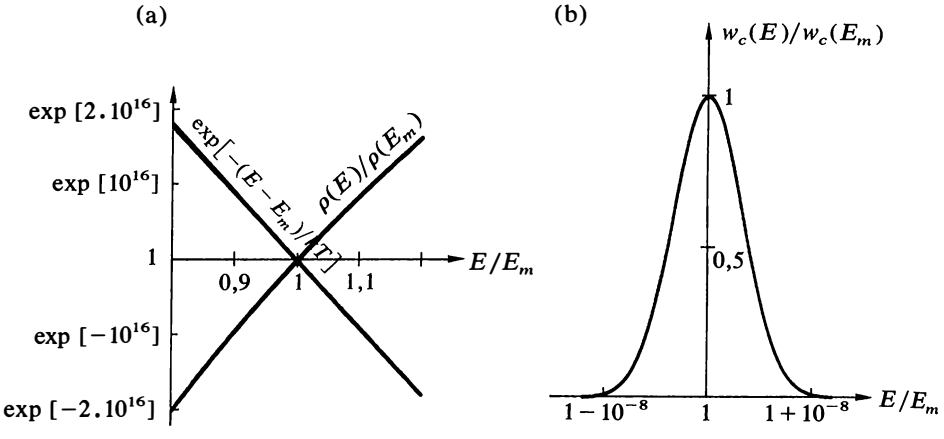


FIGURE 1

Illustration graphique de la manière dont se construit la distribution statistique de l'énergie pour un système macroscopique (on a pris l'exemple d'un gaz parfait monoatomique (cf. § IV.B) constitué de $N = 10^{-7}\ N_A$ molécules, ce qui correspond à environ 2 mm³ dans les conditions normales de température et de pression).

La figure 1.a montre les variations des deux facteurs dont la densité de probabilité $w^c(E)$ est le produit (formule (I.23)). Ces variations sont *fantastiquement rapides :* chacun des deux facteurs passe d'environ $\exp[-2\times10^{16}]$ à $\exp[+2\times10^{16}]$ lorsque l'énergie E parcourt un intervalle de seulement $\pm 20\,\%$ autour de sa valeur la plus probable E_m.

La figure 1.b représente le produit $w^c(E)$ de ces deux fonctions (lui aussi normalisé à 1 en $E = E_m$) : pour que la largeur de la courbe soit visible, il a fallu *dilater l'échelle des abscisses par un facteur* 10^7!

Pour déterminer cette valeur E_m, prenons d'abord le logarithme de $w^c(E)$[17] :

$$\ln w^c(E) = \ln \rho(E) - \frac{E}{kT} - \ln Z\,. \tag{I.26}$$

Le premier terme est égal, au facteur de Boltzmann près, à l'entropie microcanonique $S^*(E)$ du système considéré[18] (II, formule (I.8)) :

$$\ln w^c(E) = \frac{1}{k}\,S^*(E) - \frac{E}{kT} - \ln Z\,. \tag{I.27}$$

16. Les arguments de ce paragraphe β ne sont *pas valables* pour un système constitué d'une seule particule pouvant se déplacer dans un volume de dimensions macroscopiques : la densité d'états correspondante croît seulement comme $E^{1/2}$ (I, formule (III.26)).

17. ... Bien que $w^c(E)$ ait les dimensions de l'inverse d'une énergie (cf. II, note 10).

18. Ce système se trouve ici dans la situation canonique, pas dans la situation microcanonique. Mais $S^*(E)$ est l'entropie qu'il aurait s'il était isolé, à l'équilibre, et que son énergie soit fixée à la valeur E.

L'*énergie la plus probable* E_m est la valeur de E qui rend cette expression maximale; elle est donc donnée par[19]

$$\frac{\partial S^*}{\partial E}(E=E_m)-\frac{1}{T}=0\ . \tag{I.28}$$

D'après la définition de la température microcanonique $T^*(E)$ (II, formule (III.2)), cette condition s'écrit simplement

$$T^*(E_m)=T\ . \tag{I.29}$$

Ainsi, pour un système macroscopique en équilibre avec un thermostat, la valeur la plus probable E_m de l'énergie est celle qui lui donnerait, s'il était isolé et à l'équilibre (cf. note 18), une température microcanonique T^* précisément égale à la température T imposée par le thermostat[20],[21].

L'énergie du système fluctue autour de E_m; mais, s'il s'agit d'un système macroscopique, seules les valeurs très voisines de E_m ont une probabilité non négligeable. De façon plus précise[22], la densité de probabilité $w^c(E)$ se réduit pratiquement à une gaussienne centrée sur E_m (fig. 1) :

$$w^c(E)\simeq w^c(E_m)\exp\left[-(E-E_m)^2/2(\Delta E)^2\right], \tag{I.30}$$

---

19. La constante de normalisation Z est bien sûr indépendante de E (cf. formule (I.24)).

20. Cette conclusion n'est qu'une application particulière du résultat général concernant l'équilibre thermique entre deux systèmes macroscopiques (II, § III.A.2.c).

21. La valeur de E_m est macroscopique, c'est-à-dire très grande devant kT (cf. formule (I.18)). Ainsi, dans le « compromis » entre l'exponentielle décroissante de la distribution canonique et la densité d'états très rapidement croissante, seuls subsistent des états microscopiques (ℓ) dont la probabilité individuelle P_ℓ^c est extrêmement faible, parce que leur nombre $\rho(E)\mathrm{d}E$ dans un intervalle d'énergie « infinitésimal » $\mathrm{d}E$ au voisinage de E_m est prodigieusement grand.

22. Pour obtenir la formule approchée (I.30), on développe $\ln w^c(E)$ autour de son maximum :

$$\ln w^c(E)=\ln w^c(E_m)+\frac{1}{2}(E-E_m)^2\left[\frac{\partial^2}{\partial E^2}\ln w^c\right](E_m)+\frac{1}{3!}(E-E_m)^3\left[\frac{\partial^3}{\partial E^3}\ln w^c\right](E_m)+\dots\ . \tag{1}$$

Dans le premier terme

$$\ln w^c(E_m)=\frac{1}{k}S^*(E_m)-\frac{E_m}{kT}-\ln Z\ , \tag{2}$$

chacune des trois contributions est de l'ordre du nombre N de particules du système considéré (en réalité, ces trois contributions se compensent à l'ordre N — voir plus loin formule (III.3) —, laissant une valeur de $\ln w^c(E_m)$ en $\ln N$, puisque $w^c(E_m)\simeq 1/\sqrt{2\pi(\Delta E)^2}$ pour que l'expression (I.30) soit normalisée). Dans les termes suivants du développement (1), les dérivées portent en fait sur la fonction

$$\ln w^c(E)+\ln Z=\frac{1}{k}S^*(E)-\frac{E}{kT} \tag{3}$$

(car $\ln Z$ est indépendant de E), qui est d'ordre N. Ainsi, la dérivée seconde est d'ordre $1/N$ puisque E est d'ordre N comme la fonction (3), la dérivée troisième d'ordre $1/N^2$, ... Donc, pour $|E-E_m|$ de l'ordre de ΔE (donné par (I.33)), le deuxième terme de la formule (1) est d'ordre 1 et doit être conservé, le troisième (d'ordre $1/\sqrt{N}$) et les suivants pouvant être négligés (cf. note 6 du complément I.C).

où la constante ΔE est définie par

$$\frac{1}{(\Delta E)^2} \widehat{=} -\left[\frac{\partial^2}{\partial E^2} \ln w^c\right](E = E_m) . \tag{I.31}$$

L'expression (I.27) donne immédiatement

$$w^c(E_m) = \frac{1}{Z} \exp\left\{\frac{1}{k}\left[S^*(E_m) - \frac{E_m}{T}\right]\right\} \tag{I.32}$$

et

$$\frac{1}{(\Delta E)^2} = -\frac{1}{k}\frac{\partial^2 S^*}{\partial E^2}(E_m) . \tag{I.33}$$

La dérivée seconde $\partial^2 S^*/\partial E^2$ étant négative (II, formule (III.25)), $(\Delta E)^2$ est bien positif; S^* et E croissant proportionnellement au nombre de particules N du système, $(\Delta E)^2$ se comporte aussi comme N quand N devient très grand.

En résumé, *pour un système macroscopique* constitué d'un grand nombre N de particules et placé dans la situation canonique, *la distribution statistique de l'énergie est pratiquement une gaussienne* centrée sur la valeur la plus probable E_m, qui coïncide par conséquent avec la valeur moyenne $\overline{E}^c$, à cette approximation. L'écart quadratique moyen ΔE de cette gaussienne croît seulement comme $\sqrt{N}$, de sorte que

$$\frac{\Delta E}{E_m} \propto \frac{1}{\sqrt{N}} : \tag{I.34}$$

la distribution est *extrêmement étroite* (en valeur relative), et ceci *d'autant plus que N est plus grand.*

2. DISTRIBUTION STATISTIQUE D'UNE VARIABLE INTERNE QUELCONQUE

Soit y une grandeur physique qui se comporte comme une variable interne (II, § I.A.1) lorsque le système auquel elle est associée est placé dans la situation canonique. Désignons par y_ℓ la valeur prise par cette grandeur[23] dans l'état microscopique (ℓ) du système. Il existe en général *plusieurs états* (ℓ), pouvant avoir des énergies E_ℓ différentes, qui donnent la *même valeur* $y_\ell = y$ à la variable interne étudiée; la probabilité $P^c(y)$ de cette valeur y s'obtient en sommant leurs probabilités.

La formule générale s'écrit donc

$$P^c(y) = \sum_{(\ell)\in\mathcal{L}(y)} \frac{1}{Z} e^{-E_\ell/kT} , \tag{I.35}$$

où $\mathcal{L}(y)$ est l'ensemble des états microscopiques (ℓ) défini par

$$(\ell)\in\mathcal{L}(y) \quad \text{ssi} \quad y_\ell = y . \tag{I.36}$$

Cette formule est toujours valable; c'est à elle qu'il faut se référer pour savoir si telle ou telle approximation est justifiée.

Mais, si le système considéré est de taille macroscopique, on peut utiliser des formules simplifiées analogues à celles du paragraphe 1.b. Dans ce cas en effet, la variable interne y devient, comme l'énergie E, pratiquement continue. Soit alors

$$\rho(E, y)\, dE\, dy \tag{I.37}$$

le *nombre d'états dont l'énergie est comprise entre* E *et* $E + dE$ *et pour lesquels la grandeur étudiée a une valeur comprise entre* y *et* $y + dy$ (cf. II, § II.A.2). La probabilité

23. Le raisonnement simple que nous développons ici n'est en toute rigueur valable que si la grandeur considérée est représentée en mécanique quantique par un opérateur qui commute avec l'hamiltonien (cf. I, § III.A.2.a).

canonique de ces états étant la même pour tous puisqu'ils ont même énergie, la densité de probabilité $w^c(y)$ de la variable interne y est donnée par[24]

$$w^c(y)=\frac{1}{Z}\int_{E_0}^{+\infty} \mathrm{d}E\rho(E, y)\mathrm{e}^{-E/kT}. \tag{I.38}$$

Cette distribution est directement normalisée[25]. Dans la plupart des cas, elle se réduit pratiquement à une distribution gaussienne, d'autant plus étroite (en valeur relative) que le système est plus grand (cf. § II.B.2).

II. FONCTION DE PARTITION ET ÉNERGIE LIBRE

On sait déterminer les propriétés d'un système isolé à l'équilibre lorsqu'on connaît le nombre Ω de ses états accessibles, ou son entropie microcanonique S^* qui en est (à un facteur près) le logarithme. Pour un système en équilibre avec un thermostat, ce rôle fondamental est joué par la fonction de partition Z, ou par l'énergie libre F, proportionnelle à $\ln Z$.

A. Définition

1. LA FONCTION DE PARTITION

La fonction de partition Z que nous avons introduite pour normaliser les probabilités P_ℓ^c (formule (I.16)) est une « constante » en ce sens qu'elle est indépendante de (ℓ) (on somme sur tous les états (ℓ)). Elle est par contre *fonction de la température T, et des autres paramètres extérieurs* imposés au système : si l'on change les valeurs de ces paramètres, les énergies E_ℓ des divers états microscopiques (ℓ) sont bien sûr

---

24. La probabilité $w^c(E, y)\,\mathrm{d}E\,\mathrm{d}y$ pour que la variable interne considérée soit comprise entre y et $y+\mathrm{d}y$ *et* que l'énergie soit comprise entre E et $E+\mathrm{d}E$ s'obtient en multipliant la probabilité de chacun des états réalisant cette double condition par leur nombre :

$$w^c(E, y)\mathrm{d}E\,\mathrm{d}y=\rho(E, y)\mathrm{d}E\,\mathrm{d}y\times\frac{1}{Z}\mathrm{e}^{-E/kT}.$$

La probabilité $w^c(y)\mathrm{d}y$ cherchée se déduit de la précédente en sommant sur l'ensemble des valeurs possibles de l'énergie, puisqu'on ne veut pas maintenir la restriction correspondante.

25. La définition de la densité d'états partielle $\rho(E, y)$ implique

$$\int_{-\infty}^{+\infty}\mathrm{d}y\,\rho(E, y)=\rho(E), \tag{1}$$

puisque le nombre total $\rho(E)\,\mathrm{d}E$ d'états dont l'énergie vaut E à $\mathrm{d}E$ près s'obtient en sommant $\rho(E, y)\,\mathrm{d}E\,\mathrm{d}y$ sur l'ensemble des valeurs possibles de y. Si l'on intègre alors l'expression (I.38) par rapport à y, on obtient

$$\int_{-\infty}^{+\infty} w^c(y)\,\mathrm{d}y=\frac{1}{Z}\int_{E_0}^{+\infty}\mathrm{d}E\int_{-\infty}^{+\infty}\mathrm{d}y\,\rho(E, y)\mathrm{e}^{-E/kT}=\frac{1}{Z}\int_{E_0}^{+\infty}\mathrm{d}E\,\rho(E)\mathrm{e}^{-E/kT}, \tag{2}$$

qui vaut bien 1 compte tenu de la formule (I.24).

affectées, et donc à travers elles la valeur de Z. Pour un gaz pur, par exemple, Z sera fonction de la température T, du volume V et du nombre de molécules N. Plus généralement, nous noterons $\{T, x\}$ l'ensemble des paramètres extérieurs[26] :

$$Z(T, x) \widehat{=} \sum_{(\ell)} e^{-E_\ell(x)/kT} . \tag{II.1}$$

La fonction de partition est une *grandeur sans dimension.*

Dans les cas où l'on peut traiter l'énergie comme une variable continue E (cf. § I.C.1.b.α), la fonction de partition du système peut être calculée à partir de la formule (I.24) :

$$Z(T, x) = \int_{E_0}^{+\infty} dE \rho(E, x) e^{-E/kT} , \tag{II.2}$$

où $\rho(E, x)$ est la densité d'états du système et E_0 son niveau fondamental.

Il est parfois plus commode de travailler avec l'inverse de la température plutôt qu'avec T elle-même. On pose pour cela

$$\beta \widehat{=} \frac{1}{kT} \tag{II.3}$$

et l'on considère alors Z comme une fonction de la nouvelle variable β :

$$Z(\beta, x) \widehat{=} \sum_{(\ell)} e^{-\beta E_\ell(x)} . \tag{II.4}$$

Nous verrons que, suivant les cas, c'est l'une ou l'autre des variables T et β qui conduit aux formules les plus simples. Pour passer de l'une à l'autre, il suffit de connaître, outre la définition (II.3), la relation

$$\beta \frac{\partial}{\partial \beta} = -T \frac{\partial}{\partial T} , \tag{II.5}$$

qui s'en déduit immédiatement.

2. L'ÉNERGIE LIBRE

A partir de la fonction de partition Z, on définit l'énergie libre F du système par la relation

$$\boxed{F(T, x) \widehat{=} -kT \ln Z(T, x)} . \tag{II.6}$$

L'énergie libre F a les dimensions de kT, c'est-à-dire d'une énergie.

B. Rôle fondamental : calcul des propriétés du système

La connaissance de la fonction de partition Z ou de l'énergie libre F d'un système en équilibre avec un thermostat permet de déterminer très simplement ses principales caractéristiques.

1. CARACTÉRISTIQUES DE LA DISTRIBUTION STATISTIQUE DE L'ÉNERGIE

a. Énergie moyenne et écart quadratique moyen

Par définition, l'énergie moyenne $\overline{E}^c$ vaut dans la situation canonique (cf. formule (I.15))

$$\begin{aligned} \overline{E}^c &= \sum_{(\ell)} E_\ell P_\ell^c \\ &= \frac{1}{Z} \sum_{(\ell)} E_\ell e^{-E_\ell/kT} . \end{aligned} \tag{II.7}$$

26. Si le système étudié est macroscopique, les paramètres x sont les mêmes que ceux qui ont été introduits au chapitre II (§ I.B.1.a).

Or cette expression s'obtient directement à partir de la fonction de partition[27] :

$$\boxed{\bar{E}^c = -\frac{\partial}{\partial\beta}\ln Z} \quad . \tag{II.8}$$

En dérivant une deuxième fois par rapport à β, on trouve facilement[28] l'écart quadratique moyen ΔE^c :

$$(\Delta E^c)^2 = \frac{\partial^2}{\partial\beta^2}\ln Z \, . \tag{II.9}$$

Il est facile de faire apparaître dans ces formules l'énergie libre F définie par (II.6). Par exemple, l'énergie moyenne peut s'écrire *(«formule de Helmholtz»)*[29]

$$\bar{E}^c = F - T\frac{\partial F}{\partial T} \, . \tag{II.10}$$

b. Capacité calorifique et chaleur spécifique

La valeur moyenne $\bar{E}^c$ de l'énergie d'un système dépend de la température T imposée par le thermostat et des autres paramètres extérieurs x tels que volume,

---

27. En dérivant par rapport à β l'expression (II.4) de la fonction de partition, on trouve en effet

$$\frac{\partial}{\partial\beta} Z = -\sum_{(\ell)} E_\ell \, \mathrm{e}^{-\beta E_\ell} \, ,$$

c'est-à-dire le second membre de (II.7), au facteur $-1/Z$ près.

28. $$\frac{\partial^2}{\partial\beta^2}\ln Z = \frac{\partial}{\partial\beta}\left(\frac{\partial}{\partial\beta}\ln Z\right) = \frac{\partial}{\partial\beta}\left[-\frac{1}{Z}\sum_{(\ell)} E_\ell \, \mathrm{e}^{-\beta E_\ell}\right]$$

$$= \frac{1}{Z^2}\frac{\partial Z}{\partial\beta}\sum_{(\ell)} E_\ell \, \mathrm{e}^{-\beta E_\ell} + \frac{1}{Z}\sum_{(\ell)} E_\ell^2 \, \mathrm{e}^{-\beta E_\ell} \, . \tag{1}$$

Le second terme est par définition la valeur moyenne du carré de l'énergie. Quant au premier, il peut être écrit

$$\left(\frac{1}{Z}\frac{\partial Z}{\partial\beta}\right).\left(\frac{1}{Z}\sum_{(\ell)} E_\ell \, \mathrm{e}^{-\beta E_\ell}\right) = -\bar{E}^c \, . \, \bar{E}^c \tag{2}$$

d'après (II.8). Donc

$$\frac{\partial^2}{\partial\beta^2}\ln Z = -(\bar{E}^c)^2 + \overline{E^2}^c = (\Delta E^c)^2 \, . \tag{3}$$

On peut généraliser ce résultat et montrer par récurrence que

$$(-1)^n \frac{\partial^n}{\partial\beta^n}\ln Z = \overline{(E - \bar{E}^c)^n}^c \, . \tag{4}$$

29. Passons d'abord de la variable β à la variable T : d'après (II.5), on obtient à partir de (II.8)

$$\bar{E}^c = -\frac{\partial}{\partial\beta}\ln Z = +\frac{1}{\beta}\, T\frac{\partial}{\partial T}\ln Z = kT^2\frac{\partial}{\partial T}\ln Z \, . \tag{1}$$

Remplaçons ensuite $\ln Z$ en fonction de F :

$$\bar{E}^c = kT^2\frac{\partial}{\partial T}\left[-\frac{1}{kT}F\right] = -T\frac{\partial F}{\partial T} + F \, . \tag{2}$$

nombre de particules... La dérivée partielle de $\overline{E}^c$ par rapport à T est appelée *capacité calorifique à volume constant* C_V :

$$C_V \widehat{=} \frac{\partial \overline{E}^c}{\partial T} \quad (\text{à } V, N, \ldots \text{ constants}) . \tag{II.11}$$

La capacité calorifique C_V, qui dépend elle-même, *a priori*, de T et x, mesure l'accroissement $\mathrm{d}\overline{E}^c$ de l'énergie emmagasinée en moyenne dans le système lorsqu'on accroît de $\mathrm{d}T$ la température du thermostat, *en maintenant fixes les paramètres extérieurs*[30] x. S'il en est ainsi, l'accroissement de l'énergie du système ne peut être dû qu'à un apport de *chaleur* (cf. II, § III.A.1.b.α), d'où le nom de capacité « calorifique ». Ainsi, C_V caractérise la quantité de chaleur[31] $đQ_V = \mathrm{d}\overline{E}^c$ qu'il faut fournir au système lorsqu'on augmente sa température de $\mathrm{d}T$ sans changer les autres paramètres extérieurs x :

$$đQ_V = C_V \, \mathrm{d}T . \tag{II.12}$$

Pour un système macroscopique, la capacité calorifique C_V est comme $\overline{E}^c$ proportionnelle au nombre de particules[32]. C'est pourquoi l'on définit la *chaleur spécifique molaire*[33] *à volume constant :*

$$c_V \widehat{=} \frac{1}{n} \frac{\partial \overline{E}^c}{\partial T} , \tag{II.13}$$

où $$n \widehat{=} N/N_A \tag{II.14}$$

est le « nombre de moles » du système étudié; la chaleur spécifique c_V, indépendante de la taille du système, est ainsi une caractéristique du corps qui le constitue.

La capacité calorifique à volume constant se calcule simplement à partir de l'énergie libre[34] :

$$C_V = -T \frac{\partial^2 F}{\partial T^2} , \tag{II.15.a}$$

et $$c_V = -\frac{T}{n} \frac{\partial^2 F}{\partial T^2} . \tag{II.15.b}$$

---

30. La dénomination de « capacité calorifique *à volume constant* » mentionne seulement le volume pour des raisons historiques : C_V a été initialement définie et utilisée pour des fluides à nombre de particules fixé.

31. La *différentielle* de la *variable* qu'est ici la température T est notée $\mathrm{d}T$. On écrit de même $\mathrm{d}\overline{E}^c$ pour la *différentielle* de $\overline{E}^c$, qui est une *fonction* de $\{T, x\}$. Par contre, *la quantité de chaleur infinitésimale* $đQ_V$ *n'est pas en général la différentielle d'une fonction;* c'est pourquoi nous la notons de façon différente, avec un d barré. Cette distinction, fondamentale en thermodynamique, est explicitée au complément II.D (§ II.2).

32. L'énergie est une grandeur extensive, la température une grandeur intensive (complément II.A, § II).

33. On utilise aussi parfois la *chaleur spécifique massique à volume constant* $c_V^{(m)}$, définie par

$$c_V^{(m)} \widehat{=} \frac{1}{M} \frac{\partial \overline{E}^c}{\partial T} ,$$

où M est la masse totale du système.

34. Les expressions (II.15) se déduisent immédiatement de la formule de Helmholtz (II.10).

2. DISTRIBUTION D'UNE VARIABLE INTERNE

La formule générale (I.35) donnant la distribution statistique d'une variable interne y peut aussi être écrite en termes de fonctions de partition ou d'énergies libres.

En effet, si l'on pose

$$Z(T, x, y) \widehat{=} \sum_{(\ell)\in\mathcal{L}(y)} e^{-E_\ell/kT}, \tag{II.16}$$

la formule (I.35) prend la forme

$$P^c(T, x; y) = \frac{Z(T, x, y)}{Z(T, x)}. \tag{II.17}$$

Le dénominateur de cette expression est simplement la fonction de partition du système, qui dépend des paramètres extérieurs $\{T, x\}$. Dans la définition (II.16) du numérateur, la somme porte sur les états (ℓ) du système qui donnent à la variable interne étudiée la valeur y; par conséquent, $Z(T, x, y)$ est *la fonction de partition qu'aurait le système si,* outre les paramètres extérieurs T et x, *on fixait la grandeur physique considérée à la valeur* y (c'est-à-dire si l'on soumettait le système à une contrainte supplémentaire faisant de y un nouveau paramètre extérieur au même titre que x). D'ailleurs, lorsque y est libre de fluctuer, l'ensemble des états (ℓ) sur lesquels il faut sommer pour calculer $Z(T, x)$ est la réunion des divers sous-ensembles $\mathcal{L}(y)$ associés à toutes les valeurs possibles de y, de sorte que

$$Z(T, x) = \sum_y Z(T, x, y), \tag{II.18}$$

ce qui assure la normalisation de la distribution (II.17).

On peut écrire (II.17) sous la forme

$$P^c(y) = \exp\left\{-\frac{1}{kT}\left[F(T, x, y) - F(T, x)\right]\right\}, \tag{II.19}$$

où
$$F(T, x, y) \widehat{=} -kT \ln Z(T, x, y) \tag{II.20}$$

est *l'énergie libre du système dans le cas où la grandeur étudiée est fixée à la valeur* y et $F(T, x)$ son énergie libre dans le cas où cette grandeur est libre de fluctuer[35]. Dans un système maintenu à température constante T, la dépendance en y de la distribution statistique d'une variable interne est donc donnée par

$$P^c(y) \propto \exp\left[-\frac{1}{kT} F(T, x, y)\right]. \tag{II.21}$$

Pour un système macroscopique, la grandeur y prend un ensemble pratiquement continu de valeurs, et sa distribution statistique devient essentiellement gaussienne. En effet, la valeur la plus probable y_m de la grandeur est d'après (II.21) celle qui rend $F(y)$ minimum :

$$\frac{\partial F}{\partial y}(T, x, y_m) = 0. \tag{II.22}$$

35. Les formules (II.17) et (II.19) remplacent II (II.2) et II (II.8), qui étaient valables pour un système isolé : Z et F jouent ici le rôle qu'avaient respectivement Ω et S^* au chapitre II.

En développant ensuite $F(y)$ autour de y_m et en négligeant les termes d'ordre supérieur à 2, on remplace la distribution (II.21) par une gaussienne[36] avec laquelle elle coïncide pratiquement :

$$P^c(y) \simeq \frac{1}{\sqrt{2\pi(\Delta y)^2}} \exp\left[-(y-y_m)^2/2(\Delta y)^2\right], \tag{II.23}$$

et dont l'écart quadratique moyen est donné par

$$(\Delta y)^2 = kT \Big/ \frac{\partial^2 F}{\partial y^2}(T, x, y_m). \tag{II.24}$$

On constate en outre[36] que, lorsque le nombre N de particules du système croît, la largeur de cette distribution gaussienne décroît en valeur relative :

$$\frac{\Delta y}{y_m} \sim \frac{1}{\sqrt{N}}. \tag{II.25}$$

3. CALCUL DE L'ENTROPIE CANONIQUE

Si l'on reprend la définition générale de l'entropie statistique (I, formule (II.2)) et qu'on y reporte les probabilités (I.15), on trouve[37] que l'entropie canonique S^c s'exprime simplement à partir de l'énergie moyenne $\overline{E}^c$ et l'énergie libre F :

$$S^c = \frac{1}{T}\left(\overline{E}^c - F\right). \tag{II.26}$$

---

36. Pour la plupart des systèmes que nous aurons à considérer, l'énergie libre F croît comme le nombre N de particules lorsque celui-ci est grand (cf. § II.C). Si donc les valeurs possibles pour y (et en particulier sa valeur la plus probable y_m) se comportent par exemple comme N^α pour N grand, on a

$$\frac{\partial^2 F}{\partial y^2} \sim N^{1-2\alpha}, \tag{1}$$

de sorte que

$$(\Delta y)^2 \sim N^{2\alpha-1} \tag{2}$$

et

$$\frac{\Delta y}{y_m} \sim N^{-1/2}. \tag{3}$$

On montre ainsi facilement (cf. II, note 16) que les termes d'ordre supérieur à 2 dans le développement de F donnent une contribution négligeable à (II.21).

Le résultat (3) cesse cependant d'être valable aux *«points critiques»*, s'il en existe, où $\partial^2 F/\partial y^2$ s'annule : en ces points, les fluctuations de la grandeur y autour de sa valeur la plus probable jouent un rôle important, même pour N grand (c'est le cas pour certains paramètres y lors de transitions dites du «second ordre» entre deux phases d'un même corps; voir complément V.F, § II.1.c).

37. $$S^c = -k \sum_{(\ell)} P^c_\ell \ln P^c_\ell = -k \sum_{(\ell)} \frac{1}{Z} e^{-E_\ell/kT}\left[-\frac{E_\ell}{kT} - \ln Z\right]. \tag{1}$$

S^c apparaît donc comme la somme de deux termes, dont le premier s'écrit

$$\frac{1}{T}\sum_{(\ell)} \frac{1}{Z} e^{-E_\ell/kT}.E_\ell = \frac{1}{T}\overline{E}^c, \tag{2}$$

et le second

$$k \ln Z . \sum_{(\ell)} \frac{1}{Z} e^{-E_\ell/kT} = k \ln Z = -\frac{1}{T} F \tag{3}$$

(on a utilisé la normalisation des probabilités P^c_ℓ et la définition (II.6) de l'énergie libre F).

Nous retiendrons cette relation importante sous la forme

$$F = \overline{E}^c - TS^c \,. \tag{II.27}$$

L'entropie canonique S^c peut aussi se calculer directement à partir de l'énergie libre[38], [39] :

$$S^c = -\frac{\partial F}{\partial T} \tag{II.28}$$

(la dérivée partielle indique que les paramètres extérieurs x sont maintenus constants).

4. PRESSION ET POTENTIEL CHIMIQUE CANONIQUES

La formule précédente fait intervenir la dérivée de l'énergie libre par rapport à la première de ses variables, T. Ses dérivées par rapport aux autres paramètres extérieurs x correspondent également à des grandeurs physiques importantes.

Par exemple, l'énergie libre d'un *système macroscopique* est en général fonction de son volume V et de son nombre de particules N :

$$F = F(T,\ V,\ N,\ \ldots) \,. \tag{II.29}$$

On définit alors la *pression canonique* p^c et le *potentiel chimique canonique* μ^c par les relations

$$p^c \mathrel{\widehat{=}} -\frac{\partial F}{\partial V} \,, \tag{II.30.a}$$

$$\mu^c \mathrel{\widehat{=}} +\frac{\partial F}{\partial N} \,. \tag{II.30.b}$$

Évidemment, le fait que nous appelions ces grandeurs « pression » et « potentiel chimique » sous-entend qu'elles sont reliées de quelque manière aux grandeurs de même nom p^* et μ^* définies au chapitre II (§ III.B.1); nous préciserons cette relation aux paragraphes D.3.a et III.B.2 ci-dessous.

C. Système constitué de sous-systèmes macroscopiques indépendants

Lorsqu'un système macroscopique isolé est constitué par la réunion de deux (ou plusieurs) sous-systèmes macroscopiques indépendants, le nombre de ses états accessibles est le produit des nombres d'états accessibles des sous-systèmes

---

38. L'énergie moyenne $\overline{E}^c$ a déjà été exprimée à partir de la fonction de partition par la formule (II.8); en passant à la variable T, on transforme cette expression en (note 29, formule (1))

$$\overline{E}^c = kT^2 \frac{\partial}{\partial T} \ln Z \,. \tag{1}$$

La relation (II.26) donnant S^c devient alors

$$S^c = kT \frac{\partial}{\partial T} \ln Z + k \ln Z \tag{2}$$

(on a utilisé la définition (II.6) de F), c'est-à-dire

$$S^c = \frac{\partial}{\partial T} (kT \ln Z) \,, \tag{3}$$

où l'on reconnaît à nouveau F.

39. En rapprochant la formule (II.28) de l'expression (II.15.a) de la capacité calorifique à volume constant, on voit que cette dernière est également reliée de façon simple à l'entropie canonique S^c :

$$C_V = T \frac{\partial S^c}{\partial T} \,.$$

(II, § I.C.2); corrélativement, son entropie microcanonique est la somme des entropies microcanoniques des sous-systèmes. Si le système est maintenu en contact avec un thermostat, des propriétés analogues sont valables pour la fonction de partition et l'énergie libre.

Considérons (fig. 2) un système macroscopique $\mathcal{S}$ en contact avec un thermostat, et constitué de deux sous-systèmes $\mathcal{S}_1$ et $\mathcal{S}_2$ eux-mêmes macroscopiques, que nous situons pour simplifier de part et d'autre d'une paroi commune; cette cloison est ici fixe (pas d'échange de volume) et imperméable aux particules; il est sans importance qu'elle soit diatherme ou adiabatique, puisque $\mathcal{S}_1$ et $\mathcal{S}_2$ sont de toute façon maintenus à la même température T par le thermostat.

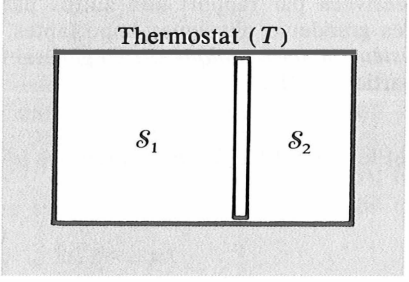


FIGURE 2
Système $\mathcal{S}$ divisé en deux sous-systèmes macroscopiques $\mathcal{S}_1$ et $\mathcal{S}_2$, et maintenu à la température T par un thermostat. Si la cloison est fixe et imperméable aux particules, les deux sous-systèmes sont indépendants.

Si l'on note (ℓ_1) les états microscopiques du sous-système $\mathcal{S}_1$ et (ℓ_2) ceux de $\mathcal{S}_2$, la définition du contact thermique (II, § III.A.1.b) permet d'écrire les états microscopiques (ℓ) du système $\mathcal{S}$ sous la forme

$$(\ell)=(\ell_1, \ell_2)\,, \tag{II.31}$$

l'énergie correspondante étant simplement la somme des énergies associées à (ℓ_1) et (ℓ_2) :

$$E_\ell = E_{\ell_1} + E_{\ell_2}\,. \tag{II.32}$$

Il en résulte[40] que *la fonction de partition du système $\mathcal{S}$ se factorise :*

$$Z = Z_1 . Z_2\,, \tag{II.33}$$

où

$$Z_i \mathrel{\widehat{=}} \sum_{(\ell_i)} \mathrm{e}^{-\beta E_{\ell_i}}\,; \qquad i=1,2 \tag{II.34}$$

40. La fonction de partition du système $\mathcal{S}$,

$$Z \mathrel{\widehat{=}} \sum_{(\ell)} \mathrm{e}^{-\beta E_\ell}\,, \tag{1}$$

s'écrit ici

$$Z = \sum_{(\ell_1, \ell_2)} \mathrm{e}^{-\beta(E_{\ell_1}+E_{\ell_2})}\,, \tag{2}$$

c'est-à-dire

$$Z = \left[\sum_{(\ell_1)} \mathrm{e}^{-\beta E_{\ell_1}}\right] \times \left[\sum_{(\ell_2)} \mathrm{e}^{-\beta E_{\ell_2}}\right]. \tag{3}$$

est la fonction de partition du sous-système $\mathcal{S}_i$. Par conséquent, *l'énergie libre F est additive :*

$$F = F_1 + F_2 \,, \tag{II.35}$$

où

$$F_i \widehat{=} -kT \ln Z_i \,; \quad i = 1,2 \tag{II.36}$$

est l'énergie libre du sous-système $\mathcal{S}_i$.

Les résultats (II.33) et (II.35) se généralisent sans difficulté au cas où le système $\mathcal{S}$ est constitué de plusieurs sous-systèmes macroscopiques $\mathcal{S}_i$ indépendants. Ils permettent de calculer les caractéristiques macroscopiques du système $\mathcal{S}$ à partir de celles des sous-systèmes : par exemple, les formules générales (II.8) et (II.28) montrent que *l'énergie moyenne et l'entropie canoniques sont additives* comme l'énergie libre.

Les propriétés précédentes de la fonction de partition et de l'énergie libre sont d'un usage très général, car les relations (II.31) et (II.32) sur lesquelles se fonde leur démonstration supposent simplement que le couplage entre les sous-systèmes est faible (II, § III.A.1.b.β) et le couplage entre systèmes macroscopiques est toujours faible si les forces entre leurs constituants sont à courte portée (cf. complément II.A). On en déduit facilement que, pour de tels systèmes, *l'énergie libre F croît comme le nombre N de particules*[41].

D. Évolution spontanée du système après relâchement d'une contrainte

Nous nous posons ici le même problème qu'au chapitre II, paragraphe II.B, mais cette fois pour un *système maintenu en contact avec un thermostat* de température T, et soumis également à des contraintes qui fixent les valeurs d'un certain nombre de paramètres extérieurs x (par exemple le volume) : ce système $\mathcal{S}$ est initialement à l'équilibre; on relâche alors une (ou plusieurs) des contraintes, en s'arrangeant pour que les transferts d'énergie se limitent aux échanges de chaleur entre le système et le thermostat[42]; peut-on caractériser simplement le « sens » dans lequel évolue spontanément le système $\mathcal{S}$, compte tenu de la présence du thermostat?

1. CALCUL DE L'ENTROPIE DU SYSTÈME GLOBAL

Nous avons vu au chapitre II (§ II.B) qu'un *système isolé* évolue spontanément *dans le « sens » qui accroît son entropie microcanonique.* Il suffit d'appliquer ce résultat au système global réunissant $\mathcal{S}$ et le thermostat $\mathcal{T}$. Calculons donc l'entropie microcanonique de ce système global en fonction des caractéristiques du système $\mathcal{S}$ et de la température T du thermostat $\mathcal{T}$.

Il nous faut pour cela préciser la définition du thermostat donnée au paragraphe I.A.1, en écrivant que la probabilité $P_{\ell, L}$ pour que le système global $\mathcal{S} \cup \mathcal{T}$ se trouve dans l'état (ℓ, L) est donnée par[43]

$$P_{\ell, L} = P^*_L(E_{\text{tot}} - E_\ell) . P_\ell \,, \tag{II.37}$$

---

41. L'énergie libre F, additive comme l'énergie et l'entropie dans le cas de couplage faible, est comme elles une grandeur *extensive* (complément II.A, § II).

42. Le système ne doit échanger d'énergie qu'avec le thermostat, et ils ne doivent échanger que de la chaleur : une augmentation de volume du système, par exemple, doit se faire comme en II, figure 2 (sauf que les parois fixes sont cette fois diathermes et que l'ensemble est plongé dans un thermostat qui fixe la température T).

43. La validité de la formule (II.37) et du raisonnement qui suit, jusqu'au résultat (II.41), n'est pas limitée au cas où le système $\mathcal{S}$, et par conséquent aussi le système global, sont à l'équilibre : elle demeure lorsque P_ℓ, et donc aussi $P_{\ell, L}$, dépendent du temps (cf. IV, § I.C.1). Si l'on se limite aux situations d'équilibre, $P_{\ell, L}$ est la probabilité microcanonique $P^*_{\ell, L}$, et P_ℓ la probabilité canonique P^c_ℓ.

où E_{tot} est l'énergie totale, fixée, du système global et $P_L^*(E_L)$ la distribution microcanonique pour le thermostat. Cette formule résume une double hypothèse physique : d'une part, le système $\mathcal{S}$ et le thermostat $\mathcal{T}$ ne sont couplés que par la conservation de l'énergie, et ce couplage est faible[44]; d'autre part, le thermostat est toujours à l'équilibre microcanonique correspondant à la part $E_{\text{tot}} - E_\ell$ de l'énergie qui lui revient. En outre, la température du thermostat demeure invariable, ce qui s'exprime par la relation (formules (I.14) et (I.6))

$$S_{\mathcal{T}}^*(E_{\text{tot}} - E_\ell) = S_{\mathcal{T}}^*(E_{\text{tot}}) - E_\ell \cdot \frac{1}{T} \cdot \tag{II.38}$$

L'entropie statistique associée à la distribution $\{P_{\ell,L}\}$ vaut par définition (I, formule (II.2))

$$S_{\mathcal{S}\cup\mathcal{T}} \widehat{=} -k \sum_{(\ell,L)} P_{\ell,L} \ln P_{\ell,L} \cdot \tag{II.39}$$

On trouve facilement[45] à partir de (II.37)

$$S_{\mathcal{S}\cup\mathcal{T}} = S + \sum_{(\ell)} P_\ell S_{\mathcal{T}}^*(E_{\text{tot}} - E_\ell) \, , \tag{II.40}$$

où S est l'entropie du système $\mathcal{S}$ associée à la distribution $\{P_\ell\}$. Si l'on tient compte de (II.38), il vient finalement[46]

$$S_{\mathcal{S}\cup\mathcal{T}} = S + S_{\mathcal{T}}^*(E_{\text{tot}}) - \frac{1}{T}\,\overline{E} \, . \tag{II.41}$$

---

44. Définition du contact thermique (II, § III.A.1.b) : l'énergie du système global est simplement la somme des énergies du système $\mathcal{S}$ et du thermostat $\mathcal{T}$, et tout état microscopique possible du système global est caractérisé par la donnée des états microscopiques respectifs de $\mathcal{S}$ et de $\mathcal{T}$.

45. $$S_{\mathcal{S}\cup\mathcal{T}} = -k \sum_{(\ell,L)} P_L^*(E_{\text{tot}} - E_\ell) \cdot P_\ell \cdot [\ln P_L^*(E_{\text{tot}} - E_\ell) + \ln P_\ell]$$

$$= -k \sum_{(\ell)} P_\ell \sum_{(L)} P_L^*(E_{\text{tot}} - E_\ell) \ln P_L^*(E_{\text{tot}} - E_\ell) - k \sum_{(\ell)} P_\ell \ln P_\ell \cdot \sum_{(L)} P_L^*(E_{\text{tot}} - E_\ell) \, . \tag{1}$$

Or $$\sum_{(L)} P_L^*(E_{\text{tot}} - E_\ell) = 1 \, , \tag{2}$$

$$-k \sum_{(L)} P_L^*(E_{\text{tot}} - E_\ell) \ln P_L^*(E_{\text{tot}} - E_\ell) = S_{\mathcal{T}}^*(E_{\text{tot}} - E_\ell) \, , \tag{3}$$

$$-k \sum_{(\ell)} P_\ell \ln P_\ell = S \, . \tag{4}$$

La formule (1) s'écrit donc

$$S_{\mathcal{S}\cup\mathcal{T}} = \sum_{(\ell)} P_\ell S_{\mathcal{T}}^*(E_{\text{tot}} - E_\ell) + S \, . \tag{5}$$

46. Lorsqu'on reporte (II.38) dans (II.40), on doit calculer

$$\sum_{(\ell)} P_\ell \left[S_{\mathcal{T}}^*(E_{\text{tot}}) - \frac{E_\ell}{T} \right] = S_{\mathcal{T}}^*(E_{\text{tot}}) \cdot \sum_{(\ell)} P_\ell - \frac{1}{T} \sum_{(\ell)} P_\ell E_\ell \, .$$

Le résultat (II.41) découle alors de la normalisation des P_ℓ et de la définition de la valeur moyenne $\overline{E}$.

2. ÉVOLUTION DU SYSTÈME $\mathcal{S}$ EN CONTACT AVEC UN THERMOSTAT

Dans une situation d'équilibre, la formule précédente relie l'entropie microcanonique $S^*_{\mathcal{S}\cup\mathcal{T}}$ du système global à l'entropie S^c et l'énergie moyenne $\overline{E}^c$ canoniques du système $\mathcal{S}$; le résultat peut s'écrire, d'après (II.26),

$$S^*_{\mathcal{S}\cup\mathcal{T}} = S^*_{\mathcal{T}}(E_{\text{tot}}) - \frac{1}{T} F \,. \tag{II.42}$$

Lorsque, après relâchement d'une contrainte, le système $\mathcal{S}$ évolue vers un nouvel état d'équilibre, celui-ci correspond à une *augmentation de l'entropie microcanonique $S^*_{\mathcal{S}\cup\mathcal{T}}$ du système global isolé*. Or le premier terme de la formule (II.42) reste inchangé, car E_{tot} ne varie pas (énergie totale du système global, qui reste constamment isolé); la température T est elle aussi fixée, de sorte que l'augmentation de $S^*_{\mathcal{S}\cup\mathcal{T}}$ ne peut venir que d'une *diminution de l'énergie libre F du système $\mathcal{S}$.*

> *L'évolution spontanée d'un système maintenu en contact avec un thermostat s'accompagne d'une diminution de son énergie libre;* le nouvel *état d'équilibre* atteint par le système après relâchement d'une contrainte correspond au *minimum de son énergie libre* compatible avec les contraintes restantes.

3. APPLICATION : ÉCHANGES ISOTHERMES DE VOLUME ET DE PARTICULES

Reprenons les systèmes macroscopiques $\mathcal{S}_1$ et $\mathcal{S}_2$ schématisés sur la figure 2. Supposons qu'ils soient constitués de particules de même nature, et que la cloison qui les sépare soit maintenant mobile et perméable à ces particules. Dans quelles conditions ces échanges de volume et de particules entre deux systèmes maintenus par un thermostat à la même température T (échanges dits « isothermes ») aboutissent-ils à un équilibre macroscopique?

a. Conditions d'équilibre

Le problème est analogue à celui du chapitre II (§ III.B) : dans le système global $\mathcal{S}$ constitué par la réunion de $\mathcal{S}_1$ et $\mathcal{S}_2$, le volume V_1 et le nombre de particules N_1 du système $\mathcal{S}_1$, ainsi que ceux V_2 et N_2 du système $\mathcal{S}_2$, sont devenus des *variables internes;* seulement deux d'entre elles sont en réalité indépendantes, puisque

$$V_1 + V_2 = V \text{ fixé} \,, \tag{II.43.a}$$

$$N_1 + N_2 = N \text{ fixé} \,, \tag{II.43.b}$$

V et N étant le volume et le nombre de particules du système global $\mathcal{S}$.

Le présent problème se distingue cependant de façon fondamentale de celui du chapitre II, par le fait que le système $\mathcal{S}$ n'est pas ici isolé, mais *maintenu à température constante T :* son état d'équilibre macroscopique correspond au minimum de son *énergie libre,* et non pas comme au chapitre II au maximum de son entropie microcanonique. Cette différence essentielle étant comprise, les raisonnements sont tout à fait semblables.

Lorsque le système $\mathcal{S}$ a atteint l'état d'équilibre compatible avec les échanges de volume et de particules entre ses sous-systèmes $\mathcal{S}_1$ et $\mathcal{S}_2$, la probabilité pour que le volume de $\mathcal{S}_1$ ait la valeur V_1 et son nombre de particules la valeur N_1 est donnée, d'après la formule générale (II.21), par

$$P^c(V_1, N_1) \propto \exp\left[-\frac{1}{kT} F(T, V, N; V_1, N_1)\right] , \tag{II.44}$$

où $F(T, V, N; V_1, N_1)$ est l'énergie libre du système $\mathcal{S}$ dans le cas où V_1 et N_1 sont

fixés aux valeurs considérées. Cette énergie libre se calcule facilement en utilisant le résultat (II.35) :

$$F(T, V, N; V_1, N_1)=F_1(T, V_1, N_1)+F_2(T, V_2=V-V_1, N_2=N-N_1)\,, \quad \text{(II.45)}$$

où F_1 et F_2 sont les énergies libres de $\mathcal{S}_1$ et $\mathcal{S}_2$.

Les *valeurs les plus probables* V_{1m} et N_{1m} des variables internes V_1 et N_1 sont donc celles qui rendent F minimum, c'est-à-dire qui vérifient

$$\frac{\partial F}{\partial V_1}(T, V, N; V_{1m}, N_{1m})=0\,, \quad \text{(II.46.a)}$$

$$\frac{\partial F}{\partial N_1}(T, V, N; V_{1m}, N_{1m})=0\,. \quad \text{(II.46.b)}$$

Compte tenu de (II.45) et des définitions (II.30), les conditions d'équilibre (II.46) s'écrivent simplement[47]

$$p_1^c(T, V_{1m}, N_{1m})=p_2^c(T, V_2=V-V_{1m}, N_2=N-N_{1m})\,, \quad \text{(II.47.a)}$$

$$\mu_1^c(T, V_{1m}, N_{1m})=\mu_2^c(T, V_2=V-V_{1m}, N_2=N-N_{1m})\,, \quad \text{(II.47.b)}$$

où p_i^c et μ_i^c sont la pression et le potentiel chimique canoniques du système $\mathcal{S}_i$ $(i=1,2)$. Si l'on connaît les caractéristiques de $\mathcal{S}_1$ et $\mathcal{S}_2$, les égalités (II.47) permettent de déterminer les deux inconnues V_{1m} et N_{1m}. On notera l'analogie avec le résultat II (III.37) trouvé pour un système global isolé.

b. Sens des échanges et stabilité de l'équilibre

Comme au chapitre II (§ III.B.3.b), supposons que V_1 et N_1 soient initialement différents de V_{1m} et N_{1m}, quoique suffisamment proches pour que

$$\delta V_1 \mathrel{\widehat{=}} V_{1m}-V_1 \quad \text{(II.48.a)}$$

$$\delta N_1 \mathrel{\widehat{=}} N_{1m}-N_1 \quad \text{(II.48.b)}$$

puissent être traités au premier ordre. Le fait que V_{1m} et N_{1m} rendent l'énergie libre globale F *minimale* implique que

$$\delta F \mathrel{\widehat{=}} \frac{\partial F}{\partial V_1}\delta V_1+\frac{\partial F}{\partial N_1}\delta N_1 \quad \text{(II.49)}$$

(où les dérivées sont prises pour les valeurs initiales de V_1 et N_1 et sont donc non nulles) doit être *négatif*, c'est-à-dire que

$$-(p_1^c-p_2^c)\delta V_1+(\mu_1^c-\mu_2^c)\delta N_1<0 \quad \text{(II.50)}$$

lorsque p_1^c et p_2^c sont différentes, ainsi que μ_1^c et μ_2^c (situation hors d'équilibre).

On en déduit comme au chapitre II le sens des échanges hors d'équilibre : si les deux systèmes échangent seulement du volume (cloison mobile mais imperméable aux particules), *celui dont la pression est la plus forte augmente de volume;* s'ils échangent seulement des *particules* (à travers une cloison fixe), celles-ci passent majoritairement *de celui qui a le potentiel chimique le plus élevé vers l'autre*[48].

---

47. Si les deux systèmes échangent du volume mais pas des particules, seul V_1 est une variable interne (N_1 et N_2 étant fixés) et la condition d'équilibre se réduit à

$$p_1^c(T, V_{1m}, N_1)=p_2^c(T, V_2=V-V_{1m}, N_2)\,. \quad (1)$$

Elle s'écrit de même

$$\mu_1^c(T, V_1, N_{1m})=\mu_2^c(T, V_2, N_2=N-N_{1m}) \quad (2)$$

s'il y a échange de particules mais que les volumes des deux systèmes restent fixés à V_1 et V_2.

48. Dans le cas d'échanges simultanés de volume et de particules, l'interprétation de l'inégalité (II.50) est moins simple, car il peut y avoir compétition entre les deux tendances que nous venons d'expliciter.

Évidemment, *l'équilibre* finalement atteint par le système global *doit être stable.* On pourrait alors reprendre le raisonnement qui nous a conduits, pour un système isolé, au résultat (III.25) du chapitre II. Nous nous contenterons ici d'arguments plus intuitifs analogues à ceux de la note 41 du chapitre II.

Supposons d'abord que les deux systèmes $\mathcal{S}_1$ et $\mathcal{S}_2$ échangent des particules mais pas de volume (cloison fixe); la condition d'équilibre est alors

$$\mu_1^c(T,\ V_1,\ N_{1m}) = \mu_2^c(T,\ V_2,\ N_2 = N - N_{1m})\ , \tag{II.51}$$

où V_1 et V_2 sont fixés. Imaginons qu'une fluctuation de la variable interne N_1 autour de sa valeur la plus probable N_{1m} l'amène par exemple à une valeur inférieure à N_{1m}. L'équilibre sera stable si les échanges de particules tendent à ramener N_1 vers N_{1m}, et non pas à l'en écarter davantage; d'après le résultat précédent, ceci ne se produira que *si la fluctuation considérée a diminué* μ_1^c *en diminuant* N_1. Cette conclusion peut être généralisée[49] : *pour un système macroscopique stable maintenu à température et volume constants, le potentiel chimique est une fonction croissante du nombre de particules :*

$$\frac{\partial \mu^c}{\partial N} > 0 \text{ à } T,\ V,\ \ldots \text{ constants}\ . \tag{II.52}$$

Cette condition de stabilité peut être exprimée directement sur l'énergie libre F, puisqu'elle s'écrit, compte tenu de la définition (II.30.b),

$$\frac{\partial^2 F}{\partial N^2} > 0\ . \tag{II.53}$$

En analysant la stabilité de l'équilibre dans le cas d'échanges de volume sans échange de particules, on aboutirait de même à la conclusion

$$\frac{\partial p^c}{\partial V} < 0 \text{ à } T,\ N,\ \ldots \text{ constants}\ : \tag{II.54}$$

pour un système macroscopique stable maintenu à température et nombre de particules constants, la pression est une fonction décroissante du volume[50]. D'après la définition (II.30.a), l'inégalité (II.54) s'écrit aussi[51]

$$\frac{\partial^2 F}{\partial V^2} > 0\ . \tag{II.55}$$

---

49. Il suffit de reprendre, en l'adaptant, le raisonnement du chapitre II (§ III.A.2.c.β). Dans le cas de systèmes plus complexes qu'un fluide pur, non seulement le volume mais tous les autres paramètres extérieurs x, à l'exception du nombre de particules, doivent être maintenus constants.

50. Pour un système ayant atteint la limite thermodynamique (voir paragraphe III), on définit le *coefficient de compressibilité isotherme* χ_T par

$$\chi_T \widehat{=} -\frac{1}{V}\left(\frac{\partial V}{\partial p}\right)_T\ . \tag{1}$$

Ce coefficient mesure la diminution relative de volume $-\mathrm{d}V/V$ associée à une augmentation $\mathrm{d}p$ de la pression, à température constante. La relation (II.54) devient alors

$$\chi_T > 0\ . \tag{2}$$

51. La positivité des dérivées partielles secondes de F par rapport à chacun des paramètres extérieurs x (formules (II.53) et (II.55)) est simplement l'expression analytique d'un des résultats obtenus au paragraphe 2 : pour un système maintenu à température constante, l'énergie libre est *minimum* à l'équilibre.

III. LE FORMALISME CANONIQUE A LA LIMITE THERMODYNAMIQUE

A. La notion de limite thermodynamique

Nous avons distingué jusqu'ici, parmi les grandeurs physiques associées à un système macroscopique placé dans une situation donnée, les *«paramètres extérieurs»* des *«variables internes»* : les premiers ont une valeur fixée de façon certaine par les conditions imposées au système, les dernières sont au contraire libres de fluctuer et doivent donc être affectées d'une distribution statistique. Mais nous avons également insisté à plusieurs reprises sur le fait que, dans un système macroscopique constitué d'un grand nombre N de particules, la distribution statistique d'une variable interne est de façon générale *gaussienne :* la valeur moyenne de la grandeur est égale à sa valeur la plus probable, et *l'écart quadratique moyen est petit en valeur relative,* puisque son rapport à la valeur moyenne est de l'ordre de $1/\sqrt{N}$. Ceci est vrai quelle que soit la situation imposée au système, qu'il soit par exemple isolé ou qu'il soit en équilibre thermique avec un thermostat, pourvu que sa taille soit suffisante. La distribution approchée que nous avons trouvée plus haut (§ I.C.1.b.β) pour l'énergie d'un système macroscopique placé dans la situation canonique illustre bien ce comportement général : $\overline{E^c} = E_m$ croît comme N, et l'écart quadratique moyen seulement comme $\sqrt{N}$.

Nous dirons qu'un système macroscopique a atteint la *«limite thermodynamique»* lorsque sa taille est suffisamment grande pour que *les fluctuations* des diverses variables internes *soient négligeables*[52],[53]. S'il en est ainsi, *chaque variable interne prend pratiquement une valeur unique,* qui est à la fois sa valeur la

---

52. Cette condition est réalisée en pratique pour la plupart des systèmes macroscopiques courants : il suffit que N soit de l'ordre du millionième du nombre d'Avogadro N_A pour que $1/\sqrt{N}$ soit de l'ordre de 10^{-9}, et les mesures macroscopiques usuelles (volume, énergie, pression, température, ...) sont très loin d'atteindre une telle précision relative.

Ceci ne signifie pas pourtant que les fluctuations accompagnant l'agitation thermique microscopique sont impossibles à mettre en évidence dans de tels systèmes : certaines expériences, très sensibles et bien choisies, permettent d'y avoir accès (voir l'exemple du mouvement brownien dans le complément I.D).

53. La véritable définition est plus précise et plus restrictive : un système possède une limite thermodynamique si

$$\lim_{\substack{N\to\infty \\ \frac{V}{N}\text{ fixé}}} \frac{1}{N} F(T, V, N, \ldots) = f\left(T, \frac{V}{N}, \ldots\right), \tag{1}$$

où f est non identiquement nulle et finie.

Les systèmes à couplage macroscopique faible (complément II.A) atteignent une limite thermodynamique (cf. § II.C) dans laquelle toutes les grandeurs intensives gardent une valeur finie. Mais il pourrait se faire que, dans d'autres types de systèmes, l'énergie libre croisse, lorsque $N \to \infty$, comme N^γ avec $\gamma > 1$, ou comme $N \ln N$, ...; il n'y aurait pas alors de limite thermodynamique au sens strict, bien que les fluctuations deviennent négligeables pour N grand (en reprenant le raisonnement de la note 36, on voit que $\Delta y / y_m$ tend toujours vers zéro quand $N \to \infty$, pourvu que F croisse avec N).

C'est l'une des tâches théoriques importantes de la mécanique statistique que d'expliciter les conditions qui rendent possible la véritable limite thermodynamique (1). Nous supposerons toujours dans la suite que les systèmes que nous aurons à considérer admettent une telle limite (voir cependant le complément VI.B, § II).

plus probable et sa valeur moyenne, puisque les écarts par rapport à celle-ci sont inappréciables. La différence de traitement statistique entre paramètres extérieurs certains et variables internes aléatoires disparaît donc dans les systèmes ayant atteint la limite thermodynamique[54], que nous appellerons *« systèmes thermodynamiques »*.

B. Limite thermodynamique de la description canonique

1. DESCRIPTION CANONIQUE D'UN GRAND SYSTÈME

Dans un système macroscopique en contact avec un thermostat, la distribution statistique de l'énergie est pratiquement une gaussienne (§ I.C.1.b), comme d'ailleurs celle de toute autre variable interne. Dans le cadre de cette approximation gaussienne, la fonction de partition $Z(T)$ du système peut être calculée à partir de la formule (I.24), c'est-à-dire de la condition de normalisation de la distribution (I.23) de l'énergie :

$$\int_{E_0}^{+\infty} w^c(E)\, \mathrm{d}E = 1 \ . \tag{III.1}$$

En utilisant pour $w^c(E)$ l'approximation gaussienne (I.30), complétée par l'expression (I.32) de $w^c(E_m)$, on voit que la condition (III.1) implique

$$Z(T) \simeq \mathrm{e}^{\frac{1}{k}[S^*(E_m) - E_m/T]} . \int_{E_0}^{+\infty} \mathrm{d}E\ \mathrm{e}^{-(E-E_m)^2/2(\Delta E)^2} , \tag{III.2}$$

c'est-à-dire finalement[55],[56]

$$Z(T) \simeq \sqrt{2\pi(\Delta E)^2}\ \exp\left\{\frac{1}{k}\left[S^*(E_m) - \frac{E_m}{T}\right]\right\} . \tag{III.3}$$

A partir de la fonction de partition, on trouve aussitôt l'énergie libre :

$$F(T) = -kT \ln Z \simeq -TS^*(E_m) + E_m - \frac{kT}{2} \ln\left[2\pi(\Delta E)^2\right] , \tag{III.4}$$

et l'entropie canonique (formule (II.26)) :

$$S^c(T) = \frac{\overline{E}^c}{T} + k \ln Z \simeq S^*(E_m) + \frac{k}{2} \ln\left[2\pi(\Delta E)^2\right] , \tag{III.5}$$

puisque, à l'approximation gaussienne, la valeur moyenne $\overline{E}^c$ coïncide avec la valeur la plus probable E_m.

---

54. Ces diverses grandeurs devenues équivalentes ne sont cependant pas toutes indépendantes. Le nombre des *« variables d'état »* indépendantes, dont les autres grandeurs sont des *fonctions*, reste à la limite thermodynamique égal à celui des paramètres extérieurs dans l'une ou l'autre des situations que nous avons étudiées, c'est-à-dire au nombre des paramètres x augmenté de 1 ($\{E, x\}$ dans la situation microcanonique, $\{T, x\}$ dans la situation canonique).

55. L'intégrale de la formule (III.2) s'étend du niveau fondamental E_0 à l'infini. Mais, comme la fonction à intégrer est très fortement piquée autour de E_m, tout se passe comme si les bornes étaient en fait $-\infty$ et $+\infty$, à condition toutefois que E_m soit suffisamment différente de E_0 :

$$E_m - E_0 \gg \Delta E ,$$

c'est-à-dire que la température ne soit pas trop basse. Si cette condition n'était pas vérifiée, il faudrait revenir aux formules exactes donnant $P^c(E_\ell)$ et $Z(T)$ (§ I.C.1.a).

56. L'expression (III.3) semble donner à $Z(T)$ les dimensions de ΔE, c'est-à-dire d'une énergie, alors que la fonction de partition devrait être sans dimensions (§ II.A.1). Cette contradiction apparente provient de la façon « cavalière » dont nous traitons l'incertitude δE sur l'énergie (II, § I.C.1) : dans la formule (III.3), ΔE doit en toute rigueur être remplacé par le rapport sans dimension $\Delta E/\delta E$.

2. PASSAGE A LA LIMITE THERMODYNAMIQUE

L'énergie d'un système placé dans la situation canonique se comporte, à la limite thermodynamique où le nombre N de particules qui le composent devient très grand, comme les autres variables internes : l'écart quadratique moyen ΔE croît seulement comme $\sqrt{N}$ (formule (I.33)) et devient négligeable devant l'énergie la plus probable E_m, qui est d'ordre N; tout se passe donc en pratique comme si l'énergie du système prenait la seule valeur E_m.

Dans la formule (III.4) donnant l'énergie libre de ce système, le dernier terme, dont la contribution principale est en ln N, devient rapidement inappréciable devant les deux premiers, où $S^*(E_m)$ et E_m croissent comme N. A la limite thermodynamique, on a donc

$$F(T) = E_m - TS^*(E_m) \text{ (lim. thermo.).} \tag{III.6}$$

De même, la formule (III.5) donne

$$S^c(T) = S^*(E_m) \text{ (lim. thermo.).} \tag{III.7}$$

Pour ce qui est de la température, nous savons déjà (formule (I.29)) que l'équilibre thermique entre le système et le thermostat implique

$$T = T^*(E_m)\ . \tag{III.8}$$

Il est également facile de montrer[57] que

$$p^c(T) = p^*(E_m) \text{ (lim. thermo.),} \tag{III.9.a}$$

$$\mu^c(T) = \mu^*(E_m) \text{ (lim. thermo.),} \tag{III.9.b}$$

où la pression et le potentiel chimique microcanoniques p^* et μ^* sont définis par les formules (III.29) du chapitre II.

C. Équivalence avec la description microcanonique dans la limite thermodynamique

Considérons un système thermodynamique *à l'équilibre.* Son état macroscopique est caractérisé par l'ensemble des paramètres extérieurs :

$\{E, x\}$ s'il est *isolé,*

$\{T, x\}$ s'il est *en contact avec un thermostat.*

---

57. La dépendance par rapport au volume V et au nombre de particules N est sous-entendue dans les formules de ce paragraphe. Par exemple, l'expression (III.6) de l'énergie libre $F(T, V, N)$ s'écrit en réalité

$$F(T, V, N) = E_m(T, V, N) - TS^*\big(E_m(T, V, N), V, N\big) \qquad \text{(lim. thermo.)}\ . \tag{1}$$

En effet, si l'on fixe T, V et N, l'énergie la plus probable E_m est déterminée par l'égalité (III.8), soit

$$T = T^*(E_m, V, N)\ , \tag{2}$$

et dépend donc de T, V et N. Il est donc nécessaire d'utiliser ici, pour les dérivées partielles, la notation où l'on précise quelles variables sont maintenues constantes.

.../...

Dans chacun de ces deux cas, ses autres propriétés (entropie, pression, valeur des variables internes, ...) sont des fonctions des paramètres extérieurs.

Prenons d'abord ce système dans la première situation (microcanonique) : son énergie vaut E (les autres paramètres extérieurs x étant fixés une fois pour toutes), sa température $T^*(E)$. Mettons-le en contact avec un thermostat dont la température T est précisément égale à la sienne :

$$T = T^*(E) \,. \tag{III.10}$$

L'équilibre du système thermodynamique n'est pas modifié : *toutes les grandeurs gardent dans la situation canonique les mêmes valeurs que dans la situation microcanonique.*

En effet, l'énergie du système reste inchangée d'après (III.8) :

$$E_m(T) = \overline{E}^c(T) = E \text{ (lim. thermo.)}. \tag{III.11}$$

Il en est de même de son entropie (formule (III.7)), de sa pression et de son potentiel chimique (formule (III.9)) :

$$S^c(T) = S^*(E) \mathrel{\widehat{=}} S \text{ (lim. thermo.)}, \tag{III.12.a}$$

$$p^c(T) = p^*(E) \mathrel{\widehat{=}} p \text{ (lim. thermo.)}, \tag{III.12.b}$$

$$\mu^c(T) = \mu^*(E) \mathrel{\widehat{=}} \mu \text{ (lim. thermo.)}. \tag{III.12.c}$$

Les définitions (II.30) s'écrivent ainsi

$$p^c \mathrel{\widehat{=}} -\left(\frac{\partial F}{\partial V}\right)_{T,N}, \tag{3.a}$$

$$\mu^c \mathrel{\widehat{=}} +\left(\frac{\partial F}{\partial N}\right)_{T,V}. \tag{3.b}$$

La formule (1) donne alors

$$p^c = -\left(\frac{\partial E_m}{\partial V}\right)_{T,N} + T\left[\left(\frac{\partial S^*}{\partial V}\right)_{E,N} + \left(\frac{\partial S^*}{\partial E}\right)_{V,N} \cdot \left(\frac{\partial E_m}{\partial V}\right)_{T,N}\right] \qquad \text{(lim. thermo.)} \tag{4}$$

et une expression analogue pour μ^c. On reconnaît dans le second terme du crochet la définition de la température microcanonique, qui est égale à T pour $E = E_m$:

$$\left(\frac{\partial S^*}{\partial E}\right)_{V,N}(E_m, V, N) = \frac{1}{T^*(E_m, V, N)} = \frac{1}{T}, \tag{5}$$

de sorte que

$$p^c(T, V, N) = T\left(\frac{\partial S^*}{\partial V}\right)_{E,N}(E_m, V, N) \qquad \text{(lim. thermo.)} \,. \tag{6}$$

Comme la pression microcanonique p^* a été définie par (formule II (III.29.a))

$$\frac{p^*}{T^*} \mathrel{\widehat{=}} \left(\frac{\partial S^*}{\partial V}\right)_{E,N}, \tag{7}$$

l'identification des températures implique celle des pressions :

$$p^c(T, V, N) = p^*(E_m, V, N) \qquad \text{(lim. thermo.)} \,. \tag{8.a}$$

Un raisonnement semblable donnerait également

$$\mu^c(T, V, N) = \mu^*(E_m, V, N) \qquad \text{(lim. thermo.)} \,. \tag{8.b}$$

Ceci permet de définir l'*entropie thermodynamique* S, la *pression thermodynamique* p et le *potentiel chimique thermodynamique* μ. La formule (III.6) fournit de façon analogue la définition de l'*énergie libre thermodynamique*[58] :

$$F \widehat{=} E - TS \,. \tag{III.13}$$

Quant aux grandeurs physiques macroscopiques qui sont des variables internes dans la situation canonique comme dans la situation microcanonique, leur valeur la plus probable reste inchangée, à la limite thermodynamique, lorsqu'on passe de l'une à l'autre situation (en maintenant la relation (III.10))[59], et devient donc leur valeur thermodynamique.

Les diverses *grandeurs thermodynamiques* ainsi définies peuvent être considérées comme des fonctions soit de l'ensemble de variables $\{E, x\}$ caractéristique de la situation microcanonique, soit de l'ensemble de variables $\{T, x\}$ associé à la situation canonique. Le passage d'un de ces ensembles de variables à l'autre est régi par les relations (III.10) et (III.11).

Ainsi, il est impossible, en examinant un système thermodynamique à l'équilibre, de déterminer à partir de ses seules propriétés s'il se trouve dans la situation microcanonique, son énergie E étant fixée, ou dans la situation canonique, maintenu à la température T (T et E étant reliées par (III.10)-(III.11)) : *à la limite thermodynamique, les descriptions canonique et microcanonique d'un système à l'équilibre sont équivalentes.*

Mais *attention :* les points de vue microcanonique et canonique ne sont équivalents que pour la description d'un *état d'équilibre,* après qu'il ait été atteint. On

---

58. La *capacité calorifique à volume constant,* que nous avons introduite dans le cadre canonique (formule (II.11)), devient à la limite thermodynamique

$$C_V = \left(\frac{\partial E}{\partial T}\right)_{V, N, \ldots} \qquad \text{(lim. thermo.)} \,, \tag{1}$$

où nous avons explicité certains des paramètres x; elle peut aussi être calculée à partir de l'énergie libre thermodynamique (cf. formule (II.15.a)) :

$$C_V = -T \left(\frac{\partial^2 F}{\partial T^2}\right)_{V, N, \ldots} \,, \tag{2}$$

ou de l'entropie thermodynamique (voir note 39) :

$$C_V = T \left(\frac{\partial S}{\partial T}\right)_{V, N, \ldots} \,. \tag{3}$$

Bien entendu, la *chaleur spécifique à volume constant* c_V (formule (II.13)) s'obtient en divisant C_V par le nombre de moles $n = N/N_A$.

L'expression (1) fait apparaître C_V comme l'inverse de la dérivée de la température par rapport à l'énergie; or nous avons montré dans le cadre microcanonique que la température est toujours une fonction croissante de l'énergie (II, § III.A.2.c.β); on en déduit

$$C_V > 0 \qquad \text{(lim. thermo.)} \,. \tag{4}$$

59. Considérons une grandeur y qui est une variable interne tant dans la situation canonique que dans la situation microcanonique.

Lorsque le système est maintenu à la température T, la distribution statistique de cette grandeur est donnée par (formule II.21))

$$P^c(y) \propto \exp\left[-\frac{1}{kT} F(T, x, y)\right] , \tag{1}$$

.../...

comprend en effet (cf. chapitre IV) que, si l'on prépare un système dans un état macroscopique hors d'équilibre, son évolution ne sera pas la même selon qu'il est isolé (auquel cas son énergie restera constante) ou qu'il est maintenu à la température T (auquel cas son énergie variera par échange de chaleur avec le thermostat, pour que sa température finale d'équilibre soit égale à T). L'état d'équilibre qu'il atteindra au bout de cette évolution sera également différent dans les deux cas.

$F(T, x, y)$ étant l'énergie libre du système dans le cas où l'on fixe la grandeur considérée à la valeur y. La valeur la plus probable y_m^c s'obtient donc en minimisant $F(T, x, y)$ par rapport à y; en particulier,

$$\left(\frac{\partial F}{\partial y}\right)_{T, x} (T, x, y_m^c) = 0 . \tag{2}$$

Lorsque le système est isolé, la densité de probabilité de la même grandeur s'écrit

$$w^*(E, x, y) \propto \exp\left[\frac{1}{k} S^*(E, x, y)\right] , \tag{3}$$

où $S^*(E, x, y)$, entropie microcanonique partielle associée à la valeur y de la grandeur, coïncide avec l'entropie microcanonique du système calculée dans la situation où la grandeur serait fixée à la valeur y (II, § II.A.1). La valeur la plus probable y_m^* rend ici S^* maximale, et vérifie donc

$$\left(\frac{\partial S^*}{\partial y}\right)_{E, x} (E, x, y_m^*) = 0 . \tag{4}$$

Il est facile de montrer que y_m^c *et* y_m^* *coïncident à la limite thermodynamique.* Dérivons en effet l'expression (III.13) de F par rapport à y, à T et x constants :

$$\left(\frac{\partial F}{\partial y}\right)_{T, x} = \left(\frac{\partial E}{\partial y}\right)_{T, x} - T\left(\frac{\partial S}{\partial y}\right)_{T, x} . \tag{5}$$

Or l'entropie thermodynamique S, si on la considère comme provenant de l'entropie microcanonique S^*, est une fonction de E, x, y; sa dérivée par rapport à y *à* T *constante,* qui apparaît dans la formule (5), s'obtient en tenant compte de la dépendance de E par rapport à T, x, y :

$$S(E, x, y) \equiv S\big(E(T, x, y), x, y\big) , \tag{6}$$

de sorte que

$$\left(\frac{\partial S}{\partial y}\right)_{T, x} = \left(\frac{\partial S}{\partial y}\right)_{E, x} + \left(\frac{\partial S}{\partial E}\right)_{x, y} \cdot \left(\frac{\partial E}{\partial y}\right)_{T, x} . \tag{7}$$

Par définition de la température microcanonique T^*,

$$\left(\frac{\partial S}{\partial E}\right)_{x, y} = \frac{1}{T^*(E, x, y)} . \tag{8}$$

Mais la température est ici maintenue constante et égale à T :

$$T^*(E, x, y) = T . \tag{9}$$

Donc, en reportant (7) dans (5), on trouve finalement, compte tenu de (8)-(9),

$$\left(\frac{\partial F}{\partial y}\right)_{T, x} = -T\left(\frac{\partial S}{\partial y}\right)_{E, x} . \tag{10}$$

La température T n'étant jamais nulle, les équations (2) et (4) ont même solution $y_m^c = y_m^*$ (étant entendu que T et E sont toujours reliées par (9)).

De plus, T étant toujours positive, les deux dérivées premières reliées par (10) sont de signe contraire quand elles ne sont pas nulles; on en déduit aisément qu'à un maximum de S correspond bien un minimum de F.

IV. DESCRIPTION CANONIQUE DES SYSTÈMES DE PARTICULES INDÉPENDANTES

Nous avons montré au paragraphe II.C que la fonction de partition d'un système formé par la réunion de sous-systèmes macroscopiques indépendants est le produit des fonctions de partition de ces sous-systèmes, et son énergie libre la somme de leurs énergies libres. Nous considérons maintenant un système $\mathcal{S}$ constitué de N *sous-systèmes microscopiques indépendants,* que l'on appelle généralement « particules ». Il est bien sûr impossible, si N est grand, que ces particules soient toutes de nature différente; pour simplifier les raisonnements, nous les supposerons toutes identiques[60].

A. Formules générales pour la fonction de partition et l'énergie libre

1. HYPOTHÈSES ET NOTATIONS

Les N particules constituant le système $\mathcal{S}$ sont *indépendantes* si l'on peut *négliger leurs interactions mutuelles.* L'hamiltonien H du système est alors la somme

$$H = H_1 + H_2 + \ldots + H_N \tag{IV.1}$$

de N hamiltoniens individuels, un pour chaque particule; il ne comporte pas de termes d'interaction entre les particules. Soulignons que *cette situation est beaucoup moins fréquente pour des particules que pour des sous-systèmes macroscopiques* (cf. § II.C)[61].

L'identité des N particules implique que les N hamiltoniens individuels ont tous même forme :

$$H_i = H\left(\vec{R}_i, \vec{P}_i, \vec{S}_i, \ldots\right), \tag{IV.2}$$

où $\vec{R}_i$, $\vec{P}_i$, $\vec{S}_i$, ... sont les opérateurs décrivant respectivement la position, l'impulsion, le spin, ... de la particule numérotée (i), la fonction $H\left(\vec{R}, \vec{P}, \vec{S}, \ldots\right)$ étant la même pour toutes les particules[62]. Nous noterons H l'hamiltonien individuel :

$$H \widehat{=} H\left(\vec{R}, \vec{P}, \vec{S}, \ldots\right), \tag{IV.3}$$

---

60. Si le système comporte N_1 particules de type (1), N_2 de type (2), ..., N_r de type (r) (toutes indépendantes), on le considère d'abord comme la réunion de r sous-systèmes indépendants, auxquels on applique les résultats du paragraphe II.C : la fonction de partition du système est le produit des fonctions de partition des r sous-systèmes. Ensuite, chacun de ces sous-systèmes, constitué de particules identiques entre elles, est traité comme nous l'indiquons dans cette partie IV.

61. On peut cependant assez souvent, pour des systèmes dans lesquels les interactions entre particules ne sont pas négligeables, se ramener à un *modèle de particules indépendantes :* on remplace en première approximation l'influence, sur une particule donnée, de l'ensemble des autres par un *« potentiel moyen »* ne dépendant plus que de la position et des caractéristiques de la particule considérée. Des exemples de cette méthode se trouvent dans les compléments III.G à III.K.

62. L'hamiltonien individuel H ne se réduit pas en général à l'énergie cinétique de la particule : il peut comprendre des termes décrivant l'énergie associée à sa structure interne (rotation et vibration des molécules polyatomiques; cf. complément III.B), ou son énergie dans un champ extérieur (électrique, magnétique, gravitationnel), ou le « potentiel moyen » créé par le reste du système (cf. note 61), ... C'est l'absence d'interactions mutuelles, c'est-à-dire de termes du type H_{ij}, qui assure l'indépendance des particules.

.../...

et (λ) un de ses états propres, d'énergie ε_λ. Par suite de l'indépendance des N particules (formule (IV.1)), un état microscopique (ℓ) du système $\mathcal{S}$ (état propre de son hamiltonien H) peut être caractérisé par la donnée de N états individuels, un pour chaque particule :

$$(\ell) \equiv (\lambda_1, \lambda_2, \ldots, \lambda_N) ; \tag{IV.4}$$

l'énergie E_ℓ de l'état (ℓ) (valeur propre correspondante) est la somme des énergies de ces N états individuels :

$$E_\ell = \varepsilon_{\lambda_1} + \varepsilon_{\lambda_2} + \ldots + \varepsilon_{\lambda_N} . \tag{IV.5}$$

Les formules (IV.4) et (IV.5) sont tout à fait semblables à celles, (II.31) et (II.32), qui ont conduit plus haut à la propriété de factorisation de la fonction de partition du système. Mais les *états quantiques des systèmes de particules identiques* doivent vérifier le *postulat de symétrisation* (appendice II, § IV). Celui-ci influe de façon essentielle sur le *décompte des états microscopiques du système.*

Il importe en réalité de distinguer deux cas fondamentalement différents.

2. PARTICULES IDENTIQUES MAIS DISCERNABLES

Dans certains systèmes, les diverses particules occupent des *sites distincts,* c'est-à-dire qu'elles sont localisées dans des régions distinctes de l'espace. L'exemple typique est celui d'un solide cristallin : les atomes, molécules ou ions qui le constituent sont répartis de façon régulière dans l'espace, aux nœuds du réseau caractérisant le cristal; celui-ci offre ainsi autant de places différentes qu'il y a de particules. Dans ces conditions, on peut numéroter les sites : $q = 1, 2, \ldots, N$.

Prenons d'abord un exemple très simple : les particules considérées ont un spin 1/2 et l'on s'intéresse uniquement à leur état de spin. L'indice λ introduit au paragraphe 1 peut alors prendre deux valeurs seulement $(\pm 1/2)$: il y a deux états individuels possibles pour chaque site. Pour caractériser complètement un état microscopique (ℓ) du système, *il faut donc se donner l'état individuel occupé sur chacun des N sites :* un état du système dans lequel l'état individuel $(+1/2)$ est occupé sur le site (q) et l'état $(-1/2)$ sur le site (q') est clairement différent d'un état où l'on aurait $(-1/2)$ sur le site (q) et $(+1/2)$ sur (q'); il existe bien sûr aussi des états du système où les sites (q) et (q') ont tous deux $(+1/2)$ ou tous deux $(-1/2)$. La situation reste fondamentalement la même si l'ensemble des états individuels (λ) offerts aux particules sur chaque site est plus vaste et même infini comme c'est le cas le plus souvent.

On dit pour un tel système que les N particules identiques et indépendantes qui le constituent sont *discernables.* Ce ne sont pas en réalité les particules qui sont discernables mais les *sites,* et donc les états individuels liés à chacun d'eux : les N états individuels (λ_q) caractérisant (formule (IV.4)) un état microscopique (ℓ) du système sont attachés chacun à un site particulier; ainsi, *à chaque état (ℓ) est associée une liste ordonnée de N états individuels,* l'ordre étant fixé par la numérotation des

---

On peut avoir par exemple

$$H(\vec{R}, \vec{P}, \vec{S}) = \frac{\vec{P}^2}{2m} + U(\vec{R}) - \gamma \vec{S} . \vec{B} ,$$

où le dernier terme décrit l'interaction entre un champ magnétique $\vec{B}$ imposé de l'extérieur et le moment magnétique $\gamma\vec{S}$ associé au spin de la particule. Chacun des termes est exactement le même pour toutes les particules, puisqu'elles ont par définition (particules *identiques*) mêmes caractéristiques (masse, charge, rapport gyromagnétique γ, ...).

sites (en particulier, deux listes comportant les mêmes N valeurs de λ, mais différant par l'ordre dans lequel ces valeurs apparaissent, définissent deux états (ℓ) et (ℓ') différents)[63].

Dans ce cas, chacun des (λ_q) de la liste (IV.4) parcourt, indépendamment des autres, l'ensemble des états individuels possibles pour une particule sur un site déterminé. Par conséquent, le calcul de la fonction de partition Z et de l'énergie libre F du système généralise simplement[64] celui du paragraphe II.C :

particules indépendantes	$Z = z^N$,	(IV.6.a)
identiques mais discernables	$F = Nf$,	(IV.6.b)

où

$$z \mathrel{\widehat{=}} \sum_{(\lambda)} e^{-\beta\varepsilon_\lambda} \tag{IV.7.a}$$

est la fonction de partition d'une particule sur un site et

$$f \mathrel{\widehat{=}} -kT \ln z \tag{IV.7.b}$$

son énergie libre. Ainsi, il suffit de déterminer la fonction de partition z d'une seule particule pour pouvoir en déduire aussitôt les diverses propriétés du système à partir des formules générales du paragraphe II.B : on voit par exemple que l'énergie moyenne canonique du système vaut simplement N fois celle d'une particule, ce qui était prévisible pour des particules indépendantes.

---

63. Le postulat de symétrisation ne restreint pas ici le nombre d'états possibles. Pour le montrer sans trop compliquer les formules, limitons-nous à $N=2$. Prenons l'état du système

$$(\ell) = \{(\lambda)^{(1)}, (\lambda')^{(2)}\}, \tag{1}$$

dans lequel la particule occupant le site n° 1 est dans l'état (λ), celle du site n° 2 dans l'état (λ'). Le ket $|(\lambda)^{(1)}, (\lambda')^{(2)}\rangle$ représentant cet état du système doit être convenablement symétrisé ou antisymétrisé par rapport à la numérotation des *particules :*

$$|(\lambda)^{(1)}, (\lambda')^{(2)}\rangle = \frac{1}{\sqrt{2}}\left[|(1):(\lambda)^{(1)}; (2):(\lambda')^{(2)}\rangle \pm |(1):(\lambda')^{(2)}; (2):(\lambda)^{(1)}\rangle\right], \tag{2}$$

la notation $|(i):(\lambda)^{(q)}\rangle$ indiquant que la particule n° i se trouve sur le site (q) dans l'état (λ) (le signe est + si les particules sont des bosons, − si ce sont des fermions). On voit sans difficulté que cette symétrisation n'affecte pas le fait que

$$(\ell') = \{(\lambda')^{(1)}, (\lambda)^{(2)}\} \tag{3}$$

est différent de (ℓ) si $(\lambda') \neq (\lambda)$.

64. La fonction de partition du système

$$Z = \sum_{(\ell)} e^{-\beta E_\ell} \tag{1}$$

s'écrit en effet ici

$$Z = \sum_{(\lambda_1, \lambda_2, \ldots, \lambda_N)} e^{-\beta(\varepsilon_{\lambda_1} + \varepsilon_{\lambda_2} + \ldots + \varepsilon_{\lambda_N})} = \sum_{(\lambda_1)} e^{-\beta\varepsilon_{\lambda_1}} \cdot \sum_{(\lambda_2)} e^{-\beta\varepsilon_{\lambda_2}} \cdot \ldots \cdot \sum_{(\lambda_N)} e^{-\beta\varepsilon_{\lambda_N}}, \tag{2}$$

puisque les divers (λ_q) peuvent varier indépendamment les uns des autres. Comme chaque somme porte sur le même ensemble d'états individuels, Z est le produit de N facteurs identiques à (IV.7.a).

3. PARTICULES INDISCERNABLES

Si les N particules (indépendantes et identiques) qui constituent le système ne sont pas astreintes à occuper des sites distincts, il n'est pas possible de les différencier les unes des autres et leurs états doivent tous être pris dans *un seul et même ensemble d'états individuels* (λ) *possibles*. On dit alors que les N particules sont *indiscernables*. C'est le cas pour un gaz, par exemple, où chacune des molécules peut se déplacer dans la totalité du volume disponible[65].

a. Caractérisation et décompte des états microscopiques du système

Dans une situation de ce type, on peut encore se donner N états individuels pour définir un état (ℓ) du système (formule (IV.4)), puisque les particules sont indépendantes. Mais on ne peut pas attribuer chaque état individuel à une particule déterminée, puisqu'il est impossible de distinguer les unes des autres des particules identiques; on ne peut pas non plus l'attribuer comme au paragraphe 2 à un site déterminé. Le postulat de symétrisation joue donc à plein : un état microscopique (ℓ) du système est caractérisé par la donnée d'une *liste non ordonnée de N états individuels* (λ_i); si les particules étudiées sont des *bosons*[66], il n'y a *pas de restriction* sur la composition de cette liste; pour des *fermions*[66] au contraire, elle ne doit comporter que des *états individuels tous différents (principe de Pauli)*[67].

On peut de façon équivalente caractériser les états du système à l'aide des *nombres d'occupation* N_λ des divers états individuels (λ) (N_λ est le nombre de particules qui se trouvent dans l'état individuel (λ); voir appendice II, § IV.3) : à chaque *état microscopique* (ℓ) du système de N particules correspond alors une *configuration* $\{N_\lambda\}$:

$$(\ell) \iff \{N_\lambda\}, \tag{IV.8}$$

telle que

$$\sum_{(\lambda)} N_\lambda = N . \tag{IV.9}$$

Rappelons qu'ici non plus il n'y a *pas de restriction* autre que (IV.9) sur la valeur des nombres d'occupation pour un système de *bosons;* pour des *fermions* en revanche, chacun des N_λ est limité à *deux valeurs possibles,* 0 et 1. Dans cette description, l'énergie de l'état (ℓ) est donnée par

$$E_\ell = \sum_{(\lambda)} N_\lambda \varepsilon_\lambda . \tag{IV.10}$$

Suivant le problème à résoudre, on pourra choisir l'une ou l'autre de ces deux descriptions équivalentes des états (ℓ) pour un système de particules indépendantes indiscernables.

---

65. *Attention :* des particules de *nature différente* (non identiques) sont *toujours discernables,* même dans des circonstances où aucun moyen n'est mis en œuvre pour les distinguer effectivement.

66. Rappelons que les particules de *spin entier* ($s = 0, 1, \ldots$) *sont des bosons,* celles de spin *demi-entier* ($s = 1/2, 3/2, \ldots$) des *fermions.*

67. Des fermions *indépendants* au sens où nous l'avons défini plus haut (formule (IV.1)) ne sont donc pas vraiment indépendants les uns des autres, puisque le principe de Pauli restreint de façon stricte le choix de leurs états individuels. Quoique moins radical, un effet analogue (mais de sens inverse) existe également dans le cas de bosons : la symétrisation augmente le « poids statistique » des états du système dans lesquels plusieurs particules occupent le même état individuel (cf. formule (IV.13)).

b. Tentative de calcul exact de la fonction de partition

Pour calculer la fonction de partition du système,

$$Z = \sum_{(\ell)} e^{-\beta E_\ell} , \tag{IV.11}$$

nous cherchons comme au paragraphe 2 à utiliser l'indépendance des particules, c'est-à-dire les formules (IV.4) et (IV.5). Cette dernière implique toujours que la fonction à sommer se décompose en un produit de N facteurs individuels :

$$e^{-\beta E_\ell} = e^{-\beta \varepsilon_{\lambda_1}} . e^{-\beta \varepsilon_{\lambda_2}} . \ldots . e^{-\beta \varepsilon_{\lambda_N}} . \tag{IV.12}$$

Mais chaque état (ℓ) possible du système est caractérisé par une liste *non ordonnée* d'états individuels (avec une restriction supplémentaire pour les fermions). Or si nous sommons l'expression (IV.12) sur toutes les valeurs possibles pour λ_1, puis indépendamment sur toutes celles de λ_2, et ainsi de suite jusqu'à λ_N, un ensemble donné de N valeurs (non ordonnées) de λ contribue un grand nombre de fois à cette somme multiple : ces contributions, toutes égales entre elles puisqu'elles correspondent à la même énergie, sont au nombre de[68]

$$N! \Big/ \Big(\prod_{(\lambda)} N_\lambda ! \Big) , \tag{IV.13}$$

où le produit du dénominateur porte sur les états individuels *distincts* de l'ensemble choisi.

Supposons d'abord que les particules étudiées soient des *bosons*. Pour retrouver la fonction de partition (IV.11), il faut diviser chaque terme (IV.12) par le nombre (IV.13) correspondant, avant d'effectuer la somme sur les λ_i :

$$\text{Bosons} : Z = \sum_{(\lambda_1)} \sum_{(\lambda_2)} \ldots \sum_{(\lambda_N)} \frac{\prod_{(\lambda)} N_\lambda !}{N!} e^{-\beta \varepsilon_{\lambda_1}} . e^{-\beta \varepsilon_{\lambda_2}} . \ldots . e^{-\beta \varepsilon_{\lambda_N}} . \tag{IV.14}$$

Comme le facteur correctif varie d'un terme à l'autre, l'expression (IV.14) ne se factorise plus en N sommes individuelles comme au paragraphe 2.

Si les particules considérées sont des *fermions*, le raisonnement est différent : seuls doivent être conservés les termes (IV.12) pour lesquels la liste $(\lambda_1, \lambda_2, \ldots, \lambda_N)$ ne comporte que des états individuels tous distincts (principe de Pauli); pour ces termes, tous les $N_\lambda !$ qui apparaissent dans (IV.13) valent 1, de sorte que

$$\text{Fermions} : Z = \sum_{(\lambda_1)} \sum_{(\lambda_2)} \ldots \sum_{(\lambda_N)} \prod_{(\lambda)} \alpha(N_\lambda) . \frac{1}{N!} e^{-\beta \varepsilon_{\lambda_1}} . e^{-\beta \varepsilon_{\lambda_2}} . \ldots . e^{-\beta \varepsilon_{\lambda_N}} , \tag{IV.15}$$

où le produit sur (λ) a la même signification qu'en (IV.13) et (IV.14), et

$$\begin{aligned} \alpha(N_\lambda) &= 1 \quad \text{si} \quad N_\lambda = 0 \text{ ou } 1 \\ &= 0 \quad \text{si} \quad N_\lambda > 1 . \end{aligned} \tag{IV.16}$$

Ici aussi, la modification apportée à la somme multiple empêche sa factorisation.

68. C'est le nombre de permutations distinctes des N états individuels de l'ensemble non ordonné choisi : le dénominateur de l'expression (IV.13) tient compte du fait que ces N états ne sont pas en général tous différents.

Ainsi, *il n'est pas possible de calculer exactement la fonction de partition d'un système de particules indépendantes indiscernables à partir de la fonction de partition d'une seule particule*[69].

c. L'approximation de Maxwell-Boltzmann

Il est cependant des situations physiques qui permettent d'effectuer *de façon approchée* le calcul de la somme (IV.14) ou (IV.15).

α. Expressions approchées pour la fonction de partition et l'énergie libre

Supposons que les nombres d'occupation N_λ des divers états individuels possibles aient tous une valeur moyenne très petite :

$$\overline{N}_\lambda^c \ll 1 \text{ pour tout } (\lambda). \tag{IV.17}$$

Cette hypothèse ne peut être vérifiée que si, dans la fonction de partition (IV.11) à calculer, les termes (IV.12) pour lesquels un ou plusieurs des N_λ prennent une valeur supérieure à 1 peuvent être négligés.

Mais alors, si les particules considérées sont des bosons, le facteur correctif (IV.13) qui figure dans l'expression (IV.14) de la fonction de partition est le même, $N!$, pour tous les termes non négligeables, de sorte que

$$Z \simeq \frac{1}{N!} \sum_{(\lambda_1)} \sum_{(\lambda_2)} \dots \sum_{(\lambda_N)} e^{-\beta\varepsilon_{\lambda_1}} . e^{-\beta\varepsilon_{\lambda_2}} \dots . e^{-\beta\varepsilon_{\lambda_N}} . \tag{IV.18}$$

Dans le cas de fermions, les termes négligeables dans la situation (IV.17) sont précisément ceux qu'il fallait éliminer de la fonction de partition, en les multipliant par $\alpha(N_\lambda)=0$ dans la somme (IV.15); celle-ci prend donc aussi la forme approchée (IV.18) lorsque la contribution de ces termes est très faible. L'approximation ainsi obtenue, dite «de Maxwell-Boltzmann», donne donc la *même expression* à la fonction de partition Z *pour des bosons et des fermions :* comme la formule (IV.18) fait apparaître le produit de N facteurs identiques, le résultat final s'écrit

particules indépendantes identiques et indiscernables (approximation de Maxwell-Boltzmann)

$$Z = \frac{1}{N!} z^N , \tag{IV.19.a}$$

$$F = Nf + kT \ln N! , \tag{IV.19.b}$$

la fonction de partition z et l'énergie libre f d'une particule étant encore définies comme en (IV.7).

Ici aussi, la connaissance de la fonction de partition z d'une seule particule suffit pour déterminer les propriétés du système grâce aux formules du paragraphe II.B.

---

69. Si l'on caractérise chaque état (ℓ) par la configuration $\{N_\lambda\}$ correspondante (cf. (IV.8)-(IV.10)), la fonction de partition du système s'écrit

$$Z = \sum_{\{N_\lambda\} : \sum_{(\lambda)} N_\lambda = N} \exp\left[-\beta \sum_{(\lambda)} N_\lambda \varepsilon_\lambda\right] .$$

La somme porte sur toutes les configurations compatibles avec le fait que le nombre total de particules est fixé à N (avec la restriction supplémentaire $N_\lambda \leqslant 1$ s'il s'agit de fermions). Cette formule n'est pas non plus calculable exactement : la condition (IV.9) qui lie les divers nombres d'occupation N_λ empêche ici aussi d'exprimer Z comme un produit de sommes indépendantes. Nous verrons au chapitre V comment, en éliminant la condition (IV.9), le formalisme grand-canonique permet de traiter sans approximation les systèmes de particules indépendantes indiscernables.

Notons en particulier que la différence entre les expressions (IV.19) et (IV.6) n'affecte pas la valeur moyenne de l'énergie, qui reste additive comme il se doit pour des particules indépendantes.

β. Domaine de validité de l'approximation

On s'attend intuitivement à ce que l'hypothèse (IV.17), sur laquelle est basée l'approximation de Maxwell-Boltzmann, soit vérifiée lorsqu'il y a « beaucoup » d'états individuels ayant une probabilité appréciable (ce qui est le cas à *température assez haute*) et relativement « peu » de particules pour les occuper (c'est-à-dire à *densité assez faible*). C'est ce que montre en effet le raisonnement quantitatif suivant.

Dans un système de particules indépendantes et indiscernables, la valeur moyenne du nombre d'occupation N_λ d'un état individuel (λ) particulier, d'énergie ε_λ, peut être obtenue par la relation formelle[70]

$$\overline{N}_\lambda^c = -kT \frac{\partial}{\partial \varepsilon_\lambda} (\ln Z) . \tag{IV.20}$$

Si l'on utilise l'expression approchée (IV.19.a) de Z, on obtient[71]

$$\overline{N}_\lambda^c = \frac{N}{z} e^{-\varepsilon_\lambda / kT} . \tag{IV.21}$$

La condition (IV.17) s'explicite alors sous la forme[72]

$$\frac{N}{z} e^{-\varepsilon_\lambda / kT} \ll 1 \text{ pour tout } (\lambda), \tag{IV.22}$$

qui se réduit en fait à

$$z e^{+\varepsilon_0 / kT} \gg N , \tag{IV.23}$$

où ε_0 est le niveau fondamental pour une particule.

---

70. Il est commode ici de caractériser un état (ℓ) du système par la configuration correspondante (formules (IV.8)-(IV.10)). La valeur moyenne de N_λ s'écrit alors

$$\overline{N}_\lambda^c = \frac{1}{Z} \sum_{(\ell)} N_\lambda e^{-\beta E_\ell} = \frac{1}{Z} \sum_{\{N_{\lambda'}\} : \sum_{(\lambda')} N_{\lambda'} = N} N_\lambda \exp \left[-\beta \sum_{(\lambda')} N_{\lambda'} \varepsilon_{\lambda'} \right] , \tag{1}$$

avec

$$Z = \sum_{\{N_{\lambda'}\} : \sum_{(\lambda')} N_{\lambda'} = N} \exp \left[-\beta \sum_{(\lambda')} N_{\lambda'} \varepsilon_{\lambda'} \right] . \tag{2}$$

La formule (IV.20) découle aussitôt des expressions (1) et (2).

71. Dans cette approximation en effet,

$$\ln Z = N \ln z - \ln N! . \tag{1}$$

De plus, la définition (IV.7.a) de z donne immédiatement

$$\frac{\partial}{\partial \varepsilon_\lambda} z = -\beta e^{-\beta \varepsilon_\lambda} . \tag{2}$$

D'où le résultat (IV.21).

72. Il s'agit seulement ici d'examiner les *conditions de cohérence* de l'approximation de Maxwell-Boltzmann : nous cherchons dans quelles situations physiques l'expression approchée de Z vérifie effectivement l'hypothèse (IV.17) qui a permis de l'obtenir. Nous pourrons étudier plus rigoureusement la validité de cette approximation lorsque nous aurons établi les formules exactes régissant les « statistiques quantiques » (V, § IV.C).

Explicitons le premier membre de l'inégalité (IV.23) :

$$z e^{+\varepsilon_0/kT} = \sum_{(\lambda)} e^{-(\varepsilon_\lambda - \varepsilon_0)/kT} . \qquad \text{(IV.24)}$$

Dans cette somme, le ou les premiers termes, ceux qui correspondent au niveau fondamental ε_0, sont constants et égaux à 1; les suivants sont inférieurs à 1, et ce d'autant plus que la température est plus basse. Pour que la condition (IV.23) puisse être satisfaite, il faut donc en premier lieu que le nombre de termes dans (IV.24) soit beaucoup plus grand que N; en pratique, ceci signifie que *l'ensemble des états individuels possibles doit être infini.* D'autre part, il faut aussi que la température soit assez haute pour que la somme (IV.24) comporte suffisamment de termes non négligeables[73]. De façon plus précise, le premier membre de (IV.23) est indépendant de N (puisque z est la fonction de partition d'une seule particule); il croît avec la température T, et dépend également du volume V et d'autres paramètres extérieurs éventuels à travers les énergies ε_λ. Donc, *à température et volume fixés,* la condition de validité de l'approximation de Maxwell-Boltzmann est satisfaite lorsque *le nombre de particules n'est pas trop grand;* inversement, *à volume et nombre de particules fixés,* il faut que *la température soit assez élevée.*

Ce que signifient quantitativement ces conditions (densité de particules assez faible et température assez élevée) dépend du système étudié, et plus précisément de la façon dont la fonction de partition individuelle z varie avec la température et le volume. Nous allons en voir ci-dessous un exemple simple (§ B.2.b).

B. Exemple : les gaz parfaits monoatomiques

1. DÉFINITION D'UN GAZ PARFAIT

On appelle gaz parfait un *ensemble de N particules sans interactions mutuelles,* pouvant se déplacer librement dans un volume V. Nous supposerons pour simplifier ces N particules identiques[74]; elles sont alors *indiscernables* (§ A.3). Nous nous limiterons ici aux *gaz parfaits classiques,* auxquels s'applique l'*approximation de Maxwell-Boltzmann* (on dit aussi *«gaz parfaits non dégénérés»*); les gaz parfaits quantiques (ou «dégénérés») feront l'objet du chapitre VI.

A chaque gaz réel on associe un gaz parfait constitué des mêmes molécules, dont on néglige les interactions mutuelles. Cette *approximation des gaz parfaits* est souvent bonne car, dans un gaz, la distance moyenne entre molécules est grande devant la portée des forces qui s'exercent entre elles, de sorte que l'énergie potentielle d'interaction est faible (cf. complément I.A, § III.3); à température donnée, l'approximation des gaz parfaits est d'autant meilleure que le gaz réel auquel on l'applique est plus dilué.

Nous distinguerons les gaz parfaits *monoatomiques,* dont les molécules sont constituées d'un seul atome (c'est le cas par exemple des «gaz rares» He, Ne, Ar, ...) des gaz parfaits *polyatomiques* (diatomiques comme O_2, N_2, HCl, ..., triatomiques comme CO_2, H_2O, ..., etc.). En effet, aux températures où on étudie habituellement

---

73. Comme il s'agit d'une seule particule, la densité d'états croît de façon modérée avec l'énergie ε (voir par exemple I, § III.B.2.a), de sorte que son influence n'est pas ici déterminante : les contributions principales à la somme (IV.24) proviennent des états (λ) tels que

$$\varepsilon_\lambda - \varepsilon_0 \lesssim kT .$$

74. On peut bien entendu étudier de manière analogue (cf. note 60) des *mélanges de gaz parfaits* (voir complément II.C, § I.3).

les gaz, l'énergie caractéristique kT est très petite devant les énergies d'excitation d'un atome[75], de sorte que les particules d'un gaz monoatomique se comportent comme des points matériels[76] sans structure; au contraire, les molécules polyatomiques possèdent des *niveaux de rotation et de vibration* qui peuvent être excités dans les conditions courantes et dont il faut par conséquent tenir compte. Nous nous restreignons ici aux gaz parfaits monoatomiques; les gaz parfaits polyatomiques font l'objet du complément III.B.

2. FONCTION DE PARTITION ET ÉNERGIE LIBRE D'UN GAZ PARFAIT MONOATOMIQUE

Considérons un gaz parfait en équilibre thermique avec un thermostat de température T. Nous saurons calculer ses propriétés si nous connaissons sa fonction de partition. Comme les particules constituant un gaz parfait sont indépendantes, le problème se ramène à évaluer la fonction de partition d'une particule.

a. Fonction de partition d'une particule du gaz

Pour un gaz monoatomique, on peut négliger la structure interne des particules qui le constituent (cf. note 75). L'hamiltonien[77] H_0 décrivant le mouvement de l'une d'entre elles se réduit alors au terme représentant son énergie cinétique de translation[78] :

$$H_0 = \vec{P}^2/2m \,, \tag{IV.25}$$

où m est la masse de l'atome considéré et $\vec{P}$ l'opérateur associé à son impulsion. Les états que nous notions (λ) au paragraphe A.1 (états propres de l'hamiltonien H_0), peuvent donc être repérés par leur impulsion $\hbar\vec{K}$, à laquelle il convient d'associer le nombre quantique m_s caractérisant la projection du spin de la particule sur un axe de quantification choisi à l'avance[79] :

$$(\lambda) \equiv \{\vec{K},\, m_s\} \,. \tag{IV.26}$$

L'énergie ε_λ correspondante est ici indépendante de m_s et vaut simplement

$$\varepsilon_\lambda = \frac{\hbar^2 \vec{K}^2}{2m} \,. \tag{IV.27}$$

---

75. Rappelons que celles-ci sont de l'ordre de quelques eV, c'est-à-dire équivalentes à plusieurs dizaines de milliers de kelvins. Jusqu'à quelques milliers de K, le cortège électronique d'un atome reste constamment dans son état fondamental, sa probabilité d'excitation étant négligeable.

76. L'approximation des gaz parfaits comporte implicitement l'hypothèse que le volume propre de chaque particule est négligeable, car sinon deux particules interagiraient forcément lorsqu'elles s'interpénétreraient. Cette hypothèse est effectivement vérifiée si le gaz est suffisamment dilué pour que la distance moyenne entre molécules soit grande devant leurs dimensions propres (complément I.A, § III.3).

77. Nous utilisons le formalisme quantique, qui est général. Un gaz parfait monoatomique pouvant également être décrit par la mécanique classique (complément I.A, § III.2), le formalisme classique que nous expliciterons au paragraphe V donnerait les mêmes résultats (voir note 92 et § V.A.2.b).

78. Nous supposons que le gaz parfait étudié n'est soumis à aucun champ extérieur, de sorte que ses molécules sont des particules libres. L'influence d'un champ extérieur (tel que le champ de pesanteur) sera examinée dans le complément V.B.

79. Rappelons que, pour une particule de spin s (s entier ou demi-entier positif ou nul), le nombre quantique m_s peut prendre les $(2s+1)$ valeurs $-s,\ -s+1,\ \ldots,\ s-1,\ s$. Pour un atome, nous appelons ici « spin » son moment cinétique total.

Nous savons (cf. I, § III.B.2.a) que les valeurs acceptables pour $\vec{K}$ sont quantifiées, puisque la particule est enfermée dans le volume V : les vecteurs d'onde acceptables sont répartis uniformément dans l'espace des $\vec{K}$, à raison d'un par cellule de « volume » $(2\pi)^3/V$. Comme V est de dimension macroscopique, la quantification correspondante de l'énergie est inappréciable par rapport à kT, de sorte que la somme sur $\vec{K}$ peut être remplacée par une intégrale[80] :

$$\sum_{(\lambda)} \Longrightarrow \sum_{m_s=-s}^{+s} \frac{V}{(2\pi)^3} \int \mathrm{d}^3K \,. \tag{IV.28}$$

La fonction de partition d'une particule, que nous noterons z_0, s'écrit donc

$$z_0 = \sum_{m_s=-s}^{+s} \frac{V}{(2\pi)^3} \int \mathrm{d}^3K \, \exp\left[-\beta \frac{\hbar^2 \vec{K}^2}{2m}\right] . \tag{IV.29}$$

La somme sur le nombre quantique de spin donne simplement un facteur $(2s+1)$, et l'intégrale se calcule sans difficulté[81]. On peut mettre le résultat final sous la forme

$$z_0 = (2s+1)\frac{V}{\Lambda^3} \,, \tag{IV.30}$$

---

80. Bien que microscopique puisqu'il comprend une seule particule, le système étudié est macroscopique par son extension spatiale (volume V), de sorte que l'intervalle entre deux niveaux consécutifs est très petit devant kT dans tout le domaine où l'approximation des gaz parfaits peut avoir quelque utilité (cf. I, note 18).

Ceci permet d'ailleurs de calculer tout aussi bien la fonction de partition z_0 à partir de la formule (I.24); compte tenu de la densité d'états (formule I (III.26)), on retrouve aussitôt le résultat (IV.30)-(IV.31) :

$$z_0 = \sum_{m_s=-s}^{+s} \int_0^{+\infty} \mathrm{d}E \frac{V\sqrt{2}}{2\pi^2\hbar^3} m^{3/2} E^{1/2} \, \mathrm{e}^{-\beta E} = (2s+1)\frac{V\sqrt{2}m^{3/2}}{\pi^2\hbar^3\beta^{3/2}} \int_0^{+\infty} y^2 \, \mathrm{d}y \, \mathrm{e}^{-y^2} \tag{1}$$

(où l'on a effectué le changement de variable $y = \sqrt{\beta E}$), c'est-à-dire finalement (voir formulaire)

$$z_0 = (2s+1)V\left(\frac{m}{2\pi\hbar^2\beta}\right)^{3/2} \tag{2}$$

81. Pour calculer

$$I \widehat{=} \frac{1}{(2\pi)^3} \int \mathrm{d}^3K \, \exp\left[-\beta \frac{\hbar^2 \vec{K}^2}{2m}\right] , \tag{1}$$

on passe en coordonnées sphériques dans l'espace des $\vec{K}$:

$$I = \frac{4\pi}{(2\pi)^3} \int_0^{+\infty} K^2 \, \mathrm{d}K \, \exp\left[-\beta \frac{\hbar^2 K^2}{2m}\right] . \tag{2}$$

Il suffit ensuite de poser

$$y = \sqrt{\frac{\hbar^2\beta}{2m}} K \tag{3}$$

pour transformer I en une intégrale figurant dans le formulaire :

$$I = \frac{4\pi}{(2\pi)^3} \left(\frac{2m}{\hbar^2\beta}\right)^{3/2} \int_0^{+\infty} y^2 \, \mathrm{d}y \, \mathrm{e}^{-y^2} \,, \tag{4}$$

ce qui donne finalement

$$I = \left(\frac{m}{2\pi\hbar^2\beta}\right)^{3/2} \tag{5}$$

où Λ, appelée «longueur d'onde thermique de de Broglie»[82], est définie par

$$\Lambda = \left[\frac{2\pi\hbar^2\beta}{m}\right]^{1/2} . \tag{IV.31}$$

b. Fonction de partition et énergie libre du gaz

Les particules qui constituent le gaz parfait étant par hypothèse *indépendantes*, identiques et *indiscernables*, il suffit maintenant d'appliquer les résultats du paragraphe A.3.c.

La condition (IV.23) de validité de l'approximation de Maxwell-Boltzmann s'écrit ici

$$(2s+1)\frac{V}{\Lambda^3} \gg N , \tag{IV.32}$$

c'est-à-dire

$$\Lambda^3 \ll \frac{V}{N}(2s+1) , \tag{IV.33}$$

ou encore

$$kT \gg \frac{2\pi\hbar^2}{m}\left(\frac{N}{V}\right)^{2/3}\frac{1}{(2s+1)^{2/3}} . \tag{IV.34}$$

L'inégalité (IV.34) précise quantitativement, dans le cas particulier étudié ici, la conclusion générale du paragraphe A.3.c.β : il faut que la température soit suffisamment élevée et/ou la densité de particules N/V suffisamment faible. Cette condition s'interprète aisément : d'après (IV.33), la longueur d'onde thermique de de Broglie Λ doit rester très petite devant $(V/N)^{1/3}$, qui n'est rien d'autre que la séparation moyenne entre deux particules; comme Λ caractérise les distances sur lesquelles pourraient se manifester des effets quantiques (cf. note 82), on voit que *l'approximation de Maxwell-Boltzmann est valable si le mouvement des particules peut être traité par la mécanique classique.* C'est effectivement le cas pour tous les gaz ordinaires[83], mais pas par exemple pour le «gaz» constitué par les électrons de conduction dans un métal (complément I.A, § III.2).

La fonction de partition Z_0 d'un gaz parfait monoatomique vérifiant la condition (IV.34) s'écrit donc

$$Z_0 = \frac{1}{N!}\left[(2s+1)\frac{V}{\Lambda^3}\right]^N = \frac{(2s+1)^N V^N}{N!}\left(\frac{mkT}{2\pi\hbar^2}\right)^{3N/2} . \tag{IV.35}$$

On en déduit aussitôt (en utilisant la formule de Stirling) l'énergie libre correspondante :

$$F_0 = -NkT\left[\ln\frac{V}{N} + \frac{3}{2}\ln\frac{mkT}{2\pi\hbar^2} + 1 + \ln(2s+1)\right] . \tag{IV.36}$$

---

82. Cette dénomination provient du fait que Λ est la longueur d'onde de de Broglie associée à une énergie de l'ordre de kT : de façon plus précise,

$$\Lambda = h/P \tag{1}$$

pour

$$\frac{P^2}{2m} = \pi kT . \tag{2}$$

83. Pour un gaz ordinaire, ce n'est pas la condition (IV.34) qui limite en fait le domaine de validité de l'approximation des gaz parfaits, mais l'hypothèse que les interactions sont négligeables (complément I.A, § III.3).

3. PROPRIÉTÉS DU GAZ PARFAIT MONOATOMIQUE

Connaissant la fonction de partition et l'énergie libre d'un gaz parfait monoatomique, on calcule facilement les grandeurs physiques qui caractérisent ses propriétés macroscopiques. Il s'agit en toute rigueur des valeurs prises par ces diverses grandeurs lorsque le gaz parfait est placé dans la situation canonique. Mais on ne parle de «gaz» que si le nombre de particules N est suffisamment grand; on se trouve alors à la *limite thermodynamique,* et les résultats que nous allons obtenir dépassent le cadre canonique adopté ici (cf. § III).

a. Énergie et chaleur spécifique

La valeur moyenne $\overline{E}_0^c$ de l'énergie d'un gaz parfait monoatomique maintenu à la température T peut être calculée par la formule générale (II.8). On trouve facilement

$$\overline{E}_0^c = \frac{3}{2} NkT . \tag{IV.37}$$

On en déduit aussitôt la *chaleur spécifique à volume constant* (formules (II.13)-(II.14)) :

$$c_V^{(0)} = \frac{3}{2} R , \tag{IV.38}$$

où $$R \mathrel{\widehat{=}} N_A k \tag{IV.39}$$

est la *«constante des gaz parfaits»,* qui vaut numériquement

$$R \simeq 8{,}31 \text{ J/K} . \tag{IV.40}$$

On constate que *l'énergie d'un gaz parfait est indépendante de son volume*[84]. Ceci provient du fait que la dépendance par rapport à V se réduit dans Z_0 à un facteur V^N, de sorte qu'elle est, dans $\ln Z_0$, complètement séparée de la dépendance par rapport à T[85]. Comme il s'agit de particules indépendantes, on a bien sûr

$$\overline{E}_0^c = N\overline{\varepsilon}_0^c , \tag{IV.41}$$

où $$\overline{\varepsilon}_0^c = -\frac{\partial}{\partial\beta} \ln z_0 = \frac{3}{2} kT \tag{IV.42}$$

est l'énergie moyenne d'une molécule.

On constate également que *la chaleur spécifique d'un gaz parfait monoatomique est une constante.* Ce résultat remarquable découle de la forme particulièrement simple de l'énergie, qui est proportionnelle à T; nous verrons plus loin (§ V.C) le lien entre cette forme simple et le fait que les mouvements de translation des molécules d'un gaz parfait peuvent tout aussi bien être décrits par la mécanique classique.

---

84. Cette affirmation reste vraie pour un gaz parfait polyatomique, bien que l'expression de son énergie soit moins simple (complément III.B, § I.3).

85. Cette dépendance très simple de Z_0 par rapport à V est la conséquence directe de l'absence d'interactions : z_0 est proportionnelle à V parce que chacune des particules est libre, et la fonction de partition globale Z_0 est proportionnelle à z_0^N parce que les diverses particules sont indépendantes.

b. Entropie

L'entropie canonique S_0^c d'un gaz parfait monoatomique se déduit de son énergie libre (IV.36) par la formule (II.28); cependant, puisque nous avons déjà calculé l'énergie moyenne $\overline{E}^c$, il est aussi simple d'utiliser la relation (II.26), qui conduit directement à la « formule de Sackur-Tétrode » :

$$S_0^c = Nk\left[\ln\frac{V}{N}+\frac{3}{2}\ln\frac{mkT}{2\pi\hbar^2}+\frac{5}{2}+\ln(2s+1)\right]. \tag{IV.43}$$

Il est intéressant de noter, sur cette expression simple, que l'entropie croît bien proportionnellement à N si la densité N/V est maintenue constante, mais qu'elle n'est pas simplement proportionnelle à N[86] pour V fixé.

c. Pression et potentiel chimique

En dérivant par rapport au volume V l'expression (IV.36) de l'énergie libre, on obtient la pression canonique p^c du gaz parfait (cf. (II.30.a)) :

$$p^c = \frac{NkT}{V}. \tag{IV.44}$$

Ce résultat remarquablement simple, *valable pour tous les gaz parfaits quelle que soit la structure de leurs molécules* (complément III.B, § I.3) constitue la célèbre *équation d'état des gaz parfaits,* que l'on écrit généralement sous la forme[87]

$$\boxed{pV = nRT}, \tag{IV.45}$$

où R est définie par (IV.39), p est la pression du gaz à la limite thermodynamique et n le nombre de moles N/N_A.

Calculons enfin le *potentiel chimique* d'un gaz parfait monoatomique, à partir de la définition générale (II.30.b). Il est commode pour cela d'écrire l'énergie libre F_0 sous la forme

$$F_0 = -NkT\,[\ln z_0 - \ln N + 1], \tag{IV.46}$$

qui se déduit directement de (IV.19.a) et de la formule de Stirling. On trouve alors

$$\mu_0 = -kT\ln\frac{z_0}{N} \tag{IV.47}$$

(pour simplifier, nous supposons atteinte la limite thermodynamique et supprimons l'indice c).

---

86. Lorsque la taille du système est multipliée par λ (complément II.A, § II), l'entropie S_0^c se comporte bien comme une grandeur *extensive :* les grandeurs extensives N et V sont multipliées par λ, leur rapport V/N restant inchangé.

Il est important de noter que l'entropie ne serait pas extensive si nous avions calculé la fonction de partition du système à partir de la formule (IV.6.a) au lieu de (IV.19.a), c'est-à-dire si nous n'avions pas tenu compte de l'indiscernabilité des particules du gaz. Le comportement anormal que nous aurions ainsi obtenu constitue le *« paradoxe de Gibbs »* (cf. § V.A.2.b et complément II.C).

87. Un gaz parfait obéit donc à la *« loi de Boyle-Mariotte »,* découverte empiriquement au dix-septième siècle : « à température constante, le volume V et la pression p d'un échantillon de gaz varient de façon inversement proportionnelle ($pV = c^{te}$) ». Les gaz réels ne vérifient cette loi que de manière approchée : le produit pV augmente ou diminue légèrement, selon le gaz étudié et la température choisie, lorsque la pression varie.

La signification de ce résultat apparaît plus clairement si l'on définit la *fugacité* φ à partir du potentiel chimique μ par la relation

$$\varphi \widehat{=} e^{\beta\mu} . \tag{IV.48}$$

On en déduit, pour le gaz parfait monoatomique,

$$\varphi_0 = \frac{N}{z_0} , \tag{IV.49}$$

c'est-à-dire, compte tenu de (IV.30)-(IV.31) et (IV.45),

$$\varphi_0 = p\, g_0(T) , \tag{IV.50}$$

avec $$g_0(T) \widehat{=} \left[kT(2s+1) \left(\frac{mkT}{2\pi\hbar^2} \right)^{3/2} \right]^{-1} . \tag{IV.51}$$

La fugacité d'un gaz parfait est proportionnelle à sa pression, le coefficient de proportionnalité étant une fonction de la température seule (mais dépendant de la nature du gaz par m et s)[88].

V. LA DISTRIBUTION CANONIQUE EN MÉCANIQUE CLASSIQUE

Les propriétés et l'évolution microscopiques d'un système quelconque sont de façon générale régies par les lois de la mécanique quantique. Il existe cependant des cas particuliers où l'on peut les traiter par la mécanique classique, ce qui simplifie souvent les raisonnements et les calculs sans que les résultats en soient affectés de façon appréciable. Cette partie V est destinée à expliciter les formules fondamentales de la mécanique statistique pour un système classique placé dans la situation canonique.

A. Formules fondamentales

Soit un système $\mathcal{S}$ qui peut être décrit par la mécanique classique. Nous notons de façon générale q_i $(i = 1, 2, \ldots, n)$ les coordonnées généralisées qui permettent de repérer sa position, et p_i leurs moments conjugués (appendice III, § II.1). L'énergie du système est donnée par la fonction de Hamilton $\mathcal{H}(q_i, p_i)$. Ce système est en équilibre thermique avec un thermostat $\mathcal{C}$, qui lui impose sa température T.

1. DENSITÉ CANONIQUE DANS L'ESPACE DES PHASES

L'état d'équilibre du système $\mathcal{S}$ est caractérisé par une densité de probabilité dans l'espace des phases $w^c(q_i, p_i)$, équivalent classique de la distribution de probabilité P_ℓ^c sur les états quantiques microscopiques (ℓ) :

$$w^c(q_1, q_2, \ldots, q_n; p_1, p_2, \ldots, p_n)\, dq_1\, dq_2 \ldots dq_n\, dp_1\, dp_2 \ldots dp_n \tag{V.1}$$

est la probabilité pour que le système se trouve dans un état microscopique dont le point représentatif appartient au volume élémentaire $dq_1\, dq_2 \ldots dq_n\, dp_1\, dp_2 \ldots dp_n$ de l'espace des phases, situé autour du point de cet espace repéré par $(q_1, q_2, \ldots, q_n; p_1, p_2, \ldots, p_n)$.

88. Ce résultat se généralise aux gaz parfaits polyatomiques (complément III.B, § I.3), le coefficient $g(T)$ ayant alors une expression plus compliquée.

La densité canonique $w^c(q_i, p_i)$ dans l'espace des phases peut être déterminée par un raisonnement en tous points semblable à celui du paragraphe I.B.1; au lieu de l'indice discret (ℓ), l'état microscopique du système est ici caractérisé par l'ensemble des $2n$ paramètres continus $\{q_i, p_i\}$, l'énergie E_ℓ étant remplacée par $\mathcal{H}(q_i, p_i)$[89] :

$$(\ell) \iff \{q_i, p_i\}\,, \tag{V.2.a}$$

$$E_\ell \iff \mathcal{H}(q_i, p_i)\,. \tag{V.2.b}$$

Cette correspondance fournit immédiatement le résultat :

$$P^c_\ell = \frac{1}{Z} e^{-E_\ell/kT} \iff w^c(q_i, p_i) = \frac{1}{A} \exp\left[-\mathcal{H}(q_i, p_i)/kT\right], \tag{V.3}$$

la constante A étant fixée par normalisation des probabilités :

$$A = \int dq_1 \ldots dq_n \int dp_1 \ldots dp_n \exp\left[-\mathcal{H}(q_i, p_i)/kT\right] \tag{V.4}$$

(l'intégrale porte sur le volume à $2n$ dimensions accessible au système dans l'espace des phases, compte tenu des contraintes extérieures).

2. FONCTION DE PARTITION CLASSIQUE

Les développements du paragraphe II.B, qui restent fondamentalement valables ici, montrent la commodité de la fonction de partition Z. Comme plus haut (formule (I.16)), son expression découle de la condition de normalisation (V.4). Il faut cependant que cette expression classique donne, dans le domaine où elle est utilisable, le même résultat que la véritable définition (I.16) de la fonction de partition, basée sur la description quantique du système.

Il se trouve que les systèmes de particules identiques et indiscernables posent un problème également en mécanique *statistique* classique. Avant d'aborder cette difficulté, traitons d'abord les cas où elle n'apparaît pas.

a. Systèmes classiques ordinaires (éléments tous discernables)

La correspondance (V.2) entre les descriptions quantique et classique montre que la « constante de normalisation » A (formule (V.4)) et la fonction de partition Z sont nécessairement proportionnelles : quand on applique cette correspondance à la définition (I.16) de Z, la somme discrète sur (ℓ) est remplacée par une intégrale sur les paramètres continus $\{q_i, p_i\}$, et l'expression obtenue ne peut différer de celle de A que par un facteur constant. Mais A et Z ne sont pas simplement égales : Z est une grandeur sans dimension, alors que A a les dimensions d'une action à la puissance n[90].

---

89. Ces modifications sont les seules qui doivent vraiment être apportées au raisonnement du paragraphe I.B.1, puisque le thermostat $\mathcal{T}$ doit être traité dans le cadre général de la mécanique quantique. On peut cependant préférer, pour plus d'unité, choisir un thermostat qui admette lui aussi un traitement classique; l'adaptation nécessaire est alors un peu plus profonde (il faut utiliser l'expression classique de l'entropie microcanonique; voir complément II.F), mais l'argument reste fondamentalement le même et le résultat est inchangé.

90. D'après la définition des moments conjugués (appendice III, § II.1), le produit $p_i q_i$ (et donc le produit $dp_i\, dq_i$ qui apparaît dans l'expression (V.4) de A) a les dimensions d'une action, c'est-à-dire d'une énergie multipliée par un temps, ou d'un moment cinétique.

Le coefficient de proportionnalité u_n entre Z et A :

$$Z = \frac{1}{u_n} A , \tag{V.5}$$

représente en quelque sorte le *volume élémentaire de l'espace des phases équivalent à un état quantique.* D'après l'argument ci-dessus, u_n est indépendant de la température et des autres paramètres extérieurs, et il doit avoir les dimensions de $\hbar^n$. Nous *admettrons* sa valeur :

$$u_n = (2\pi\hbar)^n , \tag{V.6}$$

qui correspond à *un facteur universel $2\pi\hbar$ pour chaque couple de variables conjuguées* (q_i, p_i), indépendamment du système considéré[91]. Il est facile de vérifier cette valeur sur l'exemple simple du gaz parfait monoatomique[92].

Donc, lorsque le système étudié peut être décrit par la mécanique classique (et qu'il ne comporte pas de particules indiscernables), sa fonction de partition peut être calculée selon la formule

$$Z(\beta, x) = \frac{1}{(2\pi\hbar)^n} \int \left[\prod_{i=1}^{n} \mathrm{d}q_i \, \mathrm{d}p_i \right] \mathrm{e}^{-\beta \mathcal{H}(q_i, p_i)} . \tag{V.7}$$

---

91. Que ce facteur universel soit de l'ordre de $\hbar$ n'est pas surprenant. Dans un état quantique quelconque, en effet, la position x d'une particule le long d'un axe de coordonnées et l'impulsion p_x conjuguée doivent vérifier la relation d'incertitude de Heisengerg

$$\Delta x . \Delta p_x \geqslant \hbar ,$$

où Δx et Δp_x sont les écarts quadratiques moyens de ces variables dans l'état quantique considéré. Ainsi, les lois de la mécanique quantique interdisent de découper l'espace des phases classique associé à deux variables conjuguées en « pavés » de surface inférieure à $\hbar$; inversement, cette surface ne peut pas être très supérieure à $\hbar$, puisqu'elle doit correspondre à un seul état quantique.

92. L'approximation classique étant toujours valable pour le mouvement de translation des molécules d'un gaz (complément I.A, § III.2), on doit en effet retrouver à partir de la formule (V.7) l'expression (IV.30)-(IV.31) de la fonction de partition z_0 pour une particule d'un gaz parfait monoatomique (mis à part le facteur $(2s+1)$, car le spin ne peut être traité que quantiquement).

Comme il s'agit d'une particule ponctuelle libre enfermée dans un volume V, la fonction de Hamilton est simplement

$$\mathcal{H}_0(\vec{r}, \vec{p}) = \vec{p}^{\,2}/2m , \tag{1}$$

et l'espace des phases accessible est défini par

$$\begin{cases} \vec{r} \in V , \\ \vec{p} \text{ quelconque} . \end{cases} \tag{2}$$

La formule (V.7) donne alors

$$z_0 = \frac{1}{u_3} \int_V \mathrm{d}^3 r \int \mathrm{d}^3 p \, \exp\left[-\beta \frac{\vec{p}^{\,2}}{2m} \right] = \frac{V}{u_3} . 4\pi \int_0^\infty p^2 \, \mathrm{d}p \, \exp\left[-\beta \frac{p^2}{2m} \right] , \tag{3}$$

c'est-à-dire, d'après le formulaire,

$$z_0 = \frac{V}{u_3} \left(\frac{2\pi m}{\beta} \right)^{3/2} . \tag{4}$$

La comparaison de ce résultat avec les formules (IV.30)-(IV.31) déduites du formalisme quantique (et prises pour $s = 0$) donne bien

$$u_3 = (2\pi\hbar)^3 . \tag{5}$$

Notons qu'ici la dépendance par rapport aux paramètres extérieurs x autres que la température provient non seulement de l'expression de la fonction de Hamilton mais aussi des restrictions imposées à l'espace des phases par les contraintes extérieures : l'intervalle de variation des coordonnées q_i est par exemple borné par le volume V dans lequel est enfermé le système (voir le calcul de la note 92).

b. Systèmes de particules indiscernables

Considérons maintenant un système de N particules identiques et indiscernables. Cette particularité ne devrait poser ici aucun problème : en mécanique classique, il suffit de numéroter les particules pour pouvoir les distinguer à tout instant les unes des autres, puisque chacune d'elles évolue sur une trajectoire, le long de laquelle on peut en principe la suivre pour la reconnaître. On devrait donc pouvoir appliquer la formule générale (V.7), en prenant $n = 3N$ puisque chaque particule a 3 coordonnées de position (et 3 composantes pour l'impulsion conjuguée)[93], et en utilisant la fonction de Hamilton $\mathcal{H}(\vec{r}_1, \vec{r}_2, \ldots, \vec{r}_N; \vec{p}_1, \vec{p}_2, \ldots, \vec{p}_N)$ correspondant à une numérotation quelconque, mais fixée une fois pour toutes, des particules.

Il est cependant facile de voir qu'on n'obtiendra jamais ainsi un équivalent classique du résultat quantique (approché) (IV.19.a). Supposons en effet que les N particules soient indépendantes, c'est-à-dire que la fonction de Hamilton ne comporte pas de termes d'interaction mutuelle :

$$\mathcal{H}(\vec{r}_1, \vec{r}_2, \ldots, \vec{r}_N; \vec{p}_1, \vec{p}_2, \ldots, \vec{p}_N) = \mathcal{H}(\vec{r}_1, \vec{p}_1) + \mathcal{H}(\vec{r}_2, \vec{p}_2) + \ldots + \mathcal{H}(\vec{r}_N, p_N) . \quad \text{(V.8)}$$

La fonction de partition se factorise alors en un produit de N fonctions de partition individuelles, toutes identiques et données par

$$z = \frac{1}{(2\pi\hbar)^3} \int \mathrm{d}^3 r \int \mathrm{d}^3 p \, \exp\left[-\beta \mathcal{H}(\vec{r}, \vec{p})\right] . \quad \text{(V.9)}$$

On reconnaît bien là l'expression qui, compte tenu de la correspondance (V.2) (appliquée aux états individuels) et du facteur de conversion (V.6), permet de calculer dans le domaine classique la fonction de partition z d'une particule, définie en (IV.7.a). Mais on obtient forcément ainsi la formule

$$Z = z^N , \quad \text{(V.10)}$$

valable pour des particules indépendantes identiques mais *discernables,* comme elles le sont toujours en mécanique classique.

Or, si l'on applique cette formule (V.10) à un gaz parfait, l'expression que l'on en déduit pour l'entropie du gaz parfait présente le défaut rédhibitoire de ne pas être additive : l'entropie des N particules libres occupant le volume V n'est pas égale, par exemple, au double de celle de $N/2$ particules occupant le volume $V/2$ (voir complément II.C). Cette difficulté est connue sous le nom de *«paradoxe de Gibbs».* Les physiciens l'ont rencontrée en développant la mécanique statistique classique, avant la découverte de la mécanique quantique. Ils l'ont résolue alors en décrétant (sans véritable justification dans le cadre de la mécanique classique) que *le volume de l'espace des phases d'un système de N particules identiques et indiscernables doit être réduit par un facteur N!* par rapport à celui que l'on attendrait normalement. Ils ont ainsi abouti directement à la formule (IV.19.a), dont l'origine est en réalité quantique comme nous l'avons montré au paragraphe IV.A.3.c.

Une fois que l'on a décidé de ne pas compter comme distincts deux états du système qui ne diffèrent que par une permutation des N particules (lorsqu'elles sont indiscernables), on constate que la modification des formules est plus simple à effectuer

93. Nous supposons pour simplifier que les particules considérées n'ont pas de structure interne (voir note 94).

en mécanique classique qu'en mécanique quantique. En effet, comme les paramètres $\vec{r}$ et $\vec{p}$ varient continûment, les situations dans lesquelles deux particules parmi les N occuperaient exactement le même état individuel sont infiniment rares et ne doivent pas être prises en compte : dans le domaine classique, *le facteur de réduction est N! sans approximation.* L'expression classique de la fonction de partition d'un système de N particules identiques et indiscernables s'écrit donc[94]

$$Z = \frac{1}{N!}\frac{1}{(2\pi\hbar)^{3N}}\int\left[\prod_{j=1}^{N} \mathrm{d}^3 r_j\, \mathrm{d}^3 p_j\right] \mathrm{e}^{-\beta \mathcal{H}(\vec{r}_j, \vec{p}_j)} . \tag{V.11}$$

Notons qu'*il n'est pas nécessaire que les particules soient indépendantes* pour que la formule précédente soit valable : même si la fonction de Hamilton du système comprend des termes d'interaction entre les particules, chacune d'elles possède à chaque instant une position $\vec{r}$ et une impulsion $\vec{p}$ déterminées, et ceci suffit pour que l'argument aboutissant à la formule (V.11) soit applicable[95].

Dans le cas où les N particules identiques et indiscernables sont en outre indépendantes, l'expression classique (V.11) prend la forme (IV.19.a), la fonction de partition z d'une particule étant ici donnée par (V.9).

B. Domaine de validité du traitement classique

Comment déterminer, dans chaque problème concret, s'il est absolument nécessaire de traiter par la mécanique quantique le comportement microscopique du système ou si l'on peut utiliser la mécanique classique?

On peut dire de façon générale que le traitement classique fournira une bonne approximation *si la constante fondamentale* $\hbar$, caractéristique de la mécanique quantique, *est négligeable* devant les grandeurs physiques de même dimension associées au système considéré. Par exemple, la coordonnée de position x des diverses particules d'un système peut être mesurée avec une incertitude minimale δx dont l'ordre de grandeur dépend du système et des conditions dans lesquelles on l'étudie; il en est de même de l'incertitude δp_x sur l'impulsion correspondante. Le mouvement de ces

94. Le raisonnement précédent et la formule (V.11) supposent les particules toutes identiques et sans structure interne. On les généralise aisément à des situations plus compliquées.

Si par exemple le système est un mélange de N_1 particules indiscernables d'un premier type et N_2 particules indiscernables d'un deuxième type (différent du premier), sa fonction de partition est évidemment

$$Z = \frac{1}{N_1!}\frac{1}{N_2!}\frac{1}{(2\pi\hbar)^{3(N_1+N_2)}}\int\left[\prod_{j=1}^{N_1+N_2} \mathrm{d}^3 r_j\, \mathrm{d}^3 p_j\right] \mathrm{e}^{-\beta \mathcal{H}(\vec{r}_j, \vec{p}_j)} . \tag{1}$$

Si l'on étudie N particules identiques et indiscernables présentant une structure interne (comme des molécules polyatomiques), on devra introduire pour chacune d'elles n' paramètres $q^{(i)}$ supplémentaires (accompagnés des n' moments conjugués $p^{(i)}$), et l'on écrira

$$Z = \frac{1}{N!}\frac{1}{(2\pi\hbar)^{(3+n')N}}\int\left[\prod_{j=1}^{N} \mathrm{d}^3 r_j\, \mathrm{d}^3 p_j \prod_{i=1}^{n'} \mathrm{d}q_j^{(i)}\, \mathrm{d}p_j^{(i)}\right] \mathrm{e}^{-\beta \mathcal{H}(\vec{r}_j, \vec{p}_j; q_j^{(i)}, p_j^{(i)})} . \tag{2}$$

95. En mécanique quantique, nous n'avons introduit la notion d'états individuels (λ) que pour des particules indépendantes (§ IV.A.1). Cette restriction est due au fait que, pour simplifier, nous raisonnons comme si les seuls états microscopiques possibles pour un système étaient les états propres de son hamiltonien H (I, § III.A.2.b). Cependant, dans le formalisme plus général de l'opérateur-densité (complément I.H), il n'y a aucun inconvénient de principe à utiliser une base de l'espace des états du système constituée à partir des états individuels des diverses particules (convenablement symétrisée ou antisymétrisée si ces particules sont identiques et indiscernables), même dans les cas où leurs interactions ne sont pas négligeables.

particules pourra être décrit par la mécanique classique si le produit $\delta x \,.\, \delta p_x$ de ces deux incertitudes expérimentales est grand devant $\hbar$[96] :

$$\delta x \,.\, \delta p_x \gg \hbar \,. \tag{V.12}$$

Ou bien, si l'on a affaire à un oscillateur harmonique (vibrations d'une molécule diatomique, par exemple) dont la pulsation propre est ω et qui est maintenu à la température T, on pourra le traiter par la mécanique classique si kT/ω est grand devant $\hbar$, c'est-à-dire si

$$kT \gg \hbar\omega \,. \tag{V.13}$$

Plus généralement, une première *condition nécessaire* pour que l'approximation classique puisse être valable est que *l'énergie* kT caractéristique du thermostat *soit grande devant les intervalles entre niveaux consécutifs* du système : il faut en effet que, dans l'expression (I.16) de la fonction de partition, la quantification des niveaux d'énergie ne soit pas perceptible.

Cette condition n'est cependant *pas toujours suffisante,* et il convient d'examiner de plus près chaque cas particulier. Si l'on considère par exemple le mouvement des molécules d'un gaz ordinaire dans le récipient qui les contient, à la température ambiante, on trouve facilement (complément I.A, § III.2.b) que la mécanique classique en donne une description excellente. Mais prenons les électrons de conduction d'un métal; bien que, aux températures courantes, l'énergie kT soit encore très supérieure aux intervalles de quantification de leur énergie cinétique (I, note 18), ces électrons constituent un « gaz très fortement dégénéré », c'est-à-dire que les effets quantiques associés au principe de Pauli y sont très importants (complément I.A, § III.2.c et chapitre VI, § I).

Il faut également se souvenir que l'énergie d'un système comporte souvent plusieurs contributions, associées à divers « degrés de liberté » de type différent; il arrive ainsi que, dans un domaine de températures intermédiaires, *certains degrés de liberté puissent être traités classiquement alors que d'autres nécessitent encore un traitement quantique.* L'exemple le plus simple est celui d'un gaz parfait polyatomique (complément III.B) : le mouvement de translation des molécules dans le récipient se fait à température ordinaire selon les lois de la mécanique classique; la rotation de la molécule autour de son centre de masse est en général également classique à cette température, mais pas ses vibrations, qu'il faut traiter par la mécanique quantique jusqu'à des températures nettement plus élevées.

L'utilisation de la mécanique classique conduit donc de façon générale à une *approximation valable à haute température,* mais il faut analyser chaque cas concret pour pouvoir préciser où se situe la température limite.

C. Le théorème d'« équipartition de l'énergie »

Les systèmes qui peuvent être décrits par la mécanique classique vérifient le théorème suivant, remarquable par la simplicité du résultat et important par ses applications multiples.

1. ÉNONCÉ ET DÉMONSTRATION

Supposons que l'une des impulsions généralisées, que nous noterons p_m, interviennent *uniquement par un terme quadratique* dans la fonction de Hamilton, c'est-à-dire que celle-ci soit de la forme

$$\mathcal{H}(q_i, p_i) = a p_m^2 + b \,, \tag{V.14}$$

où *a et b sont indépendants de* p_m (mais peuvent en général être fonctions des p_i pour $i \neq m$ et de tous les q_i). Cette hypothèse est couramment vérifiée : l'énergie cinétique

96. La fonction d'onde quantique décrivant le système ne permet pas de connaître la position et l'impulsion d'une particule le long d'un même axe avec une précision meilleure, respectivement, que Δx et Δp_x, qui doivent vérifier

$$\Delta x \,.\, \Delta p_x \geqslant \hbar \,. \tag{1}$$

de tout système est une fonction quadratique homogène des moments p_i; dès lors, il suffit que les autres termes de la fonction de Hamilton soient indépendants des p_i, ce qui est le cas si les forces dérivent d'une énergie potentielle.

Calculons alors la valeur moyenne canonique du terme quadratique ap_m^2; elle s'écrit, d'après l'expression (V.3)-(V.4) de la densité de probabilité canonique et la forme (V.14) de la fonction de Hamilton,

$$\overline{ap_m^2}^c = \frac{1}{A}\int\left[\prod_{i=1}^{n} \mathrm{d}q_i\ \mathrm{d}p_i\right] ap_m^2\ \mathrm{e}^{-(ap_m^2+b)/kT}\,, \tag{V.15}$$

avec
$$A = \int\left[\prod_{i=1}^{n} \mathrm{d}q_i\ \mathrm{d}p_i\right]\mathrm{e}^{-(ap_m^2+b)/kT}\,. \tag{V.16}$$

Quelles que soient les fonctions a et b qui figurent dans ces formules, on trouve[97] simplement

$$\overline{ap_m^2}^c = \frac{1}{2}kT\,. \tag{V.17}$$

Cette limitation quantique passe totalement inaperçue si les incertitudes expérimentales δx et δp_x sont telles que

$$\delta x \gg \Delta x\,, \tag{2.a}$$

$$\delta p_x \gg \Delta p_x\,, \tag{2.b}$$

ce qui est possible si

$$\delta x\,.\,\delta p_x \gg \hbar\,. \tag{3}$$

La largeur de la fonction d'onde est alors inappréciable à l'échelle considérée, tant pour la coordonnée x que pour l'impulsion conjuguée p_x. On montre en outre (C.D.L., p. 243) que le point auquel se réduit pratiquement la fonction d'onde suit les lois de la mécanique classique.

97. Supposons pour simplifier que $m = 1$:

$$\mathcal{H}(q_i, p_i) = a(q_1, q_2, \ldots, q_n; p_2, \ldots, p_n)p_1^2 + b(q_1, q_2, \ldots, q_n; p_2, \ldots, p_n)\,. \tag{1}$$

Nous voulons calculer

$$\overline{ap_1^2}^c = \frac{1}{A}\int \mathrm{d}q_1\,\mathrm{d}q_2 \ldots \mathrm{d}q_n \int \mathrm{d}p_1\,\mathrm{d}p_2 \ldots \mathrm{d}p_n\, ap_1^2\,\mathrm{e}^{-(ap_1^2+b)/kT}\,, \tag{2}$$

avec

$$A = \int \mathrm{d}q_1\,\mathrm{d}q_2 \ldots \mathrm{d}q_n \int \mathrm{d}p_1\,\mathrm{d}p_2 \ldots \mathrm{d}p_n\ \mathrm{e}^{-(ap_1^2+b)/kT}\,. \tag{3}$$

Écrivons la formule (2) sous la forme

$$\overline{ap_1^2}^c = \frac{1}{A}\int \mathrm{d}q_1\,\mathrm{d}q_2 \ldots \mathrm{d}q_n \int \mathrm{d}p_2 \ldots \mathrm{d}p_n\ \mathrm{e}^{-b/kT} a \int \mathrm{d}p_1 p_1^2\ \mathrm{e}^{-ap_1^2/kT}\,, \tag{4}$$

et effectuons une intégration par parties sur la dernière intégrale :

$$\int \mathrm{d}p_1 p_1^2\ \mathrm{e}^{-ap_1^2/kT} = \left[p_1 \cdot \left(-\frac{kT}{2a}\right)\mathrm{e}^{-ap_1^2/kT}\right]_{\text{bornes}} + \frac{kT}{2a}\int \mathrm{d}p_1\,\mathrm{e}^{-ap_1^2/kT}\,. \tag{5}$$

Si les bornes d'intégration sont $(0, +\infty)$ ou $(-\infty, +\infty)$, ce qui est toujours le cas en pratique, le terme tout intégré est nul. En reportant le second terme de (5) dans l'expression (4) de la valeur moyenne cherchée, on obtient

$$\overline{ap_1^2}^c = \frac{1}{A}\int \mathrm{d}q_1\,\mathrm{d}q_2 \ldots \mathrm{d}q_n \int \mathrm{d}p_2 \ldots \mathrm{d}p_n\ \mathrm{e}^{-b/kT} \cdot \frac{kT}{2}\int \mathrm{d}p_1\ \mathrm{e}^{-ap_1^2/kT}\,. \tag{6}$$

Il suffit de comparer avec la formule (3) qui donne A pour voir que

$$\overline{ap_1^2}^c = \frac{kT}{2}\,. \tag{7}$$

La démonstration suppose

$$a > 0 \tag{8}$$

(quels que soient q_1, q_2, .., q_n et p_2, ..., p_n). Cette condition est toujours vérifiée, car l'énergie cinétique est positive. D'ailleurs, les intégrales des formules (2) et (3) seraient dans le cas contraire violemment divergentes.

On obtient un résultat analogue dans le cas où c'est une coordonnée, disons $q_{m'}$, qui apparaît dans la fonction de Hamilton sous la forme

$$\mathcal{H}(q_i, p_i) = a' q_{m'}^2 + b' , \tag{V.18}$$

où *a' et b' sont indépendants de $q_{m'}$*. Cette situation est moins fréquente que la précédente (V.14); elle se produit lorsque la coordonnée $q_{m'}$ repère un *oscillateur harmonique,* dont l'énergie potentielle est par définition quadratique. Le même raisonnement que plus haut[98] montre que

$$\overline{a' q_{m'}^2}^{\,c} = \frac{1}{2} kT . \tag{V.19}$$

On énonce donc le théorème suivant, dit « d'équipartition de l'énergie » : *tout terme indépendant quadratique de la fonction de Hamilton a pour valeur moyenne $kT/2$.*

2. DISCUSSION PHYSIQUE

Le théorème d'équipartition de l'énergie donne immédiatement la *valeur moyenne de l'énergie cinétique* d'un système classique quelconque maintenu à la température T, ou d'une de ses parties. Ainsi, l'énergie cinétique de translation d'une particule se déplaçant dans l'espace à trois dimensions $\left(\vec{p}^{\,2}/2m = (p_x^2 + p_y^2 + p_z^2)/2m\right)$ vaut en moyenne $3kT/2$ (cf. formule (IV.42)); on obtient de même $kT/2$ pour l'énergie cinétique moyenne de rotation d'un système autour d'un axe donné ($I\dot{\theta}^2/2$, I étant le moment d'inertie du système par rapport à l'axe considéré et $\dot{\theta}$ sa vitesse angulaire).

Pour ce qui est de l'*énergie potentielle,* le théorème ne s'applique qu'à l'*oscillateur harmonique.* Ce cas est quand même assez fréquent puisqu'il se produit chaque fois qu'on étudie les petits mouvements d'un système autour d'une position d'équilibre mécanique stable; il faut cependant alors que l'approximation harmonique (V.18) reproduise correctement l'énergie potentielle du système dans un intervalle suffisamment large autour de son minimum, et que la température ne soit pas trop élevée : sinon les termes d'ordre supérieur à 2 dans le développement du potentiel ne sont plus négligeables (on dit qu'ils engendrent des *« effets d'anharmonicité »*), et les conditions du théorème ne sont plus remplies (cf. note 98). Lorsqu'elles le sont et qu'on a un oscillateur harmonique à une dimension (non couplé), la valeur moyenne de son énergie totale est kT ($kT/2$ pour l'énergie cinétique et $kT/2$ pour l'énergie potentielle); elle est $3kT$ pour un oscillateur harmonique à trois dimensions[99].

---

98. La fonction de Hamilton ne peut avoir la forme (V.18) que si

$$a' > 0 ,$$

car sinon le système serait instable du point de vue mécanique : il pourrait diminuer indéfiniment son énergie en choisissant des valeurs de $q_{m'}$ de plus en plus grandes. De toutes façons, il est nécessaire que $\exp[-\beta\mathcal{H}]$ soit intégrable.

Mais il est rare dans la pratique que la forme parfaitement harmonique (V.18) soit valable pour $q_{m'}$ variant de $-\infty$ à $+\infty$. Cependant, si l'on suit une démonstration analogue à celle de la note 97, le terme tout intégré de la formule équivalente à (5) est encore négligeable si les bornes d'intégration effectives rendent $a' q_{m'}^2$ grand devant kT.

99. Ici aussi, le résultat découle aussitôt de l'expression de l'énergie potentielle en coordonnées cartésiennes :

$$W(x, y, z) = \frac{1}{2} K(x^2 + y^2 + z^2) . \tag{1}$$

.../...

Il est essentiel de garder présent à l'esprit que *le théorème d'équipartition de l'énergie ne s'applique qu'aux systèmes pouvant être décrits par la mécanique classique.* L'exemple de l'oscillateur harmonique à une dimension illustre bien cette limitation fondamentale. On montre en effet facilement (complément III.B, § II.2.a) que la valeur moyenne de l'énergie, calculée dans le formalisme quantique, s'écrit

$$\overline{E}^c = \hbar\omega \left[\frac{1}{2} + \frac{1}{e^{\hbar\omega/kT} - 1}\right], \tag{V.20}$$

où ω est la pulsation de l'oscillateur harmonique. Cette expression ne vérifie visiblement pas, en général, le théorème d'équipartition de l'énergie; cependant, lorsque la température est suffisamment élevée pour que l'intervalle $\hbar\omega$ entre deux niveaux d'énergie consécutifs soit petit devant kT, on retrouve le résultat classique[100] :

$$\overline{E}^c \sim kT \quad \text{si} \quad \hbar\omega \ll kT . \tag{V.21}$$

Attention : En coordonnées sphériques,

$$W = \frac{1}{2} Kr^2 . \tag{2}$$

Il ne faudrait pas en déduire que la valeur moyenne de W est $kT/2$. Les hypothèses du théorème ne sont en effet pas vérifiées : la coordonnée r figure aussi dans l'expression de l'énergie cinétique (voir exercice III.26).

100. Comme $kT/\hbar\omega \gg 1$, on peut développer l'exponentielle de la formule (V.20) :

$$\overline{E}^c \simeq \hbar\omega \left[\frac{1}{2} + \frac{kT}{\hbar\omega}\right] \simeq kT .$$

COMPLÉMENTS DU CHAPITRE III

III.A. Paramagnétisme et diamagnétisme des systèmes de particules indépendantes **III.B. Les gaz parfaits polyatomiques**	Applications des notions et résultats du chapitre. Faciles dans leur principe mais les situations physiques étudiées ne sont pas toujours simples.
III.C. La théorie cinétique des gaz	Sujet intéressant du point de vue physique. Le début est très facile, les paragraphes II et III assez faciles, le paragraphe IV plus difficile.
III.D. Ensemble T-p	Reprend les raisonnements du chapitre dans un cas différent mais analogue. Les résultats obtenus permettent de résoudre simplement un problème physique dont l'étude serait difficile par une autre méthode. Difficulté moyenne si l'on a bien compris le chapitre.
III.E. Chaleur spécifique des solides	Problème important. Le paragraphe I est très facile, les suivants plus difficiles techniquement; toutefois, si l'on admet les conclusions du paragraphe II, le paragraphe III (qui contient le principal résultat physique) ne présente pas de difficulté.
III.F. Les solutions diluées	Introduit la notion d'«interaction effective» et l'applique à un cas simple. La démonstration des formules fondamentales est un peu délicate, mais leur exploitation physique (§ III) est beaucoup plus facile.
III.G. L'équation de van der Waals et la transition gaz-liquide. **III.H. Systèmes de particules chargées.** **III.J. Le ferromagnétisme** **III.K. Transition ordre-désordre dans les alliages binaires**	L'«approximation de champ moyen» appliquée à divers problèmes et vue sous différents aspects. C'est dans III.J (§ I et II, le paragraphe I.4 pouvant être sauté) que les caractéristiques de cette approximation apparaissent le plus clairement. Les sujets abordés dans III.G et III.J sont particulièrement importants du point de vue physique. III.G : assez facile au début, plus difficile dans l'étude de la transition de phase (§ III). III.H : difficulté moyenne; son intérêt réside dans l'étude d'interactions à longue portée, mais écrantées. III.J : analyse une transition de phase du second ordre très connue et qui a été beaucoup étudiée; on peut se limiter aux paragraphes I et II, de difficulté moyenne, qui donnent les idées physiques essentielles; les paragraphes III et IV pourront servir ultérieurement à un approfondissement du problème et comme introduction à la lecture d'ouvrages spécialisés. III.K : encore une transition de phase du second ordre, dans un système *a priori* très différent du précédent; le début (§ I et II), qui développe l'approximation de champ moyen, est assez facile; le paragraphe III, qui va au-delà de cette approximation, est plus complexe techniquement.
III.L. Chaleur et travail. Le «second principe» de la thermodynamique	Fait suite à II.D. Démontre le «second principe» de la thermodynamique et l'interprète physiquement. Assez facile à lire, mais plus difficile à bien comprendre car le «second principe» est plein de subtilités.

COMPLÉMENT III.A

Paramagnétisme et diamagnétisme des systèmes de particules indépendantes

I. Généralités
  1. Formulation du problème dans l'ensemble canonique
  2. Paramagnétisme et diamagnétisme
    a. Définition
    b. Origine physique

II. Étude statistique du paramagnétisme
  1. Hamiltonien du système
  2. Fonction de partition et énergie libre
  3. Propriétés du système
    a. Aimantation et susceptibilité magnétique
    b. Énergie et capacité calorifique magnétiques
    c. Entropie

III. Étude statistique du diamagnétisme
  1. Diamagnétisme des matériaux à couches électroniques complètes
    a. Hamiltonien d'un atome en présence d'un champ magnétique
    b. Aimantation et susceptibilité diamagnétiques
  2. Diamagnétisme du gaz (dilué) d'électrons libres

IV. Aimantation et mécanique classique
  1. Théorie semi-classique du paramagnétisme
  2. Théorie semi-classique du diamagnétisme
  3. Discussion et théorème de van Leeuwen

I. GÉNÉRALITÉS

Lorsqu'un système macroscopique, initialement non aimanté, est soumis à un champ magnétique $\vec{B}_0$ (que, dans tout ce complément, nous supposons uniforme et dirigé suivant l'axe Oz), il acquiert une *aimantation* $\vec{M}$, c'est-à-dire un *moment magnétique moyen par unité de volume,* parallèle à la direction du champ appliqué. C'est cette aimantation que nous allons chercher à déterminer dans le cadre de modèles microscopiques simples.

1. FORMULATION DU PROBLÈME DANS L'ENSEMBLE CANONIQUE

Soit donc un système en équilibre avec un thermostat à la température T et soumis au champ magnétique $\vec{B}_0$ dirigé suivant Oz. Les énergies E_ℓ des différents états microscopiques (ℓ) du système sont des fonctions continues de B_0 :

$$E_\ell = E_\ell(B_0) \,. \tag{A.1}$$

Lorsqu'il est dans l'état (ℓ), le système présente un moment magnétique $\mathcal{M}_\ell$, « force conjuguée » de B_0 (cf. complément II.D, § I.1) :

$$\mathcal{M}_\ell \widehat{=} -\frac{\partial E_\ell}{\partial B_0} \,. \tag{A.2}$$

La probabilité P_ℓ pour que, à l'équilibre, le système soit dans l'état (ℓ) (où son moment magnétique est $\mathcal{M}_\ell$) est (III, § I.B.1)

$$P_\ell = \frac{1}{Z} \, \mathrm{e}^{-E_\ell(B_0)/kT} \,, \tag{A.3}$$

où

$$Z \widehat{=} \sum_{(\ell)} \mathrm{e}^{-E_\ell(B_0)/kT} \tag{A.4}$$

est la fonction de partition du système, fonction des paramètres extérieurs T et B_0.

L'aimantation vaut donc

$$M \widehat{=} \frac{1}{V}\overline{\mathcal{M}} = \frac{1}{VZ}\sum_{(\ell)} \mathcal{M}_\ell \, e^{-E_\ell(B_0)/kT} . \tag{A.5}$$

L'expression de M en fonction de Z ou de l'énergie libre F s'obtient facilement en utilisant la relation (A.2) :

$$M = \frac{kT}{V}\frac{\partial \ln Z}{\partial B_0} = -\frac{1}{V}\frac{\partial F}{\partial B_0} . \tag{A.6}$$

La susceptibilité magnétique

$$\chi \widehat{=} \lim_{B_0 \to 0} \frac{\partial M}{\partial B_0} \tag{A.7}$$

s'écrit, elle aussi, à partir de Z ou F :

$$\chi = \frac{kT}{V}\lim_{B_0 \to 0}\frac{\partial^2 \ln Z}{\partial B_0^2} = -\frac{1}{V}\lim_{B_0 \to 0}\frac{\partial^2 F}{\partial B_0^2} . \tag{A.8}$$

Les relations (A.6) et (A.8) sont tout à fait générales : elles s'appliquent à n'importe quel système à l'équilibre en présence d'un champ magnétique. Dans ce complément, nous nous limitons au cas simple de systèmes constitués de particules indépendantes. Le cas de systèmes dans lesquels l'interaction entre constituants n'est pas négligeable sera traité au complément III.J.

2. PARAMAGNÉTISME ET DIAMAGNÉTISME

a. Définition

L'aimantation $\vec{M}$ est en général parallèle au champ appliqué $\vec{B}_0$; mais, suivant les matériaux, elle peut être de même sens que $\vec{B}_0$ ou de sens contraire. Si l'aimantation est de *même sens* que le champ magnétique, le matériau est dit *paramagnétique;* si elle est en *sens opposé,* il est *diamagnétique.*

Aux faibles valeurs de B_0, M est en général proportionnelle au champ appliqué et la susceptibilité magnétique (A.7) s'écrit simplement

$$\chi = \lim_{B_0 \to 0}\frac{M}{B_0} , \tag{A.9}$$

M et B_0 étant mesurés *algébriquement* sur l'axe Oz. Pour les corps paramagnétiques, la susceptibilité χ est donc positive; elle est négative pour les matériaux diamagnétiques[1].

b. Origine physique

Si les éléments constitutifs du matériau étudié (atomes, molécules ou ions) possèdent un *moment magnétique* $\vec{\mu}$ *permanent,* ce matériau est paramagnétique. En

1. Avec la définition (A.7), la susceptibilité χ a les dimensions d'une aimantation (ampère par mètre) divisée par un champ magnétique (tesla), c'est-à-dire que χ s'exprime en mètre par henry (mH^{-1}). En pratique, on utilise la grandeur sans dimension

$$\tilde{\chi} \widehat{=} \mu_0 \chi ,$$

où $\mu_0 = 4\pi . 10^{-7}$ Hm^{-1} est la perméabilité du vide. C'est directement $\tilde{\chi}$ qui intervient si, comme dans de nombreux ouvrages, on utilise le système d'unités C.G.S.

Pour les corps diamagnétiques, $|\tilde{\chi}_d|$ est toujours très faible ($\sim 10^{-5}$-10^{-6} pour les solides et les liquides) et pratiquement indépendant de la température. La susceptibilité des paramagnétiques est beaucoup plus importante ($\tilde{\chi}_p \sim 10^{-2}$ à température ordinaire) et c'est une fonction décroissante de la température.

effet, en l'absence de champ extérieur, les moments permanents sont répartis de façon aléatoire et l'aimantation résultante est nulle; mais lorsqu'on applique un champ $\vec{B}_0$, les moments $\vec{\mu}$ ont tendance à s'orienter dans le sens de $\vec{B}_0$ (malgré l'agitation thermique), créant ainsi une aimantation parallèle au champ appliqué et de même sens.

D'autre part, lorsqu'on soumet un matériau quelconque à un champ magnétique, les électrons qu'il contient modifient leur mouvement de façon à s'opposer à l'établissement du champ (loi de Lenz). Ces modifications vont donc *induire des moments magnétiques* en sens inverse du champ appliqué, donnant lieu au *diamagnétisme.*

Le diamagnétisme est donc un phénomène général, qui existe dans tous les matériaux. Dans le cas où les éléments du système portent des moments permanents, le paramagnétisme se superpose au diamagnétisme; en général, l'aimantation induite par l'effet diamagnétique est beaucoup plus faible, en valeur absolue, que celle due au paramagnétisme, de sorte que le matériau est dans ce cas globalement paramagnétique.

II. ÉTUDE STATISTIQUE DU PARAMAGNÉTISME

1. HAMILTONIEN DU SYSTÈME

On considère un système constitué de N ($\gg 1$) atomes (ou ions) fixés aux nœuds d'un réseau cristallin, en équilibre avec un thermostat à la température T. En l'absence de champ magnétique, les atomes sont pratiquement tous dans leur état fondamental; en effet, les énergies d'excitation électronique sont de l'ordre de l'eV et, sauf à des températures très élevées, la probabilité d'excitation d'un atome est tout à fait négligeable (rappelons qu'il faut $T \approx 10\,000$ K pour que kT atteigne des valeurs comparables à l'eV).

Soit $\vec{J}$ le moment cinétique total, somme des moments cinétiques orbitaux et de spin des divers électrons d'un atome dans son état fondamental (nous mesurons les moments cinétiques en unités de $\hbar$). On sait (C.D.L., p. 1045) qu'à un tel moment cinétique est associé un moment magnétique

$$\vec{\mu} = -g\mu_B \vec{J}\,, \tag{A.10}$$

où
$$\mu_B \mathrel{\widehat{=}} \frac{e\hbar}{2m} \simeq 9{,}27 . 10^{-24}\ \mathrm{A.m^2} \tag{A.11}$$

est le magnéton de Bohr et g le «facteur de Landé», de l'ordre de l'unité[2]. J ne peut prendre que des valeurs entières ou demi-entières.

Si $J \neq 0$, chaque atome possède donc un moment magnétique permanent qui, placé dans un champ magnétique, est soumis à l'hamiltonien

$$H = -\vec{\mu} . \vec{B}_0 = g\mu_B J_z B_0\,, \tag{A.12}$$

où J_z est la projection du moment cinétique $\vec{J}$ sur l'axe Oz. Les valeurs propres de J_z sont au nombre de $(2J+1)$:

$$J_z \in \{J,\ J-1,\ J-2,\ \ldots,\ -(J-1),\ -J\}\ ; \tag{A.13}$$

---

2. Si L est la valeur du moment cinétique orbital total, S celle du spin total, le facteur de Landé vaut (C.D.L., p. 1048)

$$g = \frac{3}{2} + \frac{S(S+1) - L(L+1)}{2J(J+1)}\ ;$$

les moments cinétiques L, S et J sont ici des grandeurs sans dimensions, le facteur $\hbar$ ayant été incorporé dans μ_B.

la dégénérescence de l'état fondamental de l'atome, égale à $2J+1$, est donc levée par la présence du champ magnétique : chaque atome possède alors $(2J+1)$ niveaux régulièrement espacés entre l'énergie $-g\mu_B JB_0$ et $+g\mu_B JB_0$.

Si les atomes sont suffisamment éloignés les uns des autres, les interactions entre moments magnétiques sont négligeables (voir complément III.J, § I.4); ils sont alors pratiquement indépendants, et les propriétés magnétiques du système peuvent être calculées à partir de l'hamiltonien modèle

$$H = \sum_{i=1}^{N} -\vec{\mu}_i \cdot \vec{B}_0 = \sum_{i=1}^{N} g\mu_B J_z^{(i)} B_0 \,. \tag{A.14}$$

2. FONCTION DE PARTITION ET ÉNERGIE LIBRE

Pour ce système de N particules indépendantes, la fonction de partition Z se factorise (III, § IV.A.2) :

$$Z = z^N \,, \tag{A.15}$$

où z est la fonction de partition d'un atome; celui-ci possédant $(2J+1)$ états, d'énergies $g\mu_B J_z B_0$, on a

$$z = \sum_{J_z=-J}^{+J} \exp\left(-\frac{g\mu_B B_0}{kT} J_z\right) . \tag{A.16}$$

Traitons tout d'abord le cas très simple où $J = 1/2$. Chaque atome n'a alors que deux états possibles ($J_z = +1/2$ et $J_z = -1/2$), et la fonction de partition correspondante s'écrit

$$z_{1/2} = 2 \operatorname{ch} \frac{g\mu_B B_0}{2kT} \,. \tag{A.17}$$

On en déduit immédiatement la fonction de partition $Z_{1/2}$ du système et son énergie libre :

$$\begin{aligned} F_{1/2} &\mathrel{\widehat{=}} -kT \ln Z_{1/2} = -NkT \ln z_{1/2} \\ &= -NkT \ln\left[2 \operatorname{ch} \frac{g\mu_B B_0}{2kT}\right] . \end{aligned} \tag{A.18}$$

Le cas général est à peine plus compliqué. Si l'on remarque que z est la somme de $(2J+1)$ termes qui forment une progression géométrique de raison $\exp(-g\mu_B B_0/kT)$, il vient[3]

$$z_J = \frac{\operatorname{sh}\left[\frac{g\mu_B B_0}{kT}\left(J+\frac{1}{2}\right)\right]}{\operatorname{sh}\frac{g\mu_B B_0}{2kT}} \,, \tag{A.19}$$

---

3. Une progression géométrique de raison r comprenant n termes dont le premier vaut c a pour somme

$$\sum_{j=1}^{n} c r^{j-1} = c\,\frac{1-r^n}{1-r} \,. \tag{1}$$

Appliquée à l'expression (A.16), cette formule donne

$$z = e^{\alpha J}\,\frac{1-e^{-(2J+1)\alpha}}{1-e^{-\alpha}} \,, \tag{2}$$

où $$\alpha \mathrel{\widehat{=}} \frac{g\mu_B B_0}{kT} \,. \tag{3}$$

.../...

et donc

$$F_J = -NkT \ln \frac{\operatorname{sh}\left[\frac{g\mu_B B_0}{kT}\left(J+\frac{1}{2}\right)\right]}{\operatorname{sh}\frac{g\mu_B B_0}{2kT}} . \tag{A.20}$$

3. PROPRIÉTÉS DU SYSTÈME

Comme toujours, la connaissance de la fonction de partition ou de l'énergie libre en fonction des paramètres extérieurs (ici T et B_0) permet de déterminer les propriétés du système.

a. Aimantation et susceptibilité magnétique

L'aimantation se calcule à l'aide de la relation (A.6). Pour $J = 1/2$, on obtient[4] à partir de (A.18)

$$M_{1/2} = \frac{Ng\mu_B}{2V} \operatorname{th} \frac{g\mu_B B_0}{2kT} . \tag{A.21}$$

Elle ne dépend des paramètres extérieurs T et B_0 que par le rapport

$$x_{1/2} = \frac{g\mu_B B_0}{2kT} . \tag{A.22}$$

Lorsque $x \to +\infty$, $\operatorname{th} x \to 1$: à fort champ magnétique ou à basse température,

$$M_{1/2} \simeq \frac{Ng\mu_B}{2V} \quad \text{si} \quad \frac{g\mu_B B_0}{2kT} \gg 1 ; \tag{A.23}$$

le système est alors dans son état d'aimantation maximum, tous ses moments magnétiques ($|\vec{\mu}| = g\mu_B/2$) pointant dans la direction du champ. Dans la limite inverse ($x_{1/2} \ll 1$, $\operatorname{th} x_{1/2} \sim x_{1/2}$) des faibles champs magnétiques ou des hautes températures, l'aimantation est proportionnelle au champ appliqué :

$$M_{1/2} \simeq \frac{N}{V}\left(\frac{g\mu_B}{2}\right)^2 \frac{B_0}{kT} \quad \text{si} \quad \frac{g\mu_B B_0}{2kT} \ll 1 ; \tag{A.24}$$

En multipliant haut et bas par $\exp(\alpha/2)$, on obtient

$$z = e^{\alpha\left(J+\frac{1}{2}\right)} \frac{1-e^{-(2J+1)\alpha}}{e^{\alpha/2}-e^{-\alpha/2}} = \frac{e^{\alpha\left(J+\frac{1}{2}\right)}-e^{-\alpha\left(J+\frac{1}{2}\right)}}{e^{\alpha/2}-e^{-\alpha/2}} , \tag{4}$$

d'où la formule (A.19).

4. On peut aussi calculer $M_{1/2}$ directement. La probabilité pour que le moment magnétique $\vec{\mu}$ d'un atome pointe dans la direction du champ vaut en effet

$$P_\uparrow = \frac{e^{\mu B_0/kT}}{e^{\mu B_0/kT}+e^{-\mu B_0/kT}} , \tag{1}$$

celle qu'il pointe en sens inverse

$$P_\downarrow = \frac{e^{-\mu B_0/kT}}{e^{\mu B_0/kT}+e^{-\mu B_0/kT}} . \tag{2}$$

Le moment magnétique moyen d'un atome est donc

$$\bar{\mu} = \mu P_\uparrow - \mu P_\downarrow = \mu \operatorname{th} \frac{\mu B_0}{kT} . \tag{3}$$

Pour N atomes dans le volume V, on obtient le résultat (A.21) si l'on remplace μ par sa valeur $g\mu_B/2$.

ceci correspond à une susceptibilité magnétique

$$\chi_{1/2} = \lim_{B_0 \to 0} \frac{M_{1/2}}{B_0} = \frac{N}{V} \left(\frac{g\mu_B}{2}\right)^2 \cdot \frac{1}{kT} \,. \tag{A.25}$$

Dans le cas général (J quelconque), les calculs sont un peu plus lourds, mais les conclusions restent qualitativement les mêmes. On obtient

$$M_J = \frac{N}{V} g\mu_B J \, B_J \left(\frac{g\mu_B J B_0}{kT}\right) , \tag{A.26}$$

$B_J(x)$ étant la *«fonction de Brillouin d'ordre J»*, définie par

$$B_J(x) \widehat{=} \frac{2J+1}{2J} \coth\left(\frac{2J+1}{2J}\, x\right) - \frac{1}{2J} \coth \frac{x}{2J} \,. \tag{A.27}$$

La figure 1 donne l'allure de cette fonction pour quelques valeurs de J. Ici encore, l'aimantation ne dépend de T et B_0 que par le rapport

$$x_J \widehat{=} g\mu_B \frac{JB_0}{kT} = \frac{\mu B_0}{kT} \,. \tag{A.28}$$

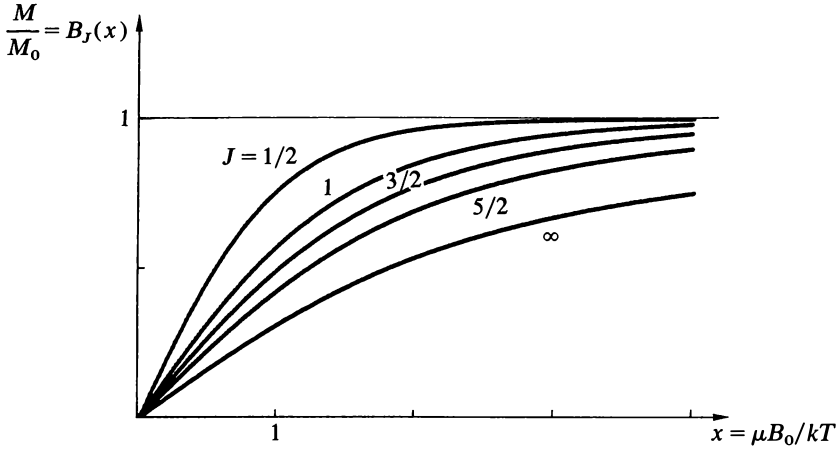


FIGURE 1
Aimantation d'un paramagnétique, rapportée à sa valeur maximum $M_0 = N\mu/V$, en fonction du paramètre $\mu B_0/kT$ pour plusieurs valeurs de J. On notera que $B_{1/2}(x) \equiv \text{th}\, x$ et $B_\infty(x) \equiv \mathcal{L}(x) \widehat{=} \coth x - 1/x$.

Quand $x \to \infty$, $\coth x \to 1$ et $B_J(x) \to 1$: à basse température ou à fort champ magnétique, le système est dans l'état d'aimantation maximum

$$M_J \simeq \frac{N}{V} g\mu_B J = \frac{N\mu}{V} \quad \text{si} \quad \frac{\mu B_0}{kT} \gg 1 \,. \tag{A.29}$$

Dans la limite inverse[5],

$$B_J(x) \simeq \frac{J+1}{3J}\, x \quad \text{si} \quad x \ll 1 \,, \tag{A.30}$$

5. Pour $x \ll 1$,

$$\coth x = \frac{1}{x} + \frac{x}{3} + \dots \,. \tag{1}$$

.../...

et l'aimantation est proportionnelle au champ appliqué :

$$M_J \simeq \frac{N}{V}(g\mu_B)^2 \frac{J(J+1)}{3}\frac{B_0}{kT} \quad \text{si} \quad \frac{\mu B_0}{kT} \ll 1 . \tag{A.31}$$

On en déduit la susceptibilité :

$$\chi_J = \frac{N}{V}(g\mu_B)^2 \frac{J(J+1)}{3kT} . \tag{A.32}$$

Elle varie en raison inverse de la température; on retrouve ainsi la *loi établie expérimentalement par P. Curie* en 1895 :

$$\chi = \frac{\mathcal{C}}{T} , \tag{A.33}$$

la « constante de Curie » $\mathcal{C}$ étant ici déterminée par le modèle :

$$\mathcal{C} = \frac{N}{V}(g\mu_B)^2 \frac{J(J+1)}{3k} . \tag{A.34}$$

Les valeurs de $\mathcal{C}$ calculées à l'aide de la formule précédente sont en bon accord avec les valeurs mesurées pour de nombreux sels comportant des ions paramagnétiques, et tout particulièrement pour ceux qui contiennent des ions de terres rares (dont la couche f, incomplètement remplie, est à l'origine des valeurs $J \neq 0$)[6].

b. Énergie et capacité calorifique magnétiques

L'énergie moyenne $\overline{E}$ s'obtient à partir de Z (formule III (II.8)) :

$$\overline{E} = -\frac{\partial \ln Z}{\partial \beta} = -N\frac{\partial \ln z}{\partial \beta} , \quad \beta = 1/kT . \tag{A.35}$$

Pour $J = 1/2$, il vient

$$\overline{E}_{1/2} = -\frac{Ng\mu_B B_0}{2} \operatorname{th} \frac{g\mu_B B_0}{2kT} , \tag{A.36}$$

et dans le cas général

$$\overline{E}_J = -NgJ\mu_B B_0 \, B_J\left(\frac{g\mu_B J B_0}{kT}\right) . \tag{A.37}$$

On retrouve la valeur moyenne de l'hamiltonien (A.14) du système :

$$\overline{E} = -VMB_0 . \tag{A.38}$$

A cette énergie magnétique est associée une capacité calorifique. Pour ne pas compliquer inutilement, nous ne donnons son expression que pour $J = 1/2$:

$$C_{1/2} \hat{=} \frac{\partial \overline{E}_{1/2}}{\partial T} = Nk\left[\frac{g\mu_B B_0}{2kT}\right]^2 \frac{1}{\operatorname{ch}^2 \dfrac{g\mu_B B_0}{2kT}} . \tag{A.39}$$

On reporte ce développement dans (A.27), les termes en $1/x$ disparaissent et on obtient finalement

$$B_J(x) = \frac{J+1}{3J} x + \ldots . \tag{2}$$

6. Pour les sels contenant des ions d'un métal de transition, l'accord est bon, à condition de ne tenir compte que du moment cinétique de spin ($J = S$). Pour plus de détails sur les comparaisons avec les résultats expérimentaux, voir les références 2 et 3.

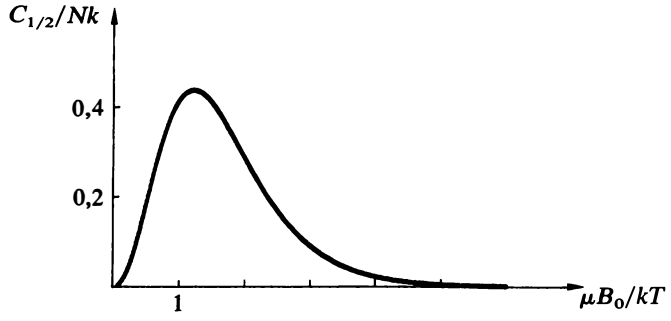


FIGURE 2
Capacité calorifique paramagnétique (pour des particules de moment cinétique 1/2).

$C_{1/2}$ est tracée sur la figure 2, en fonction du paramètre $\mu B_0/kT$.

Évidemment, pour un véritable système physique, l'énergie et la capacité calorifique totales s'obtiennent en ajoutant à $\bar{E}$ et C les contributions des autres degrés de liberté du système (supposés indépendants des moments magnétiques), et tout particulièrement celles qui proviennent des vibrations du réseau cristallin (complément III.E).

c. Entropie

L'entropie magnétique $S(T, B_0)$ se déduit elle aussi de l'énergie libre F (formule III (II.26)) :

$$S(T, B_0) = \frac{\bar{E} - F}{T} . \tag{A.40}$$

Par simple inspection des formules (A.20) et (A.37) on constate que, comme l'aimantation, l'entropie ne dépend que du rapport B_0/T :

$$S(T, B_0) \equiv S(B_0/T) . \tag{A.41}$$

Pour $J = 1/2$, on obtient explicitement

$$S_{1/2}(B_0/T) = Nk \left[\ln \left(2 \operatorname{ch} \frac{g\mu_B B_0}{2kT} \right) - \frac{g\mu_B B_0}{2kT} \operatorname{th} \frac{g\mu_B B_0}{2kT} \right] . \tag{A.42}$$

On vérifie aisément[7] que

$$\lim_{B_0/T \to \infty} S_{1/2}(B_0/T) = 0 : \tag{A.43}$$

à température nulle ou en champ infini, tous les moments magnétiques pointent dans

---

7. L'entropie s'écrit simplement en fonction de $x_{1/2}$ défini en (A.22) :

$$S_{1/2}(x_{1/2}) = Nk \left[\ln (2 \operatorname{ch} x_{1/2}) - x_{1/2} \operatorname{th} x_{1/2} \right] . \tag{1}$$

Or $$2 \operatorname{ch} x = e^x + e^{-x} = e^x [1 + e^{-2x}] , \tag{2}$$

et $$\operatorname{th} x = \frac{e^x - e^{-x}}{e^x + e^{-x}} = \frac{1 - e^{-2x}}{1 + e^{-2x}} . \tag{3}$$

On en déduit que, pour $x_{1/2} \gg 1$,

$$\ln(2 \operatorname{ch} x_{1/2}) \simeq x_{1/2} + e^{-2x_{1/2}} , \tag{4}$$

$$x_{1/2} \operatorname{th} x_{1/2} \simeq x_{1/2} - 2x_{1/2} e^{-2x_{1/2}} . \tag{5}$$

.../...

la même direction, celle du champ; le système est parfaitement ordonné et l'entropie correspondante nulle. Dans la limite inverse, on trouve[8]

$$\lim_{B_0/T \to 0} S_{1/2}(B_0/T) = Nk \ln 2 \,. \tag{A.44}$$

Ainsi, à température infinie ou en champ nul, le système est complètement désordonné : les deux directions de chaque moment individuel sont également probables; il en est donc de même des 2^N états possibles du système, ce qui conduit à la valeur maximum (A.44) de l'entropie[9].

Nous terminerons ce paragraphe en signalant que la forme (A.41) de l'entropie est à la base d'une méthode d'obtention de très basses températures par le procédé de *désaimantation adiabatique.* Considérons un système paramagnétique en équilibre à la température T_i et dans le champ B_{0i}. Isolons-le thermiquement et diminuons lentement le champ de façon qu'à chaque instant le système se trouve dans un état d'équilibre : on réalise ainsi une *transformation adiabatique réversible.* Or on sait (complément II.D, § III.3.c) que l'entropie reste constante au cours d'une telle transformation. Donc, si le champ est B_{0f} à la fin de la transformation, la température correspondante T_f sera telle que

$$S\left(\frac{B_{0i}}{T_i}\right) = S\left(\frac{B_{0f}}{T_f}\right) , \tag{A.45}$$

c'est-à-dire[10]

$$T_f = T_i \frac{B_{0f}}{B_{0i}} \,. \tag{A.46}$$

Cette technique exige que l'on parte de températures initiales suffisamment basses pour que l'entropie magnétique soit la contribution dominante à l'entropie totale du système. On a pu ainsi descendre des environs de 1 K à quelques 10^{-3} K.

En reportant ces développements dans (1), on trouve

$$S_{1/2}(x_{1/2}) \simeq Nke^{-2x_{1/2}}(2x_{1/2}+1) \quad \text{si} \quad x_{1/2} \gg 1 \,; \tag{6}$$

d'où le résultat (A.43).

8. Pour $x_{1/2} \ll 1$,

$$2 \operatorname{ch} x_{1/2} \simeq 2\left(1+\frac{x_{1/2}^2}{2}\right) , \tag{1}$$

$$x_{1/2} \operatorname{th} x_{1/2} \simeq x_{1/2}^2 \,. \tag{2}$$

En reportant dans la formule (1) de la note précédente, on obtient

$$S_{1/2}(x_{1/2}) \simeq Nk\left[\ln 2 - \frac{x_{1/2}^2}{2}\right] \quad \text{si} \quad x_{1/2} \ll 1 \,. \tag{3}$$

9. Les résultats précédents se généralisent à J quelconque :

$$\lim_{B_0/T \to \infty} S(B_0/T) = 0 \,, \tag{1}$$

et

$$\lim_{B_0/T \to 0} S(B_0/T) = Nk \ln (2J+1) \,. \tag{2}$$

10. Contrairement à ce que pourrait laisser croire la relation (A.46), il n'est pas possible d'atteindre par cette méthode (comme d'ailleurs par n'importe quelle autre) une température nulle. En effet, la relation (A.46) suppose que l'entropie est uniquement due aux moments magnétiques et que ceux-ci sont indépendants. Quand le champ appliqué devient très faible, il faut prendre en compte les contributions des autres degrés de liberté à l'entropie. En outre, les interactions entre moments magnétiques ne sont plus négligeables à très basse température, et la formule (A.41) n'est plus strictement valable.

III. ÉTUDE STATISTIQUE DU DIAMAGNÉTISME

Nous n'étudierons le diamagnétisme que dans deux cas relativement simples. Le premier est celui de cristaux isolants constitués d'atomes ou d'ions dont les couches électroniques sont complètes; il n'y a alors aucun moment magnétique permanent et le diamagnétisme est seul à se manifester. Nous considérerons ensuite l'exemple, plus intéressant du point de vue de la mécanique statistique, du gaz dilué d'électrons libres : ici, chaque électron porte un moment magnétique associé à son spin, de sorte que l'effet diamagnétique est en compétition avec l'effet paramagnétique; nous montrerons que dans ce cas, bien que le paramagnétisme reste prépondérant, les deux effets sont du même ordre de grandeur.

1. DIAMAGNÉTISME DES MATÉRIAUX A COUCHES ÉLECTRONIQUES COMPLÈTES

On considère un système de N atomes ou ions ($N \gg 1$) fixés aux nœuds d'un réseau cristallin, en contact avec un thermostat à la température T et soumis au champ magnétique B_0 dirigé suivant Oz. Les couches électroniques de chaque atome sont complètes; il s'agit donc d'atomes de gaz rares ou d'ions d'un cristal ionique (sel cristallisé, tel NaCl). Les moments cinétiques orbital et de spin de chaque atome ou ion (dans son état fondamental) sont alors nuls :

$$L = S = J = 0 \,. \tag{A.47}$$

Les particules constituant le système ne possèdent donc pas de moment magnétique permanent et l'effet diamagnétique apparaîtra seul. Pour l'étudier, il faut tout d'abord expliciter la perturbation apportée sur chaque atome par le champ magnétique.

a. Hamiltonien d'un atome en présence d'un champ magnétique

L'hamiltonien d'un atome ou d'un ion placé dans un champ magnétique $\vec{B}_0$ s'écrit (C.D.L., appendice III, § 4.b)

$$H = \sum_i \left[\frac{1}{2m} \left(\vec{p}_i + e\vec{A}(\vec{r}_i) \right)^2 + 2\mu_B \vec{B}_0 . \vec{S}_i \right] + U(\{\vec{r}_i\}) \,, \tag{A.48}$$

où $\vec{A}(\vec{r})$ est le potentiel-vecteur associé au champ $\vec{B}_0$:

$$\overrightarrow{\text{rot}}\, \vec{A} = \vec{B}_0 \,, \tag{A.49}$$

U est l'énergie potentielle d'interaction des électrons entre eux et avec le noyau atomique, et $\vec{S}_i$ le spin du i-ième électron (de masse m et de charge $-e$).

Pour un champ $\vec{B}_0$ uniforme dans la direction Oz, on peut choisir comme potentiel-vecteur

$$A_x = -\frac{1}{2} y B_0, \quad A_y = \frac{1}{2} x B_0, \quad A_z = 0 \,. \tag{A.50}$$

En reportant cette expression dans (A.48) et en développant le carré, on obtient

$$H = H_0 + \mu_B (L_z + 2S_z) B_0 + \frac{e^2}{8m} B_0^2 \sum_i (x_i^2 + y_i^2) \,, \tag{A.51}$$

où

$$H_0 = \sum_i \frac{p_i^2}{2m} + U \tag{A.52}$$

est l'hamiltonien en l'absence de champ magnétique, L_z la composante suivant Oz du moment cinétique orbital de l'atome

$$\vec{L} = \frac{1}{\hbar} \sum_i \vec{r}_i \wedge \vec{p}_i \,, \tag{A.53}$$

.../...

et S_z la composante sur le même axe du spin total

$$\vec{S} = \sum_i \vec{S}_i \,. \tag{A.54}$$

A l'échelle des excitations électroniques dans l'atome, les corrections à H_0 restent petites; on peut les traiter par la méthode des perturbations (C.D.L., p. 1091). Pour l'état fondamental, le deuxième terme de (A.51) est sans effet[11] puisque, dans cet état, $L_z = S_z = 0$. L'énergie E_0 du niveau fondamental est donc déplacée par un terme proportionnel au carré du champ :

$$E_0(B_0) = E_0 + \frac{e^2 B_0^2}{8m} \langle 0 | \sum_i (x_i^2 + y_i^2) | 0 \rangle \,, \tag{A.55}$$

où $|0\rangle$ note l'état fondamental de l'atome ou de l'ion en l'absence de champ.

Dans son état fondamental, l'atome est à symétrie sphérique; si r_i est la distance du i-ième électron au noyau, on a donc

$$\langle 0 | \sum_i x_i^2 | 0 \rangle = \langle 0 | \sum_i y_i^2 | 0 \rangle = \langle 0 | \sum_i z_i^2 | 0 \rangle = \frac{1}{3} \langle 0 | \sum_i r_i^2 | 0 \rangle \,. \tag{A.56}$$

Finalement,

$$E_0(B_0) = E_0 + \frac{e^2 B_0^2}{12m} \langle 0 | \sum_i r_i^2 | 0 \rangle \,. \tag{A.57}$$

b. Aimantation et susceptibilité diamagnétiques

Sauf à de très hautes températures (auxquelles, en général, le cristal n'existe plus), la probabilité d'exciter un atome ou un ion en dehors de son état électronique fondamental est complètement négligeable. Chaque particule a donc pratiquement un seul état accessible, d'énergie $E_0(B_0)$; l'énergie libre du système se confond alors avec son énergie[12] :

$$F(T, B_0) = N \left[E_0 + \frac{e^2 B_0^2}{12m} \langle 0 | \sum_i r_i^2 | 0 \rangle \right] . \tag{A.58}$$

Elle est indépendante de la température, et donc aussi l'aimantation

$$M = -\frac{1}{V} \frac{\partial F}{\partial B_0} = -\frac{e^2 B_0 N}{6mV} \langle 0 | \sum_i r_i^2 | 0 \rangle \,, \tag{A.59}$$

et la susceptibilité

$$\chi = -\frac{e^2 N}{6mV} \langle 0 | \sum_i r_i^2 | 0 \rangle \,. \tag{A.60}$$

La mécanique statistique ne joue donc ici aucun rôle. Comme prévu, χ est

---

11. C'est au contraire ce terme qui conduit à l'hamiltonien (A.12) et donc au paramagnétisme quand $J \neq 0$; le troisième terme, du second ordre par rapport au champ magnétique, est alors négligeable devant lui (C.D.L., p. 841).

12. C'est évident puisque, s'il n'y a qu'un état microscopique possible pour chacun des atomes, l'entropie du système est nulle. On retrouve évidemment ce résultat à partir de la fonction de partition, qui s'écrit alors

$$Z = (e^{-E_0(B_0)/kT})^N \,.$$

négative : le système est diamagnétique[13]. Lorsqu'il s'agit d'un cristal ionique, il comporte des ions de plusieurs espèces; chacune d'elles donne à M et à χ une contribution de la forme (A.59) et (A.60); l'aimantation et la susceptibilité globales sont alors obtenues en sommant les contributions des divers types d'ions.

2. DIAMAGNÉTISME DU GAZ (DILUÉ) D'ÉLECTRONS LIBRES

Nous considérons maintenant un gaz constitué de N électrons enfermés dans une boîte cubique de volume V, à la température T et soumis au champ $\vec{B}_0$. Si on néglige les interactions entre électrons, l'hamiltonien (cf. (A.48)) de ce système s'écrit

$$H = \sum_{i=1}^{N} \frac{1}{2m} \left(\vec{p}_i + e\vec{A}(\vec{r}_i)\right)^2 + \sum_{i=1}^{N} 2\mu_B S_{iz} B_0 \,. \tag{A.61}$$

La deuxième somme, qui fait intervenir explicitement le moment magnétique $2\mu_B \vec{S}_i$ associé au spin de chaque électron, donne lieu au paramagnétisme tel que nous l'avons calculé au paragraphe II; il lui correspond l'aimantation (A.21) et la susceptibilité (A.25) (le facteur de Landé g d'un électron libre est égal à 2). La première somme, c'est-à-dire le *terme orbital,* donne un effet *diamagnétique;* c'est le seul que nous analyserons ici, l'aimantation totale s'obtenant ensuite par simple addition (algébrique) des deux effets[14].

On montre[15] que les niveaux d'énergie ε, appelés *niveaux de Landau,* d'une

---

13. Il est facile d'obtenir une estimation de $\tilde{\chi} = \mu_0 \chi$ (cf. note 1). Si Z est le nombre d'électrons dans l'atome (ou l'ion),

$$\sum_i \langle 0|r_i^2|0\rangle \sim Za_0^2 \,, \tag{1}$$

où a_0 est le rayon de Bohr; donc

$$\tilde{\chi} \sim -\frac{\mu_0 e^2}{6m} \frac{Za_0^2}{d^3} \,, \tag{2}$$

où d est la distance interatomique. Pour $d \sim 5a_0$, on trouve

$$\tilde{\chi} \sim -4.10^{-6} Z \,. \tag{3}$$

Pour des valeurs expérimentales précises, voir références 2 et 3.

14. L'hamiltonien (A.61) est en effet la somme de deux termes indépendants (c'est-à-dire qui commutent l'un avec l'autre). La fonction de partition totale est donc le produit des fonctions de partition associées aux deux termes et l'énergie libre totale la somme des deux énergies libres.

15. Si l'on choisit pour potentiel vecteur

$$A_x = 0,\ A_y = B_0 x,\ A_z = 0 \,, \tag{1}$$

l'équation de Schrödinger associée à l'hamiltonien (A.62) s'écrit

$$-\frac{\hbar^2}{2m}\frac{\partial^2\psi}{\partial x^2} - \frac{\hbar^2}{2m}\frac{\partial^2\psi}{\partial z^2} + \frac{1}{2m}\left[-i\hbar\frac{\partial}{\partial y} + eB_0 x\right]^2 \psi = \varepsilon\psi \,. \tag{2}$$

Cherchons des fonctions propres de la forme

$$\psi(x, y, z) = e^{i(k_y y + k_z z)} \varphi(x) \,. \tag{3}$$

Il vient

$$-\frac{\hbar^2}{2m}\varphi'' + \frac{e^2 B_0^2}{2m}\left(x + \frac{\hbar k_y}{eB_0}\right)^2 \varphi = \left(\varepsilon - \frac{\hbar^2 k_z^2}{2m}\right)\varphi \,. \tag{4}$$

.../...

particule confinée dans une boîte cubique d'arête a et dont le mouvement est décrit par l'hamiltonien

$$H_{\text{orb}}=\frac{1}{2m}\left[\vec{p}+e\vec{A}(\vec{r})\right]^2 \tag{A.62}$$

dépendent de deux nombres quantiques, k_z et n :

$$\varepsilon_n(k_z)=\frac{\hbar^2k_z^2}{2m}+(2n+1)\mu_B B_0\ ; \tag{A.63}$$

le vecteur d'onde k_z dans la direction du champ magnétique peut prendre les mêmes valeurs qu'en l'absence de champ, et n est un entier positif ou nul :

$$k_z=\frac{2\pi}{a}n_z \quad \text{avec } n_z=0,\ \pm1,\ \pm2,\ \ldots, \tag{A.64.a}$$

$$n=0,\ 1,\ 2,\ 3,\ \ldots \tag{A.64.b}$$

Cette équation n'est autre que l'équation de Schrödinger pour une particule de masse m, se déplaçant à une dimension dans un potentiel harmonique centré au point

$$x_0=-\frac{\hbar k_y}{eB_0}\ . \tag{5}$$

La solution de ce problème est connue (appendice II, § II.3) : les valeurs propres sont de la forme

$$\varepsilon-\frac{\hbar^2k_z^2}{2m}=\left(n+\frac{1}{2}\right)\frac{\hbar eB_0}{m}\ , \tag{6}$$

avec n entier positif ou nul. Les énergies des niveaux de Landau s'écrivent donc

$$\varepsilon_n(k_z)=\frac{\hbar^2}{2m}k_z^2+(2n+1)\mu_B B_0\ . \tag{7}$$

Elles ne dépendent que des nombres quantiques n et k_z, pas de k_y.

Les conditions aux limites dans la boîte déterminent les valeurs permises de k_y et k_z; si on impose des conditions aux limites périodiques dans les directions y et z on obtient, comme pour une particule libre,

$$\left.\begin{aligned} k_y&=\frac{2\pi}{a}n_y\ ,\\ k_z&=\frac{2\pi}{a}n_z\ , \end{aligned}\right\}\quad \left.\begin{matrix} n_y\\ n_z \end{matrix}\right\}=0,\ \pm1,\ \pm2,\ \pm3\ldots\ . \tag{8}$$

Restent les conditions aux limites suivant l'axe Ox. Si a est grand devant l'extension des fonctions d'onde $\varphi_n(x-x_0)$ de l'oscillateur harmonique, il suffit d'imposer que x_0 se situe à l'intérieur de la boîte, c'est-à-dire

$$0<x_0<a\ . \tag{9}$$

En effet, les φ_n pour lesquelles x_0 se trouve nettement à l'intérieur de la boîte ne sont pas affectées par la présence des parois; inversement, il n'y a pas de solutions possibles si x_0 est nettement en dehors de la boîte. Pour les valeurs de x_0 proches de la surface, la présence des parois modifie les valeurs propres et les fonctions propres et le calcul précédent n'est pas valable; mais si l'extension des fonctions d'onde (de l'ordre du rayon de l'orbite d'une particule classique dans le même champ magnétique) est faible devant les dimensions de la boîte, cette région frontière peut être négligée (pour un électron d'énergie 1 eV, le rayon orbital dans un champ de 10^{-1} T vaut $\sim3\times10^{-5}$ m). La condition (9) détermine, compte tenu de (5) et (8), le nombre de valeurs permises pour k_y :

$$\frac{eB_0a}{\hbar}\cdot\frac{1}{2\pi/a}=\frac{eB_0a^2}{2\pi\hbar}\ , \tag{10}$$

c'est-à-dire le *nombre d'états* associés à chaque couple $(k_z,\ n)$.

Chacun de ces niveaux est hautement dégénéré; le nombre d'états est le même pour tous les couples (k_z, n) et vaut (note 15)

$$\omega(k_z, n) = \frac{eB_0 a^2}{2\pi\hbar} . \tag{A.65}$$

La fonction de partition « orbitale » d'un électron se calcule alors facilement[16] :

$$z_{\text{orb}} \widehat{=} \sum_{n,\, k_z} \omega(k_z, n)\, e^{-\varepsilon_n(k_z)/kT}$$

$$= V\left(\frac{mkT}{2\pi\hbar^2}\right)^{3/2} \frac{\mu_B B_0/kT}{\text{sh}\ \mu_B B_0/kT} . \tag{A.66}$$

Si on suppose que le gaz est suffisamment dilué (ou la température suffisamment haute) pour que l'approximation de Maxwell-Botzmann s'applique[17] (III, § IV.A.3.c), la fonction de partition orbitale des N électrons, supposés indépendants, s'écrit

$$Z_{\text{orb}} = \frac{1}{N!}(z_{\text{orb}})^N = \frac{V^N}{N!}\left(\frac{mkT}{2\pi\hbar^2}\right)^{\frac{3N}{2}} \left(\frac{\mu_B B_0/kT}{\text{sh}\ \mu_B B_0/kT}\right)^N . \tag{A.67}$$

L'aimantation diamagnétique s'en déduit immédiatement (formule (A.6)) :

$$M_d = \frac{kT}{V}\frac{\partial \ln Z_{\text{orb}}}{\partial B_0} = -\frac{N\mu_B}{V}\,\mathcal{L}\left(\frac{\mu_B B_0}{kT}\right), \tag{A.68}$$

où $\mathcal{L}(x)$ est la fonction de Langevin :

$$\mathcal{L}(x) \widehat{=} \coth x - \frac{1}{x} . \tag{A.69}$$

Remarquons (cf. formule (A.27)) que la fonction de Langevin est égale à la fonction de Brillouin d'ordre ∞ :

$$\mathcal{L}(x) \equiv B_\infty(x) . \tag{A.70}$$

---

16. Il faut sommer sur les valeurs permises de k_z et n, en tenant compte du facteur de dégénérescence (A.65) de chaque couple (k_z, n). Les valeurs de k_z étant très proches les unes des autres, on peut remplacer

$$\sum_{k_z} \quad \text{par} \quad \frac{a}{2\pi}\int_{-\infty}^{+\infty} dk_z .$$

D'autre part, à cause de la forme (A.63) de l'énergie, la somme sur n et l'intégrale sur k_z se factorisent :

$$z = \frac{eB_0 a^3}{(2\pi)^2\hbar}\int_{-\infty}^{+\infty} dk_z\, e^{-\hbar^2 k_z^2/2mkT} \times \sum_{n=0}^{\infty} e^{-(2n+1)\mu_B B_0/kT} . \tag{1}$$

L'intégrale est bien connue :

$$\int_{-\infty}^{+\infty} dk_z\, e^{-\hbar^2 k_z^2/2mkT} = \sqrt{\frac{2\pi mkT}{\hbar^2}} . \tag{2}$$

La somme sur n est celle d'une série géométrique de raison $\exp(-2\mu_B B_0/kT)$; elle donne donc (cf. note 3)

$$\sum_{n=0}^{\infty} e^{-(2n+1)\mu_B B_0/kT} = \frac{e^{-\mu_B B_0/kT}}{1 - e^{-2\mu_B B_0/kT}} = \frac{1}{2\ \text{sh}\ \mu_B B_0/kT} . \tag{3}$$

En reportant (2) et (3) dans (1), on obtient (A.66).

17. Il ne peut donc s'agir ici des électrons de conduction dans un métal : le gaz qu'ils constituent est trop dense (même pour des températures de plusieurs milliers de K) pour que l'approximation soit valable. Il faut dans ce cas prendre en compte le principe d'exclusion de Pauli et appliquer la statistique de Fermi-Dirac (voir exercice VI.18). On peut, par contre, appliquer le résultat obtenu ici aux électrons de conduction d'un semiconducteur (complément VI.A, § III).

Or $$\mathcal{L}(x) \simeq \frac{x}{3} \quad \text{pour } x \ll 1 \; ; \tag{A.71}$$

la susceptibilité diamagnétique a donc pour expression

$$\chi_d \mathrel{\widehat{=}} \lim_{B_0 \to 0} \frac{M_d}{B_0} = -\frac{N}{V}\frac{\mu_B^2}{3kT} \; . \tag{A.72}$$

Il faut comparer ces résultats aux formules (A.21) et (A.25) qui caractérisent (pour $g=2$) le paramagnétisme de ce même système (second terme de l'hamiltonien (A.61)). Les courbes de la figure 1 indiquent que l'effet paramagnétique l'emporte sur l'effet diamagnétique; mais ils sont tous deux du même ordre de grandeur. Globalement, le gaz est donc paramagnétique, avec une susceptibilité

$$\chi = \chi_{1/2} + \chi_d = \frac{2N}{V}\frac{\mu_B^2}{3kT} > 0 \; . \tag{A.73}$$

IV. AIMANTATION ET MÉCANIQUE CLASSIQUE

L'explication théorique du paramagnétisme et du diamagnétisme, accompagnée du calcul des aimantations et des susceptibilités, a été proposée en 1905 par P. Langevin, avant la découverte de la théorie quantique. Nous allons résumer ci-dessous ses raisonnements et montrer qu'ils reposent en fait sur des arguments que l'on peut qualifier de « semi-classiques ». Une application correcte de la théorie dans un cadre purement classique conduit en effet à une aimantation toujours nulle (théorème de van Leeuwen).

1. THÉORIE SEMI-CLASSIQUE DU PARAMAGNÉTISME

Reprenons le modèle du paragraphe II : N atomes indépendants, fixés aux nœuds d'un réseau cristallin, possèdent chacun un moment magnétique $\vec{\mu}$ de module constant, que nous allons traiter comme un vecteur classique.

Considérons un atome particulier. En l'absence de champ, son moment magnétique peut s'orienter dans une direction quelconque[18], repérée par les angles θ et φ (fig. 3). En présence d'un champ magnétique $\vec{B}_0$ dirigé suivant Oz, le moment magnétique $\vec{\mu}$ acquiert une énergie

$$E = -\vec{\mu} \cdot \vec{B}_0 = -\mu . B_0 \cos\theta \; . \tag{A.74}$$

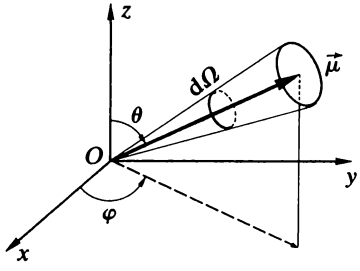


FIGURE 3
En mécanique classique, la direction du moment magnétique $\vec{\mu}$ pointant dans l'angle solide $d\Omega$ est déterminée par les deux angles θ et φ ($0 \leq \theta \leq \pi$, $0 \leq \varphi < 2\pi$).

18. C'est là la différence essentielle avec la théorie quantique : en théorie quantique, la projection de $\vec{\mu}$ sur l'axe Oz ne peut prendre que certaines valeurs discrètes du fait de la quantification des valeurs propres de J_z (formules (A.12) et (A.13)).

La probabilité dP pour que $\vec{\mu}$ pointe dans l'angle solide infinitésimal $d\Omega$ autour de la direction (θ, φ) est proportionnelle au facteur de Boltzmann $\exp -E/kT$ multiplié par le nombre d'états correspondant à ces conditions géométriques. Celui-ci est à son tour proportionnel à

$$d\Omega = \sin\theta \, d\theta \, d\varphi \,, \tag{A.75}$$

de sorte que

$$dP = \frac{1}{A} e^{\mu B_0 \cos\theta/kT} \, d\Omega \,, \tag{A.76.a}$$

avec

$$A = \int d\Omega \, e^{\mu B_0 \cos\theta/kT} \,. \tag{A.76.b}$$

Le moment magnétique moyen d'un atome a donc pour expression

$$\bar{\vec{\mu}} = \frac{1}{A} \int_0^{2\pi} d\varphi \int_0^{\pi} d\theta \vec{\mu} \sin\theta \, e^{\mu B_0 \cos\theta/kT} \,. \tag{A.77}$$

On vérifie aisément[19] que les valeurs moyennes des composantes μ_x et μ_y sont nulles, et que

$$\bar{\mu}_z = \mu \mathcal{L}\left(\frac{\mu B_0}{kT}\right) . \tag{A.78}$$

C'est ainsi qu'a été introduite pour la première fois la fonction de Langevin

$$\mathcal{L}(x) \widehat{=} \coth x - \frac{1}{x} \,. \tag{A.79}$$

L'aimantation moyenne du système est donc

$$M = \frac{N}{V} \bar{\mu}_z = \frac{N\mu}{V} \mathcal{L}\left(\frac{\mu B_0}{kT}\right) . \tag{A.80}$$

---

19. Puisque

$$\mu_x = \mu \sin\theta \cos\varphi \,, \tag{1}$$

le calcul de $\bar{\mu}_x$ fait intervenir

$$\int_0^{2\pi} d\varphi \cos\varphi = 0 \,, \tag{2}$$

et donc

$$\bar{\mu}_x = 0 \,. \tag{3}$$

On montre de même que

$$\bar{\mu}_y = \overline{\mu \sin\theta \sin\varphi} = 0 \,. \tag{4}$$

Ces résultats traduisent simplement la symétrie du problème autour de l'axe Oz du champ magnétique.

L'expression de $\bar{\mu}_z = \overline{\mu \cos\theta}$ s'obtient facilement si on remarque que

$$\bar{\mu}_z = kT \frac{\partial}{\partial B_0} \ln A \,, \tag{5}$$

avec

$$A = \int_0^{2\pi} d\varphi \int_0^{\pi} d\theta \sin\theta \, e^{\mu B_0 \cos\theta/kT} = 2\pi \int_{-1}^{+1} du \, e^{\mu B_0 u/kT} = \frac{4\pi kT}{\mu B_0} \operatorname{sh} \frac{\mu B_0}{kT} \,. \tag{6}$$

Le comportement est tout à fait semblable[20] à celui que prévoit la théorie quantique (cf. (A.26)), avec une susceptibilité

$$\chi = \frac{N}{V}\frac{\mu^2}{3kT}\,. \tag{A.81}$$

2. THÉORIE SEMI-CLASSIQUE DU DIAMAGNÉTISME

Pas plus que la théorie quantique, la théorie classique du diamagnétisme ne fait appel à la mécanique statistique.

Considérons un électron (de masse m et de charge $-e$) lié à un noyau atomique immobile au point O. Pour simplifier, nous supposons que la trajectoire classique de l'électron est un cercle du plan xOy, de rayon ρ, parcouru à la vitesse constante v (fig. 4). Le moment cinétique de cet électron par rapport à O est constant dans le temps : il a pour expression[21]

$$\vec{L} = \vec{\rho} \wedge m\vec{v}\,. \tag{A.82}$$

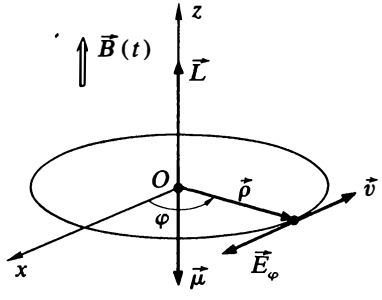


FIGURE 4
Trajectoire classique, supposée circulaire, d'un électron lié à un noyau situé au point O. $\vec{L}$ est le moment cinétique par rapport à O, $\vec{\mu}$ le moment magnétique correspondant. Le champ électrique $\vec{E}$ induit par le champ magnétique $\vec{B}(t)$ (que nous supposons perpendiculaire au plan de la trajectoire) modifie le moment cinétique de l'électron par sa composante orthoradiale $\vec{E}_\varphi$.

A ce moment cinétique est associé un moment magnétique[22]

$$\vec{\mu} = -\frac{e}{2m}\vec{L}\,. \tag{A.83}$$

---

20. Le résultat (A.80) coïncide exactement avec (A.26) si l'on y remplace $g\mu_B J$ par μ et qu'on fait ensuite tendre J vers l'infini en gardant μ fixé (puisque $\mathcal{L}(x) \equiv B_\infty(x)$). Ceci n'est guère étonnant : lorsque J augmente (à μ fixé), les valeurs permises quantiquement à μ_z occupent de façon de plus en plus serrée l'intervalle $[-\mu, \mu]$, pour former un continuum à la limite $J \to \infty$, reproduisant ainsi le comportement classique de $\vec{\mu}$.

21. Attention : $\vec{L}$ a ici les dimensions d'une action, alors que les moments cinétiques du paragraphe II étaient des grandeurs sans dimension, exprimées à l'aide de l'unité naturelle d'action quantique $\hbar$.

22. Comme la période de révolution de l'électron est extrêmement brève ($\sim 10^{-16}$ s), on peut assimiler, du point de vue électrique, la trajectoire de l'électron à une spire parcourue par le courant

$$I = \frac{ev}{2\pi\rho}\,, \tag{1}$$

circulant en sens inverse de l'électron. A cette spire d'aire $s = \pi\rho^2$ est alors associé un moment magnétique perpendiculaire au plan de la spire, de sens opposé à $\vec{L}$ et de module

$$\mu = Is = \frac{ev}{2}\rho\,. \tag{2}$$

D'autre part, $\vec{\rho}$ et $\vec{v}$ sont orthogonaux sur une trajectoire circulaire, de sorte que

$$L = mv\rho\,. \tag{3}$$

La formule (A.83) résulte de la comparaison entre (2) et (3).

Établissons alors un champ magnétique $\vec{B}_0$ dirigé suivant Oz. Ceci ne peut se faire instantanément : on a en réalité un champ dépendant du temps $B(t)$ qui croît de 0 à B_0. Ce champ magnétique variable induit un champ électrique $\vec{E}(t)$ donné par l'équation de Maxwell-Faraday

$$\overrightarrow{\text{rot}}\ \vec{E} = -\frac{\partial \vec{B}}{\partial t} . \tag{A.84}$$

Le champ électrique induit, parallèle au plan xOy, exerce sur l'électron une force $-e\vec{E}$ qui modifie son moment cinétique $\vec{L}$ selon l'équation

$$\frac{\mathrm{d}\vec{L}}{\mathrm{d}t} = \vec{\rho} \wedge (-e\vec{E}) , \tag{A.85}$$

ou encore, en projection sur Oz,

$$\frac{\mathrm{d}L_z}{\mathrm{d}t} = -\rho e E_\varphi , \tag{A.86}$$

où E_φ est la composante orthoradiale du champ, tangente à la trajectoire de l'électron (fig. 4). Le temps d'établissement du champ magnétique, c'est-à-dire le temps caractéristique de la variation de $B(t)$, est long devant la période de rotation de l'électron sur son orbite, et on peut considérer que pendant le temps

$$\delta t = \frac{2\pi\rho}{v} \tag{A.87}$$

que met l'électron à décrire son orbite, ρ et E_φ restent pratiquement constants. Pendant le temps δt, L_z varie donc de

$$\delta L_z = -\rho \frac{e}{v} \oint E_\varphi \, \mathrm{d}\ell , \tag{A.88}$$

où l'intégrale est à prendre le long de la trajectoire de l'électron, c'est-à-dire le long du cercle de rayon ρ. Or cette intégrale n'est autre que la circulation (à un instant donné) du champ électrique $\vec{E}$ le long d'une courbe fermée; cette circulation est égale au flux de $\overrightarrow{\text{rot}}\ \vec{E}$ à travers la surface s'appuyant sur cette courbe fermée. En utilisant la relation (A.84), on obtient

$$\delta L_z = \rho \frac{e}{v} \pi\rho^2 \frac{\partial B}{\partial t} = \frac{e\rho^2}{2} \frac{\partial B}{\partial t} \delta t . \tag{A.89}$$

D'après (A.83), cette modification du moment cinétique se répercute sur le moment magnétique :

$$\delta\mu_z = -\frac{e^2}{4m} \rho^2 \frac{\partial B}{\partial t} \delta t . \tag{A.90}$$

Si on néglige la variation de ρ au cours de l'opération[23], le moment magnétique induit par l'établissement du champ magnétique est indépendant de la forme précise de la fonction $B(t)$; il s'écrit simplement

$$\Delta\mu_z = \frac{-e^2}{4m} \rho^2 B_0 . \tag{A.91}$$

23. Cette approximation est équivalente à celle qui consiste, dans le calcul quantique, à se limiter à l'ordre le plus bas de la méthode des perturbations.

Sa direction est bien opposée à celle du champ appliqué : l'électron modifie son mouvement de façon à s'opposer à l'établissement du champ, comme le prévoit la loi qualitative de Lenz.

Dans le cas général où le plan de l'orbite n'est pas perpendiculaire au champ magnétique, le calcul est un peu plus compliqué; la distance ρ de l'électron à l'axe Oz varie au cours du temps. On montre (réf. 3) que la relation (A.91) reste valable à condition de remplacer ρ^2 par sa moyenne $\langle \rho^2 \rangle$ le long de la trajectoire.

Ce calcul est applicable à chacun des électrons de l'atome. Pour un cristal contenant N atomes dépourvus de moment permanent, on trouve donc finalement

$$M = -\frac{N}{V}\frac{e^2}{4m} B_0 \sum_i \langle \rho_i^2 \rangle , \tag{A.92}$$

où la somme porte sur l'ensemble des électrons d'un atome. Le résultat est le même que dans le cas quantique (formule (A.59)) si on remplace comme en (A.56) $\langle \rho_i^2 \rangle$ par $2\langle r_i^2 \rangle/3$.

3. DISCUSSION ET THÉORÈME DE VAN LEEUWEN

Les résultats précédents sont très proches de ceux que nous avons obtenus aux paragraphes II et III sur la base de la description quantique de l'atome. Il faut cependant noter que les raisonnements de Langevin postulent implicitement l'existence d'orbites privilégiées pour les électrons d'un atome, comme dans le modèle de Bohr. Or ceci sort déjà du cadre strict de la mécanique classique : en théorie purement classique, en effet, rien ne fixe les dimensions des trajectoires, ni par conséquent le module μ du moment magnétique d'un atome paramagnétique ou les distances r_i au noyau qui interviennent dans l'aimantation diamagnétique. Une analyse véritablement classique du magnétisme doit laisser ces grandeurs libres de varier continûment et les traiter elles aussi par la mécanique statistique. La théorie de Langevin est donc en réalité *«semi-classique»*, car elle incorpore certains traits caractéristiques de la description quantique de l'atome.

Or lorsqu'on étudie correctement et complètement la statistique des phénomènes dans le cadre de la théorie classique, on trouve que paramagnétisme et diamagnétisme se compensent toujours exactement et que *l'aimantation d'un système classique est toujours nulle.* Ce résultat constitue le *théorème de J.H. van Leeuwen* (1911); sa démonstration est élémentaire.

Considérons un système de N électrons soumis au champ magnétique $\vec{B}_0$ qui dérive du potentiel-vecteur $\vec{A}(\vec{r})$. La fonction de Hamilton de ce système s'écrit

$$\mathcal{H} = \sum_{i=1}^{N} \frac{1}{2m} [\vec{p}_i + e\vec{A}(\vec{r}_i)]^2 + U(\vec{r}_1, \vec{r}_2, \ldots, \vec{r}_N) , \tag{A.93}$$

où U est l'énergie potentielle totale des électrons (comprenant l'énergie d'interaction avec les noyaux supposés fixes). On en déduit la fonction de partition

$$Z(T, B_0) = \frac{1}{(2\pi\hbar)^{3N}} \int d^3p_1 \ldots d^3p_N \int d^3r_1 \ldots d^3r_N \, e^{-\mathcal{H}/kT} . \tag{A.94}$$

L'intégration sur chacune des composantes des $\vec{p}_i$ va de $-\infty$ à $+\infty$; celle sur chaque $\vec{r}_i$ est limitée au volume V de l'échantillon. Effectuons le changement de variables

$$\begin{aligned} \vec{p}_i' &= \vec{p}_i + e\vec{A}(\vec{r}_i) , \\ \vec{r}_i' &= \vec{r}_i , \end{aligned} \qquad i = 1, 2, 3, \ldots, N . \tag{A.95}$$

Le jacobien de la transformation[24] est égal à 1, de sorte que

$$Z(T, B_0)=\frac{1}{(2\pi\hbar)^{3N}}\int d^3p'_1 \ldots d^3p'_N \int d^3r'_1 \ldots d^3r'_N \exp\left\{-\frac{1}{kT}\left[\sum_{i=1}^{N}\frac{\vec{p}_i'^2}{2m}+U(\vec{r}'_1, \ldots, \vec{r}'_N)\right]\right\}$$
$$= Z(T, 0). \qquad \text{(A.96)}$$

La fonction de partition est indépendante du champ appliqué; l'aimantation du système est rigoureusement nulle (cf. (A.6)). Le magnétisme est donc un phénomène essentiellement quantique.

Références

1. Feynman, Leighton, Sands, *Le Cours de Physique de Feynman,* Tome 2, chapitres 34 et 35, Inter Éditions, Paris (1979).
2. N.W. Ashcroft et N.D. Mermin, *Solid State Physics,* Holt Saunders, New York (1981), chapitre 31.
3. A.H. Morrish, *The Physical Principles of Magnetism,* John Wiley, New York (1965).

24. Il suffit de calculer le jacobien à 6 dimensions associé à la transformation

$$\begin{cases}\vec{p}\,' = \vec{p} + e\vec{A}\,(\vec{r}\,) \\ \vec{r}\,' = \vec{r}\end{cases}$$

pour une valeur particulière de l'indice i. Tous ses éléments diagonaux sont égaux à 1; comme $\vec{r}\,'$ ne dépend pas de $\vec{p}$, les éléments non diagonaux situés d'un côté de la diagonale principale sont tous nuls. Ceci suffit pour assurer que le déterminant vaut 1, bien que les éléments non diagonaux situés de l'autre côté de la diagonale principale, tels que

$$\frac{\partial p'_x}{\partial y} = e\frac{\partial A_x}{\partial y}$$

ne soient pas nuls.

COMPLÉMENT III.B

Les gaz parfaits polyatomiques

I. Gaz parfait de particules possédant une structure interne : étude générale
   1. Énergies et états individuels
   2. Fonction de partition et énergie libre
   3. Propriétés générales des gaz parfaits classiques
   4. Le phénomène de « gel » des degrés de liberté internes

II. Structure interne des molécules polyatomiques
   1. Degrés de liberté électroniques
   2. Degrés de liberté de vibration
      a. Vibration d'une molécule diatomique
      b. Vibration d'une molécule polyatomique
   3. Degrés de liberté de rotation
      a. États et niveaux de rotation d'une molécule linéaire
      b. Rotation d'une molécule diatomique hétéropolaire
      c. Rotation d'une molécule diatomique homopolaire
      d. Rotation d'une molécule polyatomique
   4. Conclusion et discussion
      a. Résumé des résultats
      b. Anharmonicité des vibrations et couplage rotation-vibration

Nous avons étudié au chapitre III (§ IV.B) les gaz parfaits monoatomiques. Nous allons examiner ici les modifications que l'on doit apporter à cette étude lorsque le gaz est constitué de *molécules polyatomiques.* Celles-ci ne peuvent plus être considérées comme des points matériels : il faut tenir compte de leur *structure interne,* c'est-à-dire principalement des degrés de liberté associés à leur *rotation* et à leur *vibration.*

I. GAZ PARFAIT DE PARTICULES POSSÉDANT UNE STRUCTURE INTERNE : ÉTUDE GÉNÉRALE

Nous étudions un gaz constitué de N molécules identiques, *sans interactions,* enfermé dans un récipient de volume V et en équilibre avec un thermostat de température T. Nous supposons N suffisamment grand pour qu'on puisse considérer que la *limite thermodynamique* est atteinte, et T suffisamment élevée, ou N/V suffisamment faible, pour que l'*approximation de Maxwell-Boltzmann* soit valable (III, § IV.A.3.c). La fonction de partition Z du gaz prend alors la forme

$$Z = \frac{1}{N!} z^N , \tag{B.1}$$

où z est la fonction de partition pour une particule unique.

1. ÉNERGIES ET ÉTATS INDIVIDUELS

Nous reprenons donc, comme au chapitre III (§ IV.A.1), une particule unique, mais cette fois son hamiltonien H comporte, outre le terme d'énergie cinétique de translation H_0 toujours présent, diverses contributions provenant de la *structure de la molécule;* nous n'expliciterons pas ici ces nouveaux termes (voir paragraphe II), nous contentant d'en désigner la somme par H_I :

$$H = H_0 + H_I = \frac{\vec{P}^2}{2m} + H_I . \tag{B.2}$$

Les degrés de liberté internes associés à H_I *sont indépendants des degrés de liberté de translation* donnant lieu à l'énergie cinétique H_0[1]. Par conséquent, un état (λ) possible (état propre de H) peut être ici caractérisé en donnant à la fois le vecteur d'onde $\vec{K}$ qui repère l'état propre de H_0 (cf. III, § IV.B.2.a) et les nombres quantiques repérant l'état propre de H_I; si nous symbolisons par τ l'ensemble de ces nombres quantiques de structure interne (comprenant ici les nombres quantiques de spin), nous aurons

$$(\lambda) \equiv \{\vec{K}, \tau\}, \tag{B.3}$$

l'énergie correspondante ε_λ étant simplement la somme

$$\varepsilon_\lambda = \frac{\hbar^2 \vec{K}^2}{2m} + \eta_\tau \tag{B.4}$$

d'une valeur propre de H_0 et d'une valeur propre η_τ de H_I.

2. FONCTION DE PARTITION ET ÉNERGIE LIBRE

La fonction de partition z d'une molécule s'écrit donc

$$z = \frac{V}{(2\pi)^3} \int d^3K \sum_{(\tau)} \exp\left[-\beta\left(\frac{\hbar^2 \vec{K}^2}{2m} + \eta_\tau\right)\right]. \tag{B.5}$$

L'indépendance des deux types de degrés de liberté conduit à une factorisation :

$$z = z_0 \cdot \zeta, \tag{B.6}$$

où

$$z_0 = V\left(\frac{m}{2\pi\hbar^2\beta}\right)^{3/2} \tag{B.7}$$

est la fonction de partition associée au mouvement de translation de la molécule (cf. III, § IV.B.2.a) et

$$\zeta(\beta) \widehat{=} \sum_{(\tau)} e^{-\beta\eta_\tau} \tag{B.8}$$

la *fonction de partition associée à la structure interne* de la molécule.

Pour évaluer ζ, il faut expliciter l'hamiltonien interne H_I; nous le ferons au paragraphe II. Nous notons simplement ici que *cette fonction de partition interne est indépendante du volume V.*

Finalement, nous aboutissons à l'expression générale suivante pour la fonction de partition d'un gaz parfait :

$$Z = \frac{V^N}{N!}\left(\frac{m}{2\pi\hbar^2\beta}\right)^{3N/2} [\zeta(\beta)]^N, \tag{B.9}$$

et pour son énergie libre :

$$F = -NkT\left[\ln\frac{V}{N} + \frac{3}{2}\ln\frac{mkT}{2\pi\hbar^2} + 1 + \ln\zeta(\beta)\right]. \tag{B.10}$$

Il suffit de remplacer dans ces expressions ζ par $(2s+1)$ pour retrouver les formules du chapitre III, paragraphe IV.B, valables dans le cas particulier d'un gaz parfait

---

1. En termes techniques, les opérateurs H_I et H_0 commutent.

Ceci se comprend bien physiquement : ce n'est qu'au cours de collisions entre les molécules qu'une partie de leur énergie cinétique de translation peut être transférée à leurs mouvements de rotation et aux vibrations intra-moléculaires, ou vice versa (collisions dites « inélastiques »); ce couplage est donc très faible pour les gaz parfaits, où l'influence des collisions est négligeable.

monoatomique pour lequel le spin, qui n'affecte pas l'énergie, est le seul degré de liberté interne.

3. PROPRIÉTÉS GÉNÉRALES DES GAZ PARFAITS CLASSIQUES

L'énergie libre F a la même dépendance en V que dans le cas du gaz parfait monoatomique; on obtient donc ici aussi, pour la *pression,*

$$p = -\frac{\partial F}{\partial V} = \frac{NkT}{V} . \tag{B.11}$$

L'équation d'état des gaz parfaits est donc valable quelle que soit la structure interne des molécules.

Le *potentiel chimique* a toujours la forme

$$\mu = \frac{\partial F}{\partial N} = -kT \ln \frac{z}{N} , \tag{B.12}$$

mais la fonction de partition individuelle z dépend de la structure interne. *La fugacité est encore proportionnelle à la pression, le coefficient de proportionnalité étant fonction de la température seule :*

$$\varphi \widehat{=} e^{\beta\mu} = \frac{N}{z} = p \frac{1}{kT} \left(\frac{2\pi\hbar^2}{mkT}\right)^{3/2} \frac{1}{\zeta(\beta)} . \tag{B.13}$$

La valeur moyenne canonique de l'*énergie* est *proportionnelle au nombre de particules, l'énergie par particule ne dépendant que de la température :*

$$\overline{E}^c = -\frac{\partial}{\partial\beta} \ln Z = N \left[\frac{3}{2} kT - \frac{\mathrm{d}}{\mathrm{d}\beta} \ln \zeta\right] . \tag{B.14}$$

Elle est la *somme de deux termes :* l'un, toujours égal à $3NkT/2$ (cf. III, § IV.B.3.a. et § V.C), correspond aux *degrés de liberté de translation,* l'autre aux *degrés de liberté internes,* découplés des précédents. La *capacité calorifique* à volume constant d'un gaz parfait polyatomique est par conséquent *différente de celle d'un gaz parfait monoatomique,* et elle *dépend de la température* (mais pas du volume) :

$$C_v \widehat{=} \frac{\partial \overline{E}^c}{\partial T} = \frac{3}{2} Nk + \frac{N}{kT^2} \frac{\mathrm{d}^2}{\mathrm{d}\beta^2} \ln \zeta . \tag{B.15}$$

Enfin, l'expression de l'*entropie* se déduit directement de celles de l'énergie moyenne et de l'énergie libre[2] :

$$S = \frac{1}{T} (\overline{E}^c - F) = Nk \left[\ln \frac{V}{N} + \frac{3}{2} \ln \frac{mkT}{2\pi\hbar^2} + \frac{5}{2} + \ln \zeta(\beta) - \beta \frac{\mathrm{d}}{\mathrm{d}\beta} \ln \zeta(\beta)\right] . \tag{B.16}$$

4. LE PHÉNOMÈNE DE « GEL » DES DEGRÉS DE LIBERTÉ INTERNES

Examinons le comportement à basse température de la contribution des degrés de liberté internes aux propriétés du gaz. Soit η_0 la plus basse des énergies η_r possibles (niveau fondamental de structure interne) et η_1 celle qui lui est immédiatement supérieure (premier niveau excité); nous notons g_0 le degré de dégénérescence de η_0

2. L'expression trouvée pour S ne tend pas vers zéro lorsque la température T tend vers zéro, c'est-à-dire qu'elle ne vérifie pas le « troisième principe » de la thermodynamique (complément III.L, § II). Ceci n'est pas surprenant : l'approximation des gaz parfaits classiques n'est valable, pour une densité N/V fixée, qu'à des températures suffisamment hautes (cf. III, § IV.B.2.b).

(nombre d'états indépendants ayant cette énergie). Si la température T est suffisamment basse pour que

$$kT \ll \eta_1 - \eta_0 \,, \tag{B.17}$$

les divers termes de la somme (B.8) sont tous négligeables devant les g_0 premiers, associés au niveau fondamental η_0, de sorte que la fonction de partition de structure interne devient dans cette limite

$$\zeta(\beta) \simeq g_0 \, e^{-\beta\eta_0} \quad \text{si} \quad kT \ll \eta_1 - \eta_0 \,. \tag{B.18}$$

Son logarithme, qui intervient dans les formules du paragraphe 3, prend la forme simple

$$\ln \zeta(\beta) \simeq -\beta\eta_0 + \ln g_0 \quad \text{si} \quad kT \ll \eta_1 - \eta_0 \,. \tag{B.19}$$

Donc, à basse température, toutes les molécules se trouvent dans le niveau fondamental de structure interne. Les propriétés du gaz parfait polyatomique se réduisent alors, à quelques détails près, à celles du gaz parfait monoatomique : on trouve à partir de (B.19) que l'énergie moyenne $\overline{E}^c$ est simplement décalée du terme constant $N\eta_0$, et l'entropie de $Nk \ln g_0$; mais *la capacité calorifique est la même que celle d'un gaz monoatomique.* Lorsque la condition (B.17) est vérifiée, on dit que *les degrés de liberté internes sont «gelés», ou «bloqués».*

Le phénomène de «gel» des degrés de liberté internes est d'*origine quantique :* c'est l'existence d'une *énergie minimale d'excitation* $\eta_1 - \eta_0$, liée à la quantification de l'énergie, qui provoque ce phénomène. Lorsqu'il s'agit de l'énergie d'un système macroscopique, ou de l'énergie cinétique de translation d'une molécule, l'écart entre deux niveaux consécutifs est tellement faible qu'il est hors de question de le mettre en évidence. Mais nous parlons ici d'énergies individuelles, de sorte que les températures pour lesquelles se manifeste l'effet de «gel» ne sont pas *a priori* impossibles à atteindre. Nous calculerons leur ordre de grandeur au paragraphe II, et verrons que les degrés de liberté de vibration et ceux de rotation, qui interviennent de façon additive dans l'énergie η_r, se «gèlent» à des températures assez différentes. Nous concluons ici qu'*un mode d'excitation interne ne se manifeste que si kT est comparable ou supérieur à l'énergie d'excitation minimale de ce mode.*

II. STRUCTURE INTERNE DES MOLÉCULES POLYATOMIQUES

A une bonne approximation (voir plus loin paragraphe 4.b), l'hamiltonien de structure interne est la *somme de trois termes indépendants* les uns des autres[3] :

$$H_I = H_e + H_r + H_v \,; \tag{B.20}$$

H_e est l'hamiltonien du cortège électronique de la molécule, H_r décrit ses mouvements de rotation dans l'espace et H_v les vibrations de ses atomes les uns par rapport aux autres. Nous allons analyser successivement ces trois termes.

3. Il faudrait en toute rigueur ajouter un quatrième terme décrivant la structure interne des noyaux des atomes constituant la molécule. Toutefois, les énergies d'excitation d'un noyau sont de l'ordre du MeV, de sorte que les degrés de liberté nucléaires sont gelés tant que la température est petite devant

$$T_n \sim 10^6 \times 10^4 \text{ K} = 10^{10} \text{ K} \,.$$

Nous ne tenons donc pas compte ici de la structure des noyaux atomiques.

1. DEGRÉS DE LIBERTÉ ÉLECTRONIQUES

Même les molécules monoatomiques possèdent une structure interne, à laquelle sont associés des niveaux d'énergie discrets. Mais nous savons que les valeurs typiques des énergies d'excitation d'un cortège électronique atomique ou moléculaire sont de l'ordre de quelques électron-volts, ce qui correspond à des températures caractéristiques de quelles dizaines de milliers de kelvins :

$$kT \simeq 1 \text{ eV} \iff T \simeq 12000 \text{ K} . \tag{B.21}$$

Ce n'est donc qu'à très haute température que la structure interne de l'atome pourrait entrer en jeu : *jusqu'à quelques milliers de kelvins, les degrés de liberté électroniques sont gelés*[4].

Il existe toutefois des *exceptions* à cette règle. Les plus nombreuses sont créées par la *« structure fine »* du niveau fondamental de l'atome ou de la molécule[5]. Prenons par exemple un atome, et supposons que les électrons qui le constituent possèdent un moment cinétique orbital total L et un spin total S non nuls dans le niveau fondamental : pour l'atome d'aluminium, par exemple, $L = 1$ et $S = 1/2$[6]. Dans ce cas, le niveau fondamental est scindé en sous-niveaux de structure fine, dus au couplage entre L et S; le nombre de ces sous-niveaux est égal au nombre de valeurs différentes possibles pour le moment cinétique total J du cortège électronique, c'est-à-dire

$$J = L + S,\ L + S - 1,\ \ldots,\ |L - S| \tag{B.22}$$

(pour l'atome d'aluminium, il y a deux valeurs possibles : 3/2 et 1/2). Les écarts entre sous-niveaux de structure fine sont généralement de quelques centièmes d'eV, ce qui correspond à des températures caractéristiques T_e de quelques centaines de K (voir tableau I). Donc, lorsque la température T choisie est de cet ordre de grandeur, il faut tenir compte de la structure fine dans le calcul de la fonction de partition de l'atome (voir exercice III.8). Il est facile cependant de traiter le cas limite où T est

---

4. Si la température devient suffisamment élevée pour qu'on ne puisse plus négliger la population des niveaux excités électroniques (dans les étoiles, par exemple), il faut également tenir compte de la probabilité pour que l'atome soit ionisé, car l'énergie d'ionisation est du même ordre de grandeur que les énergies d'excitation. Autrement dit, le calcul de la fonction de partition doit alors faire intervenir non seulement les états liés de l'atome, mais aussi les « états du continuum », dans lesquels un ou plusieurs électrons sont détachés de l'atome.

5. Ce niveau présente également, dans de nombreux cas, une *« structure hyperfine »*, due principalement au couplage entre le spin du ou des noyau(x) et le moment cinétique du cortège électronique. Toutefois, les différences d'énergie caractéristiques de la structure hyperfine sont négligeables à l'échelle qui nous occupe ici : pour le niveau fondamental de l'atome d'hydrogène, par exemple, les deux sous-niveaux de structure hyperfine sont distants d'environ 6×10^{-6} eV, ce qui équivaut à une température caractéristique de 7×10^{-2} K; cette température est un peu plus élevée pour d'autres atomes, mais elle ne dépasse pas 1,5 K. On peut donc « oublier » totalement la structure hyperfine, c'est-à-dire considérer les sous-niveaux correspondants comme confondus.

Lorsque le noyau d'un atome possède un spin i, la fonction de partition individuelle z de cet atome est donc simplement multipliée par le nombre $(2i + 1)$ d'états distincts associés à ce spin. Pour une molécule, sauf cas très particuliers (voir exercice III.22), ce facteur est de même $(2i_1 + 1)(2i_2 + 1)\ldots$, où i_1, i_2, ... sont les spins des noyaux des atomes constituant la molécule.

6. L'atome d'aluminium possède 13 électrons. Ceux des couches fermées $1s$ (2 électrons), $2s$ (2), $2p$ (6) et $3s$ (2) combinent leurs moments cinétiques orbitaux et leurs spins pour aboutir à une somme nulle. Le 13e électron se trouve dans un état $3p$, où il est « célibataire »; c'est lui qui donne à l'atome son moment cinétique orbital et son spin.

grande devant T_e[7] : on peut alors raisonner comme si les divers sous-niveaux de structure fine étaient confondus en un seul niveau fondamental dégénéré; comme la température n'est pas suffisante pour que les autres niveaux de l'atome soient peuplés, les degrés de liberté électroniques sont bien gelés, mais avec un facteur de dégénérescence g_0 égal ici à $(2L+1)(2S+1)$[8].

TABLEAU I

Écarts entre sous-niveaux de structure fine (en électron-volts et en kelvins) pour le niveau fondamental de quelques atomes.

Atome	L, S	J	Énergie d'excitation $\Delta\eta_e$ (eV)	Température caractéristique $T_e = \Delta\eta_e/k$ (K)
F	1, 1/2	3/2	0	
		1/2	0,050 1	581
Cl	1, 1/2	3/2	0	
		1/2	0,109 2	1 267
O	1, 1	2	0	
		1	0,019 6	228
		0	0,028 1	326
Al	1, 1/2	1/2	0	
		3/2	0,013 9	161
Fe	2, 2	4	0	
		3	0,051 6	599
		2	0,087 3	1 013
		1	0,110 1	1 278
		0	0,150 3	1 744

2. DEGRÉS DE LIBERTÉ DE VIBRATION

a. Vibration d'une molécule diatomique

Considérons d'abord une molécule constituée de deux atomes seulement. Pour étudier sa structure, on utilise *l'approximation de Born-Oppenheimer :* on suppose que les deux noyaux (qui sont beaucoup plus massifs que les électrons) sont fixes, à une distance ρ l'un de l'autre, et on calcule l'énergie la plus basse du cortège électronique pour cette valeur de ρ (cf. C.D.L., complément G.XI); on obtient ainsi l'énergie (minimale) du système constitué par les deux noyaux et les électrons comme

7. Le cas contraire où $T \ll T_e$ se traite aussi facilement (exercice III.8), mais il est peu réaliste : les vapeurs d'aluminium ou de fer, ou l'oxygène monoatomique, ne sont observés qu'à haute température.

8. Des considérations analogues valent pour les molécules, mais il en existe peu qui possèdent un moment cinétique orbital non nul dans leur état fondamental. C'est le cas par exemple pour l'oxyde d'azote NO, dont le niveau fondamental est scindé en deux sous-niveaux séparés par 0,015 eV environ (ce qui équivaut à une température T_e de 175 K). Le cas de la molécule d'oxygène O_2 est un peu particulier : elle ne présente pas de structure fine dans son niveau fondamental ($S = 1$ mais moment orbital nul), mais son premier niveau électronique excité est relativement proche du précédent (à 0,97 eV); ce niveau affecte donc les propriétés de l'oxygène gazeux à haute température (quelques milliers de K).

une fonction $u(\rho)$ de la distance ρ. Si les deux atomes peuvent former une molécule stable, la fonction $u(\rho)$ présente un minimum assez profond pour qu'il existe des états liés (fig. 1). La distance d'équilibre d entre les deux noyaux est celle qui correspond à ce minimum.

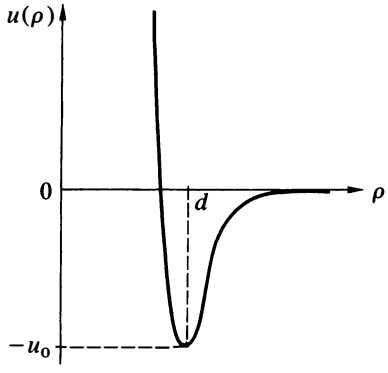


FIGURE 1
Énergie potentielle d'une molécule diatomique en fonction de la distance ρ entre les deux noyaux, dans l'approximation de Born-Oppenheimer (la répulsion électrostatique entre les noyaux est incluse dans l'énergie potentielle; c'est elle qui est responsable de la remontée rapide de la courbe aux courtes distances).

Mais les noyaux vibrent l'un par rapport à l'autre, c'est-à-dire que leur distance ρ oscille autour de d. Cette vibration est équivalente au mouvement d'une particule fictive qui aurait pour masse la masse réduite

$$m_R = \frac{m_1 m_2}{m_1 + m_2} \tag{B.23}$$

des deux noyaux, et qui serait soumise au potentiel de la figure 1 (C.D.L., p. 795). Dans la limite des petites oscillations (ou plutôt des faibles extensions de la fonction d'onde quantique), on a affaire à un *oscillateur harmonique à une dimension.* Sa pulsation ω est de la forme

$$\omega = \sqrt{\frac{\varkappa}{m_R}}\,, \tag{B.24}$$

la «constante de rappel» $\varkappa$ étant déterminée par la courbure de la fonction $u(\rho)$ à son minimum[9]. On connaît les énergies possibles, en mécanique quantique, pour un oscillateur harmonique à une dimension : si nous convenons de compter les énergies de vibration à partir du minimum de u, nous avons

$$\eta_\nu = \left(\nu + \frac{1}{2}\right)\hbar\omega\,, \tag{B.25}$$

où ν peut prendre toutes les valeurs entières positives ou nulles. Le niveau η_ν est non-dégénéré, c'est-à-dire qu'il lui correspond un seul état quantique de vibration.

9. Développons en effet la fonction potentiel $u(\rho)$ au voisinage de son minimum en $\rho = d$. Le terme du premier ordre est nul, puisque $\rho = d$ est un point stationnaire; le terme du second ordre est positif, puisqu'il s'agit d'un minimum. On peut donc écrire, si $\rho - d$ reste assez petit,

$$u(\rho) \simeq u(d) + \frac{1}{2}\varkappa(\rho - d)^2\,, \tag{1}$$

avec $$\varkappa = \frac{\mathrm{d}^2 u}{\mathrm{d}\rho^2}(\rho = d) > 0\,. \tag{2}$$

On reconnaît là l'énergie potentielle d'un oscillateur harmonique à une dimension.

Dans cette *approximation harmonique,* la *fonction de partition* ζ_v relative aux vibrations de la molécule s'écrit

$$\zeta_v(\beta)=\sum_{\nu=0}^{\infty} e^{-\beta(\nu+\frac{1}{2})\hbar\omega} . \tag{B.26}$$

Elle se calcule facilement, puisqu'il s'agit simplement de sommer une série géométrique, de raison $\exp(-\beta\hbar\omega)$:

$$\zeta_v(\beta)=e^{-\beta\hbar\omega/2}\frac{1}{1-e^{-\beta\hbar\omega}}=\frac{1}{2\ \mathrm{sh}\ \beta\hbar\omega/2} . \tag{B.27}$$

On en déduit aussitôt la contribution des vibrations à l'*énergie moyenne* d'une molécule :

$$\bar{\eta}_v=-\frac{d}{d\beta}\ln\zeta_v=\frac{\hbar\omega}{2}\coth\frac{\beta\hbar\omega}{2}=\hbar\omega\left[\frac{1}{2}+\frac{1}{e^{\beta\hbar\omega}-1}\right], \tag{B.28}$$

et le terme qu'elles ajoutent à la *capacité calorifique* du gaz :

$$C_V^{(v)}=N\frac{d\bar{\eta}_v}{dT}=Nk\frac{(\hbar\omega/2kT)^2}{\mathrm{sh}^2(\hbar\omega/2kT)} . \tag{B.29}$$

Il s'introduit ainsi de façon naturelle une *température caractéristique des vibrations* T_v, définie par

$$kT_v\widehat{=}\hbar\omega . \tag{B.30}$$

La température T_v est d'autant plus élevée que les forces interatomiques sont plus intenses[10] et que les atomes sont plus légers (cf. formule (B.24)). Elle varie selon les molécules entre quelques centaines et quelques milliers de kelvins (tableau II).

TABLEAU II

Températures caractéristiques de vibration et de rotation pour quelques molécules diatomiques simples.

Molécule	T_v (K)	T_r (K)
H_2	6215	85,3
D_2	4394	42,7
HD	5382	64
Cl_2	808	0,35
Br_2	463	0,12
$O_2^{(*)}$	2256	2,1
N_2	3374	2,9
CO	3103	2,8
$NO^{(*)}$	2719	2,5
HCl	4227	15,0
HBr	3787	12,0

* Voir note 8.

Pour des températures T très inférieures à T_v, les vibrations de la molécule sont gelées; on constate effectivement que l'expression (B.29) tend vers zéro avec T/T_v :

$$C_V^{(v)}\sim Nk\left(\frac{T_v}{T}\right)^2 e^{-T_v/T}\ \text{pour}\ T\ll T_v . \tag{B.31}$$

10. La « constante de rappel » $\varkappa$ est donnée par la dérivée seconde de $u(\rho)$ (note 9). Mais la distance d'équilibre reste à peu de chose près la même pour toutes les molécules (de 1 à 2 Å), de sorte que la dérivée seconde est d'autant plus grande que le minimum de la courbe est plus profond.

Pour T très supérieure à T_v, cette même expression tend vers une constante :

$$C_V^{(v)} \sim Nk \text{ pour } T \gg T_v \,, \tag{B.32}$$

ce qui est conforme au théorème d'équipartition de l'énergie appliqué à un oscillateur harmonique à une dimension (III, § V.C.2). Les variations de $C_V^{(v)}$ avec la température sont représentées sur la figure 2.

La connaissance de la fonction de partition ζ_v (formule (B.27)) permet de calculer la contribution des vibrations moléculaires aux propriétés du gaz (voir § I.3). On évalue également sans difficulté les populations des divers états de vibration (exercice III.6).

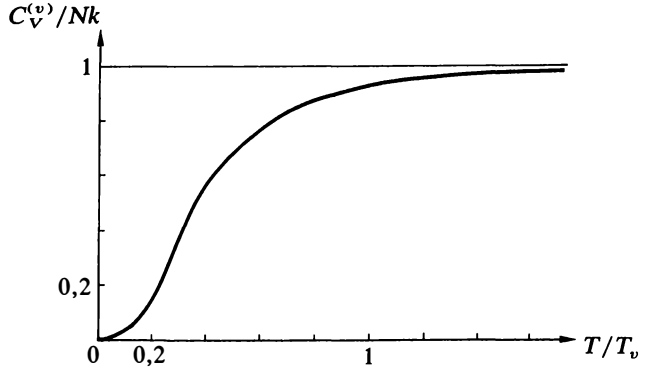


FIGURE 2

Variation avec la température T de la capacité calorifique vibrationnelle d'un gaz parfait diatomique, dans l'approximation harmonique. On note que $C_V^{(v)}$ a pratiquement atteint sa limite de haute température dès que T dépasse T_v ($C_V^{(v)} \simeq 0{,}92Nk$ pour $T = T_v$), et qu'elle reste pratiquement nulle tant que T est inférieure à $T_v/10$.

b. Vibration d'une molécule polyatomique

Considérons maintenant une molécule constituée de n atomes. Il est facile de calculer le *nombre n_v de ses degrés de liberté de vibration,* c'est-à-dire le nombre de paramètres nécessaires pour caractériser les *positions relatives* des n noyaux. Il faut en effet $3n$ coordonnées pour déterminer la position de ces n noyaux. Mais le centre de masse de la molécule est lui-même repéré par 3 coordonnées (auxquelles sont associés les 3 degrés de liberté de translation) qui sont des combinaisons linéaires des précédentes; en outre, l'orientation de la molécule dans l'espace est caractérisée par 3 paramètres, par exemple les 3 angles d'Euler donnant la position, par rapport au trièdre de référence, d'un trièdre trirectangle solidaire de la molécule (à ces 3 paramètres correspondent les degrés de liberté de rotation — voir paragraphe 3.d). Il reste donc

$$n_v = 3n - 6 \tag{B.33}$$

coordonnées repérant les positions relatives des n noyaux. Cependant, si la molécule est linéaire, c'est-à-dire si les n noyaux sont alignés, il suffit de 2 angles, et non plus de 3, pour repérer son orientation (voir plus loin la figure 3); dans ce cas,

$$n_v = 3n - 5 \text{ (molécule linéaire)} \,. \tag{B.34}$$

L'énergie potentielle de la molécule, obtenue en principe par la méthode de Born-Oppenheimer (cf. § a), est une fonction $u(\rho_i)$ des n_v paramètres ρ_i fixant les positions relatives des noyaux. Cette énergie potentielle est minimum lorsque les ρ_i prennent des valeurs particulières d_i qui correspondent à la *configuration d'équilibre*

de la molécule. Si les ρ_i s'écartent légèrement de leur valeur d'équilibre d_i, l'énergie potentielle $u(\rho_i)$ peut être approximée par son développement autour du point $\{\rho_i = d_i ;\ i = 1, 2, \ldots, n_v\}$; les dérivées premières de la fonction u sont toutes nulles en ce point, de sorte que le développement s'écrit

$$u(\rho_i) \simeq u(d_i) + \frac{1}{2} \sum_{j,k=1}^{n_v} \frac{\partial^2 u}{\partial \rho_j \partial \rho_k} (\rho_j - d_j)(\rho_k - d_k) , \tag{B.35}$$

où la dérivée doit être évaluée au minimum de u. A cette approximation, les vibrations de la molécule constituent un *système de n_v oscillateurs harmoniques couplés.*

On peut cependant montrer, grâce au fait que la forme quadratique dans les $(\rho_i - d_i)$ figurant dans (B.35) est définie positive (puisqu'on développe u autour de son minimum), que ce système de n_v oscillateurs harmoniques couplés est équivalent à un *système de n_v oscillateurs harmoniques fictifs indépendants,* que l'on appelle les *« modes normaux de vibration »* de la molécule. A chaque mode normal est associée une pulsation propre[11] ω_α ($\alpha = 1, 2, \ldots, n_v$). L'état vibratoire de la molécule est donc, à l'*approximation harmonique,* caractérisé par n_v entiers ν_α positifs ou nuls, l'énergie correspondante étant

$$\eta(\{\nu_\alpha\}) = \sum_{\alpha=1}^{n_v} \left(\nu_\alpha + \frac{1}{2}\right) \hbar\omega_\alpha . \tag{B.36}$$

Les n_v modes normaux étant *indépendants,* la fonction de partition de vibration

TABLEAU III

Données sur quelques molécules polyatomiques simples : nombre de degrés de liberté de vibration (n_v) et de rotation (n_r), et températures caractéristiques correspondantes.

Molécule	n_v	n_r	$T_\alpha^{(v)}$ (K)	$T_\alpha^{(r)}$ (K)
CO_2	4	2	954 (2 fois) 1 890 3 360	0,56 (2 fois)
H_2O	3	3	2 290 5 160 5 360	13,4 20,9 40,1
NH_3	6	3	1 360 2 330 (2 fois) 4 800 4 880 (2 fois)	8,9 13,6 (2 fois)
SO_2	3	3	750 1 660 1 960	0,42 0,49 2,9
N_2O	4	2	850 (2 fois) 1 840 3 200	0,60 (2 fois)
CH_4	9	3	1 870 (3 fois) 2 180 (2 fois) 4 170 4 320 (3 fois)	7,5 (3 fois)

11. Ce sont les pulsations propres ω_α qui se manifestent dans les spectres d'émission ou d'absorption de la molécule, et que l'on mesure donc expérimentalement.

ζ_v se factorise à raison d'un facteur par mode, chacun de ces facteurs étant calculé comme au paragraphe a :

$$\zeta_v = \prod_{\alpha=1}^{n_v} \zeta_v^{(\alpha)} = \prod_{\alpha=1}^{n_v} \frac{1}{2 \operatorname{sh} \beta\hbar\omega_\alpha/2} . \tag{B.37}$$

Comme en (B.28) et (B.29), on en déduit facilement l'énergie moyenne de vibration d'une molécule :

$$\bar{\eta}_v = \sum_{\alpha=1}^{n_v} \frac{\hbar\omega_\alpha}{2} \coth \frac{\beta\hbar\omega_\alpha}{2} , \tag{B.38}$$

et la contribution des vibrations à la capacité calorifique du gaz :

$$C_V^{(v)} = Nk \sum_{\alpha=1}^{n_v} \frac{(\hbar\omega_\alpha/2kT)^2}{\operatorname{sh}^2 (\hbar\omega_\alpha/2kT)} . \tag{B.39}$$

Les divers modes normaux ajoutent également leurs contributions aux autres grandeurs telles que l'énergie libre, l'entropie et le potentiel chimique du gaz (cf. paragraphe I.3). A chaque mode normal (α) est associée une *température caractéristique* (tableau III)

$$T_\alpha^{(v)} \widehat{=} \frac{\hbar\omega_\alpha}{k} \tag{B.40}$$

qui détermine les domaines de température où ce mode (α) est gelé ou excité (cf. (B.31) et (B.32)).

3. DEGRÉS DE LIBERTÉ DE ROTATION

Les molécules possèdent également des *modes d'excitation rotationnels,* correspondant à leur énergie cinétique de rotation d'ensemble dans l'espace. Nous supposons ici que les distances entre les noyaux restent constantes au cours de la rotation de la molécule, c'est-à-dire que celle-ci se comporte comme un *« rotateur rigide »* (voir paragraphe 4.b).

a. États et niveaux de rotation d'une molécule linéaire

Prenons d'abord le cas simple d'une *molécule linéaire,* par exemple une molécule *diatomique* telle que HCl ou O_2. Son orientation dans l'espace (la position de son centre de masse étant déjà prise en compte dans les degrés de liberté de translation) est repérée par deux angles θ et φ (fig. 3). L'énergie cinétique de rotation d'une telle molécule s'écrit

$$\eta_r = \frac{1}{2} I(\dot{\theta}^2 + \dot{\varphi}^2 \sin^2 \theta)$$

$$= \frac{1}{2I} \vec{L}^2 , \tag{B.41}$$

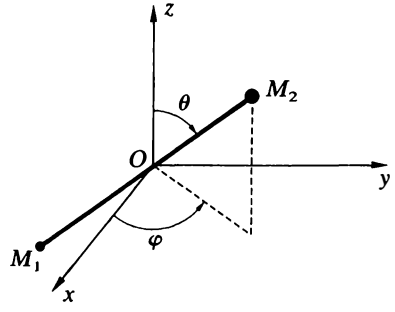


FIGURE 3

Repérage dans l'espace d'une molécule diatomique, par rapport à un système d'axes de directions fixes lié à son centre de masse *O*. On caractérise la position de la molécule par la distance $\rho = M_1M_2$ des noyaux des deux atomes, et par l'orientation de l'axe $\overrightarrow{M_1M_2}$, repérée à l'aide des deux angles θ et φ.

où $\vec{L}$ est son moment cinétique et I son moment d'inertie par rapport à son centre de masse (c'est-à-dire par rapport à un axe quelconque passant par ce point et perpendiculaire à l'axe de la molécule)[12].

L'hamiltonien associé, en mécanique quantique, à la rotation d'ensemble d'une molécule linéaire est donc simplement (C.D.L., complément C.VI)

$$H_r = \frac{1}{2I}\vec{L}^2 , \tag{B.42}$$

où $\vec{L}$ est l'opérateur de moment cinétique « orbital ». On connaît par conséquent les *états stationnaires de rotation :* les fonctions d'onde correspondantes sont les *harmoniques sphériques* $Y_j^m(\theta, \varphi)$, où j est un entier positif ou nul et m l'un des $(2j+1)$ entiers relatifs compris entre $-j$ et $+j$. L'*énergie* associée à ces états dépend seulement de j, et a pour expression

$$\eta_{j,m} = \frac{\hbar^2}{2I} j(j+1) . \tag{B.43}$$

On peut, à partir de cette expression, introduire une *température caractéristique* T_r au-dessous de laquelle les degrés de liberté de rotation seraient gelés :

$$kT_r \widehat{=} \frac{\hbar^2}{2I} . \tag{B.44}$$

Pour la molécule H_2, dont le moment d'inertie est très petit parce que les deux noyaux sont très légers, cette température caractéristique est voisine de 85 K; elle est de l'ordre de la dizaine de kelvins pour les molécules du type HA, où A est un atome plus lourd que H[13]; elle est extrêmement basse (au plus quelques kelvins) pour toutes les autre molécules (voir tableaux II et III). Donc, *sauf pour l'hydrogène* (voir exercice III.7), on est pratiquement toujours dans la *situation inverse de celle du gel* de la rotation :

$$T \gg T_r . \tag{B.45}$$

---

12. Pour une molécule diatomique, on montre facilement que le moment d'inertie peut s'écrire

$$I \widehat{=} m_1 r_1^2 + m_2 r_2^2 = m_R d^2 ,$$

où d est la distance entre les noyaux M_1 et M_2 et m_R leur masse réduite (B.23).

Nous ne tenons pas compte de la rotation de la molécule autour de l'axe M_1M_2, car le moment d'inertie par rapport à cet axe, dû uniquement aux électrons de la molécule, est totalement négligeable devant I.

13. Évaluons l'ordre de grandeur du moment d'inertie I de deux atomes d'hydrogène ($m_p \simeq 1{,}7 \times 10^{-27}$ kg) distants de $d \simeq 1$ Å :

$$I = m_R d^2 = \frac{m_p}{2} d^2 \simeq 10^{-47} \text{ kg.m}^2 . \tag{1}$$

L'énergie caractéristique correspondante vaut

$$kT_r \widehat{=} \frac{\hbar^2}{2I} \simeq 0{,}5 \times 10^{-21} \text{ J} \simeq 4 \times 10^{-3} \text{ eV} , \tag{2}$$

ce qui donne

$$T_r \simeq 50 \text{ K} \tag{3}$$

(la mesure expérimentale de I pour la molécule H_2 conduit à 85,3 K, ce qui signifie que les deux atomes sont distants de 0,74 Å).

Pour une molécule du type *HA*, la masse réduite (B.23) est voisine de m_p, ce qui multiplie I par 2 et divise donc T_r par 2. Pour une molécule comme N_2 ou O_2, l'évaluation (1) du moment d'inertie est à multiplier par 15 environ, ce qui ramène T_r à quelques kelvins.

Dans ces conditions, *les degrés de liberté de rotation peuvent être traités par la mécanique classique*[14].

b. Rotation d'une molécule diatomique hétéropolaire

Considérons d'abord une molécule constituée de *deux atomes différents,* comme HCl. Les résultats du paragraphe a permettent d'écrire la *fonction de partition* ζ_r correspondant à la rotation d'une telle molécule :

$$\zeta_r = \sum_{j=0}^{\infty} \sum_{m=-j}^{+j} e^{-\beta \eta_{j,m}}$$

$$= \sum_{j=0}^{\infty} (2j+1) \exp\left[-\beta \frac{\hbar^2}{2I} j(j+1)\right] . \qquad \text{(B.46)}$$

Il n'existe pas ici de méthode analytique simple donnant cette somme exactement. Mais la condition (B.45) montre que deux termes consécutifs de l'expression (B.46) diffèrent assez peu l'un de l'autre pour que la somme puisse être remplacée par une intégrale :

$$\zeta_r = \int_0^{+\infty} dj\,(2j+1) \exp\left[-\beta \frac{\hbar^2}{2I} j(j+1)\right] , \quad T \gg T_r . \qquad \text{(B.47)}$$

Un calcul simple[15] donne alors

$$\zeta_r = \frac{2I}{\hbar^2} kT , \quad T \gg T_r . \qquad \text{(B.48)}$$

Ce résultat peut également être obtenu à partir de l'expression classique de la fonction de partition (voir exercice III.27).

La contribution correspondante à l'*énergie individuelle moyenne* vaut simplement

$$\bar{\eta}_r = -\frac{d}{d\beta} \ln \zeta_r = kT , \quad T \gg T_r . \qquad \text{(B.49)}$$

Ceci est conforme au théorème d'équipartition de l'énergie (III, § V.C), puisqu'on a la somme de deux termes d'énergie cinétique de rotation[16], autour de deux axes

---

14. Le problème ne comporte ici qu'une grandeur dont on puisse comparer la valeur à $\hbar$ (cf. III, § V.B) : la condition de validité de la mécanique classique s'écrit

$$\sqrt{IkT} \gg \hbar ,$$

c'est-à-dire $T \gg T_r$.

15. Si l'on pose

$$y \widehat{=} j(j+1) , \qquad (1)$$

l'intégrale à calculer s'écrit

$$\zeta_r = \int_0^{+\infty} dy \exp\left[-\beta \frac{\hbar^2}{2I} y\right] , \qquad (2)$$

ce qui conduit aussitôt au résultat (B.48).

16. En termes des moments conjugués p_θ et p_φ des variables de position θ et φ (appendice III, § II.1), l'énergie cinétique de rotation (B.41) s'écrit

$$\eta_r = \frac{1}{2I}\left(p_\theta^2 + \frac{1}{\sin^2\theta} p_\varphi^2\right) .$$

On a bien la somme de deux termes quadratiques indépendants, dont chacun a pour valeur moyenne $kT/2$.

perpendiculaires à celui de la molécule (cf. note 12). Pour des molécules diatomiques, les degrés de liberté de rotation ajoutent donc à la *capacité calorifique* du gaz parfait monoatomique un *terme constant :*

$$C_V^{(r)} = Nk\,, \qquad T \gg T_r\,. \tag{B.50}$$

On peut si nécessaire (par exemple pour la molécule HD, dont la température $T_r \simeq 64$ K est relativement élevée) évaluer numériquement la somme discrète (B.46) donnant la fonction de partition ζ_r. On trouve alors que la contribution rotationnelle à la chaleur spécifique varie, dans la région des températures de l'ordre de T_r, comme l'indique la figure 4 (voir aussi l'exercice III.9).

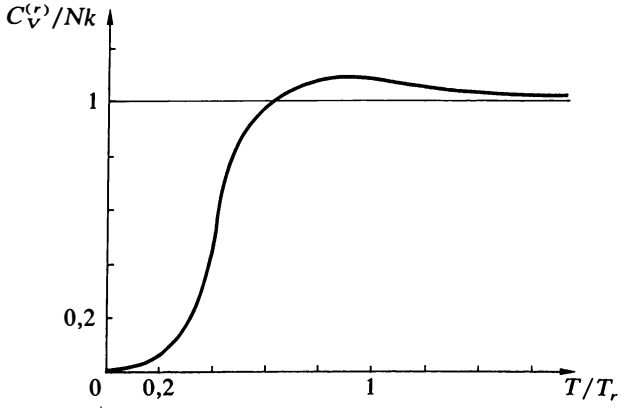


FIGURE 4

Variation, en fonction de la température T, de la contribution $C_V^{(r)}$ des degrés de liberté rotationnels à la capacité calorifique d'un gaz diatomique. La courbe est obtenue par calcul numérique. On constate que $C_V^{(r)}$ est pratiquement nulle pour $T < T_r/5$, qu'elle passe par un maximum égal à $1{,}1Nk$ pour $T \simeq 0{,}81T_r$, et qu'elle est pratiquement égale à sa valeur classique Nk dès que T dépasse $T_r/2$.

c. Rotation d'une molécule diatomique homopolaire

Lorsqu'une molécule est constituée de *deux atomes identiques,* comme H_2, N_2, O_2, ..., il faut tenir compte du *postulat de symétrisation* quantique dans l'énumération de ses états : les seuls états permis sont ceux qui sont symétriques dans l'échange des deux noyaux si ceux-ci sont des *bosons* (c'est-à-dire si leur spin i est entier), ou ceux qui sont antisymétriques si les noyaux sont des *fermions* (spin i demi-entier)[17].

Le mouvement de translation de la molécule ignore la position relative des deux noyaux, puisqu'il concerne seulement leur centre de masse. Le mouvement de vibration n'affecte que la distance ρ entre les noyaux, qui est invariante dans leur permutation. Reste le mouvement de rotation, qui est effectivement susceptible d'intervertir les positions des deux noyaux (lorsque θ change en $\pi - \theta$ et φ en $\varphi + \pi$, puisqu'ici le centre de masse est toujours situé au milieu de M_1M_2).

Analysons pour commencer la *symétrie des états de spin des deux noyaux.* Un résultat bien connu de mécanique quantique donne les valeurs i_t du spin total obtenu en composant deux spins i égaux (C.D.L., p. 1007) :

$$i_t = 2i,\ 2i-1,\ \ldots,\ 1,\ 0\,, \tag{B.51}$$

17. Bien entendu, l'état du cortège électronique de la molécule doit lui aussi, pour chaque valeur de la distance ρ entre les noyaux, être totalement antisymétrique par rapport à la permutation de deux électrons quelconques. Cette condition est déjà prise en compte dans le calcul de l'énergie potentielle $u(\rho)$ de la molécule par la méthode de Born-Oppenheimer (§ 2.a).

ainsi que la symétrie, par rapport à l'échange des deux spins individuels, des $(2i_t + 1)$ états distincts correspondant à chacune de ces valeurs (C.D.L., p. 1031) :

$$i_t = 2i,\ 2i-2,\ \ldots \quad : \text{états symétriques,} \tag{B.52.a}$$

$$i_t = 2i-1,\ 2i-3,\ \ldots : \text{états antisymétriques.} \tag{B.52.b}$$

Il y a au total $(2i+1)^2$ états indépendants pour le système des deux spins (puisque chacun d'eux peut prendre $(2i+1)$ états différents). On calcule sans difficulté[18] qu'il y a parmi eux

$$(i+1)(2i+1) \text{ états symétriques,} \tag{B.53.a}$$

$$\text{et} \quad i(2i+1) \quad \text{états antisymétriques.} \tag{B.53.b}$$

Notons que ces résultats valent aussi bien pour i demi-entier que pour i entier.

Envisageons maintenant de *permuter les positions des deux noyaux* dans l'espace. Il s'agit évidemment de le faire sans rien modifier d'autre dans la molécule. C'est possible en effectuant une *réflexion par rapport au centre de masse O* de la molécule. Dans cette transformation en effet, l'état du système électronique reste inchangé[19]; la seule modification est donc l'échange des positions des deux noyaux. Or on sait que *les harmoniques sphériques ont une parité déterminée* dans une réflexion par rapport à l'origine des coordonnées (C.D.L., p. 692) :

$$Y_j^m(\pi-\theta,\ \varphi+\pi) = (-1)^j\ Y_j^m(\theta,\ \varphi)\ . \tag{B.54}$$

Autrement dit, *les états de rotation sont symétriques ou antisymétriques* dans l'échange des positions des deux noyaux *suivant que j est pair ou impair.*

Finalement, pour satisfaire au postulat de symétrisation pour les deux noyaux identiques, il faut associer leurs états de spin et les états de rotation de la molécule de la façon suivante :

— si i est *entier* (bosons), *état de spin symétrique et j pair*
ou *état de spin antisymétrique et j impair;*

— si i est *demi-entier* (fermions), *état de spin symétrique et j impair*
ou *état de spin antisymétrique et j pair.*

---

18. Une méthode simple consiste par exemple à écrire les produits de deux états individuels $|m\rangle\,|m'\rangle$ dans un tableau carré à $(2i+1)$ lignes (m) et $(2i+1)$ colonnes (m'), ordonnées de la même façon. Les $(2i+1)$ états $|m\rangle\,|m\rangle$ situés sur la diagonale principale sont évidemment symétriques dans la permutation des deux spins individuels. Pour ceux qui sont hors de cette diagonale $(m \neq m')$, on associe $|m\rangle\,|m'\rangle$ et $|m'\rangle\,|m\rangle$ pour construire une combinaison symétrique (leur somme) et une combinaison antisymétrique (leur différence). On trouve ainsi

$$\frac{1}{2}\left[(2i+1)^2 - (2i+1)\right] \quad \text{états antisymétriques}$$

$$\text{et} \quad \frac{1}{2}\left[(2i+1)^2 + (2i+1)\right] \quad \text{états symétriques ,}$$

ce qui donne bien les résultats (B.53).

19. Pour une molécule constituée de deux atomes identiques, l'hamiltonien du système électronique, dans l'approximation de Born-Oppenheimer où les deux noyaux sont fixés, est forcément symétrique par rapport au milieu O du segment qui joint ces deux noyaux : l'énergie potentielle d'un électron est forcément la même en deux points symétriques par rapport à O. Les fonctions propres d'un tel hamiltonien sont alors soit paires soit impaires dans cette transformation. Pour la plupart des molécules, l'état fondamental électronique est pair. Nous supposons que c'est le cas pour la molécule étudiée ici; il serait facile d'adapter le raisonnement au cas contraire.

Les autres états *a priori* possibles (par exemple un état de spin symétrique associé à un j impair pour i entier) sont interdits par le postulat et ne sont pas réalisés dans la nature. Donc, *les degrés de liberté de rotation de la molécule et les degrés de liberté de spin des deux noyaux ne sont pas indépendants.*

Dans la somme sur les états qui donne la fonction de partition d'une molécule, l'énergie dépend de j, mais pas de m ni de l'état de spin des noyaux (formule (B.43)). Par conséquent, les états de spin, comme les valeurs de m, interviennent seulement par leur nombre. Si nous identifions le cas où i est entier par un indice B (pour «boson») et celui où i est demi-entier par F («fermion»), la fonction de partition associée à la rotation d'une molécule diatomique homopolaire s'écrit soit

$$\zeta_r^{(B)} = (i+1)(2i+1)\,\zeta_{\text{pair}} + i(2i+1)\,\zeta_{\text{impair}}\ , \tag{B.55.a}$$

soit

$$\zeta_r^{(F)} = (i+1)(2i+1)\,\zeta_{\text{impair}} + i(2i+1)\,\zeta_{\text{pair}}\ , \tag{B.55.b}$$

où

$$\zeta_{\text{pair}} \mathrel{\widehat{=}} \sum_{j\ \text{pair}} (2j+1)\,\mathrm{e}^{-\beta j(j+1)\hbar^2/2I}\ , \tag{B.56.a}$$

$$\zeta_{\text{impair}} \mathrel{\widehat{=}} \sum_{j\ \text{impair}} (2j+1)\,\mathrm{e}^{-\beta j(j+1)\hbar^2/2I}\ . \tag{B.56.b}$$

On n'a donc plus, comme pour les molécules hétéropolaires, factorisation des contributions provenant des degrés de liberté de rotation et des degrés de liberté de spin des noyaux, car ils ne sont plus indépendants.

Toutefois, *sauf dans des cas très particuliers comme celui de l'hydrogène et du deutérium* (exercice III.22), *les températures caractéristiques de rotation des molécules sont trop basses* (cf. tableau II) *pour que les effets quantiques que nous venons de décrire soient observables.* En effet, lorsque

$$\frac{T_r}{T} \mathrel{\widehat{=}} \beta\,\frac{\hbar^2}{2I} \ll 1\ , \tag{B.57}$$

on peut remplacer dans les expressions (B.56) les sommes discrètes par des intégrales et l'on a alors[20]

$$\zeta_{\text{pair}} \simeq \zeta_{\text{impair}} \simeq \int_0^{+\infty} (2j+1)\,\mathrm{e}^{-j(j+1)\,T_r/T}\,\frac{\mathrm{d}j}{2}\ , \tag{B.58}$$

ce qui donne tout simplement

$$\zeta_r^{(B)} \simeq \zeta_r^{(F)} \simeq (2i+1)^2\,\frac{T}{2T_r}\ . \tag{B.59}$$

Donc, les formules sont redevenues identiques pour le cas où les noyaux sont des bosons et pour celui où ce sont des fermions; en outre, on a de nouveau factorisation des deux contributions, celle des spins nucléaires étant simplement $(2i+1)^2$ (cf. note 5);

---

20. Dans chacune des deux sommes (B.56), les valeurs de j varient de 2 en 2. Lorsqu'on passe à l'intégrale, l'élément différentiel $\mathrm{d}j$ représente donc un intervalle de 2 unités; il faut le diviser par 2 pour ne pas modifier les valeurs de la fonction à sommer.

Ce facteur 2 peut aussi être compris si l'on remarque que

$$\sum_{j\ \text{pair}} \simeq \sum_{j\ \text{impair}}$$

implique que chacune de ces deux sommes vaut la moitié de celle qui porte sur toutes les valeurs entières de j (formules (B.46)-(B.47)).

enfin, la fonction de partition rotationnelle est divisée par 2 par rapport à celle d'une molécule hétéropolaire (formule (B.48)).

Ce *facteur* 2 est le seul souvenir du traitement quantique du problème. On peut, comme toujours en pareil cas (cf. III, § V.A.2.b), l'inclure dans l'expression classique de la fonction de partition si l'on veut raisonner directement dans le cadre de la mécanique classique (cf. exercice III.27) : il suffit de considérer comme non distincts l'état (classique) où la molécule pointe dans une direction (θ, φ) et celui où elle pointe dans la direction opposée $(\pi - \theta, \varphi + \pi)$; ceci a pour effet de *diviser par* 2 *l'intégrale sur l'espace des phases* classique. Cette différence entre molécules homopolaires et hétéropolaires ne se manifeste pas dans l'énergie moyenne ni la capacité calorifique (voir formules (B.14) et (B.15)); elle apparaît cependant dans l'énergie libre, l'entropie et le potentiel chimique (formules (B.10), (B.16) et (B.13)).

d. Rotation d'une molécule polyatomique

Considérons pour finir un gaz parfait dont les molécules comportent plus de deux atomes, et étudions leur mouvement de rotation.

Le cas le plus simple est celui des *molécules linéaires,* car il se traite comme celui des molécules diatomiques (§ a). L'approximation classique est pratiquement toujours valable ici, et la fonction de partition rotationnelle s'écrit (cf. (B.48))

$$\zeta_r = \frac{1}{\sigma}\frac{2I}{\hbar^2}kT = \frac{1}{\sigma}\frac{T}{T_r}, \tag{B.60}$$

où σ est appelé le *«facteur de symétrie»* : il vaut 1 pour les molécules linéaires non symétriques comme N_2O (N—N—O), et 2 pour les molécules symétriques telles que CO_2 (O—C—O) ou C_2H_2 (H—C—C—H); pour ces dernières en effet, on divise par 2 l'intégrale sur l'espace des phases classique, pour la même raison qu'au paragraphe c. L'énergie moyenne rotationnelle d'une molécule linéaire vaut kT (ce qui correspond aux deux degrés de liberté indépendants de rotation autour de deux axes perpendiculaires à celui de la molécule), et la capacité calorifique correspondante

$$C_V^{(r)} = Nk \quad \text{(molécule linéaire).} \tag{B.61}$$

Une *molécule polyatomique non linéaire* possède *trois degrés de liberté de rotation.* Nous la considérons comme tournant en bloc de façon rigide (voir § 4.b); on peut donc lui appliquer les résultats classiques de cinématique du solide[21]. On sait qu'un corps solide de forme quelconque peut être rapporté à son système d'*axes principaux d'inertie,* qui constituent un trièdre trirectangle lié au solide et ayant son origine au centre de masse. Nous noterons I_1, I_2, I_3 les *moments d'inertie principaux* du solide, c'est-à-dire ses moments d'inertie par rapport aux trois axes principaux d'inertie. L'orientation du solide par rapport à un trièdre de référence est repérée par les trois *angles d'Euler* permettant de passer du trièdre de référence au trièdre principal attaché au solide.

A chaque moment d'inertie principal on associe une *température caractéristique de rotation* $T_\alpha^{(r)}$ telle que

$$kT_\alpha^{(r)} \widehat{=} \frac{\hbar^2}{2I_\alpha}, \quad \alpha = 1, 2, 3. \tag{B.62}$$

Comme ces températures sont toujours très basses (voir tableau III), on est en pratique dans la situation où

$$T \gg T_\alpha^{(r)}, \tag{B.63}$$

21. Voir par exemple H. Goldstein, « Mécanique Classique », P.U.F., Paris (1964), chapitre IV.

c'est-à-dire où *la mécanique classique est applicable* (cf. note 14). Le calcul de la fonction de partition ne présente pas de difficulté[22]; le résultat est très simple :

$$\zeta_r = \frac{\pi^{1/2}}{\sigma}\left(\frac{2I_1kT}{\hbar^2}\right)^{1/2}\left(\frac{2I_2kT}{\hbar^2}\right)^{1/2}\left(\frac{2I_3kT}{\hbar^2}\right)^{1/2} . \qquad \text{(B.64)}$$

Cette expression reste évidemment valable si deux des moments d'inertie, ou les trois, sont égaux[23], [24]. Le *facteur de symétrie* σ est défini et déterminé selon le même principe que pour les molécules diatomiques (§ c) : on considère comme indiscernables les noyaux identiques que comporte la molécule; si elle admet un axe de symétrie unique, σ est égal au nombre de rotations différentes, autour de cet axe, qui laissent inchangée la position de la molécule (y compris la rotation d'angle 2π); s'il existe plusieurs axes de symétrie, il faut choisir parmi eux un sous-ensemble permettant d'obtenir une fois et une seule toutes les transformations de la molécule en elle-même par rotation, calculer le facteur de symétrie associé à chaque axe de ce sous-ensemble et multiplier ces nombres entre eux pour aboutir au facteur de symétrie global σ[25].

---

22. Lorsqu'on repère l'orientation du trièdre principal d'inertie de la molécule par les angles d'Euler θ, φ, ψ, son énergie de rotation a pour expression

$$\eta_r = \frac{1}{2}I_1(\dot\theta\cos\psi + \dot\varphi\sin\theta\sin\psi)^2 + \frac{1}{2}I_2(-\dot\theta\sin\psi + \dot\varphi\sin\theta\cos\psi)^2 + \frac{1}{2}I_3(\dot\psi + \dot\varphi\cos\theta)^2 . \qquad (1)$$

On en déduit la fonction de Hamilton par la méthode habituelle (appendice III, § II.2) :

$$\mathcal{H}(\theta, \varphi, \psi; p_\theta, p_\varphi, p_\psi) = \frac{p_1^2}{2I_1\sin^2\theta} + \frac{p_2^2}{2I_2\sin^2\theta} + \frac{p_3^2}{2I_3} , \qquad (2)$$

où nous avons posé

$$p_1 \mathrel{\widehat{=}} (p_\varphi - p_\psi\cos\theta)\sin\psi + p_\theta\sin\theta\cos\psi , \qquad (3.a)$$

$$p_2 \mathrel{\widehat{=}} (p_\varphi - p_\psi\cos\theta)\cos\psi - p_\theta\sin\theta\sin\psi , \qquad (3.b)$$

$$p_3 \mathrel{\widehat{=}} p_\psi . \qquad (3.c)$$

Le domaine de variation des coordonnées généralisées θ, φ, ψ est limité par les conditions

$$0 \leqslant \theta \leqslant \pi, \quad 0 \leqslant \varphi < 2\pi, \quad 0 \leqslant \psi < 2\pi ; \qquad (4)$$

les moments conjugués p_θ, p_φ, p_ψ varient de $-\infty$ à $+\infty$. La fonction de partition ζ_r se calcule facilement à partir de ces données (exercice III.28).

23. On classe les molécules polyatomiques non linéaires en trois catégories, suivant les valeurs relatives de leurs trois moments principaux d'inertie : si $I_1 = I_2 = I_3$, la molécule est une *« toupie sphérique »;* si $I_1 = I_2 \neq I_3$, c'est une *« toupie symétrique »;* enfin, si ses trois moments principaux sont différents, c'est une *« toupie asymétrique »*. Pour une toupie sphérique, tout axe passant par le centre de masse est axe principal d'inertie et le trièdre lié à la molécule peut être fixé arbitrairement. Pour une toupie symétrique, l'un des axes principaux est déterminé (c'est un axe de symétrie d'ordre supérieur ou égal à 3), les deux autres pouvant être choisis librement dans le plan perpendiculaire.

24. Par contre, la fonction de partition rotationnelle d'une molécule linéaire (formule (B.60)) ne s'obtient évidemment pas en égalant à zéro l'un des moments d'inertie : l'espace des phases de rotation d'une molécule linéaire est à 4 dimensions seulement et non plus à 6, l'angle ψ et son moment conjugué étant supprimés (cf. formules (B.41), note 16 et exercice III.27).

25. On vérifiera que $\sigma = 2$ pour H_2O (rotations de π et de 2π autour de l'axe joignant O et le milieu du segment défini par les deux H), $\sigma = 3$ pour NH_3 qui possède un axe de symétrie ternaire, $\sigma = 12$ pour le benzène C_6H_6 (6 axes de symétrie binaires dans le plan de la molécule,

.../...

On déduit facilement de ζ_r la contribution rotationnelle aux propriétés du gaz, grâce aux formules du paragraphe I.3. En particulier, l'énergie moyenne de rotation vaut $3NkT/2$, conformément au théorème d'équipartition de l'énergie (3 degrés de liberté indépendants quadratiques pour chaque molécule; voir note 22), et la capacité calorifique

$$C_V^{(r)} = \frac{3}{2} Nk \quad \text{(molécule non linéaire).} \tag{B.65}$$

4. CONCLUSION ET DISCUSSION

a. Résumé des résultats

Nous avons considéré dans ce complément des *gaz parfaits,* et analysé comment la *structure interne* des molécules qui les composent influe sur leurs propriétés. Nous avons pour cela supposé que les degrés de liberté électroniques, vibrationnels et rotationnels de la molécule sont *indépendants* les uns des autres, que les vibrations moléculaires sont *harmoniques* et que les molécules sont des *«rotateurs rigides»*. La validité et les limites de l'approximation du gaz parfait classique ont été discutées au chapitre III, paragraphe IV.B.2.b; celles des hypothèses précédentes le seront succinctement au paragraphe b ci-dessous.

Dans le cadre de ces approximations, la fonction de partition Z du gaz se calcule à partir de celle d'une molécule, z (formule (B.1)), qui est elle-même donnée comme un produit :

$$z = z_0 \cdot \zeta_e \cdot \zeta_r \cdot \zeta_v \,. \tag{B.66}$$

Le premier facteur, z_0, correspond aux *degrés de liberté de translation,* c'est-à-dire au mouvement (libre) du centre de masse de la molécule; il est le seul à dépendre du volume V; il a été calculé au chapitre III, paragraphe V.B.2.a et son expression est rappelée en (B.7).

On peut le plus souvent *négliger l'influence des états excités du système électronique de la molécule,* ce qui rend ζ_e très simple. On choisit habituellement comme zéro d'énergie celle de la configuration où les atomes formant la molécule sont séparés les uns des autres, au repos et dans leur état fondamental[26]. Si l'énergie électronique minimale de la molécule vaut alors $-u_0$ (cf. figure 1 pour une molécule diatomique), on prendra

$$\zeta_e = g_0 \, e^{+\beta u_0} \,, \tag{B.67}$$

g_0 désignant le degré de dégénérescence du niveau fondamental $-u_0$: dans le cas le plus fréquent où le cortège électronique de la molécule a un moment cinétique orbital nul, g_0 s'écrit

$$g_0 = (2S+1)(2i_1+1)(2i_2+1)\ldots \,, \tag{B.68}$$

---

ou un axe de symétrie d'ordre 6 perpendiculaire à ce plan et l'un des axes binaires précédents), 4 pour l'éthylène C_2H_4 (deux axes de symétrie binaires), et 12 pour le méthane tétraédrique CH_4 (4 axes ternaires).

26. Nous avons compté plus haut les énergies vibrationnelles à partir du minimum de l'énergie potentielle u (formule (B.25)); c'est donc ici qu'intervient la valeur $-u_0$ de ce minimum (pour les énergies de translation et de rotation, le choix du zéro s'impose de lui-même, puisque ce sont des énergies purement cinétiques). Noter que l'énergie de dissociation de la molécule (énergie minimale qu'il faut fournir pour séparer les atomes) n'est pas u_0, mais

$$u_0 - \sum_{\alpha=1}^{n_v} \frac{1}{2} \hbar\omega_\alpha$$

à cause de l'«énergie de point zéro» de l'oscillateur harmonique quantique.

où S est le spin total des électrons et i_1, i_2, ... les spins des noyaux constituant la molécule.

Pour toutes les molécules sauf celles d'hydrogène et de deutérium H_2, D_2 et HD (voir exercices III.7 et III.22), *les degrés de liberté de rotation peuvent être traités par la mécanique classique.* La fonction de partition ζ_r correspondante est donnée par la formule (B.60) pour les molécules linéaires et (B.64) pour les molécules non linéaires. Au contraire, *les degrés de liberté de vibration nécessitent un traitement quantique,* car les températures caractéristiques qui leur sont associées sont beaucoup plus élevées que pour la rotation (cf. tableaux II et III). La fonction de partition vibrationnelle ζ_v a la forme générale (B.37), qui se réduit à (B.27) pour une molécule diatomique.

Écrivons par exemple la *fonction de partition interne* ζ pour une molécule *diatomique* dont le cortège électronique a un moment cinétique orbital total et un spin total nuls ($L = S = 0$), ce qui est le cas le plus courant :

$$\zeta = (2i_1+1)(2i_2+1)\,e^{u_0/kT} \cdot \frac{1}{\sigma} \frac{2IkT}{\hbar^2} \cdot \frac{1}{2 \operatorname{sh} \hbar\omega/2kT} ; \tag{B.69}$$

i_1 et i_2 sont les spins des deux noyaux; σ vaut 1 si les noyaux sont différents (moléculaire hétéropolaire), 2 s'ils sont identiques (molécule homopolaire); I est le moment d'inertie de la molécule par rapport à son centre de masse, ω sa pulsation de vibration et u_0 son énergie de liaison. De façon plus générale, la *capacité calorifique à volume constant d'un gaz parfait classique* a pour expression

$$C_V = \frac{3}{2} Nk + C_V^{(r)} + Nk \sum_{\alpha=1}^{n_v} \frac{(\hbar\omega_\alpha/2kT)^2}{\operatorname{sh}^2(\hbar\omega_\alpha/2kT)} , \tag{B.70}$$

où $C_V^{(r)} = Nk$ pour des molécules linéaires,

$= \frac{3}{2} Nk$ pour des molécules non linéaires. (B.71)

Chacun des termes de la somme correspond à un mode normal (α) de vibration, auquel est associée une température caractéristique $T_\alpha^{(v)}$ (formule (B.40)); la contribution du mode (α) croît de 0 à Nk sur un intervalle de température assez restreint, grosso-modo lorsque T croît de $T_\alpha^{(v)}/10$ à $T_\alpha^{(v)}$ (fig. 2). Les températures caractéristiques de vibration d'une même molécule sont souvent assez différentes les unes des autres, bien qu'elles restent du même ordre de grandeur (cf. tableau III). Il s'ensuit que les divers modes normaux de vibration sont excités successivement lorsque la température croît, certains pouvant avoir atteint leur limite classique alors que d'autres sont encore presque gelés. Il arrive assez souvent que la molécule se dissocie avant que le dernier mode normal ait pu être pleinement excité.

b. Anharmonicité des vibrations et couplage rotation-vibration

Il est clair que les hypothèses sur lesquelles sont fondés les résultats que nous venons de résumer ne sont pas parfaitement vérifiées dans la réalité. Prenons le cas le plus simple d'une molécule diatomique. Il est visible sur la figure 1 que le potentiel régissant les vibrations n'est pas strictement parabolique; on s'attend donc à ce que l'approximation harmonique perde de sa validité lorsque l'amplitude des vibrations augmente. De plus, lorsqu'une molécule tourne, les noyaux ont tendance à s'écarter l'un de l'autre sous l'effet de la «force centrifuge»; ainsi, la rotation de la molécule influe sur le potentiel d'interaction entre les atomes et modifie par conséquent la fréquence de vibration : il faut remplacer $u(\rho)$ par le «potentiel effectif»

$$u_{\text{eff}}(\rho) = u(\rho) + j(j+1) \frac{\hbar^2}{2m_R\rho^2} , \tag{B.72}$$

ce qui décale légèrement la position du minimum et affecte aussi la courbe en ce point. Inversement, les vibrations de la molécule font varier la distance ρ entre les noyaux et par là le moment d'inertie, et influencent donc les mouvements de rotation.

Les effets d'anharmonicité et de couplage rotation-vibration peuvent être analysés (C.D.L., p. 867), et leur influence sur la fonction de partition interne peut être calculée (voir exercice III.10). Les corrections obtenues sont *très faibles*. On peut le comprendre simplement en donnant quelques ordres de grandeur. En premier lieu, l'extension de la fonction d'onde associée à l'énergie de vibration $(\nu + 1/2)\hbar\omega$ est de l'ordre de (C.D.L., p. 502)

$$\Delta\rho = \sqrt{\left(\nu + \frac{1}{2}\right)\frac{\hbar}{m_R\omega}} \,. \tag{B.73}$$

Il est facile de voir[27] que cette « amplitude de vibration » est toujours petite devant d :

$$\Delta\rho \ll d \,. \tag{B.74}$$

On en déduit que les effets d'anharmonicité[28] d'une part, et d'autre part les modifications du moment d'inertie induites par les vibrations, vont effectivement rester petits. En outre, le « terme centrifuge » de l'expression (B.72) n'st autre que l'énergie de rotation de la molécule. Sa valeur moyenne kT ne dépasse pas une fraction d'électron-volt si T reste inférieure à quelques milliers de kelvins, alors que l'énergie de liaison u_0 vaut typiquement entre 2 et 10 eV selon les molécules. L'influence de la rotation de la molécule sur sa fréquence de vibration est donc elle aussi très faible.

---

27. La condition

$$\Delta\rho \ll d \tag{1}$$

se transforme aisément en

$$\left(\nu + \frac{1}{2}\right)\frac{\hbar^2}{m_R d^2} \ll \hbar\omega \,, \tag{2}$$

puis, grâce à (B.30) et (B.44), en

$$(2\nu + 1) \ll \frac{T_v}{T_r} \,. \tag{3}$$

Les données du tableau II montrent que cette condition est largement satisfaite en pratique jusqu'à des valeurs de ν relativement élevées; notons que la valeur moyenne de $(2\nu + 1)$ est en réalité bornée par sa limite classique

$$\overline{2\nu + 1} \leqslant \frac{kT}{\hbar\omega} = \frac{T}{T_v} \,. \tag{4}$$

28. Pour montrer que l'anharmonicité du potentiel joue un rôle peu important, il faudrait en toute rigueur vérifier que, pour $(\rho - d)$ de l'ordre de $\Delta\rho$, le terme en $(\rho - d)^3$ que l'on néglige est petit devant le terme parabolique que l'on conserve (voir C.D.L., p. 874).

COMPLÉMENT III.C

La théorie cinétique des gaz

I. Distribution statistique des vitesses dans un fluide
   1. La distribution de Maxwell
   2. Distribution statistique d'une composante de la vitesse
   3. Distribution statistique du module de la vitesse
   4. Vérifications expérimentales

II. Pression d'un gaz
   1. Définition mécanique
   2. Méthode de calcul de la pression : pression cinétique et pression dynamique

III. Calcul de la pression cinétique d'un gaz homogène
   1. Calcul simpliste
   2. Calcul exact

IV. Pression dynamique
   1. Considérations qualitatives
   2. Calcul de la pression dynamique dans un gaz monoatomique dilué
      a. Hypothèses et approximations
      b. Force exercée sur un atome du cylindre $\mathcal{C}$ par les atomes extérieurs situés entre deux plans infiniment voisins
      c. Force totale exercée par les atomes extérieurs sur ceux du cylindre $\mathcal{C}$; pression dynamique
      d. Expression approchée de la pression du gaz

La théorie cinétique des gaz a été développée pendant la deuxième moitié du dix-neuvième siècle. Elle avait pour but d'expliquer les propriétés macroscopiques d'un gaz à partir du comportement de ses constituants microscopiques. A une époque où l'existence des atomes n'était pas universellement admise, les prédictions de la théorie cinétique concernant des phénomènes observables à l'échelle macroscopique revêtaient une importance capitale.

On peut retrouver facilement la plupart des résultats de la théorie cinétique à l'aide du formalisme général de la mécanique statistique, mais la théorie cinétique, plus proche de la mécanique, a l'avantage de mieux faire comprendre l'origine physique des phénomènes. Elle donne en particulier une interprétation intéressante de la pression dans un fluide.

I. DISTRIBUTION STATISTIQUE DES VITESSES DANS UN FLUIDE

1. LA DISTRIBUTION DE MAXWELL

C'est par des considérations générales de symétrie que Maxwell découvrit en 1860 la loi de répartition des vitesses pour les molécules d'un gaz parfait à l'équilibre (voir référence 1). Boltzmann démontra ensuite (1872) cette loi en supposant que la distribution d'équilibre n'est pas modifiée par les collisions entre molécules (cf. complément IV.E, § II.3). Elle apparaît aujourd'hui comme une conséquence directe de la distribution canonique en mécanique statistique classique. La distribution des vitesses de Maxwell est en fait valable pour *tous les fluides classiques,* gaz réels ou liquides.

Considérons un fluide de N atomes ou molécules identiques, de masse m, contenus dans un récipient de volume V et en équilibre thermique avec un thermostat à la température T. Nous supposons que l'on peut traiter le fluide dans le cadre de la *mécanique classique* (ce qui est toujours le cas, sauf pour l'hélium à basse température; cf. complément I.A, § III.2.b). Un état microscopique du système est

caractérisé par les positions $\vec{r}_i$ et les vitesses de translation $\vec{v}_i$ des N molécules[1]. Dans un tel état, le système a pour énergie :

$$E(\vec{r}_1, \vec{r}_2, \ldots, \vec{r}_N; \vec{v}_1, \vec{v}_2, \ldots, \vec{v}_N) = \sum_{i=1}^{N} \frac{1}{2} m\vec{v}_i^2 + U(\vec{r}_1, \vec{r}_2, \ldots, \vec{r}_N), \tag{C.1}$$

où $U(\vec{r}_1, \ldots, \vec{r}_N)$ est l'énergie d'interaction des molécules entre elles et éventuellement avec un champ extérieur. La probabilité de trouver le système dans cet état, ou plus précisément dans un domaine infiniment petit entourant cet état dans l'espace des phases (à $6N$ dimensions), est donnée par la distribution canonique (III, § V.A.1) :

$$\mathrm{d}P(\vec{r}_1, \vec{r}_2, \ldots, \vec{r}_N; \vec{v}_1, \vec{v}_2, \ldots, \vec{v}_N) = A \exp\left\{-\beta\left[\sum_{i=1}^{N} \frac{1}{2} m\vec{v}_i^2 + U(\vec{r}_1, \ldots, \vec{r}_N)\right]\right\} \mathrm{d}^3r_1\, \mathrm{d}^3r_2 \ldots \mathrm{d}^3r_N\, \mathrm{d}^3v_1\, \mathrm{d}^3v_2 \ldots \mathrm{d}^3v_N, \tag{C.2}$$

où A est un facteur de normalisation :

$$\frac{1}{A} = \int \mathrm{d}^3r_1 \ldots \mathrm{d}^3r_N\, \mathrm{d}^3v_1 \ldots \mathrm{d}^3v_N \exp\left\{-\beta\left[\sum_{i=1}^{N} \frac{1}{2} m\vec{v}_i^2 + U(\vec{r}_1, \ldots, \vec{r}_N)\right]\right\}. \tag{C.3}$$

En intégrant l'expression (C.2) sur toutes les valeurs possibles des positions $\vec{r}_i$ des N particules, on obtient la probabilité pour que les molécules aient des vitesses $\vec{v}_i$ déterminées (à d^3v_i près) :

$$\mathrm{d}P(\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_N) = A \exp\left\{-\beta \sum_{i=1}^{N} \frac{1}{2} m\vec{v}_i^2\right\} \mathrm{d}^3v_1 \ldots \mathrm{d}^3v_N \int \mathrm{d}^3r_1 \ldots \mathrm{d}^3r_N \exp\{-\beta U(\vec{r}_1, \ldots, \vec{r}_N)\}$$

$$= B \exp\left[-\beta \sum_{i=1}^{N} \frac{1}{2} m\vec{v}_i^2\right] \mathrm{d}^3v_1 \ldots \mathrm{d}^3v_N. \tag{C.4}$$

Une nouvelle intégration sur les vitesses de toutes les molécules sauf une fournit la probabilité pour qu'une molécule ait la vitesse $\vec{v}$ à d^3v près[2] :

$$\mathrm{d}P(\vec{v}) = \left(\frac{m}{2\pi kT}\right)^{3/2} \exp\left[-\frac{m\vec{v}^2}{2kT}\right] \mathrm{d}^3v. \tag{C.5}$$

---

1. Nous n'explicitons ici que les degrés de liberté de translation (position, vitesse) des molécules. Celles-ci possèdent également des degrés de liberté internes (cf. complément III.B), dont certains exigent d'ailleurs un traitement quantique. Mais, pourvu que les mouvements de translation puissent être décrits classiquement, la distribution des vitesses n'est pas affectée par les degrés de liberté internes des molécules.

2. En effet,

$$\mathrm{d}P(\vec{v}_1) = B \exp\left[-\beta \frac{m\vec{v}_1^2}{2}\right] \mathrm{d}^3v_1 \int \mathrm{d}^3v_2 \ldots \mathrm{d}^3v_N \exp\left[-\beta \sum_{i=2}^{N} \frac{1}{2} m\vec{v}_i^2\right]$$

$$= C \exp\left[-\beta \frac{m\vec{v}_1^2}{2}\right] \mathrm{d}^3v_1.$$

On détermine ensuite facilement la constante C à partir de la condition de normalisation

$$C \int_{-\infty}^{+\infty} \mathrm{d}v_x \exp\left[-\beta \frac{mv_x^2}{2}\right] \int_{-\infty}^{+\infty} \mathrm{d}v_y \exp\left[-\beta \frac{mv_y^2}{2}\right] \int_{-\infty}^{+\infty} \mathrm{d}v_z \exp\left[-\beta \frac{mv_z^2}{2}\right] = 1$$

(voir formulaire).

La probabilité (C.4) concernant les N molécules s'écrit donc comme le produit de N termes analogues à (C.5), un pour chaque molécule : *il n'y a pas de corrélation entre les vitesses des diverses molécules*[3].

Le nombre moyen $dN(\vec{v})$ de molécules du système ayant la vitesse $\vec{v}$ (à d^3v près) est donc :

$$dN(\vec{v}) = N\, dP(\vec{v})$$
$$= N\left(\frac{m}{2\pi kT}\right)^{3/2} \exp\left[-m(v_x^2+v_y^2+v_z^2)/2kT\right] dv_x\, dv_y\, dv_z\,. \qquad \text{(C.6)}$$

Cette formule caractérise la *distribution de Maxwell des vitesses.* Elle est valable pour les gaz et les liquides que l'on peut décrire par la *mécanique classique,* à condition qu'il n'y ait *pas d'interactions dépendant des vitesses*[4].

2. DISTRIBUTION STATISTIQUE D'UNE COMPOSANTE DE LA VITESSE

Le nombre $dN(v_x)$ de molécules dont la vitesse a une composante suivant l'axe Ox égale à v_x à dv_x près se déduit aussitôt de la formule (C.6) par intégration sur v_y et v_z :

$$dN(v_x) = N\sqrt{\frac{m}{2\pi kT}} \exp\left[-mv_x^2/2kT\right] dv_x\,. \qquad \text{(C.7)}$$

Nous avons représenté sur la figure 1 la fonction

$$\varphi(v_x) \widehat{=} \frac{1}{N}\frac{dN(v_x)}{dv_x}\,; \qquad \text{(C.8)}$$

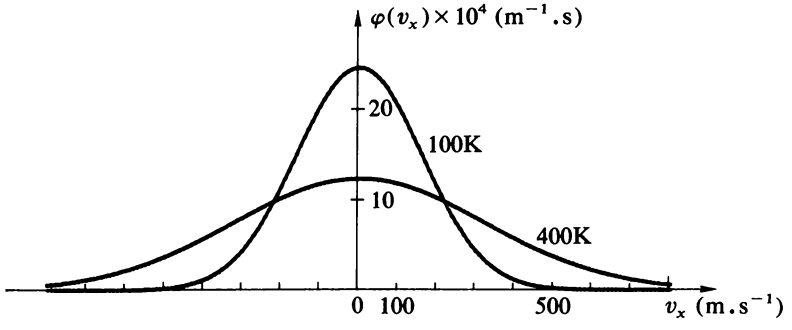


FIGURE 1

Distribution statistique d'une composante de la vitesse pour deux valeurs de la température ($T_1 = 100$ K; $T_2 = 400$ K), dans un gaz d'oxygène ($m = 32 \times 1{,}66 \times 10^{-27}$ kg).

3. Cette propriété provient du fait que l'énergie cinétique d'une molécule donnée est indépendante de la vitesse des autres. Il n'y a pas non plus de corrélation entre la position $\vec{r}$ et la vitesse $\vec{v}$ d'une molécule :

$$dP(\vec{r}, \vec{v}) = dP'(\vec{r}).dP''(\vec{v})\,.$$

En revanche, il y a en général corrélation entre les positions de deux ou plusieurs molécules (cf. paragraphe IV.2.d ci-dessous).

4. Les interactions électromagnétiques entre particules chargées relativistes dépendent de leurs vitesses. Il faut donc exclure de cette étude les plasmas relativistes.

c'est une gaussienne centrée en $v_x = 0$. La valeur moyenne et l'écart quadratique moyen de v_x s'en déduisent aussitôt :

$$\bar{v}_x = 0 \,, \tag{C.9.a}$$

$$\Delta v_x = \left[\overline{v_x^2}\right]^{1/2} = \sqrt{\frac{kT}{m}} \,. \tag{C.9.b}$$

Notons que ces résultats sont en accord avec le théorème d'équipartition de l'énergie (III, § V.C.1) :

$$\frac{1}{2} m v_x^2 = \frac{1}{2} kT \,. \tag{C.10}$$

La distribution de chacune des deux autres composantes v_y et v_z (et plus généralement celle de la projection de la vitesse sur un axe quelconque) est évidemment la même que celle de v_x, puisqu'aucune direction de l'espace n'est ici privilégiée.

3. DISTRIBUTION STATISTIQUE DU MODULE DE LA VITESSE

On peut récrire (C.6) en caractérisant le vecteur vitesse $\vec{v}$ par son module v et les angles θ et φ qui repèrent sa direction par rapport aux axes de coordonnées :

$$\mathrm{d}N(\vec{v}) = N \left(\frac{m}{2\pi kT}\right)^{3/2} \mathrm{e}^{-mv^2/2kT} v^2 \sin\theta \, \mathrm{d}v \, \mathrm{d}\theta \, \mathrm{d}\varphi \,. \tag{C.11}$$

Si l'on somme sur toutes les directions possibles, on trouve le nombre $\mathrm{d}N(v)$ de molécules dont la vitesse a un module compris entre v et $v + \mathrm{d}v$:

$$\mathrm{d}N(v) = N \left(\frac{m}{2\pi kT}\right)^{3/2} 4\pi v^2 \, \mathrm{e}^{-mv^2/2kT} \, \mathrm{d}v \,. \tag{C.12}$$

Nous avons représenté sur la figure 2

$$\psi(v) \widehat{=} \frac{1}{N} \frac{\mathrm{d}N(v)}{\mathrm{d}v} \,; \tag{C.13}$$

on appelle parfois cette fonction une « maxwellienne » : elle atteint son maximum pour $v_M = \sqrt{2kT/m}$, et ce maximum vaut $\sqrt{8m/\pi kT}/e$.

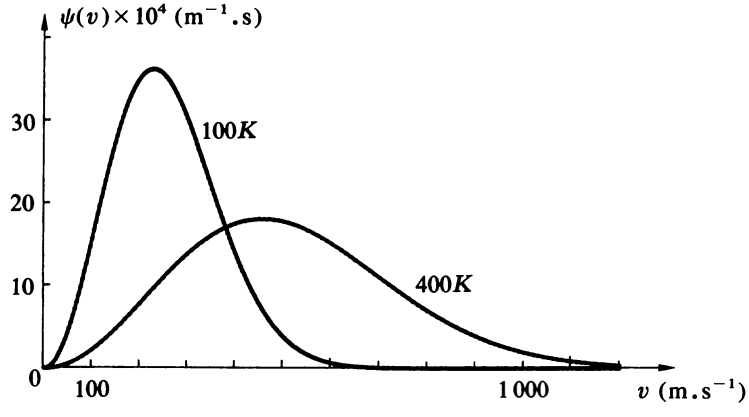


FIGURE 2

Distribution statistique du module de la vitesse pour deux valeurs de la température ($T_1 = 100$ K; $T_2 = 400$ K) dans un gaz d'oxygène.

Il est facile de calculer les valeurs moyennes de v et v^2 :

$$\bar{v}=\frac{1}{N}\int_0^\infty v\,\mathrm{d}N(v)=2\sqrt{\frac{2kT}{\pi m}}\,, \tag{C.14.a}$$

$$\overline{v^2}=\frac{1}{N}\int_0^\infty v^2\,\mathrm{d}N(v)=\frac{3kT}{m}\,. \tag{C.14.b}$$

On constate, en comparant ce dernier résultat à (C.9.b), que

$$\overline{v^2}=3\overline{v_x^2}\,; \tag{C.15}$$

ceci était attendu puisque

$$v^2=v_x^2+v_y^2+v_z^2 \tag{C.16}$$

et que toutes les directions de l'espace sont équivalentes.

4. VÉRIFICATIONS EXPÉRIMENTALES

Une vérification expérimentale indirecte de la distribution de Maxwell des vitesses est fournie par l'*élargissement Doppler des raies spectrales* émises par les molécules d'un gaz. Lorsqu'une molécule émettant une raie de fréquence ν_0 est en mouvement par rapport à l'observateur, la fréquence ν du rayonnement observé est déplacée par effet Doppler : ce déplacement $\nu-\nu_0$ est proportionnel à la composante de la vitesse dans la direction d'observation. L'intensité reçue à la fréquence ν est elle-même proportionnelle au nombre de molécules dont la vitesse provoque le déplacement $\nu-\nu_0$, de sorte que le profil de la raie observée reproduit la distribution des vitesses (C.7) (cf. exercice III.12). La première étude détaillée du profil des raies spectrales en fonction de la température a été réalisée par Michelson en 1892; les résultats sont en accord avec la prédiction tirée de la distribution de Maxwell.

On peut aussi vérifier directement la distribution de Maxwell en comptant le nombre de molécules de vitesse donnée présentes dans un « jet moléculaire » : on perce un trou fin dans la paroi d'un récipient contenant un gaz porté à une température connue, et on observe les molécules qui s'en échappent. En étudiant ce phénomène, appelé *effusion* (voir exercice III.13), on peut montrer que le nombre de molécules qui, dans le jet moléculaire, ont une vitesse de module v (à $\mathrm{d}v$ près) est proportionnel à $v^3\exp(-mv^2/2kT)$. Cette prédiction théorique est confirmée par l'expérience[5].

Ces vérifications explicites de la distribution de Maxwell, commencées vers 1920, ont atteint une grande précision dans les années cinquante. Mais bien avant cette époque existait un argument très convaincant en faveur de la théorie cinétique des gaz : comme nous allons le voir maintenant, elle permet de retrouver l'équation d'état des gaz et fournit ainsi une explication théorique à leurs propriétés.

II. PRESSION D'UN GAZ

1. DÉFINITION MÉCANIQUE

Un gaz exerce sur les parois du récipient qui le contient une force normale à la paroi, qui tend à la repousser : un élément de paroi, d'aire $\mathrm{d}s$, est soumis à la force

$$\mathrm{d}\vec{F}=p\,\mathrm{d}\vec{s}\,, \tag{C.17}$$

5. Voir R.C. Miller and P. Kusch, *J. Chem. Phys., 25,* p. 860 (1956).

où le vecteur $\overrightarrow{ds}$ (de module ds) est normal à l'élément de surface considéré et dirigé vers l'extérieur du récipient contenant le gaz; le coefficient de proportionnalité p est *la pression exercée par le gaz sur la paroi*[6]. D'après le principe des actions réciproques, la paroi exerce sur le gaz une force opposée

$$\overrightarrow{dF}' = -\overrightarrow{dF}. \qquad \text{(C.18)}$$

Pour ne pas alourdir l'écriture, nous n'expliciterons plus dans la suite le caractère infinitésimal de l'élément de surface $\overrightarrow{ds}$ et de la force $\overrightarrow{dF}$: nous considérerons une portion de paroi plane, et nous noterons Σ un disque d'aire s découpé dans cette paroi, $\vec{F}$ la force (normale) exercée par le gaz sur la surface Σ et $\vec{F}'$ la force opposée exercée par Σ sur le gaz.

2. MÉTHODE DE CALCUL DE LA PRESSION : PRESSION CINÉTIQUE ET PRESSION DYNAMIQUE

Pour calculer la pression p, il nous faut déterminer la force $\vec{F}$ qu'exerce le gaz sur l'élément de paroi Σ, ou la force $\vec{F}'$ qu'exerce Σ en retour sur le gaz. Mais ce qui se passe au voisinage immédiat de la paroi peut être complexe et mal connu; nous allons voir qu'il n'est pas nécessaire d'en tenir compte pour calculer la pression.

Imaginons un cylindre $\mathcal{C}$ normal à la paroi, situé à l'intérieur du récipient (fig. 3) : il est fermé du côté de la paroi par le disque Σ, limité de l'autre par un disque Σ' parallèle à Σ et de même aire s. Nous allons montrer que le calcul de la pression ne fait intervenir que le gaz au voisinage du disque Σ'. A l'équilibre, la quantité de mouvement moyenne des molécules contenues dans le cylindre $\mathcal{C}$ doit être nulle : si elle ne l'était pas, cette portion du gaz serait animée d'un mouvement d'ensemble et ne serait donc pas à l'équilibre. Or plusieurs causes sont susceptibles de faire varier la quantité de mouvement de la portion de gaz intérieure au cylindre $\mathcal{C}$: d'une part des molécules traversent la surface qui le délimite, apportant ou emportant leur quantité de mouvement; d'autre part, les molécules de gaz situées à l'extérieur de $\mathcal{C}$ ainsi que les parois du récipient exercent des forces sur les molécules contenues dans $\mathcal{C}$.

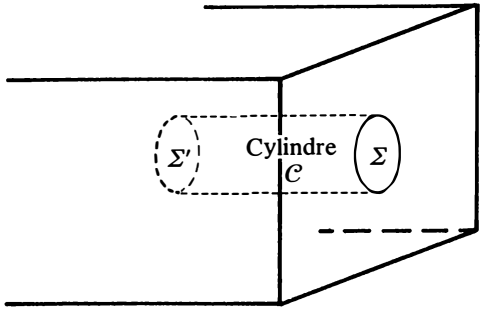


FIGURE 3
A l'équilibre, la quantité de mouvement totale du gaz contenu dans le cylindre $\mathcal{C}$ doit rester nulle en moyenne.

6. Nous avons montré dans le complément II.D (§ I.3) l'équivalence entre la formule (C.17) et la définition (III.29.a) du chapitre II :

$$\frac{p^*}{T^*} \widehat{=} \frac{\partial S^*}{\partial V}.$$

Les raisonnements qui suivent s'appuient de façon essentielle sur le fait que *les interactions entre molécules ont une portée d'ordre microscopique,* au plus une dizaine d'angströms; ceci s'applique aux interactions à l'intérieur du gaz, mais aussi à celles qui s'exercent entre le gaz et les parois. Au contraire, *les dimensions du cylindre $\mathcal{C}$ sont d'ordre macroscopique;* même si elles sont infinitésimales à cette échelle, elles sont toujours très grandes devant la portée des interactions. Cette différence d'ordre de grandeur a une première conséquence importante : on peut négliger les interactions du gaz contenu dans $\mathcal{C}$ avec les portions de paroi autres que le disque Σ, et les interactions de ce dernier avec les molécules de gaz situées hors de $\mathcal{C}$. Ainsi, la force exercée par les parois du récipient sur le gaz contenu dans $\mathcal{C}$, qui contribue à la variation de sa quantité de mouvement, s'identifie à la force $\vec{F}'$ exercée par le disque Σ sur l'ensemble du gaz, opposée à la force $\vec{F}$ qui apparaît dans la définition (C.17) de la pression. En outre, tout se passe comme si le cylindre $\mathcal{C}$ était découpé dans un récipient infini et, puisque Σ et Σ' sont des disques, le problème présente la symétrie de révolution autour de l'axe du cylindre $\mathcal{C}$.

Appelons $\mathrm{d}\vec{\mathcal{I}}_{\mathrm{cin}}$ la variation de quantité de mouvement provoquée dans $\mathcal{C}$ par les molécules qui traversent la surface du cylindre pendant un intervalle de temps dt, et $\vec{F}_g$ la résultante des forces exercées sur les molécules de $\mathcal{C}$ par le reste du gaz. La condition d'équilibre s'écrit

$$\vec{F}'\,\mathrm{d}t + \vec{F}_g\,\mathrm{d}t + \mathrm{d}\vec{\mathcal{I}}_{\mathrm{cin}} = 0\,, \tag{C.19.a}$$

ou

$$\vec{F}\,\mathrm{d}t = \mathrm{d}\vec{\mathcal{I}}_{\mathrm{cin}} + \vec{F}_g\,\mathrm{d}t\,. \tag{C.19.b}$$

Par raison de symétrie, les vecteurs intervenant dans (C.19) sont dirigés suivant la normale à Σ; nous pouvons en déduire la pression (voir (C.17)) sous la forme

$$p = \frac{1}{s}\frac{\mathrm{d}\mathcal{I}_{\mathrm{cin}}}{\mathrm{d}t} + \frac{1}{s}F_g\,, \tag{C.20}$$

où $\mathrm{d}\mathcal{I}_{\mathrm{cin}}$ et F_g sont des quantités algébriques, comptées le long de la normale extérieure à Σ. La comparaison des dimensions de $\mathcal{C}$ avec la portée des forces (voir ci-dessus) nous permettra, dans le calcul des deux contributions au second membre de (C.20), de *négliger les effets de bords,* c'est-à-dire les effets particuliers qui se produisent au voisinage immédiat des contours limitant les disques Σ et Σ' (voir plus loin la figure 4)

Nous voyons que la pression est la somme de deux termes. Le premier, qui provient du déplacement des molécules, est appelé la *«pression cinétique»;* le second est dû aux interactions entre les molécules du gaz : c'est la *«pression dynamique».*

III. CALCUL DE LA PRESSION CINÉTIQUE D'UN GAZ HOMOGÈNE

Dans un gaz homogène[7], tout élément de volume w contient un nombre de molécules égal à

$$n = N\frac{w}{V}\,, \tag{C.21}$$

dont les vitesses sont réparties suivant la distribution de Maxwell (C.6). Nous nous proposons de calculer le transfert de quantité de mouvement, de l'extérieur vers l'intérieur du cylindre $\mathcal{C}$ de la figure 3, causé par les molécules qui traversent sa surface pendant un intervalle de temps dt.

7. Nous excluons de cette étude les systèmes où plusieurs phases coexistent, ainsi que ceux où un champ extérieur modifie la répartition spatiale des molécules.

Dans ce calcul, seules interviennent en fait les molécules qui entrent ou sortent de $\mathcal{C}$ en traversant le disque Σ'. En effet, à toute molécule entrant (ou sortant) par la surface latérale de $\mathcal{C}$ on peut associer une autre molécule (fig. 4.a) qui entre (ou sort) avec l'impulsion opposée. En revanche, la contribution des molécules qui traversent Σ' n'est pas compensée, puisqu'aucune molécule ne traverse la face opposée Σ.

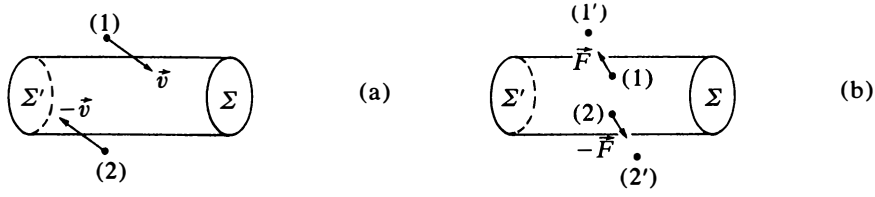


FIGURE 4
Les effets de bords étant négligeables, la quantité de mouvement totale transférée au cylindre $\mathcal{C}$ par les molécules qui traversent sa surface latérale est nulle (a), ainsi que la résultante des forces qu'exercent sur le gaz intérieur à $\mathcal{C}$ les molécules extérieures qui entourent la surface latérale (b).

1. CALCUL SIMPLISTE

Pour dégager les phénomènes physiques importants sans alourdir le raisonnement par les détails du calcul, admettons tout d'abord que les molécules ont toutes la même vitesse (nous prendrons la valeur moyenne $\bar{v}$ du module de v), et qu'un tiers d'entre elles se propagent dans la direction de chacun des trois axes choisis pour repérer l'espace; ainsi une molécule sur six se dirige parallèlement à l'axe du cylindre $\mathcal{C}$ de la gauche vers la droite avec la vitesse $\bar{v}$, et autant de la droite vers la gauche avec la même vitesse.

Appelons Oz l'axe normal à Σ et dirigé vers l'extérieur du récipient : sur la figure 5, les vitesses sont positives lorsqu'elles sont dirigées vers la droite. Lorsque la molécule (1), de quantité de mouvement $+m\bar{v}$, traverse Σ' de gauche à droite, la quantité de mouvement du gaz extérieur diminue de $m\bar{v}$ tandis que celle du cylindre $\mathcal{C}$ augmente d'autant. Lorsque la molécule (2), de quantité de mouvement $-m\bar{v}$,

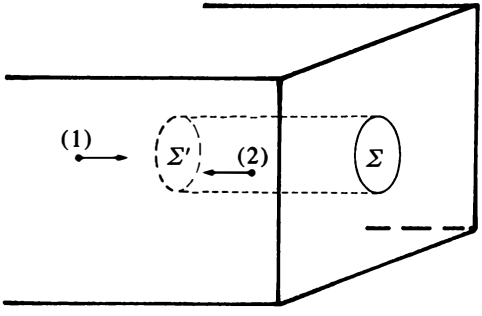


FIGURE 5
Dans un calcul simplifié, nous raisonnons comme si les seules molécules traversant Σ' avaient une vitesse normale à Σ' et de module $\bar{v}$ fixé.

traverse Σ' de droite à gauche, la quantité de mouvement du gaz extérieur « augmente » de $-m\bar{v}$, c'est-à-dire diminue de $m\bar{v}$, tandis que la quantité de mouvement de $\mathcal{C}$ augmente de $m\bar{v}$ ($\mathcal{C}$ « perd » une quantité de mouvement négative). Ainsi toute molécule qui traverse Σ', quel que soit le sens du passage, fait augmenter de $+m\bar{v}$ la quantité de mouvement du gaz contenu dans le cylindre $\mathcal{C}$ au détriment du gaz extérieur.

Cherchons maintenant quelles sont les molécules qui traversent Σ' pendant l'intervalle de temps dt. Remarquons d'abord que la distance parcourue par une molécule pendant dt est

$$d\ell = \bar{v}\, dt\,. \tag{C.22}$$

Les molécules qui traversent Σ' de gauche à droite entre les instants t et $t+dt$ sont donc celles qui se trouvent à l'instant t à gauche de Σ', avec une vitesse positive, dans le cylindre de longueur $d\ell$ s'appuyant sur Σ'. Leur nombre est

$$\frac{1}{6}\frac{N}{V}\, s\, d\ell = \frac{1}{6}\frac{N}{V}\, s\bar{v}\, dt\,. \tag{C.23}$$

La quantité de mouvement qui passe de gauche à droite avec ces molécules vaut

$$\frac{1}{6}\frac{N}{V}\, s\bar{v}\, dt\,.\, m\bar{v} = \frac{1}{6}\frac{N}{V}\, m\bar{v}^2 s\, dt\,. \tag{C.24}$$

Les molécules qui traversent Σ' de droite à gauche provoquent le même transfert de quantité de mouvement, de sorte que le transfert total à travers Σ' est

$$d\mathcal{I}_{cin} = \frac{1}{3}\frac{N}{V}\, m\bar{v}^2 s\, dt\,. \tag{C.25}$$

Dans ce modèle très simple, la pression cinétique vaut donc d'après (C.20)

$$p_{cin} \sim \frac{1}{3}\frac{N}{V}\, m\bar{v}^2\,, \tag{C.26}$$

où nous avons mis un signe $\sim$ plutôt qu'une égalité pour rappeler qu'il s'agit d'un calcul approché. Il montre cependant que *la pression cinétique est proportionnelle à la densité N/V des molécules et à leur énergie $m\bar{v}^2/2$, c'est-à-dire,* d'après (C.14.a), *à la température.* Nous ne pouvons pas prendre au sérieux le facteur 1/3 : nous avons utilisé dès le départ la valeur moyenne de la vitesse, alors que dans un calcul correct l'opération de valeur moyenne doit être effectuée seulement sur l'expression finale (ce qui remplacera $\bar{v}^2$ par $\overline{v^2}$).

2. CALCUL EXACT

Reprenons donc le calcul du transfert de quantité de mouvement, en considérant cette fois que les vitesses des molécules sont réparties suivant la distribution $dP(\vec{v})$ de Maxwell (formule (C.5)). Rappelons que, par raison de symétrie, la quantité de mouvement totale transférée est normale à Σ'.

Considérons tout d'abord les molécules qui traversent Σ' entre les instants t et $t+dt$ avec une vitesse $\vec{v}$ donnée, dirigée vers la droite (composante v_z positive); la contribution de chacune d'elles au transfert de quantité de mouvement dans la direction Oz est mv_z. A l'instant t, elles se trouvent[8] à l'intérieur du cylindre prenant appui

8. A cause des collisions entre molécules, certaines de celles qui sont contenues dans ce cylindre à l'instant t changeront de vitesse entre t et $t+dt$ et n'arriveront donc pas jusqu'à Σ' en gardant la vitesse $\vec{v}$. Mais, comme le gaz est à l'équilibre, elles seront remplacées par d'autres qui joueront exactement le même rôle. En effet, les collisions ne modifient pas la distribution de Maxwell.

sur Σ', dont les génératrices sont parallèles à la vitesse $\vec{v}$ et ont pour longueur $v\,dt$ (fig. 6). Le volume de ce cylindre est

$$dw = sv\,dt\cos\theta = sv_z\,dt\,, \tag{C.27}$$

et le nombre de molécules de vitesse $\vec{v}$ (à d^3v près) qu'il contient

$$dN = N\frac{dw}{V}\,dP(\vec{v}) = \frac{N}{V}\,dP(\vec{v})\,sv_z\,dt\,. \tag{C.28}$$

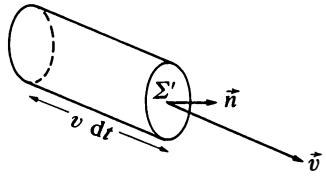


FIGURE 6
Cylindre contenant à l'instant t les molécules qui vont traverser Σ' entre les instants t et $t+dt$ avec la vitesse $\vec{v}$.

La contribution de ces molécules au transfert de quantité de mouvement à travers Σ' s'écrit donc

$$d\mathcal{I}_{cin}(\vec{v}) = \frac{N}{V}\,dP(\vec{v})\,mv_z^2 s\,dt\,. \tag{C.29}$$

On doit ensuite sommer sur toutes les valeurs possibles de $\vec{v}$ et ajouter la contribution des molécules qui traversent Σ' de droite à gauche[9]. Le transfert total de quantité de mouvement dû à l'ensemble des molécules qui traversent Σ' entre t et $t+dt$ fait donc apparaître la valeur moyenne de v_z^2 :

$$d\mathcal{I}_{cin} = \frac{N}{V}\,s\,dt\int mv_z^2\,dP(\vec{v}) = \frac{N}{V}\,\overline{mv_z^2}\,s\,dt\,. \tag{C.30}$$

La formule (C.20) donne ensuite la pression cinétique sous la forme[10] (voir (C.15))

$$p_{cin} = \frac{N}{V}\,\overline{mv_z^2} = \frac{1}{3}\frac{N}{V}\,\overline{mv^2}\,, \tag{C.31}$$

c'est-à-dire, si l'on utilise (C.14.b),

$$p_{cin} = \frac{N}{V}\,kT\,. \tag{C.32}$$

Si le fluide considéré est un *gaz parfait,* les interactions sont négligeables et *la pression se réduit à la pression cinétique.* On reconnaît effectivement dans (C.32) l'équation d'état des gaz parfaits.

9. Les composantes v_x et v_y peuvent varier de $-\infty$ à $+\infty$, tandis que v_z doit être positive pour les molécules traversant de gauche à droite. Mais les molécules qui traversent Σ' de droite à gauche ($v_z<0$) provoquent le même transfert de quantité de mouvement (positif) que celles qui traversent en sens inverse, de sorte que l'on intègre aussi de $-\infty$ à $+\infty$ sur v_z.

10. La formule (C.31) ressemble beaucoup à (C.26), que nous avons obtenue par le calcul simpliste. Mais la formule exacte comporte la valeur moyenne du carré de la vitesse tandis que c'est le carré de la valeur moyenne de la vitesse qui intervient dans (C.26). La présence du même facteur numérique 1/3 dans les deux formules ne doit pas faire illusion : elles sont vraiment différentes (cf. (C.14.a) et (C.14.b)).

IV. PRESSION DYNAMIQUE

1. CONSIDÉRATIONS QUALITATIVES

Lorsque les interactions entre molécules ne sont pas négligeables, à la pression cinétique obtenue en (C.32) s'ajoute une pression dynamique (formule (C.20)).

On sait que deux molécules très proches l'une de l'autre se repoussent fortement, tandis qu'elles s'attirent faiblement lorsqu'elles sont un peu plus éloignées, cette attraction décroissant très rapidement lorsque la distance augmente (complément I.A, § III.3). Il est facile de voir qu'aux forces attractives correspond une pression dynamique négative, appelée *«pression de cohésion»*, tandis qu'aux forces répulsives est associée une pression dynamique positive. En effet, les forces exercées par les molécules du fluide entourant la surface latérale du cylindre $\mathcal{C}$ sur celles qui se trouvent à l'intérieur ont une résultante nulle (fig. 4.b). Il suffit donc de considérer l'action sur le cylindre $\mathcal{C}$ des molécules extérieures situées à gauche de Σ'. Si ces forces sont principalement attractives, leur résultante est dirigée vers la gauche, et elle donne d'après (C.20) une contribution négative à la pression; cette contribution serait positive si les forces étaient principalement répulsives.

Dans un gaz, la distance moyenne entre molécules est supérieure à la portée des forces. Par conséquent, l'effet des interactions est peu important et dominé par la partie attractive du potentiel intermoléculaire : la pression dynamique est négative et très inférieure à la pression cinétique. Au total, la pression d'un gaz réel est donc légèrement plus faible que celle du gaz parfait de même densité (à la même température).

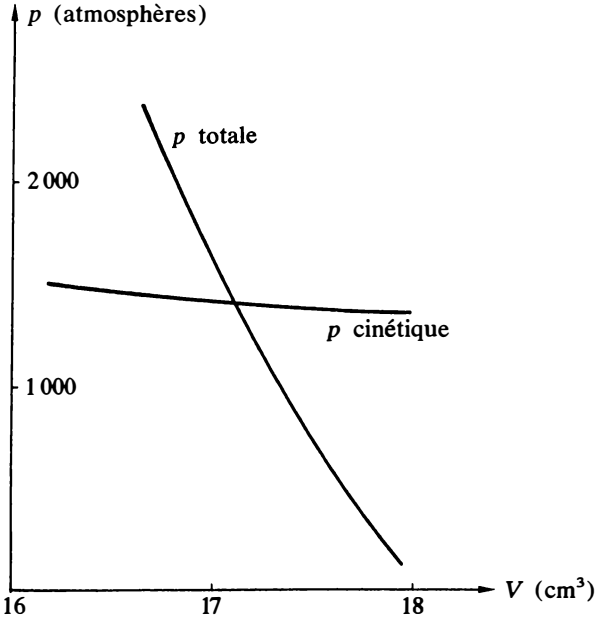


FIGURE 7

Pression cinétique et pression totale en fonction du volume dans une mole d'eau à 20 °C (cf. *American Institute of Physics Handbook*, McGraw-Hill, New York, 2e éd., pp. 2-163 à 2-165).

Les arguments et les calculs développés plus haut pour calculer la pression d'un gaz sont également valables pour un liquide homogène. Dans un liquide non comprimé, la distance moyenne entre molécules est beaucoup plus petite que dans un gaz, mais les forces sont encore principalement attractives : elles assurent la cohésion du liquide qui, contrairement aux gaz, n'occupe pas forcément tout le volume accessible. A la température ordinaire, le nombre de molécules par unité de volume dans un liquide est environ mille fois supérieur à celui d'un gaz sous la pression atmosphérique. La pression cinétique d'un liquide est donc mille fois plus grande que celle d'un gaz. Mais la pression dynamique, négative, est elle aussi très grande. La pression totale d'un liquide peut ainsi être égale à celle d'un gaz avec lequel il est en contact. Si l'on comprime un liquide pour diminuer son volume, la distance moyenne entre molécules diminue et l'influence des forces répulsives augmente : la pression dynamique négative devient moins importante et la pression totale augmente très rapidement. C'est ce qui explique qu'un liquide est peu compressible. Si l'on parvient néanmoins à diminuer son volume de façon appréciable, la contribution dynamique devient positive, de sorte que la pression totale est supérieure à la pression cinétique (voir figure 7).

2. CALCUL DE LA PRESSION DYNAMIQUE DANS UN GAZ MONOATOMIQUE DILUÉ

Nous voulons maintenant expliciter le deuxième terme de la formule (C.20), c'est-à-dire calculer la force exercée sur les molécules du cylindre $\mathcal{C}$ par celles qui sont situées à gauche du plan contenant Σ' (puisque la résultante des forces exercées sur le cylindre $\mathcal{C}$ par le gaz extérieur entourant sa surface latérale est nulle; cf. figure 4.b).

a. Hypothèses et approximations

Nous considérons un gaz *suffisamment dilué* pour que l'on puisse négliger les interactions faisant intervenir plus de deux particules à la fois : si la distance moyenne entre molécules voisines est grande devant la portée des forces qui s'exercent entre elles, la probabilité pour que trois molécules (ou plus) se trouvent assez proches les unes des autres pour interagir simultanément est très faible. Nous nous restreignons en outre à des *gaz monoatomiques*, de façon que l'énergie potentielle d'interaction entre deux molécules dépende seulement de leur distance $|\vec{r}_1 - \vec{r}_2|$; cette énergie potentielle $u(r)$ a la forme indiquée sur la figure 1 du complément I.A.

Lorsque les interactions ne sont pas négligeables, les positions des particules ne sont pas indépendantes : une molécule donnée attire ou repousse ses voisines (suivant leur distance), de sorte que son environnement n'est pas homogène sur des distances de l'ordre de la portée des forces. Pour décrire cet environnement, on introduit la *fonction de corrélation à deux particules* $g(\vec{r}, \vec{r}')$ définie comme suit : lorsqu'on sait qu'une particule se trouve au point $\vec{r}$, le nombre de particules $dN(\vec{r}, \vec{r}')$ contenues dans l'élément de volume d^3r' autour du point $\vec{r}'$ s'écrit

$$dN(\vec{r}, \vec{r}') \widehat{=} \frac{N}{V} g(\vec{r}, \vec{r}') d^3r' . \qquad \text{(C.33)}$$

En première approximation (voir note 8 du complément V.C), la fonction de corrélation est proportionnelle au facteur de Boltzmann faisant intervenir l'énergie potentielle $u(|\vec{r} - \vec{r}'|)$:

$$g(\vec{r}, \vec{r}') \simeq \exp[-\beta\, u(|\vec{r} - \vec{r}'|)] . \qquad \text{(C.34)}$$

b. Force exercée sur un atome du cylindre $\mathcal{C}$ par les atomes extérieurs situés entre deux plans infiniment voisins

Considérons un atome A particulier situé à l'intérieur du cylindre $\mathcal{C}$, à la distance ℓ de Σ' (fig. 8). Nous commençons par calculer la résultante des forces exercées sur l'atome A par les atomes extérieurs compris entre deux plans infiniment voisins P_1 et P_2 parallèles à Σ' et situés à gauche de Σ', aux distances L et $L+\mathrm{d}L$ respectivement. Nous noterons $\mathrm{d}F_g^{(1)}(L, \ell)$ cette résultante agissant sur l'atome A.

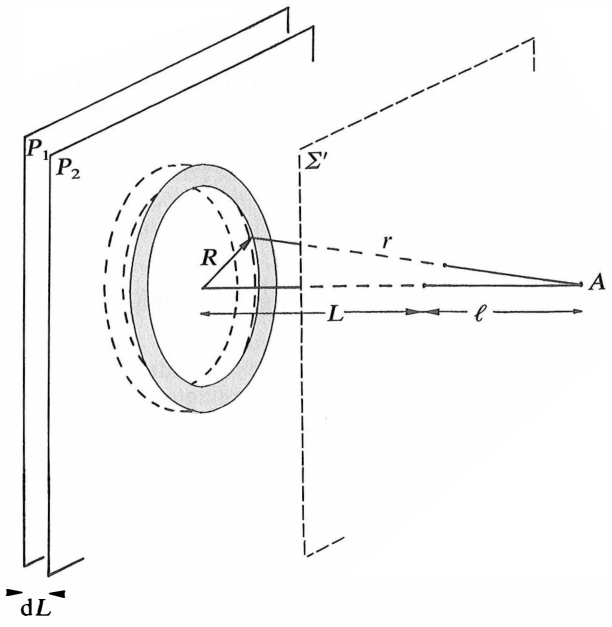


FIGURE 8
Pour calculer la force exercée sur l'atome A par les atomes contenus entre les plans P_1 et P_2, on découpe ces plans en anneaux concentriques. L'aire d'un anneau infinitésimal est $2\pi R\ \mathrm{d}R = 2\pi r\ \mathrm{d}r$.

Pour la calculer, découpons les plans P_1 et P_2 en anneaux concentriques centrés sur la projection de A. Les atomes contenus dans le volume infinitésimal délimité par P_1, P_2 et un anneau infiniment étroit sont pratiquement tous à la même distance r de l'atome A. L'utilisation du théorème de Pythagore montre facilement que ce volume infinitésimal a pour mesure $2\pi r\ \mathrm{d}r\ \mathrm{d}L$. Le nombre $\mathrm{d}N(\ell, L; r)$ d'atomes qu'il contient, compte tenu de la présence de l'atome A, s'obtient à partir de la fonction de corrélation à deux particules; dans l'approximation (C.34), ce nombre s'écrit

$$\mathrm{d}N(\ell, L; r) = \frac{N}{V} \mathrm{e}^{-\beta u(r)} 2\pi r\ \mathrm{d}r\ \mathrm{d}L\ . \tag{C.35}$$

La composante normale à Σ' de la force exercée sur A par chacun des atomes de ce volume infinitésimal a pour mesure algébrique

$$f = -\frac{\mathrm{d}u(r)}{\mathrm{d}r} \cdot \frac{L+\ell}{r}\ . \tag{C.36}$$

De (C.35) et (C.36) on déduit la résultante cherchée :

$$dF_g^{(1)}(L, \ell) = -\frac{N}{V} 2\pi(L+\ell)\,dL \int_{L+\ell}^{\infty} dr\, e^{-\beta u(r)} \frac{du}{dr}$$

$$= \frac{N}{V} 2\pi(L+\ell)\, kT\, [1 - e^{-\beta u(L+\ell)}]\, dL\,. \qquad \text{(C.37)}$$

c. Force totale exercée par les atomes extérieurs sur ceux du cylindre $\mathcal{C}$; pression dynamique

La « tranche » infinitésimale du cylindre $\mathcal{C}$, d'épaisseur $d\ell$, où se trouve A contient $(N/V).s\ d\ell$ atomes analogues. La force résultante $dF_g(L, \ell)$ qu'exercent sur eux les atomes extérieurs situés entre les plans P_1 et P_2 est donc

$$dF_g(L, \ell) = s \frac{N^2}{V^2} 2\pi kT(L+\ell)[1 - e^{-\beta u(L+\ell)}]\, dL\ d\ell\,. \qquad \text{(C.38)}$$

Pour obtenir la résultante globale F_g exercée sur l'ensemble du cylindre $\mathcal{C}$ par le gaz extérieur, il suffit d'intégrer sur ℓ et L (on peut prolonger l'intégration jusqu'à l'infini car la fonction est nulle dès que $\ell + L$ dépasse la portée des interactions) :

$$F_g = s \frac{N^2}{V^2} 2\pi kT \int_0^{\infty} d\ell \int_0^{\infty} dL\,(L+\ell)[1 - e^{-\beta u(L+\ell)}]\,. \qquad \text{(C.39)}$$

Effectuons le changement de variables

$$Z \mathrel{\widehat{=}} L + \ell\,,$$

$$z \mathrel{\widehat{=}} \ell\,; \qquad \text{(C.40)}$$

il vient[11]

$$F_g = s \frac{N^2}{V^2} 2\pi kT \int_0^{\infty} dZ\, Z\,[1 - e^{-\beta u(Z)}] \int_0^{Z} dz$$

$$= s \frac{N^2}{V^2} 2\pi kT \int_0^{\infty} dZ\, Z^2 [1 - e^{-\beta u(Z)}]\,. \qquad \text{(C.41)}$$

On en déduit la pression dynamique :

$$p_{\text{dyn}} = \frac{F_g}{s} = 2\pi kT \frac{N^2}{V^2} \int_0^{\infty} dZ\, Z^2 [1 - e^{-\beta u(Z)}]\,. \qquad \text{(C.42)}$$

d. Expression approchée de la pression du gaz

Le résultat (C.42) est en fait déjà approché, puisqu'il s'appuie sur l'expression (C.34) de la fonction de corrélation. On peut évaluer simplement l'intégrale qui y figure en prenant

$$e^{-\beta u(r)} = 0 \qquad \text{pour} \quad r < \rho_1\,,$$

$$= 1 - \frac{u(r)}{kT} \qquad \text{pour} \quad r > \rho_1\,. \qquad \text{(C.43)}$$

Ceci revient à considérer l'énergie potentielle $u(r)$ comme pratiquement infinie pour les distances inférieures au diamètre ρ_1 des atomes (sphères impénétrables), et à

11. Si l'on fixe la distance $Z = (L + \ell)$ entre les deux volumes élémentaires qui nous intéressent, la distance à Σ' de l'un d'entre eux ne peut varier qu'entre 0 et Z (puisque L et ℓ sont positifs). En outre, le jacobien du changement de variables (C.40) est égal à 1.

supposer qu'elle est petite devant l'énergie thermique kT pour les distances supérieures (atomes s'attirant faiblement). L'expression (C.42) de la pression dynamique devient alors

$$p_{\text{dyn}} = 2\pi kT \frac{N^2}{V^2} \frac{\rho_1^3}{3} + 2\pi \frac{N^2}{V^2} \int_{\rho_1}^{\infty} \mathrm{d}Z\, Z^2 u(Z) , \tag{C.44}$$

de sorte que la pression totale du gaz peut s'écrire

$$p = p_{\text{cin}} + p_{\text{dyn}} = \frac{NkT}{V} \left(1 + \frac{Nb}{V}\right) - \frac{N^2}{V^2}\, a , \tag{C.45}$$

avec $$b \mathrel{\widehat{=}} \frac{2}{3}\, \pi \rho_1^3 , \tag{C.46.a}$$

$$a \mathrel{\widehat{=}} -2\pi \int_{\rho_1}^{\infty} \mathrm{d}Z\, Z^2 u(Z) . \tag{C.46.b}$$

Nous obtenons ainsi, pour un gaz dilué (c'est-à-dire de densité N/V faible), les premiers termes correctifs, en $(N/V)^2$, à l'équation d'état des gaz parfaits. Ces termes sont exactement ceux que donne l'équation de van der Waals (complément III.G, formule (G.21)); nous les retrouverons également au complément V.C par une méthode plus systématique, capable de fournir également les termes d'ordre supérieur en (N/V).

Références

1. A. Sommerfeld, *Thermodynamics and Statistical Mechanics,* Academic Press, New York (1956), chapitre 3.
2. F. Reif, *Fundamentals of Statistical and Thermal Physics,* McGraw-Hill, New York (1965), chapitre 7.

COMPLÉMENT III.D

Ensemble *T-p*

I. Distribution statistique *T-p*
   1. Équilibre thermique et mécanique entre deux systèmes
   2. Système en contact avec un réservoir d'énergie et de volume
   3. Fonction de partition *T-p* et relations thermodynamiques
      a. Volume moyen et énergie moyenne
      b. Entropie
      c. Limite thermodynamique
   4. Évolution d'un système en contact avec un réservoir *T-p*

II. Application : l'élasticité du caoutchouc
   1. Distribution statistique d'un système maintenu à température et tension constantes
   2. Un modèle simple pour le caoutchouc
   3. Étude statistique du modèle
   4. Discussion des résultats

La distribution microcanonique s'applique aux systèmes dont l'énergie est fixée. Le passage à l'ensemble canonique se fait en relâchant cette contrainte sur l'énergie et en fixant la température à l'aide d'un thermostat. On peut procéder de façon analogue pour les autres paramètres extérieurs. On définit la variable conjuguée X du paramètre extérieur x par[1]

$$X \widehat{=} T^* \frac{\partial S^*}{\partial x}, \tag{D.1}$$

où T^* est la température et S^* l'entropie microcanonique du système. On peut relâcher les contraintes sur certains paramètres extérieurs x_i, et fixer en contrepartie les variables conjuguées X_i de ces paramètres grâce à des « réservoirs » de x_i[2]. C'est ce que nous ferons au chapitre V : la donnée du nombre N de particules y sera remplacée par celle du potentiel chimique μ, et nous obtiendrons ainsi la description grand-canonique des systèmes en contact avec un réservoir de particules. D'une façon générale, à une situation où x est fixé, on peut faire correspondre une *situation canonique généralisée* dans laquelle la contrainte sur x est remplacée par une contrainte sur la variable conjuguée X.

A titre d'exemple, nous considérons ici des systèmes dont le volume (et l'énergie) est une variable interne libre de fluctuer, et qui sont *en contact avec un réservoir leur imposant sa pression (et sa température).* On appelle *ensemble T-p* l'ensemble canonique généralisé correspondant à cette situation.

1. Certaines variables conjuguées, comme la pression, s'interprètent au niveau microscopique comme une « force » moyenne (complément II.D, § I). Le potentiel chimique, au contraire, n'a pas d'équivalent au niveau microscopique. Il est directement défini par (D.1) : $-\mu^* \widehat{=} T^* \, \partial S^* / \partial N$.

2. Il n'est cependant pas possible d'appliquer ce procédé simultanément à *tous* les paramètres extérieurs définissant l'état macroscopique d'un système isolé. Ceux-ci sont en effet extensifs, alors que leurs variables conjuguées sont intensives (complément II.A, § II); il faut conserver au moins une variable extensive car il est impossible de caractériser la taille du système si l'on ne dispose que de variables intensives.

I. DISTRIBUTION STATISTIQUE *T-p*

1. ÉQUILIBRE THERMIQUE ET MÉCANIQUE ENTRE DEUX SYSTÈMES

Considérons deux systèmes (A) et (B) (par exemple deux fluides), faiblement couplés, *séparés par un piston mobile perméable à la chaleur* (fig. 1). L'ensemble $(A) \cup (B)$ forme un système isolé dont l'énergie E_0 et le volume V_0 sont fixés (à δE_0 et δV_0 près). Les énergies E_A et E_B, et les volumes V_A et V_B des deux systèmes sont cependant libres de fluctuer, à condition de vérifier

$$E_A + E_B = E_0 \,, \tag{D.2}$$

$$V_A + V_B = V_0 \,. \tag{D.3}$$

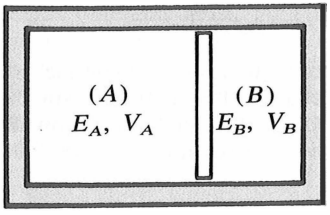


FIGURE 1
Les deux systèmes (A) et (B) sont séparés par un piston mobile diatherme. L'ensemble est isolé ($E_A + E_B = E_0$, $V_A + V_B = V_0$).

Nous avons montré (II, § III.B.3.a) que les valeurs de E_A et V_A les plus probables sont telles que la température et la pression microcanoniques sont les mêmes dans les systèmes (A) et (B); *l'équilibre thermique et mécanique* est donc caractérisé par

$$T_A^* = T_B^* \,, \tag{D.4}$$

$$p_A^* = p_B^* \,. \tag{D.5}$$

2. SYSTÈME EN CONTACT AVEC UN RÉSERVOIR D'ÉNERGIE ET DE VOLUME

Nous supposons maintenant que *le système (B) est beaucoup plus gros que le système (A)* (c'est-à-dire qu'il compte un beaucoup plus grand nombre de particules et qu'il occupe un volume beaucoup plus grand). Les variations de son énergie et de son volume, par suite des échanges avec (A), n'affectent alors pratiquement pas sa température et sa pression[3]. Il constitue un *réservoir $\mathcal{R}$ d'énergie et de volume qui impose sa température et sa pression.* Nous noterons désormais T *et* p les valeurs des nouveaux *paramètres extérieurs imposés au système* (A).

3. Supposons, par exemple, que le réservoir soit un gaz parfait monoatomique. On a alors (III, § IV.B.3)

$$E_B = \frac{3}{2} N_B k T_B \,, \tag{1}$$

$$p_B V_B = N_B k T_B \,. \tag{2}$$

.../...

Quelle est alors *la probabilité* $w_\ell(V)\,\mathrm{d}V$ *pour que le volume de* (A) *soit compris entre* V *et* $V+\mathrm{d}V$, *et que son état microscopique soit* (ℓ), d'énergie $E_\ell(V)$[4], lorsqu'il est en *contact thermique et mécanique avec le réservoir à la température* T *et à la pression* p? La réponse à cette question s'obtient de la même façon que pour la distribution canonique, mais il y a maintenant deux variables fluctuantes (énergie et volume) au lieu d'une. La probabilité d'observer pour (A) les caractéristiques indiquées, quel que soit par ailleurs l'état du réservoir, est proportionnelle au nombre d'états accessibles du système global dans lesquels (A) a le volume V (à $\mathrm{d}V$ près) et se trouve dans l'état (ℓ) (III, § I.B.1). Ce nombre est donc simplement le *nombre d'états accessibles au réservoir* lorsque son volume est (V_0-V), à $\mathrm{d}V$ près, et son énergie (E_0-E_ℓ) :

$$w_\ell(V)\,\mathrm{d}V \propto \Omega_{\mathcal{R}}\left(E_0-E_\ell(V),\ V_0-V\right)\mathrm{d}V\,, \tag{D.6}$$

ou encore, en introduisant l'entropie microcanonique $S^*_{\mathcal{R}}$ du réservoir,

$$w_\ell(V)\,\mathrm{d}V \propto \mathrm{d}V\,\exp\left[\frac{1}{k}\,S^*_{\mathcal{R}}\left(E_0-E_\ell(V),\ V_0-V\right)\right]. \tag{D.7}$$

$E_\ell(V)$ et V sont petits comparés à l'énergie E_0 et au volume V_0; on peut donc développer $S^*_{\mathcal{R}}$:

$$S^*_{\mathcal{R}}\left(E_0-E_\ell(V),\ V_0-V\right)=S^*_{\mathcal{R}}(E_0,\ V_0) - E_\ell(V)\left.\frac{\partial S^*_{\mathcal{R}}}{\partial E}\right)_{E_0,\,V_0} - V\left.\frac{\partial S^*_{\mathcal{R}}}{\partial V}\right)_{E_0,\,V_0}+\ldots\,; \tag{D.8}$$

en utilisant les relations

$$\frac{1}{T}=\frac{\partial S^*_{\mathcal{R}}}{\partial E} \quad\text{et}\quad \frac{p}{T}=\frac{\partial S^*_{\mathcal{R}}}{\partial V}\,, \tag{D.9}$$

on obtient

$$S^*_{\mathcal{R}}\left(E_0-E_\ell(V),\ V_0-V\right)=S^*_{\mathcal{R}}(E_0,\ V_0)-\frac{E_\ell(V)}{T}-p\,\frac{V}{T}\,. \tag{D.10}$$

On en déduit les variations δT_B et δp_B de T_B et p_B dues à des variations $\delta E_A(=-\delta E_B)$ et $\delta V_A(=-\delta V_B)$ de l'énergie et du volume du système (A) :

$$\frac{\delta T_B}{T_B}=\frac{\delta E_B}{E_B}=-\frac{\delta E_A}{E_B}\,, \tag{3}$$

$$\frac{\delta p_B}{p_B}=-\frac{\delta V_B}{V_B}+\frac{\delta T_B}{T_B}=\frac{\delta V_A}{V_B}-\frac{\delta E_A}{E_B}\,. \tag{4}$$

Les variations de T_B et p_B sont négligeables si E_B et V_B sont bien plus grands que E_A et V_A, et donc *a fortiori* que leurs variations δE_A et δV_A. Si ces conditions sont vérifiées, le système (B) constitue un réservoir d'énergie et de volume (cf. III, § I.A.1 et plus particulièrement la note 3).

4. Nous supposons, comme au complément II.D, que les niveaux d'énergie E_ℓ du système sont des fonctions continues du volume V.

Pour *compter les états* dans lesquels le volume du système est compris entre V et $V+\mathrm{d}V$, on peut faire comme si le volume variait de façon discontinue, par « sauts » de $\delta_0 V$ (inappréciables à l'échelle macroscopique). Le nombre d'états cherché s'obtient alors en multipliant par $\mathrm{d}V/\delta_0 V$ celui qui correspond à la valeur V du volume. Pour normaliser ensuite la distribution, on divise par la somme sur l'ensemble des valeurs V (formule (D.12) ci-dessous); cette somme est elle aussi proportionnelle à $1/\delta_0 V$, de sorte que le paramètre auxiliaire $\delta_0 V$ disparaît du résultat.

La probabilité cherchée vaut donc

$$\boxed{w_\ell(V)\,\mathrm{d}V = \frac{\mathrm{d}V}{\widetilde{Z}} \exp\left[-\frac{1}{kT}\left(E_\ell(V)+pV\right)\right]} \quad , \tag{D.11}$$

où $\widetilde{Z}$ est un facteur de normalisation, qu'on appelle *fonction de partition T-p*[5],[6] :

$$\boxed{\widetilde{Z}(T,\,p) \mathrel{\widehat{=}} \int_0^\infty \mathrm{d}V \sum_{(\ell)} \exp\left[-\frac{1}{kT}\left(E_\ell(V)+pV\right)\right]} \quad . \tag{D.12}$$

3. FONCTION DE PARTITION *T-p* ET RELATIONS THERMODYNAMIQUES

Les propriétés d'un système en équilibre thermique et mécanique avec un réservoir à la température T et à la pression p se déduisent de sa fonction de partition T-p, ou de son *enthalpie libre,* définie (par analogie avec l'énergie libre) comme

$$G(T,\,p) \mathrel{\widehat{=}} -kT \ln \widetilde{Z}(T,\,p) \; . \tag{D.13}$$

a. Volume moyen et énergie moyenne

La distribution (D.11) permet de calculer le volume moyen du système :

$$\overline{V} = \int_0^{+\infty} V \sum_{(\ell)} w_\ell(V)\,\mathrm{d}V = -kT \frac{\partial \ln \widetilde{Z}}{\partial p} \; , \tag{D.14}$$

ou encore

$$\overline{V} = \frac{\partial G}{\partial p} \; . \tag{D.15}$$

---

5. En toute rigueur, il existe toujours une valeur minimale non nulle du volume V_A, au-dessous de laquelle le système (A) ne pourrait exister; de même V_A est borné supérieurement par le volume total V_0 du système global. En pratique, pour des systèmes macroscopiques, seules les valeurs de V très proches de la valeur la plus probable contribuent de façon appréciable à l'intégrale; on peut donc prolonger le domaine d'intégration de 0 à ∞ sans modifier le résultat.

6. $\widetilde{Z}$ peut aussi s'écrire

$$\widetilde{Z}(T,\,p) = \int_0^\infty \mathrm{d}V \; \mathrm{e}^{-pV/kT} \sum_{(\ell)} \mathrm{e}^{-E_\ell(V)/kT} \; . \tag{1}$$

La somme sur les états (ℓ) à V fixé n'est rien d'autre que la *fonction de partition canonique* $Z(T,\,V)$, de sorte que

$$\widetilde{Z}(T,\,p) = \int_0^\infty \mathrm{d}V \; \mathrm{e}^{-pV/kT} Z(T,\,V) \; . \tag{2}$$

$\widetilde{Z}$ apparaît ainsi comme une transformée de Laplace de la fonction de partition canonique Z. D'ailleurs, celle-ci était déjà une transformée de Laplace de la densité d'états $\rho(E)$ (dont le logarithme donne l'entropie microcanonique) :

$$Z(T,\,V) = \int_0^\infty \mathrm{d}E \; \mathrm{e}^{-E/kT} \rho(E) \; . \tag{3}$$

L'énergie moyenne $\overline{E}$ se déduit de même de la relation

$$\overline{E}+p\overline{V}=\int_0^\infty \sum_{(\ell)}\left(E_\ell(V)+pV\right)w_\ell(V)\,\mathrm{d}V=-\frac{\partial \ln \widetilde{Z}}{\partial\beta}\,,\quad\left(\beta=\frac{1}{kT}\right), \tag{D.16}$$

complétée par (D.14).

b. Entropie

L'entropie S du système en présence du réservoir T-p s'obtient aussi à partir de (D.11)[7] :

$$S=-k\int_0^\infty \mathrm{d}V\sum_{(\ell)} w_\ell(V)\ln w_\ell(V) \tag{D.17.a}$$

$$=-k\int_0^\infty \mathrm{d}V\sum_{(\ell)} w_\ell(V)\left(-\frac{E_\ell(V)+pV}{kT}\right)$$

$$+k\int_0^\infty \mathrm{d}V\sum_{(\ell)} w_\ell(V)\ln \widetilde{Z}(T,p)\,. \tag{D.17.b}$$

On reconnaît dans la première intégrale la moyenne $\overline{E}+p\overline{V}$ (formule (D.16)); la deuxième donne simplement $k \ln \widetilde{Z}$ du fait de la normalisation des probabilités. On trouve donc

$$S=k\left[\ln \widetilde{Z}-\beta\frac{\partial \ln \widetilde{Z}}{\partial\beta}\right]=-\frac{\partial G}{\partial T}\,. \tag{D.18}$$

On peut encore écrire, en faisant intervenir explicitement $\overline{E}$ et $\overline{V}$:

$$S=-\frac{G}{T}+\frac{\overline{E}+p\overline{V}}{T}\,, \tag{D.19}$$

relation que l'on retiendra sous la forme

$$\boxed{G=\overline{E}+p\overline{V}-TS}\,. \tag{D.20}$$

---

7. La distribution de probabilité T-p (D.11) est une distribution continue, ce qui pose problème quant à la détermination de l'entropie associée (complément I.G, § III). La relation (D.17.a) ne définit en fait l'entropie qu'à une constante additive près. Il serait plus judicieux de choisir l'autre solution proposée au complément I.G en faisant intervenir un intervalle minimum $\delta_0 V$ caractérisant la meilleure précision possible sur une mesure du volume. On aurait alors

$$S=-k\int_0^\infty \mathrm{d}V\sum_{(\ell)} w_\ell(V)\ln\left[w_\ell(V)\,\delta_0 V\right], \tag{1}$$

ce qui donne

$$S=\frac{\overline{E}+p\overline{V}}{T}+k\ln\frac{\widetilde{Z}}{\delta_0 V}\,. \tag{2}$$

En fait, il faut remarquer que tel qu'il a été défini $\widetilde{Z}$ n'est pas une grandeur sans dimension (comme il devrait l'être puisqu'on en prend le logarithme pour définir G) mais s'exprime en unités de volume. La bonne définition de G devrait donc être

$$G(T,p)\mathrel{\widehat{=}}-kT\ln\frac{\widetilde{Z}}{\delta_0 V}\,, \tag{3}$$

qui, reportée dans (2), redonne la relation (D.20). En pratique on peut oublier le $\delta_0 V$ comme on le fait pour l'entropie microcanonique (II, § I.C.1) lorsqu'on écrit

$$S^*=k\ln\rho(E)\,\delta E\simeq k\ln\rho(E)\,, \tag{4}$$

et utiliser les relations (D.13) et (D.20) directement.

c. Limite thermodynamique

A la limite thermodynamique (III, § III), l'énergie et le volume du système prennent pratiquement des valeurs uniques E et V. L'enthalpie libre thermodynamique

$$G = E + pV - TS \quad \text{(lim. thermo.)} \tag{D.21}$$

s'écrit aussi (puisque dans cette limite toutes les descriptions sont équivalentes)

$$G = F + pV \quad \text{(lim. thermo.)}, \tag{D.22}$$

où F est l'énergie libre thermodynamique[8].

Dans le cas d'un *fluide simple,* l'entropie S ne dépend que de trois variables E, V et N. De sa différentielle

$$\mathrm{d}S = \frac{1}{T}\,\mathrm{d}E + \frac{P}{T}\,\mathrm{d}V - \frac{\mu}{T}\,\mathrm{d}N \quad \text{(lim. thermo.)}, \tag{D.23}$$

on déduit aisément la différentielle de G en fonction des paramètres extérieurs T, p, N :

$$\mathrm{d}G = \mathrm{d}E + p\,\mathrm{d}V + V\,\mathrm{d}p - T\,\mathrm{d}S - S\,\mathrm{d}T,$$

$$\mathrm{d}G = -S\,\mathrm{d}T + V\,\mathrm{d}p + \mu\,\mathrm{d}N \text{ (lim. thermo.)}. \tag{D.24}$$

On retrouve ainsi, dans la limite thermodynamique, les relations (D.15) et (D.18). Quant au potentiel chimique μ, il vaut

$$\mu = \frac{\partial G}{\partial N}(T, p, N)\,. \tag{D.25}$$

Or G, grandeur extensive[9], ne dépend ici que d'une seule variable extensive, N. Il s'ensuit[10] que μ ne dépend que de T et p et que

$$G(T, p, N) = N\mu(T, p)\,. \tag{D.26}$$

---

8. G est une *«transformée de Legendre»* de F, de même que F était une transformée de Legendre de E dans le passage de la description microcanonique à la description canonique. La transformation de Laplace sur les fonctions de partition (note 6) se traduit par une transformation de Legendre sur les potentiels thermodynamiques associés.

9. Si on réunit λ systèmes identiques, d'énergie E et d'entropie S, contenant chacun N particules à la température T et à la pression p, le système global sera lui aussi à la température T et à la pression p mais il contiendra λN particules, aura une énergie λE et une entropie λS. D'après (D.20), son enthalpie libre deviendra donc λG :

$$G(T, p, \lambda N) = \lambda G(T, p, N)\,.$$

10. Dérivons les deux membres de la relation de la note précédente par rapport à λ (en maintenant T, p et N constants) :

$$N\frac{\partial G}{\partial N}(N, p, \lambda N) = G(T, p, N)\,. \tag{1}$$

On en déduit

$$\mu(T, p, \lambda N) = \frac{G(T, p, N)}{N} \tag{2}$$

quel que soit λ. μ est donc indépendant de N et vérifie la relation (D.26).

4. ÉVOLUTION D'UN SYSTÈME EN CONTACT AVEC UN RÉSERVOIR *T-p*

Considérons un système en contact avec un réservoir qui fixe sa pression et sa température, et soumis également à des contraintes qui fixent les valeurs d'un certain nombre de paramètres extérieurs. Le système étant initialement à l'équilibre, relâchons une (ou plusieurs) des contraintes de façon que les échanges d'énergie et de volume aient lieu seulement avec le réservoir. *Le système évolue alors spontanément* vers un nouvel état d'équilibre correspondant au *minimum de son enthalpie libre* compatible avec les contraintes restantes (la démonstration est identique à celle que nous avons faite au chapitre III, paragraphe II.D pour un système en contact avec un thermostat).

II. APPLICATION : L'ÉLASTICITÉ DU CAOUTCHOUC

Les résultats précédents s'appliquent à n'importe quel système pour lequel la pression (ou le volume) est une variable pertinente. Pour les grands systèmes (limite thermodynamique), ils donnent les mêmes relations que les ensembles microcanonique et canonique. Nous allons voir cependant que l'ensemble *T-p* est particulièrement bien adapté à l'étude statistique des chaînes de polymères et nous allons l'appliquer à un modèle simple (pour ne pas dire simpliste) de tels systèmes. Dans le cas de fibres élastiques, ce n'est pas une pression qu'on applique de l'extérieur sur le système mais une force mécanique (tension); il nous faut donc commencer par écrire la distribution adaptée à ce cas.

1. DISTRIBUTION STATISTIQUE D'UN SYSTÈME MAINTENU A TEMPÉRATURE ET TENSION CONSTANTES

On considère une corde élastique (système $\mathcal{S}$) dont une extrémité est fixée au point O; l'autre extrémité (repérée par le vecteur $\vec{r}$) est soumise de la part d'un système mécanique à une force constante $\vec{X}$; la corde est maintenue à la température T par un thermostat. L'ensemble thermostat-système mécanique constitue donc un *réservoir $\mathcal{R}$ d'énergie et de longueur qui impose au système $\mathcal{S}$ sa température et sa tension* (fig. 2).

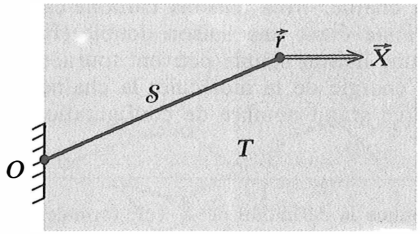


FIGURE 2
La corde élastique $\mathcal{S}$ est fixée à son extrémité O et soumise à l'autre extrémité ($\vec{r}$) à une force $\vec{X}$ constante qui lui est imposée par un « réservoir de longueur » $\mathcal{R}$. Ce dernier est aussi un « réservoir d'énergie », il fixe donc la température T du système.

La probabilité d'observer le système $\mathcal{S}$ avec son extrémité dans le volume infinitésimal d^3r autour du point $\vec{r}$ et dans l'état microscopique (ℓ), d'énergie $E_\ell(\vec{r})$ (quel que soit l'état du réservoir), est donc (cf. (D.7)) :

$$w_\ell(\vec{r})\,\mathrm{d}^3r \propto \mathrm{d}^3r \exp\left[\frac{1}{k} S^*_{\mathcal{R}}(E_0 - E_\ell(\vec{r}); \vec{r})\right] ; \qquad \text{(D.27)}$$

$S^*_{\mathcal{R}}(E_0 - E_\ell(\vec{r}); \vec{r})$ est l'entropie du réservoir correspondant à l'énergie $(E_0 - E_\ell(\vec{r}))$ et la position $\vec{r}$ de l'extrémité de la corde.

Le réservoir étant par définition beaucoup plus gros que $\mathcal{S}$, $E_\ell(\vec{r})$ et $\vec{r}$ représentent pour lui de petites perturbations par rapport à l'énergie E_0 et à la position $\vec{r}=\vec{0}$; on peut donc développer $S^*_{\mathcal{R}}$:

$$S^*_{\mathcal{R}}(E_0-E_\ell(\vec{r});\,\vec{r})=S^*_{\mathcal{R}}(E_0;\vec{0})-E_\ell(\vec{r})\,\frac{\partial S^*_{\mathcal{R}}}{\partial E}(E_0;\vec{0}) + \vec{r}\,.\overrightarrow{\text{grad}}\,S^*_{\mathcal{R}}(E_0;\vec{0})+\dots\,. \tag{D.28}$$

Or $$\frac{1}{T}\widehat{=}\frac{\partial S^*_{\mathcal{R}}}{\partial E}\,, \tag{D.29}$$

et[11] $$\frac{\vec{X}}{T}\widehat{=}\overrightarrow{\text{grad}}\,S^*_{\mathcal{R}}\,. \tag{D.30}$$

On obtient finalement

$$w_\ell(\vec{r})=\frac{1}{\widetilde{Z}}\exp\left[\frac{\vec{r}\,.\vec{X}-E_\ell(\vec{r})}{kT}\right], \tag{D.31}$$

où $\widetilde{Z}$, «*fonction de partition T-*$\vec{X}$», est un facteur de normalisation :

$$\widetilde{Z}(T,\vec{X})=\int \mathrm{d}^3r\sum_{(\ell)}\exp\left[\frac{\vec{r}\,.\vec{X}-E_\ell(\vec{r})}{kT}\right]. \tag{D.32}$$

Comme toujours, la connaissance de la fonction de partition permet de calculer les propriétés du système; en particulier, la position moyenne de l'extrémité de la corde est donnée par[12]

$$\overline{\vec{r}}=kT\,\overrightarrow{\text{grad}}_{\vec{X}}\,\ln\widetilde{Z}(T,\vec{X})\,, \tag{D.33}$$

où $\overrightarrow{\text{grad}}_{\vec{X}}$ signifie que chacune des composantes de $\vec{r}$ est obtenue en dérivant $\ln\widetilde{Z}$ par rapport à la composante correspondante de $\vec{X}$.

2. UN MODÈLE SIMPLE POUR LE CAOUTCHOUC

Une molécule de caoutchouc est un polymère linéaire d'hydrocarbures formant une longue chaîne articulée; la formule chimique de cette molécule est $(C_5H_8)_n$, avec n de l'ordre de 10^2 à 10^4. Le long de la chaîne, trois liaisons carbone-carbone sur quatre sont des liaisons simples, la quatrième étant une liaison double (fig. 3). Les deux parties d'une molécule séparées par une liaison simple peuvent tourner l'une par rapport à l'autre sans affecter beaucoup l'énergie de la molécule; la chaîne est donc extrêmement flexible et peut prendre un très grand nombre de configurations.

---

11. La relation (D.30) peut être considérée comme la définition de $\vec{X}$ (cf. formule (D.1)). Ce point de vue est justifié au complément II.D, paragraphe I.2.c.

12. Par exemple, la moyenne de la composante x de $\vec{r}$ vaut

$$\bar{x}=\iiint \mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z\,\frac{x}{\widetilde{Z}}\sum_{(\ell)}\exp\left[\frac{xX_x+yX_y+zX_z-E_\ell(\vec{r})}{kT}\right]$$

$$=\frac{kT}{\widetilde{Z}}\frac{\partial}{\partial X_x}\iiint \mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z\sum_{(\ell)}\exp\left[\frac{xX_x+yX_y+zX_z-E_\ell(\vec{r})}{kT}\right]$$

$$=kT\frac{\partial}{\partial X_x}\ln\widetilde{Z}(T,\vec{X})\,.$$

On obtient des formules analogues pour les composantes $\bar{y}$ et $\bar{z}$, d'où (D.33).

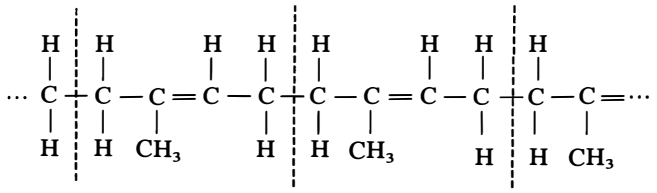


FIGURE 3
Formule développée d'un polymère $(C_5H_8)_n$ de caoutchouc.

Dans un morceau de caoutchouc, les molécules sont attachées les unes aux autres par des liaisons distribuées de façon pratiquement aléatoire, formant un réseau tridimensionnel assez lâche; ces liaisons sont relativement rares : on estime que, sur une molécule déterminée, il existe en moyenne une centaine de monomères entre deux liaisons avec des molécules voisines. Si on considère un maillon de la chaîne, c'est-à-dire une portion de polymère délimitée par deux liaisons simples successives, ce maillon interagit avec les maillons proches de la même chaîne, c'est-à-dire que l'énergie du polymère dépend de l'orientation relative de ses différents maillons. Mais il interagit aussi avec les maillons des autres chaînes et avec les maillons éloignés, soit en formant des liaisons nouvelles, soit parce que le volume de l'espace occupé par un maillon est inaccessible aux autres.

Pour étudier le comportement du caoutchouc (ou des polymères linéaires en général), nous allons négliger toutes les interactions entre maillons et choisir le *modèle très simple* suivant.

Le système est constitué de *N maillons identiques, chacun de longueur a invariable,* liés les uns aux autres de façon à former une *longue chaîne articulée. Chaque maillon peut s'orienter librement* par rapport au précédent, de sorte que la chaîne s'étend dans les trois dimensions. Elle est attachée en *O* à une de ses extrémités, et soumise à l'autre à une force $\vec{X}$ constante dont la direction est choisie comme axe *Ox* (fig. 4)[13]. Elle est d'autre part en équilibre avec un thermostat à la température *T*. Ce modèle est évidemment très grossier; il retient cependant assez de caractéristiques physiques pour permettre de comprendre *qualitativement* le comportement du caoutchouc.

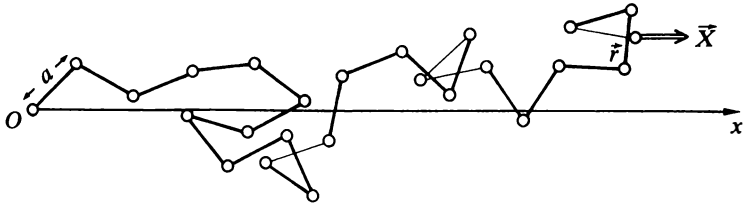


FIGURE 4
Chaîne de polymère constituée par *N* monomères de longueur fixe *a* et soumise à une force $\vec{X}$ constante dirigée suivant *Ox*.

13. Un échantillon de caoutchouc contient un grand nombre de chaînes analogues disposées plus ou moins parallèlement les unes aux autres et soumises chacune à la force $\vec{X}$. Comme le modèle suppose les chaînes indépendantes, le comportement de l'ensemble se déduit trivialement du comportement d'une seule chaîne.

3. ÉTUDE STATISTIQUE DU MODÈLE

On se propose de déterminer, en fonction de la température T et de la force $\vec{X}$ appliquée, la position moyenne de l'extrémité de la chaîne décrite au paragraphe précédent.

Le mouvement des maillons pouvant être traité *classiquement* (cf. complément III.B, § II.3), l'état de la chaîne est déterminé lorsqu'on connaît l'orientation de chacun des maillons (de longueur invariable) par rapport à un système d'axes fixes. Un état du système est donc défini par la donnée des angles polaire θ_i et azimutal φ_i de chacun des N maillons (fig. 5) : en effet θ_1 et φ_1 fixent à partir du point O la position $\overrightarrow{OM_1}$ du premier maillon, θ_2 et φ_2 fixent à partir de M_1 la position $\overrightarrow{M_1M_2}$ du second, et ainsi de suite... Puisque les maillons sont supposés pouvoir tourner librement les uns par rapport aux autres, tous les états (c'est-à-dire toutes les valeurs permises des θ_i et φ_i) correspondent à la même énergie.

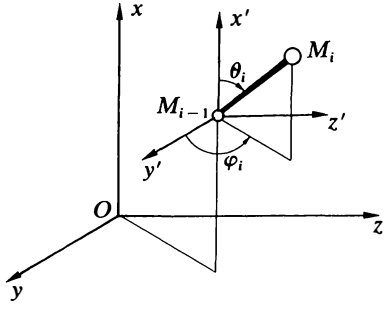


FIGURE 5
La position du i-ième maillon $M_{i-1}M_i$ est repérée par les deux angles θ_i et φ_i. L'angle polaire θ_i ($0 \leqslant \theta_i \leqslant \pi$) est défini par rapport à l'axe Ox. L'angle azimutal φ_i ($0 \leqslant \varphi_i < 2\pi$) est alors l'angle entre la direction Oy et la projection de $M_{i-1}M_i$ sur le plan yOz.

Si l'on essaye de résoudre le problème dans l'ensemble canonique, il faut fixer les extrémités O et $\vec{r}$ de la chaîne et calculer la fonction de partition $Z(T, \vec{r})$. Mais il faut alors compter le nombre d'états correspondant à chaque valeur de $\vec{r}$[14]. L'utilisation de la distribution canonique généralisée T-$\vec{X}$ qui lève la contrainte sur l'extrémité de la chaîne (très difficile à prendre en compte) permet de résoudre le problème plus simplement[15].

Le nombre d'états dans lesquels le i-ième maillon pointe dans un angle solide $d\omega_i$ autour de la direction (θ_i, φ_i) est proportionnel à

$$d\omega_i = \sin\theta_i \, d\theta_i \, d\varphi_i \, . \qquad \text{(D.34)}$$

Dans l'ensemble T-$\vec{X}$, il n'y a aucune contrainte sur la position de l'extrémité de la chaîne, et les θ_i et φ_i peuvent donc être choisis indépendamment les uns des autres.

14. Dans le cas où $\vec{r}$ est fixé, la force $\vec{X}$ est fonction de $\vec{r}$; elle est déterminée, à partir de $Z(T, \vec{r})$, par (exercice III.45)

$$\vec{X} = -kT \, \overrightarrow{\text{grad}}_{\vec{r}} \ln Z(T, \vec{r}) \, .$$

15. Sur l'expression (D.32) de $\tilde{Z}$, il semble qu'il faille calculer la somme sur les états à $\vec{r}$ *fixé* avant d'intégrer sur $\vec{r}$: on est alors ramené au même problème que dans la description canonique. La technique employée ci-dessous va tourner la difficulté en calculant la somme sur les états *sans restriction sur* $\vec{r}$, ce qui donnera *directement la valeur de* $\tilde{Z}$: on effectuera *en une seule fois* la somme sur les états (ℓ) et l'intégration sur $\vec{r}$.

Considérons la situation dans laquelle les angles polaire et azimutal du maillon (i) sont compris respectivement entre θ_i et $\theta_i + d\theta_i$, et φ_i et $\varphi_i + d\varphi_i$, avec $i = 1, 2, \ldots, N$. Dans la description T-$\vec{X}$, la probabilité dP de cet ensemble infinitésimal de configurations s'obtient en sommant (D.31) sur tous les états vérifiant les conditions énoncées[16]. L'énergie de la chaîne ne dépend pas de sa configuration; il vient donc[17]

$$dP = \frac{1}{\widetilde{Z}} e^{\vec{X}.\vec{r}/kT} \prod_{i=1}^{N} d\omega_i , \tag{D.35}$$

avec $$\widetilde{Z} = \int \ldots \int d\omega_1 \, d\omega_2 \ldots d\omega_N \, e^{\vec{X}.\vec{r}/kT} . \tag{D.36}$$

$\vec{X}$ étant dirigée suivant Ox, seule intervient la composante

$$x = \sum_{i=1}^{N} a \cos \theta_i \tag{D.37}$$

de l'extrémité $\vec{r}$ de la chaîne. Les formules (D.34) et (D.37), reportées dans (D.36), donnent

$$\begin{aligned} \widetilde{Z}(T, X) &= \prod_{i=1}^{N} \int_0^{2\pi} d\varphi_i \int_0^{\pi} d\theta_i \sin \theta_i \, e^{\frac{Xa \cos \theta_i}{kT}} \\ &= \left[\int_0^{2\pi} d\varphi \int_0^{\pi} d\theta \sin \theta \, e^{\frac{Xa \cos \theta}{kT}} \right]^N . \end{aligned} \tag{D.38}$$

L'intégrale (D.38) se calcule aisément; on obtient

$$\widetilde{Z}(T, X) = \left[\frac{4\pi \operatorname{sh} Xa/kT}{Xa/kT} \right]^N . \tag{D.39}$$

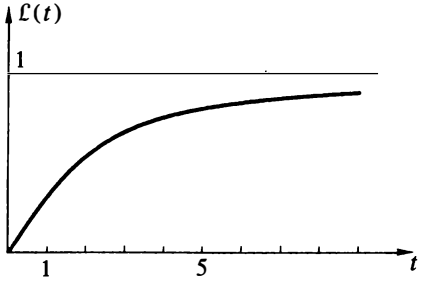


FIGURE 6
Représentation graphique de la fonction de Langevin $\mathcal{L}(t) = \coth t - 1/t$. Noter le comportement au voisinage de l'origine : $\mathcal{L}(t) \sim t/3$.

16. A une configuration $\{\theta_i, \varphi_i; i = 1, 2, \ldots, N\}$, est associée une position bien déterminée de l'extrémité de la corde : les composantes de $\vec{r}$ sont (cf. fig. 5)

$$x = \sum_{i=1}^{N} a \cos \theta_i , \quad y = \sum_{i=1}^{N} a \sin \theta_i \cos \varphi_i , \quad z = \sum_{i=1}^{N} a \sin \theta_i \sin \varphi_i .$$

Le volume infinitésimal d^3r est lui aussi déterminé lorsqu'on se donne les $d\theta_i$ et $d\varphi_i$ pour $i = 1, 2, \ldots, N$.

17. L'énergie E étant la même pour tous les états, le terme $\exp[-E/kT]$ se simplifie entre le numérateur et le dénominateur de la formule. Il revient au même de choisir E comme origine des énergies et de poser $E = 0$.

On en déduit les valeurs moyennes des projections de l'extrémité $\vec{r}$ de la chaîne sur les trois axes. Les valeurs moyennes des projections sur Oy et Oz sont évidemment nulles par raison de symétrie :

$$\bar{y} = \bar{z} = 0 \,, \tag{D.40}$$

et

$$\bar{x} = kT \frac{\partial \ln \widetilde{Z}}{\partial X} = Na \, \mathcal{L}(Xa/kT) \,, \tag{D.41}$$

où

$$\mathcal{L}(t) \widehat{=} \coth t - \frac{1}{t} \tag{D.42}$$

est la *«fonction de Langevin»* (représentée sur la figure 6).

4. DISCUSSION DES RÉSULTATS

Dans ce modèle simple, la longueur de la chaine $\bar{x}$ est fonction du seul paramètre Xa/kT; quand il tend vers zéro (tension nulle ou température infinie), l'allongement $\bar{x}$ tend vers zéro[18]; quand il tend vers l'infini (tension infinie ou température nulle), la longueur tend vers Na : tous les maillons s'orientent alors dans le sens de la force, donnant ainsi à la chaîne sa longueur maximum.

Dans la limite des faibles tensions :

$$\frac{Xa}{kT} \ll 1 \,, \tag{D.43}$$

un développement de la fonction de Langevin[19] au premier ordre donne

$$\bar{x} \simeq \frac{Na^2}{3kT} X \,. \tag{D.44}$$

On retrouve ainsi la loi bien connue de l'élasticité : *proportionnalité entre la force appliquée et l'allongement* (loi de Hooke) :

$$X = K(T) \bar{x} \,, \tag{D.45}$$

---

18. Lorsque la tension est nulle ($X = 0$), $\bar{x}$ est nul à cause de la symétrie entre les x positifs et négatifs. Il ne faudrait toutefois pas en conclure que la longueur de la chaîne est alors nulle. Elle est en effet égale en moyenne à $(\overline{r^2})^{1/2}$, c'est-à-dire à l'écart quadratique moyen sur $\vec{r}$. Ce dernier est proportionnel à $\sqrt{N}$: à tension nulle, le problème est équivalent à une marche au hasard à trois dimensions. Dès qu'une force est appliquée, la chaîne s'allonge dans la direction de cette force et sa longueur $\bar{x}$ est proportionnelle à N; les fluctuations autour de cette position, proportionnelles à $\sqrt{N}$, sont alors très faibles en valeur relative puisque $N \gg 1$.

19. Pour $t \ll 1$,

$$\coth t = \frac{1}{t} + \frac{t}{3} + \ldots \,;$$

d'où

$$\mathcal{L}(t) = \frac{1}{t} + \frac{t}{3} - \frac{1}{t} + \ldots \sim \frac{t}{3} \,.$$

$K(T)$ étant le «coefficient de raideur»[20], qui vaut ici

$$K(T)=\frac{3kT}{Na^2}\,. \tag{D.46}$$

Notons que, à tension constante, la longueur de la corde est une fonction décroissante de la température : le système, contrairement à la plupart des corps, a un *coefficient de dilatation négatif;* on retrouve là une propriété bien connue des polymères.

Évidemment, le modèle est trop simpliste pour donner autre chose que des résultats qualitatifs sur les propriétés des chaînes de polymères. Pour améliorer la description, il faudrait tenir compte des interactions entre maillons et des empêchements stériques aux rotations. Mais ceci est une autre affaire et nous renvoyons pour cela aux ouvrages spécialisés.

Références

1. H.M. James et E. Guth, *J. Polymer Sc. 4,* 153 (1949).
2. P.G. de Gennes, *Scaling Concepts in Polymer Physics,* Cornell University Press, Ithaca (1979).
3. P.J. Flory, *Statistical Mechanics of Chain Molecules,* Interscience Publishers, New York (1979).

20. Dans le modèle considéré, l'énergie de la chaîne est la même pour tous les états du système; l'énergie moyenne (égale nécessairement à cette valeur unique) est donc indépendante de la longueur moyenne. La forme $Kx^2/2$ de l'«énergie» que l'on utilise en mécanique correspond à l'*énergie libre F.* En effet, (D.22) appliquée ici donne

$$F=E-TS=G+Xx\,, \tag{1}$$

avec $$G\widehat{=}-kT\ln\widetilde{Z}=-NkT\ln\frac{4\pi\ \mathrm{sh}\ Xa/kT}{Xa/kT}\,. \tag{2}$$

Dans la limite des faibles tensions,

$$G=-NkT\ln 4\pi-\frac{NkT}{6}\left(\frac{Xa}{kT}\right)^2+\dots\,. \tag{3}$$

On en déduit

$$F(T,x)=F(T,0)+K(T)x^2-\frac{N}{6kT}(Xa)^2+\dots\,, \tag{4}$$

c'est-à-dire, en utilisant (D.46),

$$F(T,x)=F(T,0)+\frac{1}{2}K(T)x^2+\dots\,. \tag{5}$$

Dans notre modèle, le terme $Kx^2/2$ provient uniquement du terme $-TS$ de F (puisque E est indépendante de la longueur) : il est donc d'*origine purement entropique.*

COMPLÉMENT III.E

Chaleur spécifique des solides

I. Le modèle d'Einstein
   1. Présentation du modèle
   2. Fonction de partition et propriétés du système
   3. Discussion

II. Vibrations collectives. Modes normaux
   1. Exemple simple : le cristal à une dimension
      a. Description du système. Équations du mouvement
      b. Solutions en ondes progressives
      c. Modes normaux
   2. Modes normaux d'un cristal à trois dimensions
      a. Approximation harmonique et modes normaux
      b. Caractérisation des modes normaux
      c. Relation de dispersion pour les modes de petit vecteur d'onde

III. Quantification des modes normaux de vibration
   1. Expression générale des propriétés du cristal à l'approximation harmonique
   2. L'approximation de Debye
      a. Expression approchée des propriétés du système
      b. Discussion

IV. Cristaux à maille polyatomique
   1. Un exemple très simple
   2. Généralisation
   3. Calcul approché de la chaleur spécifique

V. Les phonons

Les solides cristallins sont constitués par un arrangement régulier, périodique dans l'espace, de motifs microscopiques identiques. Mais ce sont en fait les positions d'équilibre des divers atomes ou ions du cristal qui sont agencées de façon parfaitement régulière[1]; leurs positions réelles s'en écartent légèrement, car ils oscillent constamment. Nous nous proposons d'analyser ici ces *vibrations du réseau cristallin,* qui fournissent la contribution dominante à la chaleur spécifique d'un solide.

Les données expérimentales qu'il s'agit de comprendre sont schématisées sur la figure 1. *A haute température* (c'est-à-dire au-dessus de quelques centaines de kelvins), la capacité calorifique d'un solide obéit à la *loi de Dulong et Petit :* elle est *indépendante de T et du corps considéré,* et vaut 3 Nk, où N est le nombre total

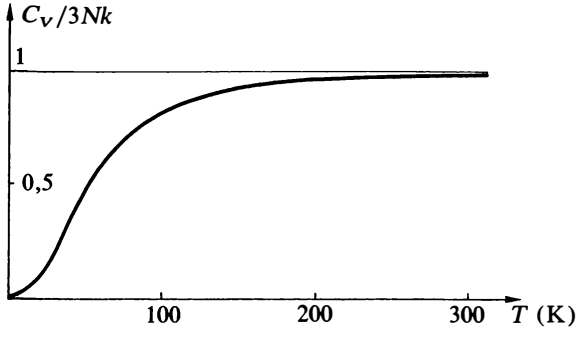


FIGURE 1
Allure des variations de la capacité calorifique d'un solide avec la température (exemple de l'argent).

1. En réalité, les cristaux présentent le plus souvent des défauts : impuretés, dislocations ou lacunes. Nous ne considérons ici que le cas limite de cristaux parfaits.

d'atomes du cristal. Lorsque la température décroît puis tend vers zéro, la capacité calorifique fait de même; *elle se comporte en T^3 à basse température*[2] (c'est-à-dire au-dessous d'une dizaine de kelvins). La mécanique statistique classique permet d'expliquer la loi de Dulong et Petit à partir du théorème d'équipartition de l'énergie. Mais la décroissance de la chaleur spécifique vers les basses températures est un *phénomène d'origine quantique.* C'est ce qu'a montré Einstein, dans le cadre d'un modèle très simple, dès 1907, alors que les idées quantiques en étaient à leurs tout premiers développements. Il faut cependant une théorie plus précise, développée par Debye en 1912, pour prédire correctement le comportement à basse température de la chaleur spécifique des solides.

Nous nous limiterons dans la suite presque exclusivement aux cas où la maille cristalline se réduit à un seul atome ou ion. Nous indiquerons dans la partie IV comment se généralisent les résultats aux solides dont la maille cristalline est polyatomique.

I. LE MODÈLE D'EINSTEIN

1. PRÉSENTATION DU MODÈLE

Dans un solide, les interactions entre particules ne sont jamais négligeables. Le modèle proposé par Einstein suppose néanmoins que chacun des atomes ou ions du cristal vibre autour de sa position d'équilibre indépendamment des autres. Le principal intérêt du modèle d'Einstein réside dans sa très grande *simplicité;* mais, de façon plus précise, la possibilité de ramener l'étude d'un solide à celle d'un système de *particules indépendantes* découle de ce que l'on nomme une *« approximation de champ moyen »*[3].

Considérons en effet l'un quelconque des N atomes ou ions qui constituent le cristal. En première approximation, il ressent l'influence des $(N-1)$ autres particules de façon globale, à travers une *énergie potentielle moyenne* $u_m(\vec{r})$, dépendant de sa position $\vec{r}$, mais pas de celle des autres particules. Repérons $\vec{r}$ par rapport à la position d'équilibre de l'atome considéré, et supposons pour simplifier que le potentiel moyen est isotrope, c'est-à-dire fonction seulement de la distance r à la position d'équilibre. Comme l'atome s'en écarte peu, nous développons le potentiel moyen $u_m(r)$ au voisinage de $r=0$; ce point correspond à un *minimum* du potentiel moyen, de sorte que

$$\frac{\mathrm{d}u_m}{\mathrm{d}r}(0)=0 \; ; \quad \frac{d^2u_m}{\mathrm{d}r^2}(0)\widehat{=}K>0 \; . \tag{E.1}$$

Nous avons donc, à l'ordre non trivial le plus bas en r,

$$u_m(r)\simeq -u_0+(K/2)\,r^2 \; , \tag{E.2}$$

où u_0 est l'énergie de liaison (positive) par atome dans le cristal.

---

2. Pour les métaux, il faut ajouter à la capacité calorifique du réseau cristallin (constitué d'ions positifs) celle du gaz d'électrons quasi-libres qui le baigne. Cette dernière est en fait négligeable devant la précédente, sauf à très basse température (inférieure à quelques kelvins) où elle est proportionnelle à T (complément VI.A, § II.2.a).

3. La méthode du « champ moyen » est très souvent utilisée en mécanique statistique comme première tentative pour prendre en compte de façon approchée les interactions entre constituants microscopiques d'un système. Nous en verrons d'autres exemples dans les compléments III.G à III.K.

A cette approximation, chaque atome ou ion du réseau constitue un oscillateur harmonique à trois dimensions, puisque son énergie s'écrit (m étant sa masse)

$$\varepsilon = \frac{\vec{p}^{\,2}}{2m} - u_0 + \frac{K}{2} r^2 = \frac{1}{2m}(p_x^2 + p_y^2 + p_z^2) + \frac{K}{2}(x^2 + y^2 + z^2) - u_0 \,. \tag{E.3}$$

Les vibrations du cristal sont alors décrites comme un *système de N oscillateurs harmoniques à 3 dimensions indépendants,* ayant tous même pulsation

$$\omega_E = \sqrt{K/m} \,. \tag{E.4}$$

2. FONCTION DE PARTITION ET PROPRIÉTÉS DU SYSTÈME

Les N atomes ou ions du cristal, bien qu'identiques, sont *discernables* puisqu'ils occupent des sites différents (III, § IV.A.2); dans ce modèle, on les considère comme *indépendants;* la fonction de partition Z du système est donc de la forme

$$Z = z^N \,, \tag{E.5}$$

où z est la fonction de partition d'un atome.

On connaît (appendice II, § II.3) les états propres quantiques d'un oscillateur harmonique à une dimension et les énergies associées; comme il s'agit ici d'oscillateurs harmoniques à 3 dimensions, leurs états sont caractérisés par *trois nombres entiers positifs ou nuls* n_x, n_y et n_z, l'énergie de l'état (n_x, n_y, n_z) ayant pour expression

$$\varepsilon_{n_x, n_y, n_z} = -u_0 + + \left[\left(n_x + \frac{1}{2}\right) + \left(n_y + \frac{1}{2}\right) + \left(n_z + \frac{1}{2}\right)\right] \hbar\omega_E \,, \tag{E.6}$$

version quantique de la formule (E.3). La fonction de partition z d'un atome s'écrit donc

$$z = \sum_{n_x,\, n_y,\, n_z = 0}^{\infty} \exp \beta \left[u_0 - \left(n_x + n_y + n_z + \frac{3}{2}\right) \hbar\omega_E\right] . \tag{E.7}$$

Elle se calcule facilement[4] :

$$z = e^{\beta u_0} \left(\frac{e^{-\beta\hbar\omega_E/2}}{1 - e^{-\beta\hbar\omega_E}}\right)^3 = e^{\beta u_0} \left(\frac{1}{2 \operatorname{sh} \beta\hbar\omega_E/2}\right)^3 . \tag{E.8}$$

On trouve ainsi :

$$Z = z^N = e^{N\beta u_0} \left(\frac{1}{2 \operatorname{sh} \beta\hbar\omega_E/2}\right)^{3N} . \tag{E.9}$$

---

4. L'expression (E.7) de z se factorise :

$$z = e^{\beta u_0} \times \left(\sum_{n_x=0}^{+\infty} e^{-(n_x + \frac{1}{2})\beta\hbar\omega_E}\right) \times \left(\sum_{n_y=0}^{+\infty} e^{-(n_y + \frac{1}{2})\beta\hbar\omega_E}\right) \times \left(\sum_{n_z=0}^{+\infty} e^{-(n_z + \frac{1}{2})\beta\hbar\omega_E}\right) . \tag{1}$$

Les trois sommes à calculer sont identiques. Chacune d'elles est la somme d'une progression géométrique de premier terme égal à $\exp[-\beta\hbar\omega_E/2]$ et de raison $\exp[-\beta\hbar\omega_E]$; on a donc par exemple

$$\sum_{n_x=0}^{+\infty} e^{-(n_x + \frac{1}{2})\beta\hbar\omega_E} = \frac{e^{-\beta\hbar\omega_E/2}}{1 - e^{-\beta\hbar\omega_E}} = \frac{1}{2 \operatorname{sh} \beta\hbar\omega_E/2} \,. \tag{2}$$

Le résultat (E.8) s'en déduit aussitôt.

On en déduit toutes les propriétés du système[5], en particulier son énergie moyenne :

$$\overline{E} = -\frac{\partial}{\partial\beta}\ln Z = -N\frac{\partial}{\partial\beta}\ln z$$

$$= N\left\{-u_0 + \frac{3}{2}\hbar\omega_E \coth\frac{\hbar\omega_E}{2kT}\right\} = N\left\{-u_0 + \frac{3}{2}\hbar\omega_E + \frac{3\hbar\omega_E}{e^{\hbar\omega_E/kT}-1}\right\}, \qquad (E.10)$$

et sa capacité calorifique :

$$C_V = \frac{\partial\overline{E}}{\partial T} = 3Nk\left[\frac{\hbar\omega_E}{2kT}\right]^2 \frac{1}{\mathrm{sh}^2\,\hbar\omega_E/2kT} = 3Nk\left(\frac{\hbar\omega_E}{kT}\right)^2 \frac{e^{\hbar\omega_E/kT}}{(e^{\hbar\omega_E/kT}-1)^2}. \qquad (E.11)$$

Il est commode de définir la *« température d'Einstein »* T_E par

$$kT_E \widehat{=} \hbar\omega_E. \qquad (E.12)$$

Comme ω_E, la température d'Einstein dépend du matériau considéré : la formule (E.4) indique que c'est une fonction décroissante de la masse des atomes et une fonction croissante de la constante de rappel K, c'est-à-dire de l'intensité[6] du potentiel moyen u_m. La formule (E.11) devient alors :

$$C_V = 3Nk\left(\frac{T_E}{2T}\right)^2 \frac{1}{\mathrm{sh}^2\,T_E/2T}. \qquad (E.13)$$

A *haute température,* on peut remplacer le sinus hyperbolique par son argument, et la capacité calorifique C_V tend vers une *constante, indépendante du corps considéré :*

$$C_V \simeq 3Nk \quad \text{si} \quad T \gg T_E. \qquad (E.14)$$

Ce résultat remarquable, conforme à la loi empirique de Dulong et Petit (1819), est en réalité une *conséquence du théorème d'équipartition de l'énergie* (III, § V.C), valable lorsque le système peut être décrit par la *mécanique classique :* l'énergie moyenne d'un oscillateur harmonique classique à 3 dimensions vaut $3kT$, ce qui donne aussitôt $3Nk$ pour la capacité calorifique des N particules du cristal.

---

5. L'énergie libre a pour expression

$$F = -kT\ln Z = -NkT\ln z$$

$$= N\{-u_0 + 3kT\ln(2\,\mathrm{sh}\,\hbar\omega_E/2kT)\} = N\left\{-u_0 + \frac{3}{2}\hbar\omega_E + 3kT\ln(1-e^{-\hbar\omega_E/kT})\right\}, \qquad (1)$$

et l'entropie

$$S = -\frac{\partial F}{\partial T} = 3Nk\left\{-\ln(2\,\mathrm{sh}\,\hbar\omega_E/2kT) + \frac{\hbar\omega_E}{2kT}\coth\frac{\hbar\omega_E}{2kT}\right\}$$

$$= 3Nk\left\{-\ln(1-e^{-\hbar\omega_E/kT}) + \frac{\hbar\omega_E}{kT}\frac{1}{e^{\hbar\omega_E/kT}-1}\right\}. \qquad (2)$$

6. La constante de rappel est donnée par la dérivée seconde de la fonction u_m en son minimum (formule (E.1)). Cette dérivée seconde a pour ordre de grandeur

$$\frac{d^2u_m}{dr^2}(0) \sim \frac{u_0}{\rho_0^2},$$

où ρ_0 est la distance entre deux atomes voisins dans le cristal. Comme ρ_0 reste, à un facteur 2 ou 3 près, la même pour tous les corps (quelques angströms), K est surtout sensible à la profondeur u_0 du « puits de potentiel » où se déplace l'atome.

A *basse température,* la formule (E.13) devient

$$C_V \simeq 3Nk(T_E/T)^2 \, e^{-T_E/T} \quad \text{si} \quad T \ll T_E \, . \tag{E.15}$$

Cette expression tend vers zéro avec T de façon exponentielle, car c'est le facteur $\exp(-T_E/T)$ qui domine la dépendance en T.

3. DISCUSSION

L'importance du modèle d'Einstein est considérable du point de vue historique : il a montré que la décroissance de la chaleur spécifique des solides vers les basses températures, inexplicable dans le cadre de la mécanique classique, est un phénomène d'origine quantique. En outre, la courbe représentant l'expression (E.13) de C_V en fonction de la température est semblable à celle de la figure 1, qui donne l'allure des résultats expérimentaux : en ajustant pour chaque corps la température d'Einstein T_E, on reproduit approximativement la variation de sa chaleur spécifique avec la température.

Ce modèle présente cependant des insuffisances et des défauts : s'il donne l'allure générale de la variation de C_V avec la température, le comportement qu'il prédit à basse température (formule (E.15)) est nettement différent de celui qui est observé ($C_V \propto T^3$). La présence de l'exponentielle, dans la formule (E.15), est une conséquence directe de *l'existence d'un quantum d'excitation minimum $\hbar\omega_E$ non nul.* En effet, si nous notons E_0 et E_1 les deux niveaux d'énergie les plus bas du système, ils sont ici séparés par l'intervalle $E_1 - E_0 = \hbar\omega_E$; pour T tendant vers zéro, la fonction de partition prend donc la forme

$$Z \simeq g_0 \, e^{-\beta E_0} + g_1 \, e^{-\beta E_1} = e^{-\beta E_0}(g_0 + g_1 \, e^{-\beta\hbar\omega_E}) \, , \tag{E.16}$$

où g_0 et g_1 notent les degrés de dégénérescence respectifs de E_0 et E_1. L'énergie moyenne est alors

$$\overline{E} = -\frac{\partial}{\partial\beta} \ln Z \simeq E_0 + g_1\hbar\omega_E \frac{1}{g_0 \, e^{\hbar\omega_E/kT} + g_1} \, , \tag{E.17}$$

et la capacité calorifique

$$C_V = \frac{d\overline{E}}{dT} = g_0 g_1 \left(\frac{\hbar\omega_E}{kT}\right)^2 \frac{e^{\hbar\omega_E/kT}}{(g_0 \, e^{\hbar\omega_E/kT} + g_1)^2} \underset{T\to 0}{\sim} \frac{g_1}{g_0}\left(\frac{\hbar\omega_E}{kT}\right)^2 e^{-\hbar\omega_E/kT} \, . \tag{E.18}$$

On ne peut donc espérer reproduire le comportement en puissance de T observé que si le spectre d'énergie du système est pratiquement continu au voisinage du niveau fondamental E_0. Remarquons pour finir que l'existence d'une pulsation caractéristique ω_E unique est liée à l'*indépendance* des divers oscillateurs harmoniques : lorsqu'on tient compte des termes de couplage, les fréquences propres d'un système de deux ou plusieurs oscillateurs harmoniques identiques s'écartent de leur valeur commune initiale.

II. VIBRATIONS COLLECTIVES. MODES NORMAUX

Les vibrations des divers atomes d'un cristal sont évidemment couplées : lorsque l'un des atomes s'écarte de sa position d'équilibre, il exerce des forces sur ses voisins, repoussant ceux dont il s'approche et attirant ceux dont il s'éloigne; ceux-ci agissent à leur tour sur leurs propres voisins, et ainsi de suite de proche en proche. *Les vibrations du cristal sont donc collectives :* elles intéressent non pas chaque atome indépendamment des autres, mais l'ensemble des atomes constituant le réseau. En outre, *elles se propagent sous forme d'ondes,* puisque chaque atome transmet le mouvement à ses voisins.

1. EXEMPLE SIMPLE : LE CRISTAL A UNE DIMENSION

a. Description du système. Équations du mouvement

Nous considérons une chaîne linéaire de N atomes identiques M_q ($q = 1, 2, \ldots, N$), de masse m, régulièrement espacés. Pour n'avoir pas à nous préoccuper des effets de bords et pouvoir appliquer commodément les conditions aux limites périodiques, nous imaginons que ces atomes sont répartis sur un cercle de longueur L, de sorte que M_{N+1} coïncide avec M_1 (fig. 2). La position d'équilibre de l'atome M_q est située au point d'abscisse (curviligne) $q\ell$, où $\ell = L/N$ est le pas du réseau; nous ne considérons que des mouvements longitudinaux et notons x_q l'écart (algébrique) de M_q par rapport à sa position d'équilibre.

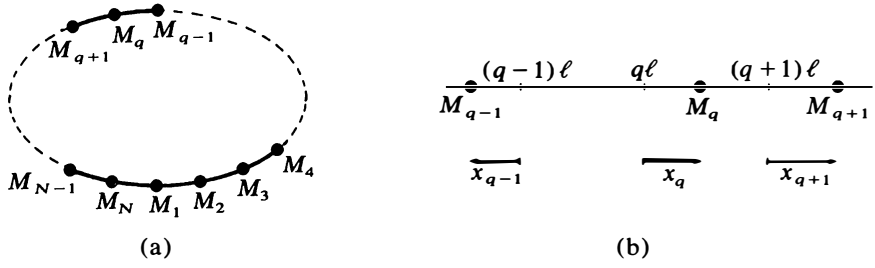


FIGURE 2
Chaîne linéaire fermée de N atomes M_q, constituant un cristal à une dimension sans effets de bords. La figure 2.a indique les positions d'équilibre des atomes, la figure 2.b leurs écarts x_q par rapport à ces positions, à un instant donné.

Nous supposons que chaque atome M_q de la chaîne est soumis, de la part de chacun de ses deux voisins M_{q-1} et M_{q+1}, à une force de rappel de type harmonique, c'est-à-dire proportionnelle à l'écart relatif $(x_q - x_{q\pm1})$:

$$F_q = -K(x_q - x_{q+1}) - K(x_q - x_{q-1}) , \tag{E.19}$$

où F_q est la mesure algébrique de la force le long de la chaîne orientée et K une constante de rappel positive. L'ensemble des forces F_q dérive de l'énergie potentielle

$$U = -U_0 + \frac{1}{2} K \sum_{q=1}^{N} (x_q - x_{q+1})^2 , \tag{E.20}$$

c'est-à-dire que

$$F_q = -\frac{\partial}{\partial x_q} U . \tag{E.21}$$

Les équations classiques du mouvement se déduisent immédiatement de (E.19) :

$$\frac{d^2 x_q}{dt^2} = -\omega_0^2 (2x_q - x_{q+1} - x_{q-1}) , \tag{E.22}$$

avec $\omega_0 \widehat{=} \sqrt{K/m}$; (E.23)

elles forment un *système de N équations couplées* pour les N variables x_q.

b. Solutions en ondes progressives

Le mouvement le plus général du cristal linéaire que nous étudions peut être décomposé en une *superposition d'ondes progressives.* Dans une onde progressive

(longitudinale) de vecteur d'onde k (positif ou négatif) et de pulsation ω (positive), le déplacement $x_q(t)$ de l'atome M_q, d'abscisse à l'équilibre $q\ell$, est de la forme

$$x_q(t) = a\ e^{i[k\cdot q\ell - \omega t]} . \tag{E.24}$$

On vérifie en effet facilement[7] que cette expression est solution des équations (E.22), à condition que ω et k vérifient la *« relation de dispersion »*

$$\omega(k) = 2\omega_0 \left|\sin(k\ell/2)\right| . \tag{E.25}$$

Le vecteur d'onde k ne peut pas être quelconque. En premier lieu, ses valeurs permises doivent vérifier la *condition aux limites périodique :*

$$x_{N+1} \equiv x_1 . \tag{E.26}$$

Cette condition donne comme toujours

$$k = n\frac{2\pi}{N\ell} = n\frac{2\pi}{L} , \tag{E.27}$$

où n est entier, positif, négatif ou nul. Mais en outre deux vecteurs d'onde k et k' différant par un nombre entier de fois $2\pi/\ell$ donnent la même pulsation (formule (E.25)) et le même déplacement $x_q(t)$ à chacun des atomes M_q (formule (E.24))[8] :

$$k' = k + p\frac{2\pi}{\ell}\ (p \text{ entier relatif}) \implies \begin{cases} \omega(k') = \omega(k) \\ e^{ik'\cdot q\ell} = e^{ik\cdot q\ell} \quad \forall q . \end{cases} \tag{E.28}$$

On obtient donc une seule fois toutes les ondes progressives distinctes physiquement en restreignant la variation de k à un intervalle de largeur $2\pi/\ell$, par exemple à la *« première zone de Brillouin »* (complément I.F, § I.3) :

$$-\frac{\pi}{\ell} \leqslant k < \frac{\pi}{\ell} . \tag{E.29}$$

Les *vecteurs d'onde permis* ont donc la forme (E.27), avec

$$-\frac{N}{2} \leqslant n < \frac{N}{2} ; \tag{E.30}$$

ils sont au nombre de N.

La relation de dispersion (E.25) est représentée, pour la première zone de Brillouin (E.29), sur la figure 3. Noter que, dans un domaine assez large autour de $k = 0$, la pulsation ω de l'onde est pratiquement proportionnelle au module du vecteur d'onde :

$$\omega \simeq c\left|k\right| \quad \text{pour } \left|k\right| \text{ petit} , \tag{E.31}$$

où $\quad c \mathrel{\widehat{=}} \omega_0\ell$ (E.32)

est la *vitesse du son* dans le cristal.

---

7. Lorsqu'on reporte (E.24) dans (E.22) on obtient, après division des deux membres par x_q,

$$-\omega^2 = -\omega_0^2[2 - e^{ik\ell} - e^{-ik\ell}] , \tag{1}$$

soit $\quad \omega^2 = 2\omega_0^2(1 - \cos k\ell) = 4\omega_0^2 \sin^2 k\ell/2 . \tag{2}$

8. Ceci est dû au fait que le système considéré est *discret :* l'onde ne se manifeste véritablement, par un déplacement des atomes, qu'aux points d'abscisse $q\ell$. Les oscillations de l'onde entre deux atomes consécutifs, qui se produisent si la longueur d'onde $2\pi/\left|k\right|$ est inférieure à la distance ℓ, n'ont pas de signification physique.

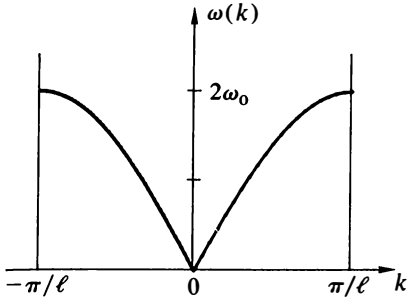


FIGURE 3
Relation de dispersion (ω en fonction de k) pour un cristal à une dimension.

c. Modes normaux

Les équations (E.22) donnant les déplacements x_q en fonction du temps sont des équations d'oscillateurs harmoniques, mais ces oscillateurs sont couplés. Les résultats du paragraphe b permettent de trouver facilement des *«variables normales»*, combinaisons linéaires des x_q, qui obéissent à des équations d'*oscillateurs harmoniques découplés,* que l'on appelle les *«modes normaux»* ou «modes propres» de vibration du système.

Posons en effet

$$\xi(k;\, t) \widehat{=} \sum_{q=1}^{N} e^{ikq\ell} x_q(t) \,. \tag{E.33}$$

En prenant successivement pour k les N valeurs permises données par (E.27) et (E.30), on définit ainsi N variables ξ; il est facile de voir[9] que chacune d'elles est régie par l'équation

$$\frac{\partial^2}{\partial t^2}\, \xi(k;\, t) = -\left[\omega(k)\right]^2 \xi(k;\, t) \,, \tag{E.34}$$

où $\omega(k)$ est donné par (E.25). On a bien N équations d'oscillateurs harmoniques découplées.

---

9. Multiplions les deux membres de l'équation (E.22) par exp ($ikq\ell$) et sommons sur q de 1 à N. Le premier membre et le premier terme du second membre donnent directement $\xi(k;\, t)$. Pour les deux derniers termes, on a à calculer

$$\sum_{q=1}^{N} e^{ikq\ell} x_{q+1} = e^{-ik\ell} \sum_{q=1}^{N} e^{ik(q+1)\ell} x_{q+1} \tag{1}$$

et

$$\sum_{q=1}^{N} e^{ikq\ell} x_{q-1} = e^{ik\ell} \sum_{q=1}^{N} e^{ik(q-1)\ell} x_{q-1} \,. \tag{2}$$

Nous avons imposé que

$$x_{N+1} \equiv x_1 \,, \tag{3}$$

et le fait que k soit de la forme (E.27) implique

$$e^{ikN\ell} = 1 \,, \tag{4}$$

de sorte que, dans l'expression (1),

$$\sum_{q=1}^{N} e^{ik(q+1)\ell} x_{q+1}(t) = \xi(k;\, t) \,. \tag{5}$$

.../...

2. MODES NORMAUX D'UN CRISTAL A TROIS DIMENSIONS

Les notions et résultats que nous venons d'introduire sur un exemple simple se généralisent aux cristaux à trois dimensions.

a. Approximation harmonique et modes normaux

Convenons de repérer chacun des N atomes ou ions du cristal par un vecteur $\vec{r}_i (i = 1, 2, \ldots, N)$ caractérisant son écart par rapport à sa position d'équilibre; nous noterons $r_{i\alpha} (\alpha = x, y, z)$ les composantes cartésiennes de $\vec{r}_i$. Comme les atomes s'écartent peu de leur position d'équilibre, nous développons l'énergie potentielle $U(\vec{r}_1, \vec{r}_2, \ldots, \vec{r}_N)$ du cristal en puissances des $3N$ variables $r_{i\alpha}$. Les termes du premier ordre sont nuls, car le point $\vec{r}_1 = \vec{r}_2 = \ldots = \vec{r}_N = \vec{0}$ autour duquel s'effectue le développement est un minimum de U :

$$\frac{\partial U}{\partial r_{i\alpha}} (\vec{0}, \vec{0}, \ldots, \vec{0}) = 0 \quad \forall i \in (1, 2, \ldots, N); \ \alpha \in (x, y, z) . \tag{E.35}$$

Si nous posons

$$\frac{\partial^2 U}{\partial r_{i\alpha} \partial r_{j\beta}} (\vec{0}, \vec{0}, \ldots, \vec{0}) \widehat{=} K_{i\alpha, j\beta} , \tag{E.36}$$

nous pouvons écrire

$$U(\vec{r}_1, \vec{r}_2, \ldots, \vec{r}_N) \simeq - U_0 + \frac{1}{2} \sum_{i, j = 1}^{N} \sum_{\alpha, \beta = x, y, z} K_{i\alpha, j\beta} r_{i\alpha} r_{j\beta} , \tag{E.37}$$

où U_0 (positif) est l'énergie de liaison totale du cristal.

A cette *« approximation harmonique »*, les vibrations des N atomes du cristal sont décrites comme un *système de* $3N$ *oscillateurs harmoniques à une dimension couplés :* les termes « diagonaux » de l'expression (E.37), en $K_{i\alpha, i\alpha} r_{i\alpha}^2 / 2$, correspondent aux forces de rappel élastiques de ces $3N$ oscillateurs, et les termes « non diagonaux » dans lesquels $(i, \alpha) \neq (j, \beta)$ sont responsables du couplage.

On peut montrer de façon générale qu'il existe toujours des *variables normales* $\rho_{i\alpha}$, combinaisons linéaires des $r_{j\beta}$, qui permettent d'écrire la formule (E.37) sous la forme[10]

$$U = - U_0 + \frac{1}{2} \sum_{i=1}^{N} \sum_{\alpha = x, y, z} \widetilde{K}_{i\alpha} \rho_{i\alpha}^2 . \tag{E.38}$$

Comme $-U_0$ est le *minimum* de U, les coefficients $\widetilde{K}_{i\alpha}$ sont tous *positifs*. On a donc maintenant, dans le cadre de la *même approximation* que ci-dessus, un *système de* $3N$ *oscillateurs harmoniques à une dimension découplés,* c'est-à-dire indépendants les uns des autres; ce sont les *modes normaux de vibration* du cristal.

La somme à calculer dans (2) fait intervenir x_0, déplacement de l'atome précédant immédiatement M_1 dans la chaîne, c'est-à-dire

$$x_0 \equiv x_N . \tag{6}$$

Le coefficient de x_0 est égal à 1, comme celui de x_N d'après (4). On obtient donc ici aussi $\xi(k; t)$. L'équation (E.22) donne en définitive

$$\frac{\partial^2}{\partial t^2} \xi(k; t) = - \omega_0^2 (2 - e^{-ik\ell} - e^{ik\ell}) \xi(k; t) , \tag{7}$$

c'est-à-dire (E.34) (cf. note 7).

10. La matrice $3N \times 3N$ constituée par les coefficients $K_{i\alpha, j\beta}$ est réelle et symétrique. Elle est donc diagonalisable, et ses valeurs propres sont réelles (C.D.L., p. 136).

b. Caractérisation des modes normaux

Les modes normaux correspondent à des *vibrations collectives* de l'ensemble des atomes du cristal. Si l'on impose des conditions aux limites périodiques, à chaque mode normal est d'abord associé un *vecteur d'onde* $\vec{k}$ de la forme (complément I.F, § II.2)

$$\vec{k} = \frac{p_1}{N_1}\vec{b}_1 + \frac{p_2}{N_2}\vec{b}_2 + \frac{p_3}{N_3}\vec{b}_3 \ : \tag{E.39}$$

$\vec{b}_1$, $\vec{b}_2$, $\vec{b}_3$ sont les vecteurs de base du *réseau réciproque* du cristal; N_1, N_2, N_3 sont des nombres très grands, d'ordre $N^{1/3}$, tels que $N_1 . N_2 . N_3 = N$; p_1, p_2, p_3 sont des entiers positifs, négatifs ou nuls. Si l'on restreint en outre $\vec{k}$ à la *première zone de Brillouin* du réseau réciproque, *le nombre de vecteurs* $\vec{k}$ *permis est égal au nombre N de nœuds* du réseau cristallin.

Mais, dans un cristal à trois dimensions, *à chaque vecteur d'onde permis sont associés trois modes normaux,* différant les uns des autres par la *polarisation* de la vibration, c'est-à-dire la direction dans laquelle se fait le déplacement des atomes au passage de l'onde, par rapport au vecteur d'onde $\vec{k}$ (réf. 2, p. 439). Nous les repèrerons par un indice λ prenant les trois valeurs 1, 2, 3.

La pulsation ω associée à un mode normal dépend en général du vecteur d'onde $\vec{k}$ et de la polarisation (λ) de ce mode :

$$\omega = \omega_\lambda(\vec{k}) \ ; \quad \lambda = 1, \ 2, \ 3 \ . \tag{E.40}$$

On dit que *la relation de dispersion possède trois «branches» distinctes.*

c. Relation de dispersion pour les modes de petit vecteur d'onde

Lorsque le vecteur d'onde $\vec{k}$ est suffisamment petit en module, plus précisément lorsque

$$|\vec{k}| \ll \pi/\ell \ , \tag{E.41}$$

où ℓ est l'ordre de grandeur du pas du réseau, le mode normal correspondant est pratiquement insensible à la structure discontinue du cristal, car sa longueur d'onde $2\pi/|\vec{k}|$ est grande devant les distances entre atomes voisins. Pour de tels modes normaux, le cristal se comporte comme un milieu continu. Les ondes de vibration d'un milieu continu sont des *ondes sonores,* qui obéissent à l'équation de propagation[11]

$$\Delta n - \frac{1}{c^2}\frac{\partial^2 n}{\partial t^2} = 0 \ : \tag{E.42}$$

$n(\vec{r}, t)$ désigne ici par exemple l'écart (algébrique) de la densité de particules, au point $\vec{r}$ et à l'instant t, par rapport à sa valeur au repos; c est la *vitesse du son* dans le cristal (que nous prenons, pour simplifier, isotrope).

---

11. Reprenons pour simplifier le modèle à une dimension du paragraphe 1. La densité de particules (par unité de longueur) y est égale à $1/\ell$ en valeur moyenne. L'écart par rapport à cette valeur moyenne s'écrit, au voisinage du point d'abscisse $q\ell$,

$$n(q\ell; t) \simeq \frac{1}{\ell + x_q - x_{q-1}} - \frac{1}{\ell} \simeq -\frac{x_q - x_{q-1}}{\ell^2} \ . \tag{1}$$

Les équations (E.22) donnent alors

$$\frac{\partial^2}{\partial t^2} n(q\ell; t) = \omega_0^2 \{[n((q+1)\ell; t) - n(q\ell; t)] - [n(q\ell; t) - n((q-1)\ell; t)]\} \ . \tag{2}$$

.../...

L'équation de propagation (E.42) admet pour solutions des *ondes progressives sinusoïdales* de la forme

$$n(\vec{r}, t) = n_0 \, e^{i(\vec{k}.\vec{r} - \omega t)} , \tag{E.43}$$

à condition que la pulsation ω soit liée au vecteur d'onde $\vec{k}$ par

$$\omega = c\,|\vec{k}| . \tag{E.44}$$

En réalité, un solide peut être le siège d'ondes sonores *transversales* (oscillations du milieu dans une direction perpendiculaire au vecteur $\vec{k}$ qui donne la direction de propagation de l'onde) aussi bien que *longitudinales* (oscillations parallèles à $\vec{k}$). Les vitesses de propagation c_t et c_ℓ des ondes transversales et longitudinales sont le plus souvent différentes, de sorte que la relation de dispersion comporte *deux «branches acoustiques» différentes,* caractérisées respectivement par

$$\omega_t(\vec{k}) = c_t\,|\vec{k}| , \tag{E.45.a}$$

$$\omega_\ell(\vec{k}) = c_\ell\,|\vec{k}| . \tag{E.45.b}$$

La branche transversale est double, car il existe deux directions indépendantes dans le plan perpendiculaire au vecteur d'onde $\vec{k}$; on retrouve donc les *trois branches* introduites de façon générale au paragraphe b.

III. QUANTIFICATION DES MODES NORMAUX DE VIBRATION

1. EXPRESSION GÉNÉRALE DES PROPRIÉTÉS DU CRISTAL A L'APPROXIMATION HARMONIQUE

Dans le domaine des faibles écarts par rapport à leur position d'équilibre (approximation harmonique), l'étude des N atomes constituant un cristal à trois dimensions se ramène à celle d'un système de $3N$ oscillateurs harmoniques indépendants, les modes normaux de vibration.

Comme le nombre d'oscillateurs indépendants reste toujours égal à $3N$, le comportement de la chaleur spécifique à haute température, c'est-à-dire dans le domaine classique, est le même que dans le modèle d'Einstein : on retrouve la loi de Dulong et Petit. Mais le comportement à basse température va être très différent, puisqu'on a ici un spectre de pulsations ω possibles, et non plus une pulsation unique ω_E.

Chaque mode normal est repéré par un vecteur d'onde $\vec{k}$ et une polarisation (λ) (§ II.2.b). Un état quantique microscopique du système est donc caractérisé par $3N$ entiers positifs ou nuls $n_{\vec{k}\lambda}$, et son énergie s'écrit

$$E_{\{n_{\vec{k}\lambda}\}} = -U_0 + \sum_{\vec{k},\,\lambda} \left(n_{\vec{k}\lambda} + \frac{1}{2}\right) \hbar\omega_\lambda(\vec{k}) : \tag{E.46}$$

Pour des vibrations dont la longueur d'onde $2\pi/|k|$ est grande devant ℓ, on peut considérer que

$$(1/\ell)[n((q+1)\ell; t) - n(q\ell; t)] \tag{3}$$

est la dérivée par rapport à l'abscisse x de la fonction $n(x; t)$, où x est une variable pratiquement continue. C'est donc la dérivée seconde de n par rapport à x qui apparaît au second membre de l'égalité (2), qui s'écrit donc

$$\frac{\partial^2}{\partial t^2} n(x; t) = \omega_0^2 \ell^2 \frac{\partial^2}{\partial x^2} n(x; t) . \tag{4}$$

Compte tenu de (E.32), on retrouve l'équation de propagation (E.42).

la somme sur $\vec{k}$ porte sur les N vecteurs d'onde permis, et la somme sur λ porte sur les 3 polarisations possibles. La fonction de partition Z du système prend alors la forme

$$Z = e^{\beta U_0} \prod_{\vec{k},\,\lambda} z_{\vec{k}\lambda} , \tag{E.47}$$

où
$$z_{\vec{k}\lambda} \widehat{=} \sum_{n_{\vec{k}\lambda}=0}^{\infty} e^{-\beta\left(n_{\vec{k}\lambda}+\frac{1}{2}\right)\hbar\omega_\lambda(\vec{k})} \tag{E.48}$$

est la fonction de partition du mode normal $(\vec{k}, \lambda)$. Le calcul est le même que dans la note 4 :

$$z_{\vec{k}\lambda} = \frac{\exp\left[-\beta\hbar\omega_\lambda(\vec{k})/2\right]}{1-\exp\left[-\beta\hbar\omega_\lambda(\vec{k})\right]} = \frac{1}{2\,\mathrm{sh}\left[\beta\hbar\omega_\lambda(\vec{k})/2\right]} . \tag{E.49}$$

On en déduit les propriétés du système[12], en particulier son énergie moyenne :

$$\bar{E} = -\frac{\partial}{\partial\beta}\ln Z = -U_0 + \frac{1}{2}\sum_{\vec{k},\,\lambda}\hbar\omega_\lambda(\vec{k}) + \sum_{\vec{k},\,\lambda}\hbar\omega_\lambda(\vec{k})\,\frac{1}{\exp\left[\hbar\omega_\lambda(\vec{k})/kT\right]-1} , \tag{E.50}$$

et sa capacité calorifique :

$$C_V = \frac{\partial\bar{E}}{\partial T} = k\sum_{\vec{k},\,\lambda}\left(\frac{\hbar\omega_\lambda(\vec{k})}{kT}\right)^2 \frac{\exp\left[\hbar\omega_\lambda(\vec{k})/kT\right]}{\left[\exp\left(\hbar\omega_\lambda(\vec{k})/kT\right)-1\right]^2} \tag{E.51}$$

(ne pas confondre la constante de Boltzmann k et le vecteur d'onde $\vec{k}$).

Comme les vecteurs $\vec{k}$ permis sont très proches les uns des autres, on peut remplacer la somme discrète sur $\vec{k}$ par une intégrale, à condition de diviser le volume infinitésimal d^3k de l'espace des $\vec{k}$ par celui de la maille élémentaire du réseau réciproque, qui vaut $(2\pi)^3/V$ où V est le volume (dans l'espace ordinaire) du cristal (complément I.F, § II.2) :

$$\sum_{\vec{k}} \longrightarrow \frac{V}{(2\pi)^3}\int d^3k \; ; \tag{E.52}$$

l'intégrale est limitée à la première zone de Brillouin. On peut également introduire et utiliser la *densité* $\rho(\omega)$ *de modes normaux :* par définition, $\rho(\omega)\,d\omega$ est le nombre de

12. Comme dans la note 5, on tire facilement de la fonction de partition (E.47)-(E.49) l'expression (E.50) de l'énergie moyenne, ainsi que celles de l'énergie libre :

$$F = -kT\ln Z = -U_0 + \frac{1}{2}\sum_{\vec{k},\,\lambda}\hbar\omega_\lambda(\vec{k}) + kT\sum_{\vec{k},\,\lambda}\ln\left\{1-\exp\left[-\hbar\omega_\lambda(\vec{k})/kT\right]\right\} , \tag{1}$$

et de l'entropie :

$$S = -\frac{\partial F}{\partial T} = \frac{1}{T}\left(\bar{E}-F\right) = k\sum_{\vec{k},\,\lambda}\left\{-\ln\left[1-e^{-\hbar\omega_\lambda(\vec{k})/kT}\right] + \frac{\hbar\omega_\lambda(\vec{k})}{kT}\,\frac{1}{\exp\left[\hbar\omega_\lambda(\vec{k})/kT\right]-1}\right\} . \tag{2}$$

modes normaux $(\vec{k}, \lambda)$ dont la pulsation $\omega_\lambda(\vec{k})$ est comprise entre ω et $\omega + d\omega$. L'énergie moyenne (E.50) prend alors la forme

$$\overline{E} = -E_0 + \int_0^{\omega_M} \rho(\omega)\, d\omega \, \frac{\hbar\omega}{e^{\hbar\omega/kT} - 1}, \tag{E.53}$$

où $$E_0 \hat{=} U_0 - \frac{1}{2} \sum_{\vec{k},\,\lambda} \hbar\omega_\lambda(\vec{k}) \tag{E.54}$$

est une constante indépendante de T qui prend en compte l'« énergie de point zéro » des $3N$ modes normaux[13]; le nombre de modes normaux étant fini, l'intégrale sur la pulsation ω est bornée supérieurement à une valeur ω_M donnée par

$$\int_0^{\omega_M} \rho(\omega)\, d\omega = 3N. \tag{E.55}$$

La densité $\rho(\omega)$ de modes normaux peut être déterminée expérimentalement par diffusion inélastique de rayons X ou de neutrons. La figure 4 indique l'allure de la courbe ainsi obtenue. Sa forme, relativement compliquée, présente le plus souvent deux maximums : le plus étroit, qui se produit pour ω proche de la borne supérieure ω_M du spectre, correspond aux modes longitudinaux; l'autre, plus large et situé vers $\omega \sim \omega_M/2$ ou $\omega_M/3$, provient principalement des modes transversaux. Pour les petites valeurs de ω, la courbe démarre comme ω^2 (voir formule (E.59) page suivante).

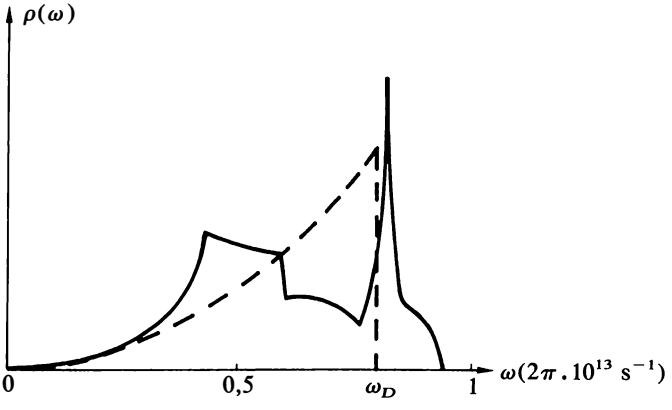


FIGURE 4

Densité de modes normaux dans un cristal tel que l'aluminium. La courbe en trait plein schématise la densité déterminée expérimentalement, la courbe pointillée correspond à l'approximation de Debye (§ 2).

2. L'APPROXIMATION DE DEBYE

A *haute température,* chacun des termes de la somme (E.51) tend vers 1 (cf. formules (E.11)-(E.14)); comme ils sont au nombre de $3N$, on retrouve comme on s'y attendait *la loi de Dulong et Petit.* D'autre part, la discussion du paragraphe I.3 permet

13. E_0 est l'*énergie de dissociation* du cristal, c'est-à-dire l'énergie minimale qu'il faut fournir (à température nulle) pour séparer les N atomes qui le constituent.

de comprendre que *ce sont les modes normaux de faible pulsation ω qui dominent à basse température :* lorsque l'énergie caractéristique kT décroît, la contribution des modes normaux dont la pulsation devient grande devant $kT/\hbar$ chute exponentiellement (formule (E.18)). Or la relation de dispersion est connue dans ce domaine : elle est donnée par les formules simples (E.45).

a. Expression approchée des propriétés du système

L'approximation de Debye consiste à étendre à tous les modes normaux une relation de dispersion linéaire de la forme (E.45). On obtient ainsi une *interpolation approchée* entre les hautes températures, où la forme de la relation de dispersion est sans importance, et les basses températures où les formules (E.45) dominent.

Avec la relation de dispersion simple

$$\omega_\lambda(\vec{k}) = c_\lambda |\vec{k}| , \tag{E.56}$$

la densité $\rho_\lambda(\omega)$ de modes normaux de polarisation (λ) se calcule aisément[14] :

$$\rho_\lambda(\omega) = \frac{4\pi V}{(2\pi c_\lambda)^3} \omega^2 . \tag{E.57}$$

Dans la formule (E.53), la fonction à intégrer ne dépend pas de la polarisation des modes normaux; la densité totale de modes normaux s'écrit alors :

$$\rho(\omega) = 2\rho_t(\omega) + \rho_\ell(\omega) , \tag{E.58}$$

c'est-à-dire

$$\rho(\omega) = \frac{3V}{2\pi^2 \tilde{c}^3} \omega^2 , \tag{E.59}$$

où $\tilde{c}$ est définie par

$$\frac{3}{\tilde{c}^3} \widehat{=} \frac{2}{c_t^3} + \frac{1}{c_\ell^3} . \tag{E.60}$$

La « pulsation de coupure » ω_M, déterminée par la relation générale (E.55), devient ici la *« pulsation de Debye »* ω_D, qui se déduit aussitôt de la forme (E.59) de la densité de modes normaux (cf. figure 4) :

$$\int_0^{\omega_D} \frac{3V}{2\pi^2 \tilde{c}^3} \omega^2 \, d\omega = 3N \implies \omega_D = \tilde{c} . \left(\frac{6\pi^2 N}{V}\right)^{1/3} . \tag{E.61}$$

Il est habituel d'introduire la *température de Debye* T_D par

$$kT_D \widehat{=} \hbar\omega_D . \tag{E.62}$$

L'expression (E.53) de l'*énergie moyenne* du cristal s'écrit, à l'approximation de Debye,

$$\overline{E} = -E_0 + \frac{3V}{2\pi^2 \tilde{c}^3} \int_0^{\omega_D} \frac{\hbar\omega^3 d\omega}{e^{\hbar\omega/kT} - 1} . \tag{E.63}$$

---

14. Le nombre $\rho_\lambda(\omega)\, d\omega$ de modes normaux de polarisation (λ) dont la pulsation est comprise entre ω et $\omega + d\omega$ est donné par

$$\rho_\lambda(\omega)\, d\omega = \frac{V}{(2\pi)^3} . 4\pi k^2 \, dk ,$$

où ω et $k = |\vec{k}|$ sont liés par l'égalité (E.56). La formule (E.57) en découle immédiatement.

Introduisons la variable sans dimension

$$x \widehat{=} \hbar\omega/kT \; ; \tag{E.64}$$

si l'on utilise la définition (E.62) et la formule (E.61), on aboutit à

$$\bar{E} = -E_0 + 9NkT\left(\frac{T}{T_D}\right)^3 \int_0^{T_D/T} \frac{x^3\,dx}{e^x - 1} \, . \tag{E.65}$$

La *capacité calorifique* se calcule aussi très facilement dans le modèle de Debye : il suffit de dériver par rapport à T l'expression (E.63) puis d'effectuer le même changement de variable (E.64). Il vient ainsi

$$C_V = 9Nk\left(\frac{T}{T_D}\right)^3 \int_0^{T_D/T} \frac{x^4\,e^x}{(e^x - 1)^2}\,dx \, . \tag{E.66}$$

b. Discussion

A basse température, c'est-à-dire pour $T \ll T_D$, *l'approximation de Debye devient pratiquement exacte.* En effet, la contribution des modes normaux pour lesquels la relation de dispersion s'écarte de la forme simple (E.45) décroît exponentiellement avec T. Corrélativement, la borne supérieure des intégrales figurant dans les formules (E.65) et (E.66) devient très grande, et comme la fonction à intégrer est très rapidement décroissante à cause de l'exponentielle du dénominateur, on peut sans inconvénient prolonger l'intégration jusqu'à l'infini; les intégrales définies correspondantes figurent dans les tables : il nous suffit de savoir que

$$\int_0^{+\infty} \frac{x^3 dx}{e^x - 1} = \frac{\pi^4}{15} \, . \tag{E.67}$$

L'énergie moyenne s'écrit alors

$$\bar{E} \simeq -E_0 + \frac{3}{5}\,\pi^4 NkT\left(\frac{T}{T_D}\right)^3 \quad \text{pour} \quad T \ll T_D \, , \tag{E.68}$$

et l'on en déduit la capacité calorifique

$$C_V \simeq \frac{12}{5}\,\pi^4 Nk\left(\frac{T}{T_D}\right)^3 \simeq 233{,}8Nk\left(\frac{T}{T_D}\right)^3 \quad \text{pour} \quad T \ll T_D \, . \tag{E.69}$$

Le comportement en T^3 de la chaleur spécifique aux basses températures est donc parfaitement reproduit par le modèle de Debye. En ajustant les données expérimentales dans ce domaine, on peut déterminer pour chaque corps sa température de Debye T_D. Le tableau I donne quelques-unes des valeurs ainsi obtenues.

TABLEAU I

Température de Debye (en kelvins) pour quelques corps solides, déterminée à partir du comportement de la chaleur spécifique à basse température (formule (E.69))

Na	150	Zn	234
K	100	Cd	120
Mg	318	Cr	460
Ca	230	Mn	400
B	1250	Fe	420
Al	394	Pt	230
Si	625	NaCl	321
Pb	88	KCl	231
Ne	63	NaBr	224
Cu	315	KBr	173
Ag	215	KI	131

A haute température, lorsque la borne d'intégration T_D/T devient petite devant 1, on peut développer la fonction à intégrer au voisinage de $x=0$, et on retrouve[15] évidemment la *loi de Dulong et Petit :*

$$C_V \simeq 3Nk \quad \text{pour} \quad T \gg T_D \,. \tag{E.70}$$

Aux températures intermédiaires, c'est-à-dire pour T du même ordre que T_D, le modèle de Debye est seulement *approché.* Mais, comme les comportements à haute température et à basse température sont tous deux correctement reproduits, l'interpolation approchée entre ces deux domaines est forcément proche de la réalité. La figure 5 représente la variation avec T de la chaleur spécifique dans l'approximation de Debye. Même si une confrontation précise des formules de Debye avec les données expérimentales fait apparaître des différences sensibles aux températures intermédiaires, il est clair que l'essentiel des phénomènes est compris et que la température de Debye, déterminée à partir du comportement à basse température, donne correctement l'échelle des variations de la chaleur spécifique d'un solide cristallin avec la température dans tout le domaine accessible.

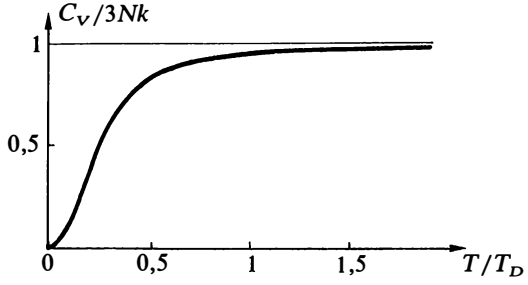


FIGURE 5
Variation avec la température de la capacité calorifique d'un solide dans l'approximation de Debye.

Pour préciser la signification de l'approximation de Debye, on peut comparer (fig. 4) la densité de modes normaux déterminée expérimentalement avec la densité approchée du modèle de Debye, parabolique pour $\omega < \omega_D$ et nulle au-delà de ω_D. On pourrait évidemment améliorer les résultats en utilisant la densité $\rho(\omega)$ expérimentale pour effectuer un calcul numérique à partir des formules générales du paragraphe 1.

IV. CRISTAUX A MAILLE POLYATOMIQUE

Nous avons jusqu'à présent considéré uniquement des cristaux dont la maille élémentaire est constituée d'un seul atome ou ion. Nous allons indiquer ici les modifications qu'apporte une maille polyatomique aux propriétés thermiques du cristal.

15. En effet,

$$\frac{x^4 \, e^x}{(e^x-1)^2} \sim x^2 \quad \text{pour } x \ll 1 \,. \tag{1}$$

L'intégrale donne donc bien

$$\int_0^{T_D/T} x^2 \, dx = \frac{1}{3}\left(\frac{T_D}{T}\right)^3 . \tag{2}$$

1. UN EXEMPLE TRÈS SIMPLE

Reprenons, comme au paragraphe II.1, un « cristal » à une dimension. Les atomes qui le constituent, toujours identiques, sont maintenant au nombre de $2N$ et disposés de la façon suivante (fig. 6) : la moitié d'entre eux a pour positions d'équilibre les points d'abscisse $q\ell (q = 1, 2, \ldots, N)$, où ℓ est le pas du réseau, et les N autres les points d'abscisse $q\ell + d$; la distance d étant supposée inférieure à $\ell/2$, les atomes sont groupés deux par deux, formant N « mailles » diatomiques. Nous ne prenons en compte que les interactions entre plus proches voisins; mais la force qui s'exerce à l'intérieur d'une maille (c'est-à-dire entre deux atomes dont la distance à l'équilibre est d) est différente de celle qui s'exerce entre des atomes appartenant à deux mailles voisines (dont la distance est ℓ-d à l'équilibre)[16].

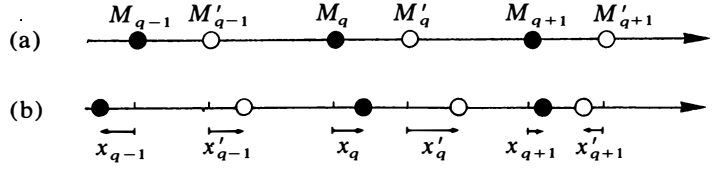


FIGURE 6
« Cristal » à une dimension constitué de « mailles » diatomiques : positions d'équilibre (a) et déplacements des atomes (b).

Si x_q et x'_q désignent les écarts à l'équilibre des deux atomes de la maille (q), les équations du mouvement s'écrivent (cf. figure 6)

$$m \frac{d^2 x_q}{dt^2} = -K_1(x_q - x'_q) - K_2(x_q - x'_{q-1}) , \tag{E.71.a}$$

$$m \frac{d^2 x'_q}{dt^2} = -K_1(x'_q - x_q) - K_2(x'_q - x_{q+1}) , \tag{E.71.b}$$

avec $K_1 \neq K_2$. Elles admettent encore des solutions en ondes progressives, caractérisées ici par

$$x_q(t) = a\, e^{i[k.q\ell - \omega t]} , \tag{E.72.a}$$

$$x'_q(t) = a'\, e^{i[k.q\ell - \omega t]} \tag{E.72.b}$$

(dans la notation complexe que nous utilisons, le rapport a'/a est *a priori* un nombre complexe, dont le module et l'argument donnent respectivement le rapport des amplitudes et le déphasage entre les oscillations des deux atomes d'une même maille). En effet, si l'on reporte les expressions (E.72) dans les équations du mouvement (E.71), on aboutit à un système de deux équations couplées pour a et a' :

$$[m\omega^2 - (K_1 + K_2)]a + [K_1 + K_2\, e^{-ik\ell}]a' = 0 , \tag{E.73.a}$$

$$[K_1 + K_2\, e^{ik\ell}]a + [m\omega^2 - (K_1 + K_2)]a' = 0 . \tag{E.73.b}$$

16. On obtient des résultats analogues en étudiant une chaîne constituée de deux sortes d'atomes disposés en alternance, de masses m_1 et m_2 différentes, interagissant par des forces de même constante de rappel K (réf. 1, p. 165).

Ce système homogène n'admet de solution non triviale que si le déterminant de ses coefficients est nul :

$$(m\omega^2-(K_1+K_2))^2=|K_1+K_2\,e^{-ik\ell}|^2\,. \tag{E.74}$$

Cette condition, considérée comme une équation en ω, admet deux solutions positives :

$$\omega=\left(\frac{1}{m}(K_1+K_2)\pm\frac{1}{m}\sqrt{K_1^2+K_2^2+2K_1K_2\cos k\ell}\right)^{\frac{1}{2}}\,. \tag{E.75}$$

Les conditions aux limites périodiques donnent encore ici N valeurs permises pour le vecteur d'onde k, les mêmes qu'au paragraphe II.1 (formules (E.27) et (E.30)). Mais, pour chacune de ces valeurs de k, il existe ici deux pulsations différentes : *la relation de dispersion comporte deux branches distinctes,* représentées sur la figure 7.

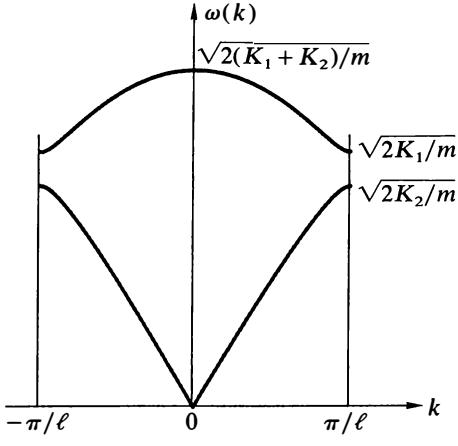


FIGURE 7
Les deux branches de la relation de dispersion pour un « cristal » à une dimension constitué de « mailles » diatomiques (formule (E.75); nous avons pris $K_1>K_2$).

La branche inférieure, dite *« branche acoustique »,* est très semblable à la courbe unique que nous avions obtenue au paragraphe II.1 (fig. 3) : en particulier, pour $|k|\ll\pi/\ell$, la pulsation ω est sur cette branche proportionnelle à $|k|$ comme en (E.31)[17]. La branche supérieure est appelée *« branche optique »,* car les modes correspondants peuvent, dans les cristaux ioniques notamment, être excités par des ondes électromagnétiques; sur cette branche, la pulsation ω ne s'annule jamais et la tangente en $k=0$ est horizontale, c'est-à-dire que ω est pratiquement indépendante de k pour k voisin de zéro[17].

17. Pour $|k|\ll\pi/\ell$,

$$\cos k\ell\simeq 1-\frac{1}{2}k^2\ell^2\,, \tag{1}$$

ce qui donne pour les deux pulsations

$$\omega_-\simeq\sqrt{\frac{K_1K_2}{2m(K_1+K_2)}}\,.\,|k|\,\ell\,, \tag{2.a}$$

$$\omega_+\simeq\sqrt{\frac{2(K_1+K_2)}{m}}+\mathcal{O}(k^2\ell^2)\,. \tag{2.b}$$

.../...

2. GÉNÉRALISATION

Dans un cristal à trois dimensions, une maille polyatomique a pour principal effet d'induire des « branches optiques » dans la relation de dispersion des modes normaux. Leur nombre doit évidemment, ajouté à celui des « branches acoustiques », redonner le nombre total de degrés de liberté du réseau cristallin.

Si nous notons toujours N le nombre de mailles du cristal, il existe N vecteurs d'onde $\vec{k}$ permis par les conditions aux limites périodiques, et non équivalents. Si ν est le nombre d'atomes ou d'ions constituant une maille, le nombre total de degrés de liberté du réseau vaut $3\nu N$. Il y a donc 3ν modes normaux distincts pour chaque vecteur d'onde $\vec{k}$ permis, c'est-à-dire 3ν *branches* pour la relation de dispersion : $\omega_\lambda(\vec{k})$, avec $\lambda = 1, 2, \ldots, 3\nu$. *Trois d'entre elles sont des branches acoustiques,* sur lesquelles ω est linéaire en $|\vec{k}|$ pour $|\vec{k}|$ petit; les $(3\nu - 3)$ autres branches sont du type optique : la pulsation ω y est partout non nulle et elle devient indépendante de $|\vec{k}|$ pour $|\vec{k}|$ suffisamment petit. On peut se représenter les 3 modes normaux acoustiques comme des vibrations déplaçant en bloc les diverses mailles; les $(3\nu - 3)$ modes optiques correspondent aux vibrations à l'intérieur de la maille (cf. note 17).

3. CALCUL APPROCHÉ DE LA CHALEUR SPÉCIFIQUE

On peut envisager deux façons différentes d'appliquer l'approximation de Debye à un cristal constitué de mailles polyatomiques.

La méthode la plus simple consiste à ignorer la différence entre les branches optiques et les portions non linéaires des branches acoustiques (pour les vecteurs d'onde $\vec{k}$ se rapprochant des bords de la zone de Brillouin). On se contente alors de remplacer N, dans les formules du paragraphe III.2, par le nombre total νN d'atomes du cristal. A température suffisamment élevée, les $3\nu N$ degrés de liberté sont classiques, de sorte que la capacité calorifique devient constante et égale à $3\nu Nk$, en accord avec la loi de Dulong et Petit. D'autre part, *le comportement de la chaleur spécifique à basse température est inchangé* par rapport au cas $\nu = 1$ (maille monoatomique). En effet, le facteur ν que l'on rajoute dans la formule (E.69) est compensé par un facteur égal dans T_D^3 au dénominateur (formules (E.62) et (E.61)). Ceci est compréhensible physiquement : à basse température, seules les trois branches acoustiques contribuent de façon significative à la capacité calorifique.

Dans la seconde méthode, on traite de façon différente (et plus appropriée) les *branches optiques.* On garde sans la modifier la contribution des branches acoustiques calculée au paragraphe III.2, avec le nombre N de mailles du cristal. On calcule celle des branches optiques dans le *modèle d'Einstein,* c'est-à-dire que l'on approxime la fonction $\omega(k)$ caractérisant une branche optique par une constante ω_{opt} indépendante de k, dont la valeur (intermédiaire entre le minimum et le maximum de ω sur la branche) est déterminée ensuite par ajustement des données expérimentales pour le cristal considéré. Chaque branche optique ajoute alors à la capacité calorifique un

---

Quant au rapport a'/a des amplitudes complexes de x'_q et x_q, il a pour expression générale

$$\frac{a'}{a} = \mp \frac{K_1 + K_2\, e^{ik\ell}}{|K_1 + K_2\, e^{ik\ell}|}, \tag{3}$$

et devient donc, pour $|k|$ petit,

$$\left(\frac{a'}{a}\right)_{\mp} \simeq \pm 1 \quad \text{pour } |k| \ll \pi/\ell. \tag{4}$$

La branche acoustique (signe supérieur) correspond alors à des vibrations en phase pour les deux atomes d'une même maille, qui oscille « en bloc » sans se déformer; pour la branche optique (signe inférieur), les deux atomes vibrent en opposition de phase, c'est-à-dire que la maille se dilate et se rétrécit périodiquement.

terme de la forme (E.11); comme il y a $(3\nu-3)$ branches optiques, la capacité calorifique totale est la somme de l'expression (E.66) et de

$$C_V^{(\mathrm{opt})} = (3\nu-3)Nk\left(\frac{\hbar\omega_{\mathrm{opt}}}{2kT}\right)^2 \frac{1}{\mathrm{sh}^2\,\hbar\omega_{\mathrm{opt}}/2kT} \tag{E.76}$$

(nous avons pris pour simplifier la même pulsation ω_{opt} pour les $(3\nu-3)$ branches optiques). A basse température, $C_V^{(\mathrm{opt})}$ tend vers zéro exponentiellement (cf. formule (E.15)) et la contribution acoustique est seule présente. A haute température, l'expression (E.76) tend vers $(3\nu-3)Nk$ (cf. formule (E.14)); en ajoutant $3Nk$ provenant des branches acoustiques (formule (E.70)), on retrouve à nouveau $3\nu Nk$, c'est-à-dire la loi de Dulong et Petit puisque νN est le nombre total d'atomes ou d'ions du cristal. C'est évidemment aux températures intermédiaires que cette seconde méthode s'avère meilleure que la précédente.

V. LES PHONONS

On interprète souvent les formules du paragraphe III, provenant de la quantification des modes normaux, en disant que les excitations collectives d'un réseau cristallin mettent en jeu des *«quasiparticules»*, nommées *«phonons»* puisqu'elles sont liées à la propagation du son dans le cristal (du moins celles qui correspondent aux grandes longueurs d'onde — cf. § II.2.c). Il ne s'agit pas de véritables particules, car elles n'existent que dans la mesure où le cristal leur fournit son support matériel; c'est pourquoi l'on parle de «quasiparticules».

Voici comment s'expriment, dans ce langage particulaire, les résultats du paragraphe III.1. Le cristal contient un *gaz parfait de phonons.* Un mode normal de vecteur d'onde $\vec{k}$, de polarisation (λ) et de pulsation $\omega_\lambda(\vec{k})$ est un *état individuel possible* pour les phonons; un phonon se trouvant dans cet état possède l'*impulsion* $\vec{p}=\hbar\vec{k}$, (λ) caractérise son *état de spin,* et son *énergie* est $\varepsilon_\lambda(\vec{p})=\hbar\omega_\lambda(\vec{k})$. Si le mode normal $(\vec{k},\lambda)$ est excité de telle sorte qu'il ait l'énergie $(n+1/2)\hbar\omega_\lambda(\vec{k})$, on dira que l'état individuel correspondant est occupé par n phonons. La *densité d'états individuels* $\rho_\phi(\varepsilon)$ est, à un facteur trivial près[18], égale à la densité de modes normaux :

$$\rho_\phi(\varepsilon)=\frac{1}{\hbar}\,\rho\left(\frac{\varepsilon}{\hbar}\right). \tag{E.77}$$

Les phonons étant indiscernables, un état microscopique du gaz qu'ils constituent est caractérisé par un ensemble de nombres d'occupation $\{n_{\vec{p}\lambda}\}$ pour les divers états individuels possibles ($n_{\vec{p}\lambda}$ est le nombre de phonons se trouvant dans l'état individuel $(\vec{p},\lambda)$); comme les phonons sont indépendants (gaz parfait), l'énergie correspondant à cet état est de la forme (cf. formules (E.46) et (E.54)) :

$$E_{\{n_{\vec{p}\lambda}\}}=-E_0+\sum_{\vec{p},\,\lambda} n_{\vec{p}\lambda}\varepsilon_\lambda(\vec{p}). \tag{E.78}$$

Mais *les états individuels possibles pour les phonons sont en nombre limité,* contrairement à ce qui se passe pour un gaz de particules ordinaires : il y a seulement N valeurs permises pour l'impulsion $\vec{p}$ (autant que de nœuds dans le cristal), et donc $3N$ états individuels. Ceci fixe pour l'énergie individuelle une borne supérieure ε_M donnée par (cf. (E.55))

$$\int_0^{\varepsilon_M}\rho_\phi(\varepsilon)\,\mathrm{d}\varepsilon=3N. \tag{E.79}$$

18. Le nombre $\rho_\phi(\varepsilon)\,\mathrm{d}\varepsilon$ d'états individuels dont l'énergie est comprise entre ε et $\varepsilon+\mathrm{d}\varepsilon$ est simplement égal à $\rho(\omega)\,\mathrm{d}\omega$. Le changement de variable $\varepsilon=\hbar\omega$ donne aussitôt (E.77).

Pour retrouver les formules du paragraphe III.1 et notamment l'expression (E.53) de l'énergie moyenne, il nous faut admettre que le *nombre d'occupation moyen,* à la température T, d'un état individuel d'énergie ε a pour expression

$$\bar{n}(\varepsilon, T) = \frac{1}{e^{\varepsilon/kT} - 1} . \tag{E.80}$$

La signification de ce résultat apparaîtra aux chapitres V (§ IV.B.2) et VI (§ III.A.3) : la formule (E.80) caractérise la *distribution de Bose pour un potentiel chimique nul;* les phonons sont, comme les photons, des *bosons dont le nombre total n'est pas conservé.* Nous montrerons alors comment s'expriment les diverses propriétés du système à partir du nombre d'occupation moyen (E.80), ce qui redonne les formules du paragraphe III.1.

Bien qu'il s'agisse ici d'une simple reformulation des résultats, la notion de phonon s'avère très utile car elle fournit une interprétation commode de la quantification des modes normaux de vibration du cristal. Elle facilite en outre la compréhension de phénomènes plus compliqués : on peut traiter les *corrections à l'approximation harmonique* comme résultant d'*interactions entre les phonons* (les effets d'anharmonicité couplent les modes normaux) et décrire les transferts d'énergie entre le gaz d'électrons d'un métal et le réseau d'ions comme provenant d'*interactions entre les électrons et les phonons* (celles-ci sont à l'origine de la résistance électrique du métal). Il est même possible de mettre expérimentalement en évidence les phonons, par exemple par diffusion inélastique de neutrons lents : lorsqu'un neutron incident excite un mode normal de vecteur d'onde $\vec{k}$ et de pulsation ω (c'est-à-dire crée un phonon d'impulsion $\vec{p} = \hbar\vec{k}$ et d'énergie $\varepsilon = \hbar\omega$), son impulsion diminue de $\hbar\vec{k}$ et son énergie de $\hbar\omega$.

Dans l'approximation de Debye (§ III.2), l'énergie ε d'un phonon est simplement proportionnelle à son impulsion (cf. formule (E.56)). Les phonons apparaissent alors comme très semblables aux photons (VI, § III.A.2), la vitesse du son dans le cristal remplaçant la vitesse de la lumière. Cependant, cette proportionnalité entre énergie et impulsion, exacte et toujours valable pour les photons (qui sont de véritables particules de masse nulle), est seulement approchée pour les phonons : sa validité est restreinte aux « phonons acoustiques » de faible impulsion (c'est-à-dire aux quasiparticules associées aux modes normaux acoustiques de petit vecteur d'onde). Nous retrouverons les phonons, dans un contexte différent, au complément VI.D (§ II.2).

Références

1. D.L. Goodstein, *States of Matter,* Prentice-Hall, Englewood Cliffs (1975), chapitre 3.
2. N.W. Ashcroft et N.D. Mermin, *Solid State Physics,* Holt-Saunders, Philadelphia (1981), chapitres 22 et 23.

COMPLÉMENT III.F

Les solutions diluées

I. Notions préliminaires
  1. Enthalpie libre et potentiel chimique d'un fluide pur
  2. Énergie d'interaction effective entre molécules du soluté dans le solvant
    a. Fonction de partition et intégrale de configuration
    b. Analyse des interactions. Notion d'interaction effective

II. Enthalpie libre et potentiels chimiques d'une solution diluée
  1. Fonction de partition de la solution
    a. Évaluation de l'énergie d'interaction effective du soluté
    b. Expression approchée de la fonction de partition et de l'énergie libre
  2. Enthalpie libre approchée
  3. Expression approchée des potentiels chimiques du solvant et du soluté

III. Quelques applications
  1. Pression osmotique d'une solution
    a. Description du phénomène
    b. Loi de van't Hoff pour les solutions diluées
  2. Équilibre entre deux phases diluées
    a. Conditions d'équilibre
    b. Lois de Raoult
    c. Concentration des deux phases en équilibre

Nous étudions ici les fluides (liquides ou gaz) *homogènes,* constitués de *deux sortes de molécules :* N molécules d'un premier corps pur appelé le *solvant* et N' molécules d'un autre corps pur, le *soluté* ($N > N' \gg 1$).

Après une étude préliminaire du solvant pur, nous introduirons la notion d'*interaction effective* entre les molécules du soluté compte tenu de la présence du solvant. Nous utiliserons ensuite cette notion pour analyser le cas simple des *solutions diluées,* et terminerons par quelques *applications* des résultats ainsi obtenus.

I. NOTIONS PRÉLIMINAIRES

1. ENTHALPIE LIBRE ET POTENTIEL CHIMIQUE D'UN FLUIDE PUR

Revenons tout d'abord au cas d'un *fluide pur homogène,* par exemple le solvant en l'absence de soluté.

Dans la description microcanonique, l'état macroscopique d'un fluide pur homogène est caractérisé par les trois variables E (énergie totale), V (volume) et N (nombre de particules). Ses propriétés macroscopiques sont complètement déterminées si l'on réussit à calculer son entropie S comme fonction de ces trois variables :

$$S = S(E, V, N)\,. \tag{F.1}$$

Il s'agit en principe de l'entropie microcanonique, mais nous nous plaçons d'emblée dans la *limite thermodynamique* où cette précision n'est pas nécessaire (III, § III.C). La définition de la température T, de la pression p et du potentiel chimique μ (II, § III.A.1 et § III.B.1) permet d'écrire la différentielle de la fonction S :

$$\mathrm{d}S = \frac{1}{T}\,\mathrm{d}E + \frac{p}{T}\,\mathrm{d}V - \frac{\mu}{T}\,\mathrm{d}N\,, \tag{F.2}$$

relation que nous utiliserons plutôt sous la forme

$$\mathrm{d}E = T\,\mathrm{d}S - p\,\mathrm{d}V + \mu\,\mathrm{d}N\,. \tag{F.3}$$

Lorsqu'on souhaite remplacer la variable énergie E par la variable

température T, on utilise la description canonique, où les propriétés du système sont déduites de son énergie libre

$$F = F(T, V, N) . \tag{F.4}$$

On sait que les deux formalismes sont équivalents à la limite thermodynamique. La relation

$$F = E - TS \tag{F.5}$$

permet de passer de l'un à l'autre, c'est-à-dire des variables $\{E, V, N\}$ aux variables $\{T, V, N\}$. En différentiant cette relation et en remplaçant dE par l'expression (F.3), on trouve aussitôt

$$dF = -S\, dT - p\, dV + \mu\, dN . \tag{F.6}$$

Il sera commode dans la suite d'utiliser les variables $\{T, p, N\}$. Il suffit, pour les introduire de façon naturelle, de répéter le procédé qui, à partir de (F.5), a conduit à (F.6) : on définit l'*enthalpie libre* du système

$$G \widehat{=} E - TS + pV = F + pV \tag{F.7}$$

(cf. complément III.D). Si l'on calcule la différentielle de cette nouvelle fonction et qu'on y reporte l'expression (F.3) de dE, il vient

$$dG = -S\, dT + V\, dp + \mu\, dN . \tag{F.8}$$

Il est donc naturel de considérer G comme fonction des variables T, p et N qui nous intéressent :

$$G = G(T, p, N) , \tag{F.9}$$

avec
$$\left(\frac{\partial G}{\partial T}\right)_{p,\,N} = -S , \tag{F.10.a}$$

$$\left(\frac{\partial G}{\partial p}\right)_{T,\,N} = V , \tag{F.10.b}$$

$$\left(\frac{\partial G}{\partial N}\right)_{T,\,p} = \mu \tag{F.10.c}$$

(les indices précisent les variables maintenues constantes dans chaque dérivation).

Rappelons (complément III.D, § I.3.c) que, pour un fluide pur homogène,

$$G(T, p, N) = N\mu \tag{F.11}$$

et que le potentiel chimique μ, grandeur intensive, ne dépend que des variables intensives T et p :

$$\mu = \mu(T, p) , \tag{F.12}$$

avec, d'après les relations (F.10),

$$\left(\frac{\partial \mu}{\partial T}\right)_{p} = -\frac{S}{N} , \tag{F.13.a}$$

$$\left(\frac{\partial \mu}{\partial p}\right)_{T} = \frac{V}{N} . \tag{F.13.b}$$

2. ÉNERGIE D'INTERACTION EFFECTIVE ENTRE MOLÉCULES DU SOLUTÉ DANS LE SOLVANT

Entreprenons maintenant l'étude du mélange homogène de N molécules de solvant et N' molécules de soluté. L'état de ce système est caractérisé par quatre variables, par exemple T, p, N et N'; il est souvent plus commode d'utiliser T, p, N et la *concentration* c de la solution, définie par

$$c \widehat{=} \frac{N'}{N + N'} . \tag{F.14}$$

Noter que c est intensive : elle reste inchangée si l'on multiplie la taille du système par un facteur quelconque λ, car N et N' sont alors multipliés tous deux par ce même facteur.

a. Fonction de partition et intégrale de configuration

Raisonnons dans le cadre de la *mécanique classique,* qui décrit correctement les fluides ordinaires. Nous notons $\vec{r}_i$ et $\vec{p}_i$ les positions et impulsions des particules du solvant ($i = 1, 2, \ldots, N$), $\vec{r}'_j$ et $\vec{p}'_j$ celles des particules du soluté ($j = 1, 2, \ldots, N'$). Nous ignorons pour simplifier la structure interne de ces particules : il serait possible d'en tenir compte, au moins formellement, sans changer les résultats de l'analyse. En outre, pour rendre plus compacte l'écriture des formules, nous symboliserons par $\vec{r}$ l'ensemble des variables $\vec{r}_i$, par $\vec{p}$ l'ensemble des $\vec{p}_i$, et de même par $\vec{r}'$ et $\vec{p}'$ l'ensemble des $\vec{r}'_j$ et $\vec{p}'_j$.

La fonction de partition du système a alors pour expression (III, § V.A.2)[1]

$$Z(T, V, N, N') = \frac{1}{N!}\frac{1}{N'!}\frac{1}{(2\pi\hbar)^{3N+3N'}}\int\left[\prod_{i=1}^{N} \mathrm{d}^3 r_i\, \mathrm{d}^3 p_i\right] \int\left[\prod_{j=1}^{N'} \mathrm{d}^3 r'_j\, \mathrm{d}^3 p'_j\right] \exp\left[-\mathcal{H}(\vec{r}, \vec{p}\,; \vec{r}', \vec{p}\,')/kT\right]. \qquad \text{(F.15)}$$

De façon générale, la fonction de Hamilton $\mathcal{H}$ qui figure dans cette expression est de la forme

$$\mathcal{H}(\vec{r}, \vec{p}\,; \vec{r}', \vec{p}\,') = \sum_{i=1}^{N} \frac{\vec{p}_i^{\,2}}{2m} + \sum_{j=1}^{N'} \frac{\vec{p}_j^{\,\prime 2}}{2m'} + U(\vec{r}, \vec{r}'), \qquad \text{(F.16)}$$

où U désigne l'énergie potentielle totale du mélange (m est la masse d'une molécule de solvant, m' celle d'une molécule de soluté). On peut comme toujours effectuer sans peine les intégrations sur les impulsions (voir chapitre III, note 92), ce qui donne

$$Z(T, V, N, N') = \left(\frac{mkT}{2\pi\hbar^2}\right)^{3N/2}\left(\frac{m'kT}{2\pi\hbar^2}\right)^{3N'/2} \times \frac{1}{N!}\cdot\frac{1}{N'!}\, Y(T, V, N, N'), \qquad \text{(F.17)}$$

où
$$Y(T, V, N, N') \widehat{=} \int \mathrm{d}^3 r \int \mathrm{d}^3 r'\, \mathrm{e}^{-U(\vec{r}, \vec{r}')/kT}. \qquad \text{(F.18)}$$

Le problème est ainsi ramené au calcul de l'*« intégrale de configuration »* Y du système.

b. Analyse des interactions. Notion d'interaction effective

L'énergie potentielle U du système des $N + N'$ molécules peut être décomposée en la somme de trois termes :

$$U(\vec{r}, \vec{r}') = U_0(\vec{r}) + U'(\vec{r}, \vec{r}') + U''(\vec{r}'). \qquad \text{(F.19)}$$

Le premier est l'*énergie potentielle des N molécules du solvant,* qui dépend seulement de leurs positions $\vec{r}_i (i = 1, 2, \ldots, N)$; de même, le troisième terme est l'*énergie potentielle des N' molécules de soluté,* fonction de leurs seules positions $\vec{r}'_j (j = 1, 2, \ldots, N')$. Quant au deuxième terme, c'est l'énergie associée aux *interactions dans lesquelles interviennent*

1. L'homogénéité du mélange se traduit concrètement par le fait que les deux types de particules se déplacent dans le *même volume V,* domaine d'intégration pour chacun des $\vec{r}_i$ et des $\vec{r}'_j$.

à la fois des molécules du solvant et des molécules du soluté[2]. L'indice zéro que nous affectons au premier terme rappelle qu'il reste seul présent lorsque le solvant est pur.

Si l'énergie potentielle «croisée» U' était nulle, l'intégrale de configuration (F.18) (et donc aussi la fonction de partition (F.17)) se factoriserait : elle serait égale au produit de l'intégrale de configuration Y_0 associée au solvant pur :

$$Y_0(T, V, N) \widehat{=} \int d^3r\, e^{-U_0(\vec{r})/kT} \tag{F.20}$$

par celle Y'_0 correspondant au soluté pur. L'interaction entre solvant et soluté empêche évidemment cette factorisation. Cependant, si l'on définit l'*énergie effective d'interaction* U'_e *du soluté dans le solvant* par la formule

$$e^{-U'_e(\vec{r}')/kT} \widehat{=} \frac{1}{Y_0} \int d^3r\, e^{-[U_0(\vec{r})+U'(\vec{r},\vec{r}')+U''(\vec{r}')]/kT}\,, \tag{F.21}$$

on peut écrire

$$Y = Y_0 . Y'_e\,, \tag{F.22}$$

où

$$Y'_e(T, V, N, N') = \int d^3r'\, e^{-U'_e(\vec{r}')/kT} \tag{F.23}$$

est l'*intégrale de configuration du soluté lorsqu'il est mélangé au solvant*. En définitive, la fonction de partition Z du système se met ainsi sous la forme

$$Z(T, V, N, N') = Z_0(T, V, N) \times Z'_e(T, V, N, N')\,, \tag{F.24}$$

où Z_0 est la fonction de partition du solvant pur et

$$Z'_e(T, V, N, N') = \frac{1}{N'!}\left(\frac{m'kT}{2\pi\hbar^2}\right)^{3N'/2} Y'_e(T, V, N, N') \tag{F.25}$$

celle du soluté mélangé au solvant.

L'énergie d'interaction effective U'_e entre les molécules du soluté s'obtient d'après la formule (F.21) en fixant leurs positions $\vec{r}'_1, \vec{r}'_2, \ldots, \vec{r}'_{N'}$ et en sommant sur toutes les positions possibles $\vec{r}_1, \vec{r}_2, \ldots, \vec{r}_N$ des molécules du solvant. Son interprétation physique découle des formules (F.23) et (F.25) : si l'on s'intéresse seulement au soluté, sa fonction de partition Z'_e se calcule à partir de l'énergie d'interaction effective $U'_e(\vec{r}')$; ainsi, *le solvant est alors considéré comme un milieu continu dans lequel sont plongées les molécules du soluté et qui modifie leurs interactions.* Évidemment, si l'interaction $U'(\vec{r}, \vec{r}')$ entre soluté et solvant est négligeable, l'énergie d'interaction U'_e se réduit à l'énergie d'interaction $U''(\vec{r}')$ entre les molécules du soluté.

L'énergie d'interaction effective est difficile à évaluer dans le cas général. C'est

---

2. Il n'est pas nécessaire ici de supposer que les interactions à trois particules ou plus sont négligeables. Mais si elles le sont, les termes U_0, U' et U'' s'écrivent

$$U_0 = \frac{1}{2} \sum_{\substack{i,\, i'=1 \\ i \neq i'}}^{N} u_0(\vec{r}_i - \vec{r}_{i'})\,,$$

$$U' = \sum_{i=1}^{N} \sum_{j=1}^{N'} u'(\vec{r}_i - \vec{r}'_j)\,,$$

$$U'' = \frac{1}{2} \sum_{\substack{j,\, j'=1 \\ j \neq j'}}^{N} u''(\vec{r}'_j - \vec{r}'_{j'})\,,$$

où u_0, u' et u'' sont les énergies potentielles d'interaction entre deux molécules du solvant, une molécule de solvant et une de soluté, et deux molécules du soluté, respectivement.

cependant une notion très utile[3], comme le montre par exemple l'étude théorique des solutions diluées, que nous allons développer au paragraphe II.

II. ENTHALPIE LIBRE ET POTENTIELS CHIMIQUES D'UNE SOLUTION DILUÉE

Nous considérons maintenant une solution *diluée,* c'est-à-dire un mélange homogène dans lequel

$$N \gg N' \gg 1 , \tag{F.26}$$

ce qui implique que la concentration c définie en (F.14) est très faible :

$$c \simeq \frac{N'}{N} \ll 1 . \tag{F.27}$$

Les propriétés d'une telle solution vont être proches de celles du solvant pur; nous voulons évaluer les différences à l'ordre dominant en c : *nous négligerons les termes en* c^2 et d'ordre supérieur.

1. FONCTION DE PARTITION DE LA SOLUTION

Dans une solution très diluée, une molécule de soluté est constamment entourée de solvant et ne rencontre pratiquement jamais dans son mouvement une autre molécule de soluté. Nous pouvons donc en première approximation *négliger les interactions entre les N′ molécules de soluté,* qui seront ainsi indépendantes[4] : l'énergie $U''(\vec{r}\,')$ associée à ces interactions est d'ordre c^2 par rapport à $U_0(\vec{r}\,)$[5].

a. Évaluation de l'énergie d'interaction effective du soluté

Écrivons, comme au paragraphe I.2.b, la fonction de partition de la solution sous la forme

$$Z = Z_0 . Z'_e . \tag{F.28}$$

Z_0 est la fonction de partition du solvant pur :

$$Z_0 = \frac{1}{N!} \left(\frac{mkT}{2\pi\hbar^2} \right)^{3N/2} Y_0(T, V, N) , \tag{F.29}$$

avec $Y_0(T, V, N) = \int d^3r \, e^{-U_0(\vec{r})/kT}$.

Z'_e est obtenue à partir de l'énergie d'interaction effective $U'_e(\vec{r}\,')$, donnée ici par

$$e^{-U'_e(\vec{r}\,')/kT} = \frac{1}{Y_0} \int d^3r \, e^{-[U_0(\vec{r})+U'(\vec{r},\,\vec{r}\,')]/kT} . \tag{F.31}$$

---

3. Ce point de vue est utilisé dans d'autres situations. Par exemple, dans les solutions d'électrolytes (complément III.H), l'énergie d'interaction effective entre deux ions s'obtient en divisant l'énergie d'interaction de leurs deux charges dans le vide par la constante diélectrique ε_r du solvant.

4. Nous excluons le cas où le soluté serait un électrolyte : les interactions entre ions chargés ont une portée trop longue pour que l'hypothèse d'indépendance des particules de soluté entre elles puisse être vérifiée. Ce cas spécifique est étudié dans le complément III.H.

5. Dans un fluide comportant N particules, l'énergie potentielle totale est, très grossièrement, proportionnelle au nombre $N(N-1)/2 \simeq N^2/2$ de paires de particules (cf. complément I.A, formule (A.53)). L'ordre de grandeur du rapport U''/U_0 est donc $N'^2/N^2 \simeq c^2$.

Pour évaluer l'interaction effective U'_e, nous allons utiliser un raisonnement du même type que celui que nous avons introduit au complément II.A (§ II). Il repose sur l'hypothèse fondamentale que *les interactions intermoléculaires sont à courte portée* (cf. note 4).

Dans la formule (F.31), *les positions* $\vec{r}'_1, \vec{r}'_2, \ldots, \vec{r}'_{N'}$ *des molécules du soluté sont fixées.* Comme la solution est très diluée, les distances entre ces molécules sont très grandes devant la portée des forces intermoléculaires (c'est ce qui nous a déjà permis de négliger les interactions directes entre molécules du soluté). Divisons alors le système global, de volume V, en N' *sous-systèmes* $\delta\mathcal{S}_j$, de volume δV_j, *dont chacun contient seulement une molécule de soluté.* Ces sous-systèmes sont certes petits à l'échelle macroscopique, mais ils contiennent cependant, si N/N' est très grand, un nombre moyen δN_j de molécules du solvant suffisamment important pour que les fluctuations puissent y être négligées; en outre, leurs dimensions sont suffisamment grandes par rapport à la portée des forces intermoléculaires pour que leur couplage soit faible. A une bonne approximation, les sous-systèmes $\delta\mathcal{S}_j$ sont donc *indépendants,* de sorte que *la fonction de partition du système est le produit des fonctions de partition des sous-systèmes* $\delta\mathcal{S}_j$ (cf. III, § II.C)[6]; les intégrales de configuration se factorisent évidemment de façon analogue.

Cette méthode permet de simplifier considérablement l'expression de l'énergie d'interaction effective. Dans la formule (F.31), le numérateur est le produit de N' intégrales portant chacune sur un sous-système $\delta\mathcal{S}_j$ de volume δV_j centré en $\vec{r}'_j$, c'est-à-dire sur les δN_j molécules de solvant que contient ce sous-système, autour d'une molécule unique de soluté située au point $\vec{r}'_j$. Le dénominateur Y_0 de la formule (F.31) peut être calculé à partir du même découpage du volume total du système; il est alors lui aussi le produit de N' intégrales correspondant chacune à un sous-système $\delta\mathcal{S}_j$. On aboutit ainsi à une expression de la forme

$$e^{-U'_e(\vec{r}')/kT} \simeq e^{-u'_e(\vec{r}'_1)/kT} . e^{-u'_e(\vec{r}'_2)/kT} \ldots . e^{-u'_e(\vec{r}'_{N'})/kT} , \tag{F.32}$$

où

$$e^{-u'_e(\vec{r}'_j)/kT} \widehat{=} \frac{\displaystyle\int_{\delta\mathcal{S}_j} d^3r\, e^{-[U_0(\vec{r})+U'(\vec{r},\vec{r}'_j)]/kT}}{\displaystyle\int_{\delta\mathcal{S}_j} d^3r\, e^{-U_0(\vec{r})/kT}} . \tag{F.33}$$

On retrouve évidemment que les molécules du soluté sont indépendantes les unes des autres (formule (F.32)). Chacune d'elles est soumise, de la part des molécules de solvant qui l'entourent, à une énergie potentielle d'interaction u'_e donnée par (F.33). Pour expliciter l'expression de cette énergie potentielle, il faudrait préciser le modèle microscopique, c'est-à-dire se donner les potentiels intermoléculaires. Nous nous contenterons ici de montrer que l'énergie potentielle effective u'_e est d'une part *indépendante de la position* $\vec{r}'$ de la molécule de soluté considérée dans le récipient de volume V, et d'autre part *fonction de la densité* N/V *des molécules du solvant* (ainsi bien sûr que de la température). En effet, si l'on néglige, comme nous l'avons toujours fait jusqu'ici, les effets de surface, la solution étudiée est *homogène.* Donc, sauf dans une zone d'extension négligeable (à l'échelle macroscopique) le long des parois du récipient, une molécule de soluté subit partout l'influence du même environnement : le potentiel effectif u'_e est uniforme à l'intérieur du volume V. D'autre part, la valeur de ce potentiel uniforme dépend évidemment du nombre de molécules de solvant qui,

6. L'approximation est évidemment d'autant meilleure que les sous-systèmes $\delta\mathcal{S}_j$ sont plus grands. Elle est parfaitement justifiée si chacun d'eux a une taille suffisante pour que la limite thermodynamique y soit atteinte. Concrètement, on peut considérer que le calcul sera valable si $\delta N_j \sim N/N'$ est supérieur à 10^4, ou au moins à 10^3 si l'on accepte des incertitudes de quelques pour cent (ces incertitudes sont de l'ordre de $1/\sqrt{\delta N_j}$). Si le solvant est de l'eau, une concentration $c \simeq N'/N$ de 10^{-3} signifie environ 5/100 de mole de soluté par litre.

en moyenne, se trouvent dans la « sphère d'influence » de la molécule de soluté considérée (c'est-à-dire dans la sphère centrée sur cette molécule et de rayon égal à la portée maximum des forces intermoléculaires); ce nombre moyen est directement proportionnel à N/V puisque la sphère d'influence a un volume fixé. Nous changerons donc de notation, et écrirons

$$u'_e \equiv \varphi(N/V)\,, \tag{F.34}$$

sans chercher à déterminer la forme de la fonction φ.

b. Expression approchée de la fonction de partition et de l'énergie libre

En définitive, la fonction de partition Z de la solution s'écrit à une bonne approximation

$$Z \simeq Z_0 \cdot \frac{1}{N'!}\, z^{N'}\,, \tag{F.35}$$

où $$Z_0 = Z_0(T,\ V,\ N) \tag{F.36}$$

est la *fonction de partition du solvant pur* et z la *fonction de partition d'une molécule de soluté plongée dans le solvant :*

$$z = \left(\frac{m'kT}{2\pi\hbar^2}\right)^{3/2} \int_V \mathrm{d}^3 r'\, \mathrm{e}^{-u'_e(\vec{r}')/kT}\,. \tag{F.37}$$

Comme l'énergie d'interaction effective u'_e est indépendante de $\vec{r}'$, l'intégrale se calcule aussitôt; la dépendance de u'_e par rapport à N/V (formule (F.34)) se répercute dans z, qui prend la forme

$$z\left(T,\ V,\ \frac{N}{V}\right) = \left(\frac{m'kT}{2\pi\hbar^2}\right)^{3/2} V \exp\left[-\varphi(N/V)/kT\right]. \tag{F.38}$$

De la fonction de partition on déduit l'énergie libre F de la solution : les formules (F.35) et (F.38) donnent facilement

$$F(T,\ V,\ N,\ N') \simeq F_0(T,\ V,\ N) + N'\left[\varphi\left(\frac{N}{V}\right) - \frac{3}{2}\, kT \ln \frac{m'kT}{2\pi\hbar^2} + kT\left(\ln \frac{N'}{V} - 1\right)\right], \tag{F.39}$$

où F_0 est l'énergie libre du solvant pur. Les variables macroscopiques sont au nombre de quatre : dans l'énergie libre F, ce sont T, V, N et N'. Le calcul approché que nous avons effectué, valable pour N' très petit devant N, a eu pour résultat d'*expliciter la dépendance par rapport à* N'[7].

2. ENTHALPIE LIBRE APPROCHÉE

Nous voulons, comme dans le cas du solvant pur (§ I.1), remplacer la variable V par sa variable conjuguée, la pression p. La méthode est la même qu'au paragraphe I.1 : nous introduisons l'enthalpie libre G, et nous allons l'exprimer comme fonction de ses variables « naturelles » T, p, N et N' :

$$G \widehat{=} F + pV = G(T,\ p,\ N,\ N')\,. \tag{F.40}$$

7. Comme les N' molécules du soluté sont indépendantes, leur contribution à la fonction de partition (F.35) et à l'énergie libre (F.39) est semblable à celle d'un gaz parfait (cf. III, § IV.B). L'existence de l'énergie d'interaction u'_e avec le solvant apporte cependant une différence cruciale : dans l'énergie libre apparaît un terme faisant intervenir à la fois N' et la densité N/V des molécules du solvant.

L'expression (F.39) de l'énergie libre F donne la pression p sous la forme

$$p = -\frac{\partial F}{\partial V} = -\frac{\partial F_0}{\partial V} + \frac{NN'}{V^2}\,\varphi'\left(\frac{N}{V}\right) + \frac{N'kT}{V}\,, \tag{F.41}$$

où φ' désigne la dérivée de la fonction φ par rapport à la variable N/V. Pour effectuer le changement de variables qui, dans (F.40), remplacera V par p, nous devons inverser la relation (F.41), c'est-à-dire en déduire V comme fonction de p (et T, N, N'); mais, pour rester compatible avec les approximations précédentes, cette opération doit être menée en négligeant les termes d'ordre supérieur ou égal à 2 en N'/N.

L'énergie libre F et la pression p comportent chacune un terme d'ordre zéro et un terme correctif petit; l'enthalpie libre obtenue par le changement de variable $V \to p$ aura la même structure, le terme correctif étant du même ordre que les précédents. On peut montrer de façon générale[8] que le terme correctif de G s'obtient très simplement à partir de celui de F : il suffit de remplacer dans ce dernier V par la fonction $V_0(T, p, \ldots)$ obtenue en inversant la formule (F.41) *prise à l'ordre zéro :*

$$p = -\frac{\partial F_0}{\partial V}\,. \tag{F.42}$$

Ce procédé donne ici une fonction V_0 proportionnelle à N :

$$V_0(T, p, N) = N v(T, p)\,, \tag{F.43}$$

---

8. Supposons que l'énergie libre d'un système soit de la forme

$$F(T, V) \simeq F_0(T, V) + \alpha F_1(T, V)\,, \tag{1}$$

où α est un paramètre petit que l'on traite au premier ordre (nous sous-entendons les autres variables dont peut dépendre F). La pression p est alors donnée par

$$p \simeq -\frac{\partial F_0}{\partial V} - \alpha\,\frac{\partial F_1}{\partial V}\,; \tag{2}$$

si l'on s'en tient aux variables canoniques $\{T, V, \ldots\}$, elle est la somme d'un terme d'ordre zéro et d'une correction du premier ordre en α :

$$p(T, V) \simeq p_0(T, V) + \alpha p_1(T, V)\,. \tag{3}$$

Mais si l'on passe aux variables $\{T, p, \ldots\}$, adaptées à l'enthalpie libre

$$G(T, p) \widehat{=} F + pV$$

$$\simeq F_0 + \alpha F_1 - V\frac{\partial F_0}{\partial V} - \alpha V\frac{\partial F_1}{\partial V}\,, \tag{4}$$

c'est maintenant V, fonction de T, p, ..., qui est de la forme

$$V(T, p) \simeq V_0(T, p) + \alpha V_1(T, p)\,. \tag{5}$$

Cette fonction est déterminée par inversion de la relation (2) entre p, V et T. A l'ordre zéro en α,

$$p = -\frac{\partial F_0}{\partial V}(T, V) \implies V = V_0(T, p)\,. \tag{6}$$

Au premier ordre en α, $V_1(T, p)$ doit être tel que

$$p = -\frac{\partial F_0}{\partial V}(T, V_0 + \alpha V_1) - \alpha\,\frac{\partial F_1}{\partial V}(T, V_0)\,, \tag{7}$$

c'est-à-dire, si l'on développe au premier ordre et que l'on tient compte de (6),

$$V_1\,\frac{\partial^2 F_0}{\partial V^2}(T, V_0) + \frac{\partial F_1}{\partial V}(T, V_0) = 0\,. \tag{8}$$

.../...

car le volume est une grandeur extensive comme N, alors que T et p sont intensives (le raisonnement est le même que pour G; cf. formule (F.11)). Comme l'enthalpie libre G_0 du solvant pur a elle-même la forme (F.11)-(F.12), l'expression de G s'écrit finalement[9]

$$G(T, p, N, N') \simeq N\mu_0(T, p) + N'\left[\psi(T, p) + kT\left(\ln\frac{N'}{N} - 1\right)\right], \tag{F.44.a}$$

où $$\psi(T, p) \widehat{=} \varphi\left[\frac{1}{v(T, p)}\right] - \frac{3}{2}\,kT\,\ln\frac{m'kT}{2\pi\hbar^2} - kT\,\ln v(T, p) \tag{F.44.b}$$

et μ_0 est le potentiel chimique du solvant pur. La fonction $\psi(T, p)$ peut être calculée en principe si l'on connaît les propriétés du solvant pur et la fonction $\varphi(N/V)$ qui caractérise l'interaction effective d'une molécule de soluté avec le solvant environnant; il faudrait pour cela développer un modèle microscopique explicite pour cette interaction; nous n'essaierons pas ici de le faire et garderons $\psi(T, p)$ sans préciser sa forme.

Calculons alors l'enthalpie libre G à partir de (4), en nous limitant toujours au premier ordre :

$$G(T, p) = F_0(T, V_0 + \alpha V_1) - (V_0 + \alpha V_1)\frac{\partial F_0}{\partial V}(T, V_0 + \alpha V_1) + \alpha F_1(T, V_0) - \alpha V_0 \frac{\partial F_1}{\partial V}(T, V_0), \tag{9}$$

où V_0 et V_1 sont les fonctions de T et p définies par (6) et (8). Le terme d'ordre zéro est évidemment

$$G_0(T, p) = F_0(T, V_0) - V_0\frac{\partial F_0}{\partial V}(T, V_0). \tag{10}$$

Le terme d'ordre un s'écrit

$$G_1(T, p) = V_1\frac{\partial F_0}{\partial V}(T, V_0) - V_0V_1\frac{\partial^2 F_0}{\partial V^2}(T, V_0) - V_1\frac{\partial F_0}{\partial V}(T, V_0) + F_1(T, V_0) - V_0\frac{\partial F_1}{\partial V}(T, V_0), \tag{11}$$

c'est-à-dire tout simplement, compte tenu de (8),

$$G_1(T, p) = F_1[T, V_0(T, p)]. \tag{12}$$

9. On aurait pu envisager, sachant que N' est petit devant N, d'écrire directement pour G une formule approchée au premier ordre en N'/N :

$$G(T, p, N, N') \simeq N\left[\mu_0(T, p) + \frac{N'}{N}\,g_1(T, p)\right]. \tag{1}$$

Nous voyons sur le résultat (F.44) qu'une telle formule serait *erronée,* puisqu'elle ne comporterait pas le terme en $\ln N'/N$. Autrement dit, le raisonnement que nous avons mené montre que G n'est pas analytique en N' au voisinage de $N' = 0$: ses dérivées par rapport à N' sont infinies en ce point (cf. formule (F.47.b)). Notons que ceci découle de l'*identité des molécules du soluté,* puisque le terme en $N' \ln N'$ provient du $N'!$ qui figure dans Z (formule (F.35)).

Malgré ce défaut d'analyticité en $N' = 0$, G tend bien vers G_0 quand N' tend vers zéro. En outre, étant donné son origine, le terme en $N' \ln N'$ n'affecte que le terme du premier ordre en N'/N; si l'on pousse à l'ordre suivant le développement de G par rapport au paramètre N'/N, on obtiendra une expression de la forme

$$G \simeq N\left\{\mu_0(T, p) + \frac{N'}{N}\left[\psi(T, p) + kT\left(\ln\frac{N'}{N} - 1\right)\right] + \left(\frac{N'}{N}\right)^2 g_2(T, p)\right\}. \tag{2}$$

3. EXPRESSION APPROCHÉE DES POTENTIELS CHIMIQUES DU SOLVANT ET DU SOLUTÉ

Un calcul simple, analogue à celui du paragraphe I.1, donne la différentielle de l'enthalpie libre de la solution :

$$dG = -S\,dT + V\,dp + \mu\,dN + \mu'\,dN' \,. \tag{F.45}$$

Il y a ici deux potentiels chimiques, puisqu'il y a deux espèces de particules : μ' est le potentiel chimique du soluté dans le solvant, et μ celui du solvant *en présence du soluté*. De façon générale, on s'attend à ce que μ et μ', intensifs, soient l'un et l'autre fonctions de T, p, et du rapport intensif N'/N, c'est-à-dire de la concentration c (formule (F.27))[10]. Ces deux potentiels chimiques peuvent, d'après (F.45), se calculer facilement à partir de G[11] :

$$\mu(T, p, c) = \left(\frac{\partial G}{\partial N}\right)_{T, p, N'} , \tag{F.46.a}$$

$$\mu'(T, p, c) = \left(\frac{\partial G}{\partial N'}\right)_{T, p, N} . \tag{F.46.b}$$

De l'expression approchée (F.44) de G on déduit donc celle des potentiels chimiques de la solution diluée :

$$\mu(T, p, c) \simeq \mu_0(T, p) - ckT \,, \tag{F.47.a}$$

$$\mu'(T, p, c) \simeq \psi(T, p) + kT \ln c \,. \tag{F.47.b}$$

Ces formules ne sont valables que si

$$c \ll 1 \,. \tag{F.48}$$

Lorsque la concentration c de la solution tend vers zéro, le potentiel chimique μ du solvant tend comme il se doit vers sa valeur μ_0 en l'absence de soluté. Le comportement de la formule (F.47.b) est moins simple (cf. note 9) : le potentiel chimique μ' du soluté ne tend pas vers $\psi(T, p)$, mais vers $-\infty$[12].

---

10. Cette affirmation, sans doute évidente intuitivement, peut être démontrée en généralisant le raisonnement du complément III.D, notes 9 et 10. Dans le système de variables T, p, N, c, seule N est extensive. Si l'on multiplie par λ la taille du système, l'un des potentiels chimiques, μ par exemple, doit vérifier la condition

$$\mu(T, p, \lambda N, c) = \mu(T, p, N, c) \,. \tag{1}$$

Il suffit de dériver cette égalité par rapport à λ pour obtenir

$$N \frac{\partial \mu}{\partial N} = 0 \,; \tag{2}$$

μ ne dépend donc que de T, p et c.

11. On pourrait évidemment calculer μ et μ' directement à partir de l'énergie libre (F.39); on les obtiendrait ainsi comme fonctions de T, V, N et N'. Pour aboutir à leur expression simple (F.47) en fonction de T, p et c, il faudrait alors effectuer le changement de variable faisant passer de V à p (cf. formule (5) de la note 8).

12. On peut s'étonner que le potentiel chimique du soluté tende vers $-\infty$ lorsque la concentration c tend vers 0. Du point de vue physique, ceci signifie que, si l'on met en contact le solvant pur avec un corps quelconque contenant des molécules du soluté, certaines d'entre elles passeront toujours en solution, quel que soit leur potentiel chimique μ_{ex} dans le corps considéré : μ_{ex} sera toujours supérieur à la valeur de (F.47.b) pour $c = 0$. La concentration c finalement obtenue sera telle que

$$\mu_{ex} = \psi(T, p) + kT \ln c$$

(relation valable seulement si $c \ll 1$); elle dépend donc fortement des valeurs de ψ et de μ_{ex} et sera très faible si ψ est très grand positif (auquel cas le soluté sera pratiquement insoluble dans le solvant).

III. QUELQUES APPLICATIONS

1. PRESSION OSMOTIQUE D'UNE SOLUTION

Pour la plupart des solutions, on connaît ou on sait construire des *membranes semi-perméables,* qui laissent librement passer les molécules du solvant mais sont imperméables à celles du soluté[13]. Ceci permet de définir et de mesurer la *«pression osmotique»* de la solution.

a. Description du phénomène

Mettons en contact, à travers une membrane semi-perméable, une solution de concentration c avec du solvant pur (fig. 1). Les deux liquides échangent de la chaleur et des molécules de solvant; donc, *à l'équilibre, les températures et les potentiels chimiques du solvant sont égaux de part et d'autre de la cloison semi-perméable;* par contre, le potentiel chimique du soluté reste évidemment différent dans les deux liquides.

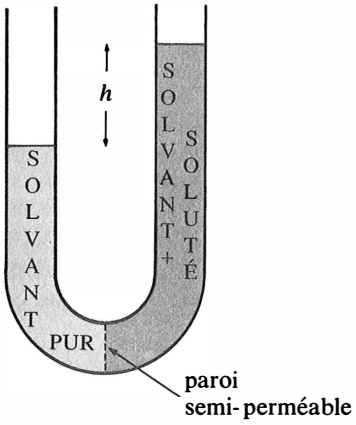


FIGURE 1
Mise en évidence et mesure de la pression osmotique p_s d'une solution : $p_s = h\rho_s g$, où ρ_s est la masse volumique de la solution.

Supposons que l'on connaisse les expressions $\mu_0(T, p)$ et $\mu(T, p, c)$ du potentiel chimique pour le solvant pur et pour le solvant dans une solution de concentration c. La température T étant la même des deux côtés, l'égalité des potentiels chimiques pour le solvant n'est possible que si *la pression est différente de part et d'autre de la membrane :*

$$\mu_0(T, p) = \mu(T, p', c) \qquad \text{(F.49)}$$

13. Le mécanisme le plus simple que l'on puisse envisager pour expliquer la propriété des membranes semi-perméables est fondé sur la taille des molécules : pour une solution de sucre ou de protéines dans l'eau, on conçoit qu'une membrane percée de pores microscopiques puisse être perméable à l'eau mais pas aux molécules plus volumineuses du soluté. En réalité, le fonctionnement des membranes semi-perméables est beaucoup plus complexe, notamment pour les membranes biologiques.

donne la relation entre p' et p si l'on se fixe T et c. En pratique[14], p' est supérieur à p; l'excès de pression

$$p_s \widehat{=} p' - p \tag{F.50}$$

dû à la présence de soluté est appelé *«pression osmotique»* de la solution.

b. Loi de van't Hoff pour les solutions diluées

Nous connaissons effectivement, pour une solution diluée, l'expression du potentiel chimique μ du solvant (formule (F.47.a)). L'égalité (F.49) s'écrit dans ce cas

$$\mu_0(T, p) \simeq \mu_0(T, p') - ckT \quad \text{si} \quad c \ll 1 . \tag{F.51}$$

On s'attend à ce qu'ici p' et p soient très voisines. Développons donc $\mu_0(T, p')$ autour de $p' = p$:

$$\mu_0(T, p') \simeq \mu_0(T, p) + (p' - p)\left(\frac{\partial \mu_0}{\partial p}\right)_T , \tag{F.52}$$

soit, d'après la relation (F.13.b),

$$\mu_0(T, p') \simeq \mu_0(T, p) + (p' - p)\frac{V}{N} \tag{F.53}$$

(V est ici le volume de la solution, qui diffère très peu de celui du solvant pur pour les mêmes valeurs de N, T et p). En comparant (F.51) et (F.53), on trouve aussitôt

$$p' - p \simeq \frac{N}{V} ckT , \tag{F.54}$$

c'est-à-dire, compte tenu de (F.27),

$$p_s \simeq \frac{N'kT}{V} \quad \text{si} \quad c \ll 1 . \tag{F.55}$$

Ce résultat est connu sous le nom de *loi de van't Hoff : dans une solution diluée, la pression osmotique créée par le soluté est la même que celle d'un gaz parfait* qui comporterait le même nombre N' de molécules dans le même volume V que celui de la solution, et à la même température T.

2. ÉQUILIBRE ENTRE DEUX PHASES DILUÉES

Nous allons étudier pour terminer des situations où le système solvant-soluté est séparé en *deux phases distinctes,* liquide et gaz, chacune d'elles étant très diluée :

$$c_\ell \ll 1 , \quad c_g \ll 1 , \tag{F.56}$$

où c_ℓ est la concentration de la phase liquide, c_g celle de la phase gazeuse.

---

14. On peut comprendre intuitivement comme suit l'origine de la pression osmotique. La pression sur une paroi est la manifestation macroscopique des chocs très nombreux des molécules sur la paroi. Dans la situation envisagée ici, la pression exercée par les molécules du solvant est la même sur les deux faces de la membrane puisque (égalité des potentiels chimiques) leurs flux dans les deux sens s'équilibrent. Mais il y a en plus, du côté de la solution, les molécules du soluté qui exercent une pression supplémentaire.

Cette explication doit cependant être prise avec précaution : elle suppose que le comportement des molécules du solvant n'est pas affecté par la présence du soluté, c'est-à-dire que la solution est très diluée; pour des concentrations plus fortes, il n'est pas possible de séparer la pression p' dans la solution en une somme de deux termes, dont l'un proviendrait seulement du solvant et l'autre seulement du soluté.

a. Conditions d'équilibre

Les deux phases en présence peuvent échanger de l'énergie et du volume; par conséquent, *à l'équilibre, la température et la pression sont les mêmes dans les deux phases.* Elles peuvent également échanger librement des molécules de solvant et des molécules de soluté, ce qui conduit à l'*égalité des potentiels chimiques* pour le solvant d'une part, et pour le soluté d'autre part, entre les deux phases à l'équilibre :

$$\mu_\ell(T, p, c_\ell) = \mu_g(T, p, c_g)\,, \tag{F.57.a}$$

$$\mu'_\ell(T, p, c_\ell) = \mu'_g(T, p, c_g)\,. \tag{F.57.b}$$

Pour des *mélanges dilués,* on peut utiliser les expressions trouvées en (F.47), ce qui donne[15]

$$\mu_{0\ell}(T, p) - c_\ell kT = \mu_{0g}(T, p) - c_g kT\,, \tag{F.58.a}$$

$$\psi_\ell(T, p) + kT \ln c_\ell = \psi_g(T, p) + kT \ln c_g\,. \tag{F.58.b}$$

L'équilibre entre les phases liquide et gazeuse du *solvant pur* n'est possible que si les valeurs T_0 et p_0 de la température et de la pression vérifient la condition

$$\mu_{0\ell}(T_0, p_0) = \mu_{0g}(T_0, p_0)\,. \tag{F.59}$$

Choisissons alors un « point d'équilibre » (T_0, p_0) entre les phases liquide et vapeur du solvant pur, et cherchons comment les paramètres de la transition sont modifiés par la présence du soluté. Soit

$$T = T_0 + \Delta T \tag{F.60.a}$$

$$p = p_0 + \Delta p \tag{F.60.b}$$

un point d'équilibre entre les deux phases du mélange. Si les concentrations c_ℓ et c_g sont petites, on s'attend à ce qu'un faible écart de température ΔT corresponde à un faible écart de pression Δp, et vice versa. On peut donc, dans la condition d'équilibre (F.58.a), développer $\mu_{0\ell}$ et μ_{0g} au premier ordre en ΔT et Δp au voisinage de (T_0, p_0). Si l'on se rappelle que T_0 et p_0 satisfont la relation (F.59), on obtient

$$(v_{0g} - v_{0\ell})\,\Delta p - (s_{0g} - s_{0\ell})\,\Delta T = (c_g - c_\ell)\,kT\,, \tag{F.61}$$

où $$v_{0\alpha} = \frac{\partial \mu_\alpha}{\partial p} \tag{F.62.a}$$

et $$s_{0\alpha} = -\frac{\partial \mu_{0\alpha}}{\partial T} \tag{F.62.b}$$

sont le volume et l'entropie par molécule du solvant pur dans la phase (α). La quantité

$$\ell_0 = T_0(s_{0g} - s_{0\ell}) \tag{F.63}$$

n'est autre que la *chaleur latente de vaporisation* du solvant pur, rapportée à une molécule (complément III.G, § III.4.b); cette grandeur est positive. La condition d'équilibre (F.58.a) s'écrit donc finalement[16]

$$(v_{0g} - v_{0\ell})\,\Delta p - \frac{\ell_0}{T_0}\,\Delta T = (c_g - c_\ell)\,kT_0\,. \tag{F.64}$$

---

15. Le potentiel chimique du solvant pur prend des formes différentes dans les deux phases, qui ont des propriétés physiques différentes : $\mu_{0\ell}$ et μ_{0g}. Il en est de même pour la fonction ψ : ψ_ℓ et ψ_g sont des fonctions différentes de T et p.

16. Si $c_\ell = c_g = 0$ (solvant pur), la relation (F.64) s'identifie à la célèbre *formule de Clapeyron.*

b. Lois de Raoult

α. Déplacement de la température d'équilibre à pression constante

A pression constante p_0, la présence du soluté provoque un déplacement ΔT_{p_0} de la température de transition $T_0(p_0)$ définie par (F.59), égal à

$$\Delta T_{p_0} = \frac{kT_0^2}{\ell_0}(c_\ell - c_g) . \tag{F.65}$$

Dans le cas particulier où le soluté est insoluble dans la phase gazeuse ($c_g = 0$)[17], la température de transition est plus élevée que pour le solvant pur, cette élévation de température étant proportionnelle à la concentration c_ℓ de la solution liquide :

$$\Delta T_{p_0} = c_\ell \frac{kT_0^2}{\ell_0} \quad \text{si} \quad c_g = 0 . \tag{F.66}$$

On démontre ainsi la « loi de l'ébullioscopie », établie expérimentalement par F. Raoult en 1882[18]. Pour l'eau, le rapport sans dimension kT_0/ℓ_0 vaut environ 0,076 à $T_0 = 373$ K et $p_0 = 1$ atmosphère; l'élévation relative de température $\Delta T_{p_0}/T_0$ est donc très faible.

β. Déplacement de la pression d'équilibre à température constante

On peut aussi fixer la température T_0 et calculer le déplacement Δp_{T_0} de la pression de vapeur saturante $p_0(T_0)$ définie par (F.59). La relation (F.61) donne alors

$$\Delta p_{T_0} = \frac{c_g - c_\ell}{v_{0g} - v_{0\ell}} kT_0 . \tag{F.67}$$

Loin du point critique (complément III.G, § III.1), $v_{0\ell}$ est négligeable devant v_{0g}, et la phase gazeuse peut être considérée comme un gaz parfait :

$$v_{0\ell} \ll v_{0g} \simeq \frac{kT_0}{p_0} . \tag{F.68}$$

On a dans ce cas

$$\frac{\Delta p_{T_0}}{p_0} \simeq c_g - c_\ell . \tag{F.69}$$

En particulier, lorsque le soluté est insoluble en phase gazeuse ($c_g = 0$), on retrouve la « loi de la tonométrie » de Raoult : la pression de vapeur saturante est plus faible au-dessus d'une solution qu'au-dessus du solvant pur; son abaissement relatif est égal à la concentration c_ℓ de la solution.

c. Concentration des deux phases en équilibre

Nous n'avons jusqu'ici exploité qu'une des deux conditions d'équilibre; les formules du paragraphe b sont correctes, mais les concentrations c_ℓ et c_g n'y sont pas indépendantes l'une de l'autre. On peut considérer que l'ensemble des deux conditions

17. Ceci se produit lorsque solvant et soluté ont des caractéristiques très différentes : pour une solution de sucre dans l'eau, par exemple, la vapeur en équilibre avec le liquide ne contient pratiquement aucune molécule de sucre.

18. On peut analyser de façon analogue l'équilibre entre une solution liquide de concentration c_ℓ faible et la phase solide, *dans le cas où celle-ci est constituée de solvant pur.* On aboutit alors à la « loi de la cryoscopie » de Raoult : à pression constante p_o, le « point de solidification » est abaissé par la présence du soluté dans le liquide, la variation ΔT_{p_0} de la température de transition étant encore proportionnelle à la concentration c_ℓ de la solution :

$$\Delta T_{p_0} = -c_\ell \frac{kT_0^2}{\ell_0^{(f)}} ,$$

où $\ell_0^{(f)}$ est ici la *chaleur latente de fusion* par molécule pour le solvant pur.

(F.57) détermine c_ℓ et c_g comme fonctions des paramètres T et p. Si l'*on fixe la pression p*, on peut représenter c_ℓ et c_g en fonction de la température T; il est commode d'utiliser le même graphique pour ces deux fonctions, et de porter T en ordonnée et c_ℓ ou c_g en abscisse. On obtient ainsi ce que l'on appelle un *« diagramme de phases »* pour le mélange binaire étudié. On trace un tel diagramme pour différentes valeurs de la pression p, échelonnées de façon régulière dans le domaine d'équilibre entre les deux phases. Les concentrations c_ℓ et c_g peuvent varier sur tout l'intervalle [0, 1]; mais nous nous cantonnons ici dans la région des solutions diluées (c_ℓ, $c_g \ll 1$) où les approximations (F.58) sont valables[19].

Le rapport c_ℓ/c_g est déterminé par la relation (F.58.b) :

$$\frac{c_\ell}{c_g} = \exp\left[\psi_g(T, p) - \psi_\ell(T, p)\right]/kT . \tag{F.70}$$

D'autre part, comme nous fixons la pression, la condition (F.58.a) s'écrit sous la forme (F.65) :

$$\Delta T_{p_0} \widehat{=} T - T_0 = (c_\ell - c_g)\,\frac{kT_0^2}{\ell_0} . \tag{F.71}$$

Les concentrations étant faibles, T est voisine de la température T_0 d'équilibre liquide-vapeur pour le solvant pur sous la pression p_0. On peut donc remplacer T et p par T_0 et p_0 dans (F.70). Le système d'équations (F.70)-(F.71) se résout alors aisément et donne

$$c_\ell(T, p_0) = \frac{\ell_0}{kT_0^2}(T - T_0)\left\{1 - \exp\left[\psi_\ell(T_0, p_0) - \psi_g(T_0, p_0)\right]/kT_0\right\}^{-1} , \tag{F.72.a}$$

$$c_g(T, p_0) = \frac{\ell_0}{kT_0^2}(T - T_0)\left\{\exp\left[\psi_g(T_0, p_0) - \psi_\ell(T_0, p_0)\right]/kT_0 - 1\right\}^{-1} . \tag{F.72.b}$$

Dans le domaine de validité de l'approximation des solutions diluées, les deux courbes d'équilibre représentant $c_\ell(T, p_0)$ et $c_g(T, p_0)$ sont donc des droites issues du même point $T = T_0$ de l'axe des ordonnées. Deux configurations sont possibles (fig. 2), suivant le signe de $\psi_\ell(T_0, p_0) - \psi_g(T_0, p_0)$.

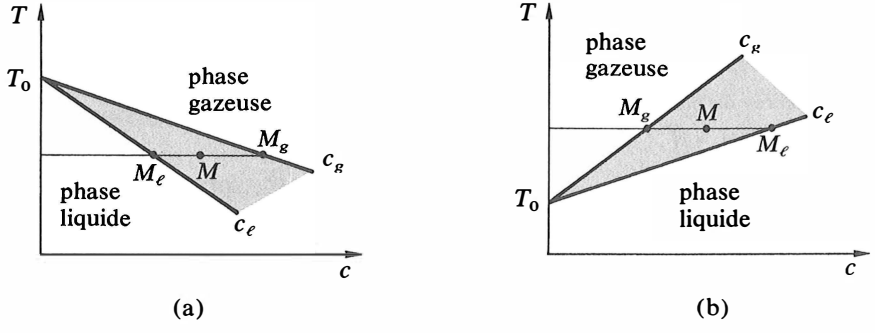


FIGURE 2

Les deux configurations possibles pour le diagramme de phases d'un mélange binaire dans le domaine des faibles concentrations : la configuration (a) correspond au cas où ψ_ℓ est supérieur à ψ_g, la configuration (b) au cas contraire.

19. Les résultats que nous allons obtenir sont aussi applicables à la région des concentrations voisines de 1, puisque les rôles du solvant et du soluté y sont simplement échangés.

Dans le diagramme de phases, un mélange de concentration globale c maintenu à la température T (la pression p_0 étant fixée) est représenté par le point d'abscisse c et d'ordonnée T. Si ce point se trouve au-dessus de la courbe associée à c_g, le mélange est homogène en phase gazeuse; il est homogène en phase liquide si le point se trouve au-dessous de la courbe associée à c_ℓ. Lorsque le point représentatif se situe entre les deux courbes en M, le mélange est séparé en deux phases coexistant à l'équilibre (exercice III.25); les proportions (en nombre total de molécules) des phases liquide et vapeur sont données par le rapport des longueurs MM_ℓ et MM_g[20].

20. Un mélange constitué de N molécules de solvant et N' molécules de soluté a pour *concentration globale*

$$c \widehat{=} \frac{N'}{N+N'} . \tag{1}$$

Supposons qu'il soit séparé en deux phases distinctes, liquide et vapeur, en équilibre; la phase (α) comprend N_α molécules de solvant et N'_α molécules de soluté, de sorte que sa concentration vaut

$$c_\alpha \widehat{=} \frac{N'_\alpha}{N_\alpha + N'_\alpha} , \quad \alpha = \ell \text{ ou } g . \tag{2}$$

L'importance relative des deux phases est caractérisée par le rapport

$$x \widehat{=} \frac{N_\ell + N'_\ell}{N_g + N'_g} \tag{3}$$

des nombres de molécules qu'elles contiennent. On a évidemment

$$N_\ell + N_g = N , \tag{4.a}$$

$$N'_\ell + N'_g = N' . \tag{4.b}$$

D'après les égalités (4), la concentration globale (1) peut être écrite

$$c = \frac{N'_\ell + N'_g}{N_\ell + N_g + N'_\ell + N'_g} , \tag{5}$$

ce qui donne, compte tenu des définitions (2),

$$(N_\ell + N_g + N'_\ell + N'_g)\, c = (N_\ell + N'_\ell)\, c_\ell + (N_g + N'_g)\, c_g . \tag{6}$$

On en déduit facilement la relation

$$x = \frac{c_g - c}{c - c_\ell} : \tag{7}$$

l'importance quantitative de la phase liquide est proportionnelle à la longueur du segment MM_g de la figure 2, celle de la phase gazeuse à la longueur MM_ℓ.

COMPLÉMENT III.G

L'équation de van der Waals et la transition gaz-liquide

I. Étude statistique d'un fluide
  1. Formalisme général
  2. Interaction entre les molécules

II. L'équation de van der Waals
  1. Fonction de partition dans l'approximation de van der Waals
    a. Approximation des particules indépendantes
    b. Forme du potentiel effectif
    c. Évaluation des paramètres
  2. Grandeurs thermodynamiques
    a. Énergie libre
    b. Pression
    c. Énergie moyenne
    d. Entropie
    e. Potentiel chimique

III. La transition gaz-liquide
  1. Les isothermes de van der Waals
  2. Étude de l'énergie libre
  3. L'équilibre liquide-vapeur
  4. Caractéristiques de la transition
    a. Coexistence des deux phases
    b. La transition s'accompagne d'une variation d'entropie
    c. Existence d'états métastables

IV. Conclusion

Dans un gaz, l'énergie potentielle d'interaction des molécules est faible comparée à leur énergie cinétique; elle est même négligeable à haute température et faible densité, domaine où l'approximation du gaz parfait est valable (III, § IV.B). Cependant, lorsque la température diminue ou que la densité augmente, la part de l'énergie d'interaction dans l'énergie totale devient plus importante et les propriétés du gaz s'éloignent de celles du gaz parfait. Lorsque l'énergie d'interaction devient comparable à l'énergie cinétique, le gaz se condense et se transforme en liquide.

L'équation d'état de van der Waals permet de comprendre ce passage du gaz au liquide. Bien qu'elle repose sur une théorie de champ moyen très fruste, qui ne devrait donner que des résultats qualitatifs, l'équation de van der Waals possède les propriétés nécessaires pour décrire l'essentiel des phénomènes physiques mis en jeu dans un gaz réel et dans la transition gaz-liquide.

I. ÉTUDE STATISTIQUE D'UN FLUIDE

1. FORMALISME GÉNÉRAL

Considérons un fluide monoatomique[1] constitué de N ($N \gg 1$) particules identiques de masse m, contenu dans un récipient de volume V et maintenu à la température T. Les propriétés macroscopiques de ce système peuvent être calculées dans le cadre de la mécanique classique (voir complément I.A, § III.2) et donc déduites de la fonction de partition (cf. III, § V.A.2).

$$Z = \frac{1}{N!}\frac{1}{h^{3N}} \int \dots \int \mathrm{d}^3p_1 . \mathrm{d}^3p_2 \dots \mathrm{d}^3p_N . \mathrm{d}^3r_1 . \mathrm{d}^3r_2 \dots \mathrm{d}^3r_N \, \mathrm{e}^{-\mathcal{H}/kT} , \qquad \text{(G.1)}$$

1. Il n'est pas essentiel que le fluide soit monoatomique pour établir l'équation d'état de van der Waals, mais cela simplifie considérablement l'exposé (voir plus bas note 3).

où $$\mathcal{H} = \frac{1}{2m} \sum_{i=1}^{N} \vec{p}_i^{\,2} + U(\vec{r}_1, \vec{r}_2, \ldots, \vec{r}_N) , \tag{G.2}$$

somme des énergies cinétique et potentielle, est la fonction de Hamilton du système. L'intégration sur les moments $\vec{p}_i$ est immédiate[2]; on obtient

$$Z = \frac{1}{N!} \left(\frac{2\pi mkT}{h^2} \right)^{3N/2} Y , \tag{G.3}$$

avec

$$Y = \int \ldots \int \mathrm{d}^3 r_1 . \mathrm{d}^3 r_2 \ldots \mathrm{d}^3 r_N \; \mathrm{e}^{-U(\vec{r}_1, \vec{r}_2, \ldots, \vec{r}_N)/kT} . \tag{G.4}$$

Lorsqu'on néglige U ($U/kT \to 0$), Y est égal à V^N et on retrouve la fonction de partition du gaz parfait. Si on veut faire mieux et tenir compte de l'interaction entre les molécules, la première difficulté que l'on rencontre est la détermination de la fonction U.

2. INTERACTION ENTRE LES MOLÉCULES

Nous supposons que l'énergie d'interaction totale est la somme des énergies d'interaction $u(|\vec{r}_i - \vec{r}_j|)$ des atomes[3] pris deux à deux; autrement dit,

$$U(\vec{r}_1, \vec{r}_2, \ldots, \vec{r}_N) = \frac{1}{2} \sum_{\substack{i,j \\ i \neq j}} u(|\vec{r}_i - \vec{r}_j|) \tag{G.5}$$

(le facteur 1/2 évite que l'énergie d'interaction entre les particules i et j soit comptée deux fois).

Cette relation, qui peut sembler aller de soi, n'est en fait pas du tout évidente. En effet l'interaction entre deux molécules est principalement due au déplacement des charges électroniques qu'elles induisent l'une dans l'autre lorsqu'elles sont proches. Ce déplacement de charge, et donc l'énergie d'interaction u, peuvent être ainsi modifiés de façon appréciable par la présence d'une troisième molécule. La formule (G.5) suppose donc que la probabilité pour que deux molécules se trouvent à une distance inférieure à la portée de u est faible : les situations où trois molécules sont proches les unes des autres sont alors suffisamment rares pour être négligées. Cette condition est bien satisfaite dans un gaz, beaucoup moins dans un liquide.

Cette approximation étant cependant admise, reste à déterminer la fonction $u(\rho)$. En principe on peut l'obtenir à partir de la mécanique quantique, mais c'est un problème très complexe. On démontre[4] cependant qu'à longue distance, l'énergie d'interaction entre deux molécules est attractive et varie comme ρ^{-6} (ce qui correspond

2. Le calcul est le même que pour le gaz parfait (III, § V, note 92).

3. Des complications de deux sortes s'introduisent dans l'étude des fluides constitués de molécules polyatomiques. En premier lieu apparaissent dans l'énergie potentielle U des termes correspondant aux degrés de liberté internes des molécules (rotation, vibration). Ces termes sont analysés, dans le cadre de l'approximation des gaz parfaits, au complément III.B; si les interactions entre les molécules ne sont plus négligeables, les degrés de liberté internes se couplent aux degrés de liberté de translation (position, vitesse), car les collisions intermoléculaires peuvent transformer de l'énergie cinétique de translation en énergie de rotation ou de vibration et vice-versa (collisions inélastiques). D'autre part, les molécules n'étant plus à symétrie sphérique, les forces qui s'exercent entre elles dépendent alors, non seulement de leur distance, mais aussi de leur orientation relative.

à une force d'attraction, dite de van der Waals, en ρ^{-7}). Aux courtes distances, lorsque ρ devient de l'ordre des dimensions propres des molécules, celles-ci exercent l'une sur l'autre des forces répulsives très intenses; c'est essentiellement le principe de Pauli auquel obéissent les électrons qui rend ainsi les molécules pratiquement impénétrables les unes aux autres. La fonction $u(\rho)$ a donc l'allure représentée sur la figure 1. Ses caractéristiques importantes sont la position ρ_0 du minimum et sa valeur $(-u_0)$. Leur ordre de grandeur est l'angström pour ρ_0 et quelques centièmes d'eV pour u_0 (c'est-à-dire quelques centaines de K pour u_0/k).

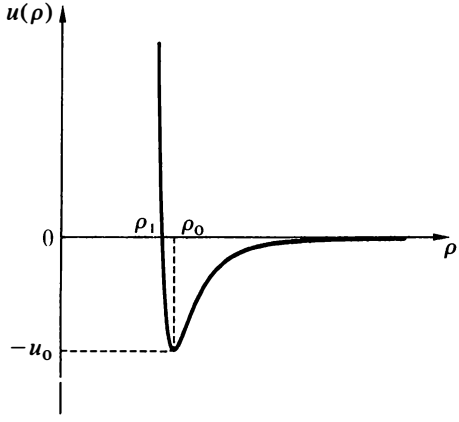


FIGURE 1
Allure générale de l'énergie d'interaction $u(\rho)$ entre deux molécules en fonction de leur distance ρ. La valeur ρ_1 de la distance qui annule $u(\rho)$ est très voisine de ρ_0 (pour le potentiel de Lennard-Jones (G.6), $\rho_1 = 2^{-1/6}\,\rho_0 \simeq 0{,}89\,\rho_0$).

En pratique on utilise des formes analytiques empiriques qui reproduisent l'allure de la figure 1 à l'aide de quelques paramètres. L'expression la plus utilisée pour $u(\rho)$ est probablement celle de Lennard-Jones :

$$u_{\text{L. J.}}(\rho) = u_0 \left[\left(\frac{\rho_0}{\rho}\right)^{12} - 2 \left(\frac{\rho_0}{\rho}\right)^{6} \right], \qquad \text{(G.6)}$$

qui ne fait intervenir que les deux paramètres u_0 et ρ_0.

II. L'ÉQUATION DE VAN DER WAALS

1. FONCTION DE PARTITION DANS L'APPROXIMATION DE VAN DER WAALS

Le calcul de l'intégrale (G.4) est rendu difficile par la forme compliquée (G.5) de l'énergie potentielle. Nous allons nous limiter à une évaluation grossière, fondée sur des arguments physiques simples.

4. L'origine physique de cette « interaction de van der Waals » entre les molécules est la suivante : une fluctuation de charge dans une molécule crée un dipôle électrique, qui polarise une molécule voisine; le dipôle induit sur cette deuxième molécule agit à son tour sur la polarisation de la première. Pour un calcul explicite, voir par exemple C.D.L., p. 1120.

a. Approximation des particules indépendantes

Considérons une molécule particulière. Les $(N-1)$ autres, très nombreuses et réparties quasi-uniformément dans le volume du récipient, lui apparaissent, en première approximation, comme un nuage pratiquement continu. Si on néglige les perturbations apportées à ce « nuage » par la présence et le mouvement de la particule considérée, celle-ci aura une énergie potentielle $U_{\text{eff}}(\vec{r})$ dépendant de sa seule position $\vec{r}$, et non plus de celle des autres particules. Dans cette approximation, où chaque particule se déplace dans le *champ moyen* créé par toutes les autres, le système se comporte comme un ensemble de *particules indépendantes;* son énergie potentielle est une somme de termes dont chacun fait intervenir les coordonnées d'une seule particule[5] :

$$U \simeq \frac{1}{2} \sum_{i=1}^{N} U_{\text{eff}}(\vec{r}_i) \,. \tag{G.7}$$

L'intégrale (G.4) se factorise alors, et la fonction de partition prend la forme

$$Z = \frac{1}{N!} \left(\frac{2\pi mkT}{h^2} \right)^{3N/2} \left[\int \mathrm{d}^3 r \; \mathrm{e}^{-U_{\text{eff}}(\vec{r})/2kT} \right]^N . \tag{G.8}$$

b. Forme du potentiel effectif

L'approximation précédente néglige les corrélations entre les diverses molécules. On considère donc qu'elles se répartissent de façon uniforme dans le volume du récipient. On pourrait penser que dans ces conditions le potentiel effectif U_{eff} ne peut être que constant sur l'ensemble du volume. Mais cette approximation cesse d'être valable lorsque la molécule choisie se rapproche d'une autre à des distances de l'ordre de ρ_1 (cf. figure 1). Ce n'est plus alors une attraction moyenne créée par un nuage homogène qu'elle ressent, mais bien une très forte répulsion de la part de la molécule voisine. Pour tenir compte grossièrement de ce phénomène, on considère qu'il existe une portion V_0 du volume du récipient qui est interdite à la molécule étudiée à cause de la présence des autres; V_0 est appelé le *« volume exclu »*. L'intégrale de la formule (G.8) porte alors sur le volume accessible $V - V_0$, où le potentiel effectif $U_{\text{eff}}(\vec{r})$ est égal à une constante U_0. On obtient donc

$$Z = \frac{1}{N!} \left(\frac{2\pi mkT}{h^2} \right)^{3N/2} \left[(V - V_0) \mathrm{e}^{-U_0/2kT} \right]^N . \tag{G.9}$$

c. Évaluation des paramètres

Reste à déterminer V_0 et U_0, et particulièrement leur dépendance par rapport à N, V, et T.

---

5. Si l'on se réfère à l'expression (G.5) que l'on veut approximer, on voit que le potentiel $U_{\text{eff}}(\vec{r})$ ressenti par une particule donnée au point $\vec{r}$ s'obtient en moyennant

$$\sum_{i=2}^{N} u(|\vec{r} - \vec{r}_i|)$$

sur les positions $\vec{r}_2$, $\vec{r}_3$, ..., $\vec{r}_N$ des $(N-1)$ autres molécules. Ainsi $u(|\vec{r}_i - \vec{r}_j|)$ intervient dans le calcul de $U_{\text{eff}}(\vec{r}_i)$ et dans celui de $U_{\text{eff}}(\vec{r}_j)$; c'est pourquoi il est commode d'introduire un facteur 1/2 dans l'expression approchée (G.7)

Le *volume exclu* est facile à évaluer. La distance minimum permise entre deux molécules est pratiquement égale à ρ_1; lorsque deux molécules se rapprochent, le volume interdit à l'une d'elles du fait de la présence de l'autre est égal à[6] $4\pi\rho_1^3/3$. Or il y a ainsi $N(N-1)/2$ paires de molécules; le volume exclu total[7] est donc

$$\frac{N(N-1)}{2}\frac{4}{3}\pi\rho_1^3 \simeq \frac{2N^2}{3}\pi\rho_1^3 , \tag{G.10}$$

d'où un volume exclu *par particule*[8]

$$V_0 \simeq \frac{2N}{3}\pi\rho_1^3 \widehat{=} bN . \tag{G.11}$$

V_0 est donc proportionnel à N; le coefficient de proportionnalité b dépend uniquement de la nature du fluide[9]; en ordre de grandeur il représente quatre fois le « volume propre » v_0 d'une molécule[6],[10].

Évaluons maintenant le *potentiel moyen* U_0. Prenons comme origine des coordonnées la position d'une particule déterminée. Les autres, dans le modèle simple que nous avons décrit, sont réparties de façon uniforme dans l'espace, sauf qu'aucune ne peut s'approcher de l'origine à des distances inférieures à ρ_1. Le nombre $n(\vec{r})\mathrm{d}^3r$ de particules contenues dans un volume infinitésimal d^3r autour du point $\vec{r}$ est donc tel que

$$n(\vec{r}) = \begin{cases} 0 & \text{si} \quad r < \rho_1 , \\ \dfrac{N-1}{V-v_0} \simeq \dfrac{N}{V} & \text{si} \quad r > \rho_1 . \end{cases} \tag{G.12}$$

---

6. Si les particules sont des sphères dures de rayon $R = \rho_1/2$, chaque molécule exclut ainsi de son voisinage (fig. 2) un volume

$$\frac{4}{3}\pi(2R)^3 = \frac{4}{3}\pi\rho_1^3$$

égal à 8 fois son volume propre $v_0 = 4\pi R^3/3$.

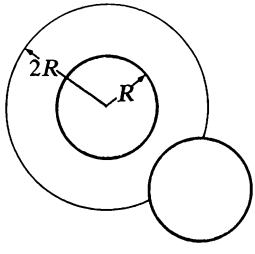


FIGURE 2
Dans un gaz de sphères dures de rayon *R*, aucune autre particule ne peut avoir son centre à l'intérieur d'une sphère de rayon 2*R* centrée sur une particule donnée.

7. Nous faisons l'hypothèse que les diverses sphères d'exclusion ne se recoupent pratiquement jamais; ce qui suppose, là encore, que $V_0/V \ll 1$.

8. Si on considère une molécule particulière, le volume exclu à cette particule du fait de la présence des $(N-1)$ autres est $(N-1)4\pi\rho_1^3/3$. Mais si on considère alors une deuxième particule, le volume exclu supplémentaire n'est plus que $(N-2)4\pi\rho_1^3/3$, puisque l'exclusion de la première molécule a déjà été comptée; le volume exclu associé à la troisième n'est plus que $(N-3)4\pi\rho_1^3/3$, etc. On retrouve ainsi le résultat (G.10) pour le volume exclu à l'ensemble des particules.

9. En toute rigueur, la distance minimum d'approche entre deux particules dépend de leur énergie; cependant $u(\rho)$ varie très rapidement pour $\rho \lesssim \rho_1$, et cette distance peut être considérée comme une constante pour un fluide donné.

10. La situation est moins simple pour un gaz polyatomique. Néanmoins V_0 reste proportionnel à N, et b peut alors être considéré comme un paramètre phénoménologique.

L'énergie potentielle de la particule située à l'origine vaut alors

$$U_0 = \int u(\vec{r})\, n(\vec{r})\, \mathrm{d}^3 r = \frac{N}{V} \int_{\rho_1}^{+\infty} 4\pi r^2\, \mathrm{d}r\, u(r)\,. \tag{G.13}$$

$u(r)$ étant négatif dans la région d'intégration (cf. figure 1), U_0 est négatif et, comme il se doit, indépendant de la position $\vec{r}$ de la particule. En ordre de grandeur, U_0 est égal à N fois la valeur au minimum $(-u_0)$ du potentiel entre deux particules, multiplié par le rapport du « volume d'interaction » (c'est-à-dire le volume d'une sphère dont le rayon est égal à la portée du potentiel $u(\rho)$) au volume total V; si v_0 est le volume propre d'une molécule, ce rapport est égal à quelques v_0/V, le coefficient de proportionnalité dépendant évidemment du modèle utilisé[11]. Par exemple, avec le potentiel de Lennard-Jones (G.6), on obtient

$$U_0^{(\mathrm{L.\,J.})} = -\frac{8}{3} N u_0 \frac{4\pi\rho_1^3}{3V} = -\frac{64}{3} N u_0 \frac{v_0}{V}\,. \tag{G.14}$$

De façon générale, on pose

$$U_0 = -2a \frac{N}{V}\,, \tag{G.15}$$

où a est une contante positive, d'ordre quelques $u_0 v_0$ et caractéristique du fluide.

Reportées dans (G.9), les formules (G.11) et (G.15) donnent finalement l'expression de la fonction de partition dans l'approximation de van der Waals :

$$Z = \frac{1}{N!} \left(\frac{2\pi mkT}{h^2}\right)^{3N/2} \left[(V - Nb)\, \mathrm{e}^{aN/kTV}\right]^N\,, \tag{G.16}$$

ou encore, si l'on fait intervenir la fonction de partition Z_p du gaz parfait correspondant aux mêmes valeurs des paramètres[12],

$$Z = Z_p \left[\left(1 - \frac{Nb}{V}\right) \mathrm{e}^{aN/kTV}\right]^N\,. \tag{G.17}$$

2. GRANDEURS THERMODYNAMIQUES

Les grandeurs thermodynamiques du fluide, dans l'approximation de van der Waals, se déduisent de la fonction de partition (G.16) par les relations usuelles.

a. Énergie libre

$$F = -kT \ln Z = -NkT \left\{\ln\left[\left(\frac{2\pi mkT}{h^2}\right)^{3/2} \left(\frac{V}{N} - b\right)\right] + 1 + \frac{aN}{kTV}\right\}\,, \tag{G.18}$$

ou encore, si F_p est l'énergie libre du gaz parfait correspondant,

$$F = F_p - N\left[\frac{aN}{V} + kT \ln\left(1 - \frac{bN}{V}\right)\right]\,. \tag{G.19}$$

Comme il se doit, F est une grandeur extensive : l'énergie libre par particule F/N ne fait intervenir que la densité de particules N/V (et non pas N et V séparément).

---

11. Le point important est que $u(\rho)$ tende vers zéro suffisamment vite aux grandes distances pour que l'intégrale (G.13) soit définie. C'est le cas pour l'interaction de van der Waals (en ρ^{-6}). Le problème est radicalement différent pour des particules chargées dont le potentiel d'interaction varie en $1/\rho$, et qui demandent un traitement spécial (complément III.H).

12. C'est la formule (G.17) que l'on utilisera dans le cas d'un fluide polyatomique : Z_p est alors donnée par la formule (B.9) du complément III.B; a et b sont dans ce cas deux coefficients phénoménologiques qui ne dépendent que de la nature du fluide.

b. Pression

$$p = -\frac{\partial F}{\partial V} = \frac{NkT}{V - bN} - a\frac{N^2}{V^2} . \tag{G.20}$$

Cette relation constitue *l'équation d'état de van der Waals,* qu'on retiendra sous la forme[13]

$$\boxed{\left(p + a\frac{N^2}{V^2}\right)(V - bN) = NkT} \tag{G.21}$$

c. Énergie moyenne

$$\overline{E} = -\frac{\partial}{\partial\beta} \ln Z = \frac{3}{2} NkT - a\frac{N^2}{V} . \tag{G.22}$$

Ce résultat n'est pas étonnant : au terme $3NkT/2$ correspondant à l'énergie cinétique moyenne s'ajoute un terme d'énergie potentielle $U_0/2 = -Na/V$ par particule.

La capacité calorifique à volume constant ne dépend pas de la température et coïncide avec celle du gaz parfait[14] :

$$C_V = \frac{\partial \overline{E}}{\partial T} = \frac{3}{2} Nk . \tag{G.23}$$

d. Entropie

$$S = -\frac{\partial F}{\partial T} = Nk\left[\frac{5}{2} + \ln\left(\frac{V}{N} - b\right) + \frac{3}{2} \ln \frac{2\pi mkT}{h^2}\right] . \tag{G.24}$$

C'est la même relation que pour un gaz parfait, à condition de tenir compte de l'effet du volume exclu : l'entropie est plus faible que celle du gaz parfait correspondant, simplement parce que le volume réellement offert aux particules est plus faible :

$$S = S_p + Nk \ln\left(1 - \frac{bN}{V}\right) . \tag{G.25}$$

e. Potentiel chimique

$$\mu = \frac{\partial F}{\partial N} = \frac{F}{N} + kT\frac{V}{V - bN} - \frac{aN}{V} . \tag{G.26}$$

---

13. L'équation de van der Waals ressemble beaucoup à celle du gaz parfait, il y a cependant deux corrections tout à fait essentielles :

— Le volume offert aux particules est le volume V du récipient diminué du terme bN proportionnel au nombre de particules présentes. C'est la taille non nulle des molécules qui est responsable de cette correction.

— La pression est la somme de deux termes : au terme d'origine cinétique $NkT/(V - bN)$ s'ajoute $-aN^2/V^2$. Ce terme correctif, nécessairement négatif puisque dû à l'attraction entre molécules à moyenne distance, est proportionnel au carré de la densité N/V car il fait intervenir les interactions entre les molécules prises deux à deux (cf. complément III.C).

14. L'énergie moyenne d'un gaz parfait ne dépend que de sa température (en l'absence de champ extérieur). En effet, si les molécules n'interagissent pas, la variation de leur distance moyenne (c'est-à-dire du rapport V/N) ne peut modifier l'énergie du gaz. Dans le cas du fluide de van der Waals, on tient compte des interactions par le terme correctif $-aN^2/V$. A N et V fixés, cette énergie potentielle est indépendante de T; elle ne contribue donc pas à la capacité calorifique à volume constant (il n'en serait pas de même pour la capacité calorifique à pression constante).

A l'aide de l'expression (G.20) de la pression p, μ peut se mettre sous la forme[15]

$$\mu = \frac{1}{N}(F + pV) . \tag{G.27}$$

Si on veut le comparer au potentiel chimique μ_p du gaz parfait correspondant, on trouve facilement

$$\mu = \mu_p + kT \ln \frac{V}{V - Nb} + kT \frac{bN}{V - Nb} - 2a \frac{N}{V} . \tag{G.28}$$

La propriété simple du gaz parfait, dont la fugacité $\varphi = \exp(\mu/kT)$ est proportionnelle à la pression (III, § IV.B.3), n'est plus valable ici.

III. LA TRANSITION GAZ-LIQUIDE

1. LES ISOTHERMES DE VAN DER WAALS

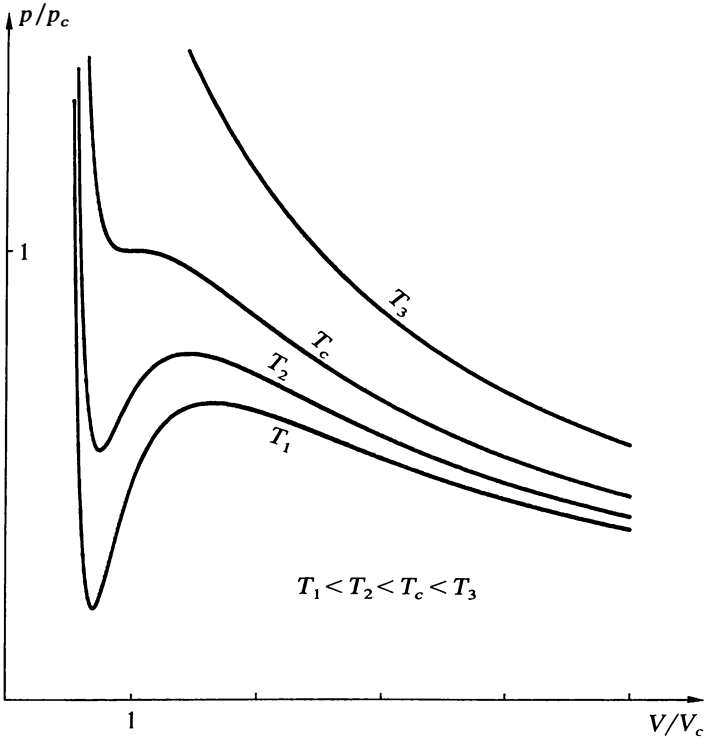


FIGURE 3

Isothermes de van der Waals dans le plan (p, V). Pour $T > T_c$, les isothermes sont toujours décroissantes, pour $T < T_c$ chaque isotherme présente deux extremums.

15. La « fonction de Gibbs », ou « enthalpie libre »

$$G \widehat{=} F + pV \tag{1}$$

s'introduit naturellement lorsqu'on prend comme paramètres extérieurs la température T et la pression p, imposées au système par un réservoir d'énergie et de volume (complément III.D, § I.3).

.../...

La figure 3 représente l'équation d'état (G.21) dans un diagramme pression-volume, pour une série de valeurs de la température T. Chaque courbe constitue ce qu'on appelle une *isotherme*[16].

Pour un fluide, dont l'état macroscopique est caractérisé par trois paramètres extérieurs (E, V et N si le fluide est isolé, T, p et N s'il est maintenu à température et pression constantes, ...), le potentiel chimique est toujours égal à l'enthalpie libre par particule (complément III.D, § I.3.c) :

$$\mu = \frac{G}{N} . \tag{2}$$

16. Les isothermes du gaz parfait ($pV = NkT$ = constante) sont des hyperboles. Celles de van der Waals ont essentiellement la même allure. Ce sont cependant des cubiques et elles peuvent donc présenter des extremums. En effet

$$\left(\frac{\partial p}{\partial V}\right)_T = -\frac{NkT}{(V-Nb)^2} + \frac{2aN^2}{V^3} ; \tag{1}$$

cette dérivée s'annule lorsque

$$V - Nb = \left(\frac{kT}{2aN}\right)^{1/2} V^{3/2} . \tag{2}$$

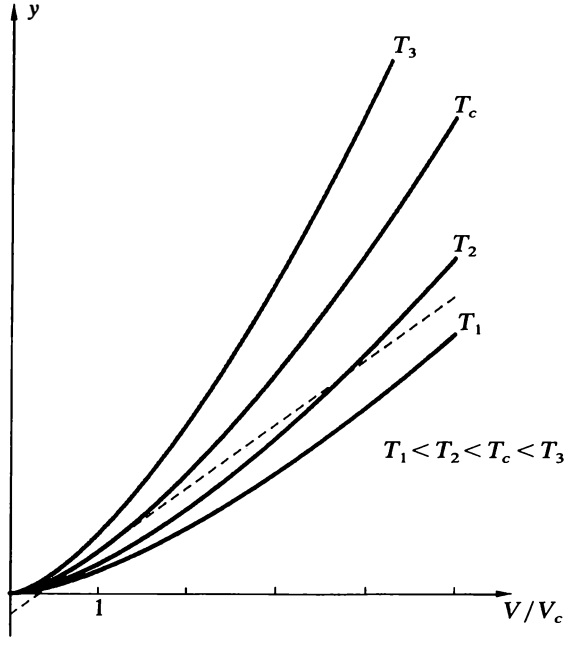


FIGURE 4

Courbe représentant $y_1 = V - Nb$ (tirets) et série de courbes représentant $y_2 = (kT/2aN)^{1/2} V^{3/2}$ (trait plein) pour différentes valeurs de T. Pour $T > T_c$, la courbe tiretée ne coupe pas la courbe en trait plein; pour $T < T_c$ elles ont deux points d'intersection.

En traçant les courbes $y_1 = V - Nb$ et $y_2 = (kT/2aN)^{1/2} V^{3/2}$ sur un même graphique (fig. 4), on se convainc aisément que l'équation (2) possède zéro ou deux racines en V suivant que T est

.../...

Considérons une isotherme particulière à la température T. Pour les grandes valeurs de V, la pression est une fonction décroissante de V ($\partial p/\partial V<0$), à variation relativement lente; ceci correspond à une forte compressibilité, caractéristique d'un *gaz*[17]. Dans l'autre limite, celle des petites valeurs de V, la pression est encore une fonction décroissante de V, mais cette fois à variation très rapide : il faut une très forte augmentation de pression pour faire diminuer le volume de façon appréciable. Le fluide est alors presque incompressible, c'est un *liquide*[18]. Ce passage du gaz au liquide est tout à fait compréhensible. Aux grands volumes en effet, les molécules sont en moyenne très éloignées les unes des autres et l'énergie potentielle d'interaction (G.15) est négligeable; l'énergie cinétique domine, le système est gazeux. Lorsque le volume diminue, l'énergie potentielle augmente; aux faibles volumes elle est du même ordre de grandeur que l'énergie cinétique : le système devient liquide.

Les choses se compliquent pour les valeurs intermédiaires du volume. On distingue en effet sur la figure 3 deux domaines de température. Pour les hautes températures ($T>T_c$), p reste une fonction décroissante de V pour toutes les valeurs de V. Lorsqu'on suit l'isotherme, c'est-à-dire lorsqu'on augmente V à température constante, on passe de façon progressive d'un régime de faible compressibilité (liquide) à un régime de forte compressibilité (gaz). Il n'y a aucune discontinuité entre les deux régimes : le système reste homogène tout au long de l'opération; on dit qu'on a affaire à un *« fluide supercritique »*. Pour les basses températures ($T<T_c$), il existe un intervalle de valeurs de V où $\partial p/\partial V>0$, c'est-à-dire un domaine à compressibilité négative. Nous savons (III, § II.D.3) qu'un système n'est pas stable dans ces conditions. Dans cette région, l'isotherme de van der Waals ne peut pas représenter un état d'équilibre du fluide; une analyse plus détaillée est nécessaire.

La température qui sépare les deux domaines s'appelle la *température critique,* l'isotherme correspondante *l'isotherme critique;* les coordonnées V_c et p_c du point à tangente horizontale sur l'isotherme critique sont *le volume et la pression critiques.* Les valeurs des coordonnées critiques T_c, V_c, p_c sont reliées aux paramètres a et b de l'équation de van der Waals[19],[20] :

$$T_c=\frac{8a}{27bk}\,,\quad V_c=3Nb\,,\quad p_c=\frac{a}{27b^2}\,. \tag{G.29}$$

---

supérieure ou inférieure à une valeur limite T_c. Cette valeur limite s'obtient lorsque les deux courbes sont tangentes; elles vérifient alors

$$1=\frac{3}{2}\left(\frac{kT}{2aN}\right)^{1/2}V^{1/2} \tag{3}$$

et l'équation (2), ce qui donne

$$T_c=\frac{8a}{27bk}\,. \tag{4}$$

Notons que $(\partial p/\partial V)_T$ est toujours négatif, sauf entre les deux extremums lorsqu'ils existent. Ceux-ci sont d'autant plus écartés que la température est plus basse. On déduit facilement de ces résultats l'allure des isothermes de la figure 3.

17. Dans la limite des grands volumes, l'équation de van der Waals s'identifie en fait avec l'équation des gaz parfaits.

18. Dans ce domaine les approximations que nous avons faites pour établir l'équation de van der Waals sont mal justifiées. Elle prédit pourtant de façon qualitativement correcte le comportement du liquide, car elle prend en compte les caractéristiques essentielles de l'interaction entre molécules.

19. Pour déterminer les coordonnées critiques, on peut utiliser les résultats de la note 16. On peut aussi remarquer que T_c, V_c, p_c doivent vérifier l'équation d'état (G.20) et annuler les dérivées $\partial p/\partial V$ et $\partial^2 p/\partial V^2$.

.../...

Le rapport $p_c V_c / NkT_c$, appelé parfois « facteur de compression au point critique », est une grandeur sans dimension. Dans l'approximation de van der Waals, il est indépendant de a et b et a donc la même valeur pour tous les fluides :

$$\frac{p_c V_c}{NkT_c} = \frac{3}{8} = 0{,}375 \,. \tag{G.30}$$

Les valeurs expérimentales (tableau I) donnent un rapport pratiquement constant d'un fluide à l'autre (plus faible cependant pour les fluides polaires, tels que l'eau ou l'ammoniac), mais plutôt de l'ordre de 0,30.

TABLEAU I

Coordonnées et facteur de compression critiques pour quelques fluides (valeurs tirées de N. Lange, *Lange's Handbook of Chemistry*, McGraw-Hill (1967), p. 506, où on trouvera de nombreux autres exemples). On note que chacune des coordonnées critiques varie considérablement d'un fluide à l'autre, alors que le facteur de compression critique reste remarquablement stable.

fluide	T_c (K)	v_c (cm³/mole)	p_c (atm)	$p_c V_c / NkT_c$
He	5,2	57,8	2,26	0,30
Ne	44,4	41,7	26,9	0,31
H_2	33,1	65,0	12,8	0,31
N_2	126,1	90,1	33,5	0,29
O_2	154,4	74,4	49,7	0,29
CH_4	191,1	98,8	45,8	0,29
CO_2	304,2	94,0	72,9	0,27
H_2O	647,4	56,3	218,3	0,23
NH_3	405,5	72,3	111,3	0,24

Pour étudier le domaine $T < T_c$, il nous faut revenir à la fonction de partition, ou, ce qui est équivalent, à l'énergie libre.

Les coordonnées critiques sont donc déterminées par les trois équations

$$p_c = \frac{NkT_c}{V_c - Nb} - \frac{N^2 a}{V_c^2} \,, \tag{1}$$

$$0 = -\frac{NkT_c}{(V_c - Nb)^2} + \frac{2N^2 a}{V_c^3} \,, \tag{2}$$

$$0 = \frac{2NkT_c}{(V_c - Nb)^3} - \frac{6N^2 a}{V_c^4} \,. \tag{3}$$

20. Si on définit les coordonnées réduites (sans dimension)

$$\widetilde{T} \widehat{=} T/T_c, \quad \widetilde{V} \widehat{=} V/V_c, \quad \widetilde{p} \widehat{=} p/p_c \,, \tag{1}$$

l'équation d'état de van der Waals (G.21) s'écrit

$$\left(\widetilde{p} + \frac{3}{\widetilde{V}^2}\right)(3\widetilde{V} - 1) = 8\widetilde{T} \,. \tag{2}$$

Cette équation réduite est la même pour tous les fluides de van der Waals : les paramètres a et b, caractéristiques du fluide, n'y apparaissent plus explicitement. Deux fluides différents (c'est-à-dire dont les paramètres critiques T_c, V_c, p_c ont des valeurs différentes) qui sont tous deux correctement décrits par l'équation d'état de van der Waals vérifient la *« loi des états correspondants »* : si l'on fixe pour les deux fluides les mêmes valeurs de $\widetilde{T}$ et $\widetilde{V}$, la valeur de $\widetilde{p}$ est aussi la même.

2. ÉTUDE DE L'ÉNERGIE LIBRE

Examinons comment varie l'énergie libre (G.18) lorsqu'on fait varier le volume à température constante. Puisque la pression est une grandeur positive, $\partial F/\partial V = -p$ est toujours négative : à température fixée, F est une fonction constamment décroissante du volume. La décroissance est rapide aux faibles volumes, qui correspondent aux hautes pressions, et lente aux grands volumes, qui correspondent aux basses pressions.

Si $T > T_c$, la pression p trouvée en (G.20) est une fonction partout décroissante de V, de sorte que

$$\frac{\partial^2 F}{\partial V^2} = -\frac{\partial p}{\partial V} > 0 \quad (T > T_c)\,. \tag{G.31}$$

La courbe représentant F en fonction de V à T constante tourne donc sa concavité vers le haut; en chaque point la tangente se situe au-dessous de la courbe, et toute corde joignant deux points de la courbe se trouve au-dessus de l'arc de courbe correspondant (fig. 5).

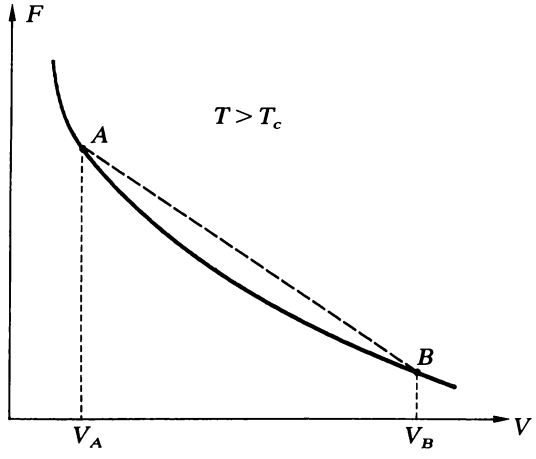


FIGURE 5

Courbe représentative de l'énergie libre F en fonction de V à T constante, pour $T > T_c$. La corde AB joignant deux points quelconques de la courbe se trouve au-dessus de l'arc $\widehat{AB}$ correspondant.

Pour $T < T_c$ au contraire, la courbe $F(V)$ présente deux points d'inflexion correspondant aux deux extremums de l'isotherme, puisque $\partial^2 F/\partial V^2 = -\partial p/\partial V$. Entre ces deux points, la concavité de la courbe est tournée vers le bas (fig. 6); dans cette région, la condition de stabilité $\partial p/\partial V < 0$ n'est pas satisfaite; la courbe $F(V)$ ne peut donc en aucun cas représenter l'énergie libre d'un système en équilibre (cf. III, § II.D.3).

Cependant nous n'avons considéré jusqu'à présent que des systèmes homogènes. Rien n'empêche *a priori* le système de se séparer en deux sous-systèmes, chacun d'eux étant homogène, mais correspondant à deux états macroscopiques différents. Lorsqu'une telle séparation se produit, on dit que le système est constitué de *deux phases en équilibre.* On sait que l'état d'équilibre d'un système, à température et volume fixés, est celui qui rend l'énergie libre minimum (III, § II.D.2). La solution diphasique

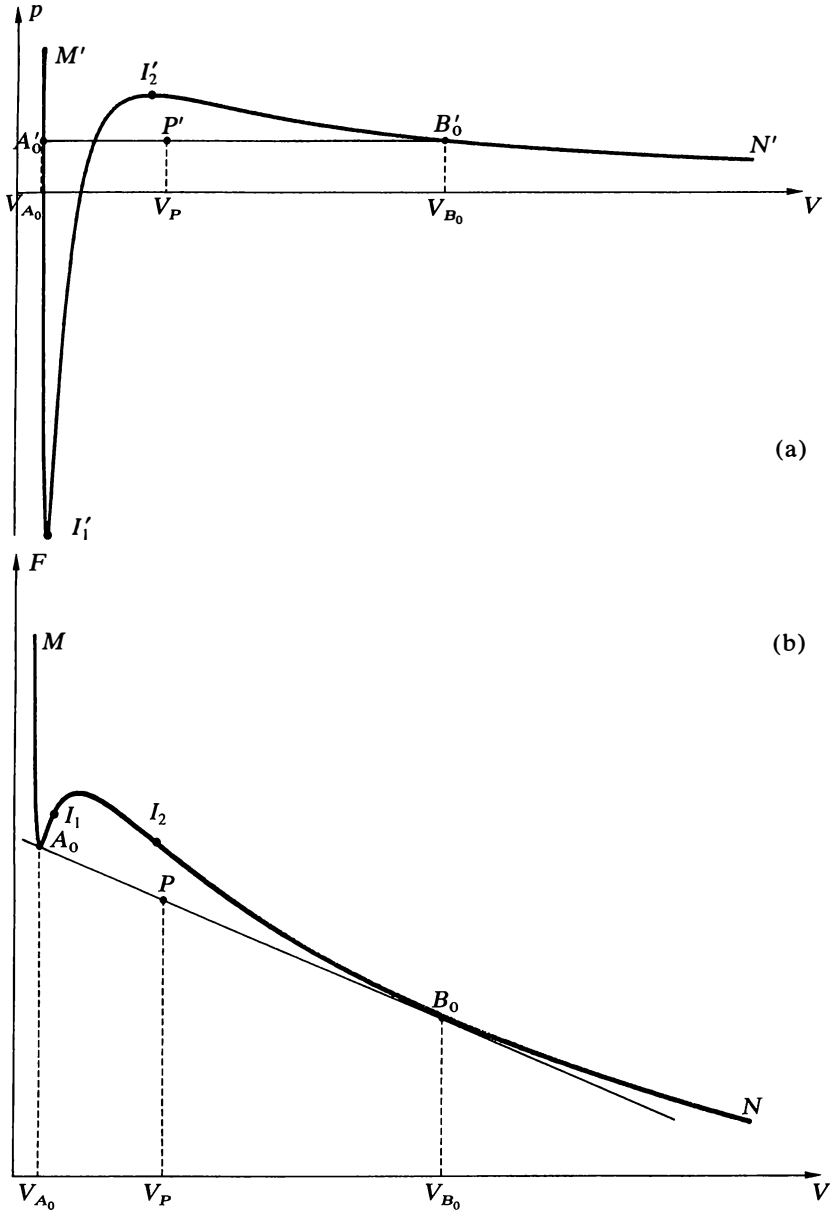


FIGURE 6

Isotherme de van der Waals (a) et énergie libre correspondante (b) à température fixée inférieure à T_c. Parmi toutes les cordes joignant deux points de la courbe représentative de $F(V)$, la plus basse est la bitangente A_0B_0; c'est donc elle qui correspond à l'énergie libre $F_d(V)$ du système séparé en deux phases.

sera donc choisie par le système si l'énergie libre résultante, somme des énergies libres des deux phases (III, § II.C), est plus faible que celle du système homogène. De toutes les séparations possibles en deux phases, c'est évidemment celle qui aura l'énergie libre la plus basse qui constituera l'état d'équilibre du système.

Soit $F_A = F(V_A, T, N)$ l'énergie libre du système homogène des N molécules lorsque, à température T, son volume est V_A (état A) et soit $F_B = F(V_B, T, N)$ son énergie libre lorsque son volume est V_B (état B). Si nous formons un système diphasique comprenant N_A molécules dans l'état A et $N_B = N - N_A$ dans l'état B, le volume occupé sera

$$V = N_A \frac{V_A}{N} + N_B \frac{V_B}{N} , \tag{G.32}$$

et l'énergie libre

$$F_d = N_A \frac{F_A}{N} + N_B \frac{F_B}{N} . \tag{G.33}$$

En éliminant N_A/N $(=1 - N_B/N)$ entre (G.32) et (G.33), on exprime F_d en fonction des paramètres extérieurs V, T et N et des variables internes V_A et V_B :

$$F_d = F_A + \frac{V - V_A}{V_B - V_A} (F_B - F_A) . \tag{G.34}$$

V étant compris entre V_A et V_B (formule (G.32)), (G.34) est l'équation de la corde qui joint le point (V_A, F_A) au point (V_B, F_B) de la courbe $F(V)$ représentant l'énergie libre de van der Waals du système.

Si $T > T_c$, pour V donné, la corde est toujours au-dessus de la courbe (fig. 5); $F_d(V)$ est donc toujours supérieure à $F(V)$ et le système reste homogène. Si au contraire $T < T_c$, du fait du changement de concavité de la courbe $F(V)$, il existe des valeurs de V pour lesquelles on peut trouver des états A et B tels que $F_d(V) < F(V)$. Pour ces valeurs de V, le système se sépare en deux phases; de toutes les cordes possibles, celle qui donne la valeur la plus faible de $F_d(V)$ est la bitangente A_0B_0 (fig. 6). Ainsi, lorsque $T < T_c$, deux situations peuvent se présenter :

— pour $V < V_{A_0}$ ou $V > V_{B_0}$, le système à l'équilibre est homogène et son énergie libre égale à l'énergie libre de van der Waals $F(V)$;

— pour $V_{A_0} < V < V_{B_0}$, le système se sépare en deux phases, chacune correspondant à un état de van der Waals (états A_0 et B_0); l'énergie libre est alors, à température donnée, une fonction linéaire du volume[21] :

$$F_d(V) = F_{A_0} + \frac{V - V_{A_0}}{V_{B_0} - V_{A_0}} (F_{B_0} - F_{A_0}) . \tag{G.35}$$

3. L'ÉQUILIBRE LIQUIDE-VAPEUR

Il est maintenant facile, en analysant la figure 6, de décrire le comportement du système à l'équilibre pour des températures inférieures à la température critique. Plaçons-nous à température fixée $(T < T_c)$ et augmentons le volume de façon suffisamment lente pour que le système passe pratiquement par une succession d'états d'équilibre (cf. complément II.D, § II.3.b).

(i) Tant que V reste inférieur à $V_{A_0}(T)$, le système est dans une phase liquide homogène, son énergie libre suit la courbe en trait plein $F(V)$ (de M à A_0) et sa pression l'isotherme de van der Waals (de M' à A_0').

(ii) Lorsque V devient supérieur à $V_{A_0}(T)$, et jusqu'à la valeur $V_{B_0}(T)$, le système se sépare en deux phases : une *phase liquide* ayant les caractéristiques d'un système homogène occupant un volume V_{A_0}/N par molécule (état A_0) et une *phase*

21. F_d est aussi une fonction de la température : elle en dépend explicitement à travers F_{B_0} et F_{A_0}, mais aussi implicitement par V_{A_0} et V_{B_0}.

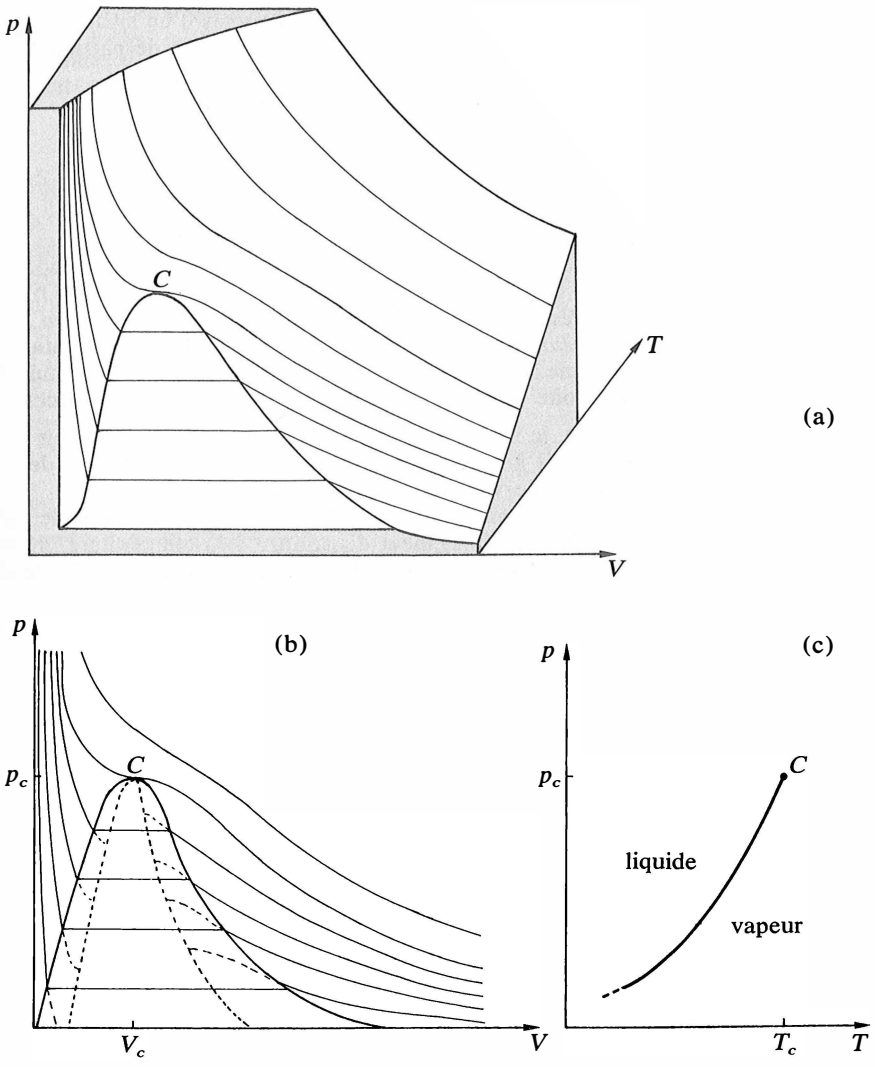


FIGURE 7

Diagramme d'état d'un fluide, résumant son comportement lors de la transition liquide-gaz.

a) Surface d'*équilibre* dans l'espace (p, V, T), ensemble des points représentant les états d'équilibre permis au fluide. Diverses isothermes ont été tracées sur cette surface, ainsi que la courbe (non plane) joignant les extrémités des paliers de liquéfaction.

b) Projection de la surface d'équilibre sur le plan (p, V). La courbe en cloche, dite « *courbe de saturation* » (en trait gras) sépare le domaine où coexistent liquide et vapeur du domaine monophasique, liquide à gauche, gaz à droite. Les courbes en trait plein sont les isothermes. Entre la courbe de saturation et la « courbe spinodale » (en pointillé) se situe la région où peuvent être obtenus des états homogènes métastables (les isothermes y sont représentées en tirets).

c) Projection de la surface d'équilibre sur le plan (p, T). Les paliers de liquéfaction se projettent chacun en un point dont l'ensemble constitue la « *courbe de vaporisation* » (en trait plein). Celle-ci sépare le domaine du liquide homogène (au-dessus et à gauche) de celui du gaz homogène (au-dessous et à droite). Les deux phases ne coexistent à l'équilibre que sur la courbe de vaporisation. Elle est limitée vers le haut par le point critique et vers le bas par le « point triple » (qui n'est pas représenté sur la figure), lié à l'existence d'une phase solide qui apparaît aux basses températures.

gazeuse (on dit aussi *phase vapeur*) ayant les caractéristiques d'un système homogène occupant un volume V_{B_0}/N par molécule (état B_0). Le nombre de particules est

$$N_\ell = N \frac{V_{B_0} - V}{V_{B_0} - V_{A_0}} \tag{G.36}$$

dans la phase liquide et

$$N_v = N \frac{V - V_{A_0}}{V_{B_0} - V_{A_0}} \tag{G.37}$$

dans la phase vapeur[22]. L'énergie libre est alors $F_d(V)$; c'est une fonction linéaire de V, le point représentatif du système suit la corde A_0PB_0 de la figure 6.b. Comme $p = -\partial F_d/\partial V$, *la pression reste constante tant que les deux phases sont en présence;* on l'appelle *pression de vapeur saturante,* elle ne dépend que de la température. Entre A'_0 et B'_0 (fig. 6.a), l'isotherme réelle ne suit pas la courbe déduite de la formule (G.21), mais bien le segment de droite $A'_0P'B'_0$, que l'on appelle *«palier de liquéfaction».*

(iii) Pour $V > V_{B_0}(T)$, le système redevient homogène, mais cette fois en phase gazeuse. Son énergie libre suit à nouveau la courbe en trait plein $F(V)$ (de B_0 vers N) et sa pression l'isotherme de van der Waals (de B'_0 vers N').

Lorsque la température croît, les points A_0 et B_0 se rapprochent; le palier de liquéfaction $A'_0B'_0$ se rétrécit pour finalement disparaître à la température critique T_c, au-dessus de laquelle la distinction entre liquide et vapeur n'a plus véritablement de sens. Les figures 7 résument le comportement général du fluide.

4. CARACTÉRISTIQUES DE LA TRANSITION

La transition liquide-vapeur présente trois propriétés essentielles, qui sont caractéristiques de toute une classe de transitions de phases dites *«transitions du premier ordre».*

a. Coexistence des deux phases

Pour T fixée, inférieure à T_c, les phases gazeuse et liquide coexistent, en équilibre, dans le récipient pour toutes les valeurs de V correspondant au palier de liquéfaction. Si la liquéfaction ou la vaporisation se fait à pression (ou température) donnée, la température (ou la pression) reste constante tout au long du processus[23], tant que les deux phases sont en présence à l'équilibre.

---

22. Le rapport

$$\frac{N_\ell}{N_v} = \frac{V_{B_0} - V}{V - V_{A_0}} \tag{1}$$

peut se lire directement sur l'isotherme de la figure 6.a :

$$\frac{N_\ell}{N_v} = \frac{P'B'_0}{A'_0P'} . \tag{2}$$

Le nombre de particules dans une phase est inversement proportionnel à la distance qui sépare le point représentatif du système du point représentatif de la phase considérée. C'est ce qu'on appelle la «règle du bras de levier».

23. La fonction thermodynamique associée aux variables p et T est (complément III.D) l'enthalpie libre

$$G \widehat{=} F + pV . \tag{1}$$

Dans le domaine de coexistence des deux phases, les variables p et T sont fonctions l'une de l'autre; l'enthalpie libre ne dépend plus que de l'une d'elles, T ou p. Comme T et p, G reste

.../...

b. La transition s'accompagne d'une variation d'entropie

Lorsque n molécules passent, à température et pression constantes, de l'état liquide (état A_0) à l'état vapeur (état B_0), l'entropie du système (cf. (G.24)) augmente de

$$\Delta S \widehat{=} \frac{n}{N}\left[S(B_0) - S(A_0)\right] = nk \ln \frac{V_{B_0} - Nb}{V_{A_0} - Nb} . \tag{G.38}$$

Cet accroissement est positif et proportionnel au nombre n de particules vaporisées. Cette variation d'entropie s'accompagne d'un échange de chaleur avec le thermostat : celui-ci doit fournir au système la *« chaleur latente de transformation »*[24]. On définit, en général, la chaleur latente L *par mole;* elle vaut ici (N_A étant le nombre d'Avogadro, et $R = N_A k$ la constante des gaz parfaits)

$$L = \frac{N_A}{n} T\Delta S = R \ln \frac{V_{B_0} - Nb}{V_{A_0} - Nb} > 0 . \tag{G.39}$$

---

constante tout le long d'un palier de liquéfaction. On vérifie aisément ces propriétés à partir des formules précédentes. Dans le régime diphasique, en effet,

$$G_d = F_d + pV \tag{2}$$

où p est la pression de vapeur saturante

$$p = -\frac{\partial F_d}{\partial V} . \tag{3}$$

En utilisant l'expression (G.35) de F_d, on obtient

$$G_d = \frac{F_{A_0} V_{B_0} - F_{B_0} V_{A_0}}{V_{B_0} - V_{A_0}} ; \tag{4}$$

V_{A_0} et V_{B_0} étant des fonctions de la température seule, G_d ne dépend que de T. Donc, lorsque les deux phases sont en équilibre, l'enthalpie libre par particule dans la phase liquide est égale à l'enthalpie libre par particule dans la phase vapeur; à l'équilibre les potentiels chimiques des deux phases sont égaux (cf. (G.27)) :

$$\mu_\ell = \mu_v . \tag{5}$$

C'est souvent sous cette forme qu'on utilise la condition d'équilibre mutuel de deux phases. Si, par exemple, on exprime μ_ℓ et μ_v en fonction des variables T et p, l'équation (5) détermine la pression de vapeur saturante en fonction de la température.

24. L'enthalpie libre

$$G \widehat{=} F + pV = \overline{E} - TS + pV \tag{1}$$

reste constante au cours de la vaporisation à température et pression constantes (note 23). On a donc ici

$$\Delta G = \Delta \overline{E} - T\Delta S + p\Delta V = 0 . \tag{2}$$

Si W et Q sont le travail et la chaleur fournis de l'extérieur au fluide au cours de l'opération, l'accroissement de son énergie vaut (complément II.D, § II.2)

$$\Delta \overline{E} = W + Q . \tag{3}$$

La vaporisation à T et p constantes est une succession d'états d'équilibre, la transformation est réversible, de sorte que (complément II.D, § III.2)

$$W = -p\Delta V \tag{4}$$

et

$$Q = T\Delta S . \tag{5}$$

Il faut *fournir de la chaleur au système pour le faire passer, à température et pression constantes, de l'état liquide à l'état vapeur.* V_{B_0} et V_{A_0} étant des fonctions de la température seule, la chaleur latente ne dépend que de la température; elle s'annule à la température critique (puisqu'alors $V_{B_0} = V_{A_0}$).

c. Existence d'états métastables

Reprenons les courbes de la figure 6. Sur l'isotherme de van der Waals, entre le point A'_0 et le minimum I'_1, et entre le maximum I'_2 et B'_0, la pression est une fonction décroissante du volume; entre ces points, la compressibilité est positive et la condition de stabilité $\partial p/\partial V < 0$ est satisfaite. Bien sûr nous avons montré que l'état d'équilibre du système ne suivait pas l'isotherme de van der Waals entre A'_0 et B'_0, néanmoins les points des arcs d'isotherme $\widehat{A'_0 I'_1}$ et $\widehat{I'_2 B'_0}$ peuvent être atteints par le système : ils correspondent à des *états métastables.* En effet, comme on le voit sur la figure 6.b, entre A_0 et I_1 d'une part et entre I_2 et B_0 d'autre part, la courbe représentant l'énergie libre de van der Waals tourne encore sa concavité vers le haut. Toute « corde locale », c'est-à-dire toute corde joignant deux points situés entre A_0 et I_1, ou entre I_2 et B_0, passe donc au-dessus de la courbe. Ce sont les cordes joignant un point de $\widehat{A_0 I_1}$ à un point de $\widehat{I_2 B_0}$ qui donnent des énergies libres toujours inférieures à celles de la courbe, montrant ainsi que les arcs $\widehat{A_0 I_1}$ et $\widehat{I_2 B_0}$ ne correspondent pas à des états d'équilibre vraiment stable. Si donc le système parvient, d'une manière ou d'une autre, dans un état correspondant à un point de l'arc d'isotherme $\widehat{A'_0 I'_1}$ ou $\widehat{I'_2 B'_0}$, les fluctuations de faible amplitude dont il est constamment le siège ne l'éloigneront pas de cet état, qui est « localement stable ». Par contre, une fluctuation d'amplitude suffisante pourra « percevoir » la possibilité de réaliser un état plus stable par séparation en deux sous-systèmes, l'un décrit par un point de $\widehat{A'_0 I'_1}$, l'autre par un point de $\widehat{I'_2 B'_0}$; le système quittera alors son état métastable. Mais, comme la probabilité d'une fluctuation décroît très rapidement quand son amplitude croît, un état métastable peut se maintenir longtemps.

En pratique, lorsqu'on augmente lentement le volume offert au liquide à température constante, le système pourra rester homogène au-delà de A'_0, dans un état métastable appelé *« liquide surchauffé »;* mais en aucun cas on ne pourra dépasser le point I'_1 de l'isotherme, limite de métastabilité, au-delà duquel le système devient vraiment instable ($\partial p/\partial V > 0$). De la même façon, si on diminue lentement le volume du gaz à température fixée, on pourra garder le système à l'état gazeux pour des volumes inférieurs à V_{B_0}, dans un état métastable appelé *« vapeur sursaturée ».* Évidemment, toute perturbation importante (par exemple l'introduction d'une goutte de liquide dans une vapeur sursaturée ou d'une bulle de vapeur dans un liquide surchauffé, la présence d'impuretés, ou même un simple choc) fera disparaître la métastabilité et déclenchera l'évolution rapide du système vers son véritable état d'équilibre, l'état diphasique.

Les limites de métastabilité sont déterminées par les extremums des isothermes de van der Waals; ces extremums sont situés sur la « courbe spinodale », représentée en pointillés sur la figure 7.b

IV. CONCLUSION

Ainsi, l'approximation de van der Waals permet de décrire les propriétés des fluides, et en particulier la transition liquide-gaz. Les fluides réels ont un comportement semblable à celui du fluide de van der Waals que nous avons analysé ci-dessus. Évidemment, comme toute théorie basée sur une approximation de champ moyen, la théorie de van der Waals simplifie considérablement la réalité physique et on ne peut espérer qu'elle donne des résultats quantitativement corrects. Néanmoins, elle prend

en compte les deux principales caractéristiques des forces intermoléculaires, répulsion à très courte distance et attraction à moyenne distance, qui sont essentielles (et suffisantes) pour la description qualitative de la transition. Comme nous l'avons vu, c'est plus particulièrement du côté du liquide que les approximations sont mal justifiées, et c'est dans cette région que l'on doit s'attendre à ce que les valeurs expérimentales s'écartent le plus des valeurs calculées; le comportement qualitatif y reste cependant correct.

Il existe une autre méthode simple (au moins dans son principe) pour l'étude des fluides; c'est une méthode de perturbation basée sur un développement en puissances de la densité (complément V.C). Les approximations y sont beaucoup mieux contrôlées, mais ne sont justifiées qu'à faible densité, c'est-à-dire, là encore, dans la région du gaz.

COMPLÉMENT III.H

Systèmes de particules chargées

I. Position du problème
1. Les systèmes physiques concernés
2. Fonction de partition et énergie libre

II. Potentiel électrostatique moyen autour d'un ion dans l'approximation de Debye-Hückel
1. Mise en équations
2. Solution
3. Discussion

III. Propriétés thermodynamiques
1. Énergie moyenne
2. Énergie libre
3. Pression

I. POSITION DU PROBLÈME

Les méthodes que nous avons développées pour étudier les gaz réels (complément III.G; voir aussi complément V.C) ne sont pas utilisables pour les systèmes constitués de particules chargées : leurs interactions (loi de Coulomb) sont à trop longue portée pour que les intégrales à calculer convergent. De tels systèmes nécessitent donc une étude particulière.

1. LES SYSTÈMES PHYSIQUES CONCERNÉS

Le cas le plus simple est celui d'un *plasma complètement ionisé,* c'est-à-dire d'un gaz constitué uniquement de particules chargées, ou ions[1]. Celles-ci peuvent être de plusieurs types : nous noterons N_a le nombre de particules de type (a) ($N_a \gg 1$), et q_a la charge de chacune d'elles (q_a, positif ou négatif, est un multiple entier de la charge élémentaire e). Le plasma, dans son ensemble, est *électriquement neutre,* c'est-à-dire que

$$\sum_{(a)} N_a q_a = 0 \,. \tag{H.1}$$

Les recombinaisons éventuelles entre particules de charges opposées sont supposées improbables et seront négligées; les N_a sont donc des paramètres extérieurs fixés, au même titre que le volume V du récipient et la température T du thermostat avec lequel le plasma est en équilibre thermique. Deux particules (i) et (j) du plasma, situées aux points $\vec{r}_i$ et $\vec{r}_j$ et portant les charges q_i et q_j, interagissent suivant la loi de Coulomb : leur énergie d'interaction[2] vaut

$$u_{ij} = \frac{1}{4\pi\varepsilon_o} \frac{q_i q_j}{|\vec{r}_i - \vec{r}_j|} \,. \tag{H.2.a}$$

---

1. Certains des constituants du plasma peuvent être des électrons. Dans tout ce qui suit nous supposons cependant que le plasma est *non dégénéré,* c'est-à-dire que les densités y sont suffisamment faibles (ou la température suffisamment haute) pour que le système puisse être correctement décrit par la mécanique classique. Ceci suppose, pour chaque type (a) de particules (III, § IV.B.2.b),

$$kT \gg \frac{\hbar^2}{m_a} \left(\frac{N_a}{V}\right)^{2/3}$$

C'est en général pour les électrons, dont la masse est la plus faible, que cette condition est le plus difficile à satisfaire.

2. Nous négligeons les interactions autres qu'électrostatiques. Elles sont à courte portée et négligeables aux faibles densités que nous considérons ici.

La méthode que nous allons développer s'applique aussi aux *solutions d'électrolytes forts,* c'est-à-dire aux solutions des bases fortes, des acides forts ou de leurs sels. De tels corps, en solution dans l'eau, sont totalement ionisés et, du point de vue des particules chargées, le problème est analogue à celui du plasma : en première approximation le rôle du solvant consiste simplement à réduire l'énergie d'interaction (H.2.a) entre deux ions d'un facteur égal à sa permittivité relative ε_r (pour l'eau, $\varepsilon_r = 80{,}4$ à 20°C).

Dans toute la suite, nous écrirons donc

$$u_{ij} = \frac{1}{4\pi\varepsilon} \frac{q_i q_j}{|\vec{r}_i - \vec{r}_j|} , \tag{H.2.b}$$

remplaçant la constante diélectrique ε par ε_0 dans le cas d'un plasma, et par $\varepsilon_0 \varepsilon_r$ pour une solution d'électrolyte fort.

2. FONCTION DE PARTITION ET ÉNERGIE LIBRE

La fonction de partition Z du système peut se calculer dans le cadre de la mécanique classique. L'intégration sur les impulsions des N particules se découple de l'interaction sur leur position; la première donne la fonction de partition Z_p du gaz parfait correspondant[3], de sorte que

$$Z = \frac{Z_p}{V^N} \iint \dots \int \mathrm{d}^3 r_1 \, \mathrm{d}^3 r_2 \dots \mathrm{d}^3 r_N \, \mathrm{e}^{-U(\vec{r}_1, \vec{r}_2, \dots, \vec{r}_N)/kT} , \tag{H.3}$$

où U est l'énergie potentielle d'interaction coulombienne entre les particules. Il est commode de l'écrire

$$U(\vec{r}_1, \vec{r}_2, \dots, \vec{r}_N) = \frac{1}{2} \sum_{i=1}^{N} q_i \varphi_i(\vec{r}_i) , \tag{H.4}$$

où $\varphi_i(\vec{r}_i)$ est le potentiel électrostatique créé à la position $\vec{r}_i$ de l'ion (i) par les $(N-1)$ autres :

$$\varphi_i(\vec{r}_i) = \frac{1}{4\pi\varepsilon} \sum_{j \neq i} q_j \frac{1}{|\vec{r}_i - \vec{r}_j|} ; \tag{H.5}$$

le facteur 1/2 qui apparaît dans (H.4) évite de compter deux fois l'énergie d'interaction de chaque paire d'ions.

Calculer Z directement, comme nous l'avons fait pour le gaz de van der Waals, conduit à une impasse. En effet le potentiel effectif moyen que l'on introduit fait intervenir l'intégrale, sur toutes les positions $\vec{r}_j$, du potentiel d'interaction entre deux particules (formule (G.13) du complément III.G); du fait de la longue portée de l'interaction coulombienne (H.2), l'intégrale correspondante est ici divergente. Mais la présence de charges électriques de signes contraires permet de résoudre le problème : une particule donnée attire dans son voisinage des particules de charge opposée à la sienne et ces dernières « écrantent » l'interaction avec les particules plus éloignées; le potentiel effectif moyen est donc, ici aussi, à courte portée. Cependant il n'est plus possible de considérer seulement deux particules pour calculer l'énergie d'interaction moyenne.

La *méthode de Debye et Hückel* permet de déterminer directement, de *façon approchée,* la valeur moyenne des $\varphi_i(\vec{r}_i)$ et par conséquent l'énergie moyenne $\overline{E}$. Pour

3. Les degrés de liberté internes éventuels de chaque type d'ions se découplent aussi des variables de position et contribuent à Z_p (cf. complément III.B).

une fois, nous procéderons donc « à l'envers » : nous obtiendrons l'énergie libre F à partir de $\bar{E}$, ces deux quantités étant, comme toujours, reliées par

$$\bar{E} = -\frac{\partial \ln Z}{\partial \beta} = \frac{\partial}{\partial \beta}(\beta F) = -T^2 \frac{\partial}{\partial T}\left(\frac{F}{T}\right) . \tag{H.6}$$

II. POTENTIEL ÉLECTROSTATIQUE MOYEN AUTOUR D'UN ION DANS L'APPROXIMATION DE DEBYE-HÜCKEL

L'approximation de Debye et Hückel permet de résoudre le problème lorsque le système s'écarte peu de l'état parfait, c'est-à-dire dans le cas où l'énergie cinétique moyenne du système est bien supérieure à son énergie d'interaction moyenne[4].

1. MISE EN ÉQUATIONS

Considérons un ion particulier, de charge q, et choisissons sa position comme origine des coordonnées. Cet ion attire dans son voisinage les ions de charge opposée et repousse ceux de même signe. Il crée ainsi autour de lui un *« nuage ionique »* à symétrie sphérique, en moyenne, mais de charge non uniforme. Nous allons traiter ce nuage comme une *distribution continue de charge,* caractérisée par sa *densité volumique moyenne* $\rho(r)$. Nous appellerons $\psi(r)$ le *potentiel électrostatique moyen* créé par l'ion central et son nuage.

Pour déterminer les deux fonctions inconnues $\rho(r)$ et $\psi(r)$, il nous faut deux équations. La première est purement électrostatique; c'est l'équation de Poisson qui relie, en chaque point, potentiel et densité de charge :

$$\Delta\psi(r) + \frac{\rho(r)}{\varepsilon} = 0 , \tag{H.7}$$

où Δ est l'opérateur laplacien.

La deuxième est statistique; nous allons *supposer que chaque ion du nuage évolue dans le potentiel moyen* $\psi(r)$. Une particule chargée de type (a), située au point $\vec{r}$, possède alors l'énergie potentielle $q_a\psi(r)$. La probabilité de trouver un ion de ce

---

4. L'énergie cinétique moyenne par particule est de l'ordre de kT, alors que l'énergie moyenne d'interaction entre deux ions voisins, séparés en moyenne par la distance d, vaut environ

$$|\bar{u}_{ij}| \sim \frac{q^2}{4\pi\varepsilon d} . \tag{1}$$

La validité de l'approximation de Debye-Hückel suppose donc

$$kT \gg \frac{q^2}{4\pi\varepsilon d} . \tag{2}$$

Si le système comporte au total N ions dans le volume V,

$$d = \left(\frac{V}{N}\right)^{1/3} , \tag{3}$$

et la condition (2) devient

$$\frac{N}{V} \ll \left(\frac{4\pi\varepsilon kT}{q^2}\right)^3 . \tag{4}$$

Le système s'écartera donc d'autant moins de l'état parfait que sa densité sera plus faible ou que sa température sera plus haute.

type au point $\vec{r}$, avec une vitesse quelconque, est donc proportionnelle à $\exp[-q_a\psi(r)/kT]$. On en déduit immédiatement la densité moyenne $n_a(r)$ d'ions du type (a) au point $\vec{r}$:

$$n_a(r)=\frac{N_a}{V}\,e^{-q_a\psi(r)/kT}\;; \tag{H.8}$$

le « préfacteur » N_a/V a été choisi égal à la densité moyenne des ions de type (a) dans le récipient, car loin de l'origine $\psi(r)$ tend vers zéro (l'influence de l'ion central ne se fait plus sentir) et on doit retrouver une distribution de charge homogène, c'est-à-dire pour les ions (a) une densité uniforme égale à N_a/V[5].

La densité totale de charge $\rho(r)$ est donc donnée en fonction du potentiel par

$$\rho(r)=\sum_{(a)} q_a n_a(r)=\sum_{(a)} q_a\,\frac{N_a}{V}\,e^{-q_a\psi(r)/kT}\,. \tag{H.9}$$

Les équations (H.7) et (H.9) constituent le *système « autocohérent » d'équations* que doivent satisfaire le potentiel et la distribution ionique dans l'approximation de Debye-Hückel.

2. SOLUTION

L'énergie d'interaction étant faible devant l'énergie d'agitation thermique, on peut développer les exponentielles de l'équation (H.9). En se limitant au premier ordre en $q_a\psi/kT$, on obtient[6]

$$\rho(r)\simeq\sum_{(a)} q_a\,\frac{N_a}{V}\left(1-\frac{q_a\psi(r)}{kT}\right). \tag{H.10}$$

La condition de neutralité électrique (H.1) implique que le premier terme est nul, de sorte que

$$\rho(r)\simeq-\frac{\psi(r)}{VkT}\sum_{(a)} q_a^2 N_a\,. \tag{H.11}$$

On reporte cette expression dans (H.7) pour obtenir

$$\Delta\psi-\frac{\psi}{\ell_D^2}=0\,, \tag{H.12}$$

---

5. En toute rigueur, cette densité n'est plus tout à fait égale à N_a/V puisque l'ion origine crée un excès ou un défaut (suivant le signe de q_a) de particules de type (a) dans son voisinage. Mais cet excès, ou ce défaut, reste de l'ordre de l'unité (cf. note 8) et donc négligeable devant $N_a(\gg 1)$.

6. L'équation de Poisson entre potentiel et densité est valable pour n'importe quelle distribution de charge; étant linéaire, elle reste vraie pour les valeurs moyennes $\rho(r)$ et $\psi(r)$; la relation (H.7) est donc exacte. Par contre l'équation (H.9) est approchée puisque, schématiquement, on y a remplacé la valeur moyenne d'une exponentielle par l'exponentielle de la valeur moyenne. Cependant, comme on linéarise ensuite, la formule (H.10) donne correctement le premier terme du développement de $\rho(r)$ en puissances de $q_a\psi(r)/kT$; les résultats que nous allons obtenir sont donc exacts dans la limite où les densités N_a/V sont faibles; plus précisément, nous cherchons le premier terme correctif en N_a/V à la loi des gaz parfaits, et la méthode que nous utilisons fournit ce premier terme correctif exactement.

Il est intéressant de noter qu'on aboutirait à des incohérences si on résolvait les équations (H.7) et (H.9) sans linéariser la seconde (voir par exemple J.O'M. Bockris et A.K.N. Reddy, *Modern Electrochemistry*, volume 1 (1977), paragraphe 3.7, Plenum, New York.

où nous avons posé

$$\ell_D \widehat{=} \left[\frac{\varepsilon V k T}{\sum_{(a)} q_a^2 N_a} \right]^{1/2} ; \tag{H.13}$$

ℓ_D, qui a les dimensions d'une longueur, s'appelle la *longueur de Debye* ou *longueur d'écran*.

Nous cherchons des solutions à symétrie sphérique : ψ ne dépend que de la distance r à l'origine. L'équation (H.12) se réduit alors à l'équation différentielle

$$\frac{1}{r} \frac{\mathrm{d}^2}{\mathrm{d}r^2} [r\psi] = \frac{\psi}{\ell_D^2} , \tag{H.14}$$

dont la solution générale est de la forme[7]

$$\psi(r) = \frac{A}{r} \mathrm{e}^{-r/\ell_D} + \frac{B}{r} \mathrm{e}^{r/\ell_D} . \tag{H.15}$$

Pour $r \to \infty$, $\psi(r)$ doit tendre vers zéro, ce qui implique $B = 0$. A l'autre limite, $r \to 0$, on doit retrouver le potentiel d'une charge ponctuelle q (non écrantée) située au point O. En définitive,

$$\psi(r) = \frac{q}{4\pi\varepsilon r} \mathrm{e}^{-r/\ell_D} , \tag{H.16}$$

d'où l'on déduit par (H.11) ou (H.7) la densité de charge associée :

$$\rho(r) = -\frac{q}{4\pi\ell_D^2} \frac{1}{r} \mathrm{e}^{-r/\ell_D} . \tag{H.17}$$

3. DISCUSSION

Ainsi, l'influence d'un ion particulier se fait sentir sur une distance de quelques longueurs de Debye; au-delà, sa charge est totalement « écrantée » par le nuage ionique qu'il attire autour de lui; ℓ_D représente donc l'ordre de grandeur du rayon de ce nuage. On vérifie aisément[8] que la charge totale portée par le nuage est exactement égale et opposée à celle de l'ion central.

Reste à discuter la validité du calcul. L'approximation qui consiste à traiter les

---

7. Il suffit de poser

$$\psi(r) = \frac{u(r)}{r} \tag{1}$$

pour mettre (H.14) sous la forme

$$u'' - \frac{u}{\ell_D^2} = 0 . \tag{2}$$

La solution générale de cette équation est

$$u = A \mathrm{e}^{-r/\ell_D} + B \mathrm{e}^{r/\ell_D} , \tag{3}$$

d'où le résultat (H.15).

8. La charge totale Q du nuage entourant l'ion de charge q s'obtient immédiatement à partir de la formule (H.17) :

$$Q = \int \rho(r) \mathrm{d}^3 r = -\frac{q}{4\pi\ell_D^2} \int_0^\infty \mathrm{d}r \, 4\pi r^2 \frac{\mathrm{e}^{-r/\ell_D}}{r} = -q .$$

particules chargées comme une distribution volumique continue n'a évidemment de sens que si la longueur caractéristique ℓ_D sur laquelle varie $\rho(r)$ est grande devant la distance moyenne d entre deux ions. La méthode, pour être cohérente, suppose donc

$$\ell_D \gg d \,. \tag{H.18}$$

Or
$$d = \left(\frac{V}{N}\right)^{1/3} \tag{H.19}$$

et, si $\tilde{q}$ est l'ordre de grandeur des charges présentes, (H.13) donne

$$\ell_D \simeq \frac{(\varepsilon kT)^{1/2}}{\tilde{q}}\, d^{3/2} \,. \tag{H.20}$$

En reportant dans (H.18), on obtient finalement

$$d \gg \frac{\tilde{q}^2}{\varepsilon kT} \,, \tag{H.21.a}$$

ou
$$kT \gg \frac{\tilde{q}^2}{\varepsilon d} \,. \tag{H.21.b}$$

La condition (H.21.b) exprime simplement que l'énergie cinétique ($\sim kT$) d'une particule est grande devant l'énergie potentielle d'interaction avec l'ion voisin ($\sim q^2/4\pi\varepsilon d$); autrement dit, l'énergie cinétique du système doit être grande devant son énergie potentielle, c'est-à-dire que le système doit s'écarter peu de l'état parfait (voir paragraphe III.1 ci-dessous et note 4).

Prenons l'exemple d'un système constitué de deux types d'ions seulement, de charges $+q$ et $-q$. Il est alors commode d'écrire ℓ_D sous l'une des deux formes suivantes :

$$\ell_D = \frac{d}{\sqrt{4\pi}} \sqrt{\frac{d}{q^2/4\pi\varepsilon kT}} \,, \tag{H.22.a}$$

ou

$$\ell_D = \frac{d}{\sqrt{4\pi}} \sqrt{\frac{kT}{q^2/4\pi\varepsilon d}} \,. \tag{H.22.b}$$

La deuxième expression fait apparaître le rapport de deux énergies : énergie thermique sur énergie d'interaction coulombienne. ℓ_D est une fonction croissante de ce rapport : les charges sont d'autant mieux réparties que leur énergie cinétique est plus grande. Dans l'expression (H.22.a) c'est le rapport de deux longueurs qui intervient : distance moyenne entre les ions sur *« longueur de Landau »* $q^2/4\pi\varepsilon kT$. La longueur de Landau est la distance minimum d'approche de deux ions de même charge q lancés l'un vers l'autre avec l'énergie kT.

Terminons ce paragraphe par quelques estimations numériques. Prenons tout d'abord le cas d'un plasma contenant deux types de particules de charges $+e$ et $-e$. En faisant intervenir l'unité atomique naturelle de longueur, le rayon de Bohr a_0, la longueur de Landau peut s'écrire

$$\frac{e^2}{4\pi\varepsilon_0 kT} = 2a_0 \times \frac{e^2}{8\pi\varepsilon_0 a_0} \times \frac{1}{kT} \,. \tag{H.23}$$

Le facteur $e^2/8\pi\varepsilon_0 a_0$ est l'énergie de Rydberg, qui vaut 13,6 eV; à la température ordinaire $T_0 \simeq 300$ K, kT_0 vaut 1/40 eV, de sorte que

$$\frac{e^2}{4\pi\varepsilon_0 kT} \simeq 1 \times 13{,}6 \times 40 \times \frac{300}{T} \text{ Å} \simeq \frac{160\,000}{T} \text{ Å} \,. \tag{H.24}$$

Si on se rappelle (complément I.A, § III.2.b) que la distance moyenne entre particules

d'un gaz pris dans les conditions normales est de l'ordre de 30 Å, on voit que la condition (H.21.a) ne sera vérifiée qu'à très haute température ou à très faible densité (à 3 000 K, la longueur de Landau est encore supérieure à 50 Å!).

Prenons maintenant le cas d'une solution électrolytique (toujours avec deux types d'ions de charges $+e$ et $-e$). La longueur de Landau est considérablement réduite du fait de la forte permittivité relative de l'eau. A 300 K on trouve[9]

$$\frac{e^2}{4\pi\varepsilon_0\varepsilon_r kT} \simeq \frac{160\,000}{300\times 80}\ \text{Å} \simeq 7\ \text{Å}\ . \tag{H.25}$$

Une solution molaire contient $2N_A$ ions par litre, ce qui correspond à

$$d \simeq 10\ \text{Å}\ . \tag{H.26}$$

Pour de telles concentrations, la condition (H.21) n'est pas vérifiée[10]. L'approximation ne devient véritablement justifiée que pour des solutions millimolaires ($d \sim 100$ Å) ou plus diluées encore.

III. PROPRIÉTÉS THERMODYNAMIQUES

1. ÉNERGIE MOYENNE

Nous sommes maintenant en mesure de calculer la valeur moyenne de l'énergie d'interaction coulombienne (H.4) du système dans l'approximation de Debye-Hückel. Il faut pour cela déterminer la valeur moyenne $\overline{\varphi}_i$ du potentiel électrostatique (H.5) qui agit sur l'ion (i) (dû à la présence de tous les autres).

Le potentiel électrostatique moyen $\psi_i(r)$ au voisinage de l'ion (i), de charge q_i, est donné par la relation (H.16). Ce potentiel est créé par l'ion (i) mais aussi par l'ensemble des autres ions répartis dans le « nuage »; pour calculer le potentiel moyen $\overline{\varphi}_i$ vu par l'ion (i), il faut retrancher de ψ_i la contribution de l'ion (i) lui-même :

$$\overline{\varphi}_i = \lim_{r\to 0}\left[\psi_i(r) - \frac{q_i}{4\pi\varepsilon r}\right]. \tag{H.27}$$

La formule (H.16) donne alors

$$\overline{\varphi}_i = -\frac{q_i}{4\pi\varepsilon\ell_D}\ . \tag{H.28}$$

On en déduit la valeur moyenne de l'énergie d'interaction coulombienne (H.4) :

$$\overline{U} = -\frac{1}{8\pi\varepsilon\ell_D}\sum_{i=1}^{N} q_i^2 = -\frac{1}{8\pi\varepsilon\ell_D}\sum_{(a)} q_a^2 N_a\ , \tag{H.29}$$

ou encore, en utilisant la définition (H.13) de ℓ_D,

$$\overline{U} = -\frac{VkT}{8\pi\ell_D^3} = -\frac{1}{8\pi\varepsilon^{3/2}}\times\left(\frac{1}{VkT}\right)^{1/2}\times\left(\sum_{(a)} q_a^2 N_a\right)^{3/2}\ . \tag{H.30}$$

Cette énergie potentielle est toujours négative : deux ions de signes opposés sont en effet en moyenne plus proches l'un de l'autre que deux ions de même signe. Mais

9. Dans le cas d'une solution aqueuse, la température ne peut varier qu'entre les températures de solidification et d'ébullition de la solution.

10. Ne pas oublier le facteur 4π par lequel il faut multiplier le résultat (H.25) avant de le comparer à d (cf. condition (H.21.a)).

elle diminue en valeur absolue lorsque la température augmente, car l'agitation thermique tend à uniformiser la distribution des ions[11].

Enfin on vérifie directement sur (H.30) que l'énergie potentielle est faible comparée à l'énergie cinétique ($\sim NkT$) si

$$\ell_D^3 \gg \frac{V}{N} = d^3 \,. \tag{H.31}$$

On retrouve la condition (H.18), et le calcul est bien cohérent dès que cette condition est satisfaite.

L'énergie moyenne totale s'obtient en ajoutant à $\overline{U}$ l'énergie moyenne du gaz parfait correspondant (provenant du facteur Z_p dans (H.3)) : pour un plasma constitué de particules monoatomiques, elle est simplement égale à $3NkT/2$; si certains ions sont polyatomiques, il faut y ajouter les contributions des degrés de liberté de rotation et de vibration (complément III.B). Dans le cas d'une solution électrolytique, se rajoutent encore des termes propres à la solution (voir complément III.F). Si l'on note $E_0(T, V, \{N_a\})$ l'ensemble de ces contributions, on a donc

$$\overline{E} = E_0(T, V, \{N_a\}) - \frac{V}{8\pi(kT)^{1/2}} \left(\sum_{(a)} q_a^2 \frac{N_a}{\varepsilon V} \right)^{3/2} \,. \tag{H.32}$$

2. ÉNERGIE LIBRE

L'énergie libre F peut être déduite de $\overline{E}$ par intégration de la relation (H.6). Pour $T \to +\infty$, la contribution des termes d'interaction coulombienne doit devenir négligeable, ce qui détermine sans ambiguïté la constante d'intégration. On obtient donc[12], si $F_0(T, V, \{N_a\})$ est l'énergie libre associée à $E_0(T, V, \{N_a\})$,

$$F = F_0(T, V, \{N_a\}) - \frac{V}{12\pi} \times \frac{1}{(kT)^{1/2}} \left(\sum_{(a)} q_a^2 \frac{N_a}{\varepsilon V} \right)^{3/2} \,. \tag{H.33}$$

3. PRESSION

Une fois l'énergie libre connue en fonction des paramètres extérieurs V, T, $\{N_a\}$, toutes les autres grandeurs macroscopiques en découlent. Par exemple, la pression

---

11. Malgré la longue portée de l'interaction coulombienne, l'effet d'écran maintient le caractère extensif de l'énergie : si V et tous les N_a sont multipliés par le même facteur λ, $\overline{U}$ est elle-même multipliée par λ.

On notera cependant que le résultat fait intervenir les densités N_a/V à une puissance fractionnaire (3/2); on comprend pourquoi les méthodes basées sur des développements en puissances entières de la densité ne sont pas applicables.

12. La relation (H.6)

$$\frac{\partial}{\partial T} \left(\frac{F}{T} \right) = -\frac{\overline{E}}{T^2} \tag{1}$$

donne, pour le second terme de (H.32),

$$\frac{\partial}{\partial T} \left(\frac{F_c}{T} \right) = \frac{V}{8\pi k^{1/2}} \left(\sum_{(a)} q_a^2 \frac{N_a}{\varepsilon V} \right)^{3/2} \frac{1}{T^{5/2}} \,. \tag{2}$$

L'intégration est immédiate : puisque $F_c \to 0$ si $T \to +\infty$, on trouve

$$F_c = -\frac{V}{12\pi(kT)^{1/2}} \left(\sum_{(a)} q_a^2 \frac{N_a}{\varepsilon V} \right)^{3/2} \,. \tag{3}$$

.../...

(c'est-à-dire l'équation d'état) est donnée par

$$p = -\frac{\partial F}{\partial V} = p_0 - \frac{1}{24\pi} \times \frac{1}{(kT)^{1/2}} \times \left(\sum_{(a)} q_a^2 \frac{N_a}{\varepsilon V} \right)^{3/2} ; \tag{H.34}$$

p_0 est la pression qu'exercerait un système équivalent, mais non chargé, dans les mêmes conditions de volume et de température.

Dans le cas d'un *plasma*, p_0 est simplement la pression d'un mélange de gaz parfaits :

$$p_0 = \frac{kT}{V} \sum_{(a)} N_a , \tag{H.35}$$

et l'équation d'état du plasma s'écrit donc, à l'approximation de Debye-Hückel,

$$p = \frac{kT}{V} \sum_{(a)} N_a - \frac{1}{24\pi} \frac{1}{(kT)^{1/2}} \left(\sum_{(a)} q_a^2 \frac{N_a}{\varepsilon_0 V} \right)^{3/2} . \tag{H.36}$$

Rappelons que ce résultat n'est valable qu'aux faibles valeurs des rapports N_a/V. Il est clair que (H.36) est le début d'un développement de p faisant intervenir les concentrations N_a/V. Ce que montre le calcul basé sur l'approximation de

---

Il existe une autre méthode pour calculer directement l'énergie libre F à partir des $\bar{\varphi}_i$. De l'expression

$$\bar{\varphi}_1 = \frac{\int d^3r_1 \, d^3r_2 \ldots d^3r_N \varphi_1 \, e^{-U/kT}}{\int d^3r_1 \, d^3r_2 \ldots d^3r_N \, e^{-U/kT}} , \tag{4}$$

on déduit par exemple, si on traite formellement q_1 comme une variable continue,

$$\bar{\varphi}_1 = -kT \frac{\partial \ln Z}{\partial q_1} = \frac{\partial F}{\partial q_1} , \tag{5}$$

puisque

$$U = \frac{1}{4\pi\varepsilon} q_1 \sum_{q_j=2}^{N} \frac{q_j}{|\vec{r}_1 - \vec{r}_j|} + \text{termes indépendants de } q_1 . \tag{6}$$

De façon générale, on a

$$\bar{\varphi}_i = \frac{\partial F}{\partial q_i} , \tag{7}$$

de sorte que, à V et T constants, la différentielle de F prend la forme

$$dF \cong \sum_{i=1}^{N} \frac{\partial F}{\partial q_i} dq_i = \sum_{i=1}^{N} \bar{\varphi}_i \, dq_i . \tag{8}$$

Pour déterminer F à partir des $\bar{\varphi}_i$, le plus simple est d'imaginer que la charge de chacun des ions (j) varie de 0 à sa véritable valeur q_j; on pose

$$q_j(\lambda) = \lambda q_j , \quad j = 1, 2, \ldots, N , \tag{9}$$

λ variant de 0 à 1. La valeur moyenne $\bar{\varphi}_i$ est donnée en fonction de λ par (H.28) et (H.13) :

$$\bar{\varphi}_i(\lambda) = -\frac{q_i}{4\pi\varepsilon\ell_D} \lambda^2 . \tag{10}$$

La relation (8) s'écrit donc

$$dF = -\frac{1}{4\pi\varepsilon\ell_D} \left(\sum_{i=1}^{N} q_i^2 \right) \lambda^2 \, d\lambda . \tag{11}$$

On retrouve ainsi la formule (H.33) :

$$F = F_0 - \frac{1}{4\pi\varepsilon\ell_D} \left(\sum_{(a)} q_a^2 N_a \right) \int_0^1 \lambda^2 \, d\lambda = F_0 - \frac{1}{12\pi\varepsilon\ell_D} \left(\sum_{(a)} q_a^2 N_a \right) . \tag{12}$$

Debye-Hückel, c'est que la première correction significative au gaz parfait intervient à la puissance fractionnaire 3/2. Les corrections que nous avons calculées dans l'approximation de van der Waals (complément III.G) ou dans le cadre du développement du viriel (complément V.C), et qui sont dues aux interactions à courte portée, commencent par des termes en $(N/V)^2$; elles sont donc petites, pour N/V faible, devant celles qui interviennent ici, à cause de la longue portée de l'interaction coulombienne.

Dans le cas d'une *solution électrolytique,* il faut ajouter à la pression (H.36) provenant de la présence des substances dissoutes celle qui correspond au solvant. La quantité

$$\Delta p = \frac{kT}{V} \sum_{(a)} N_a - \frac{1}{24\pi} \times \frac{1}{(kT)^{1/2}} \left(\sum_{(a)} q_a^2 \frac{N_a}{\varepsilon V} \right)^{3/2} , \tag{H.37}$$

qui s'ajoute à la pression qu'aurait le solvant pur pour le même volume et la même température, s'appelle la *pression osmotique*[13].

13. La pression osmotique existe dans toutes les solutions, pas seulement celles des électrolytes. Un soluté dont les molécules ne s'ionisent pas se comporte en première approximation (s'il est suffisamment dilué) comme un gaz parfait (premier terme de la formule (H.37), cf. complément III.F, § III.1). Dans le cas des électrolytes, la dissociation en ions a pour effet de diminuer la pression osmotique puisque le deuxième terme de (H.37) est négatif.

COMPLÉMENT III.J

Le ferromagnétisme

I. Le phénomène physique et son origine
   1. Les résultats expérimentaux
      a. Aimantation et température de Curie
      b. La structure en domaines
      c. Le voisinage du point critique
   2. Le problème à résoudre
   3. L'hamiltonien de Heisenberg
   4. Origine physique du ferromagnétisme

II. L'approximation de champ moyen
   1. Principe de l'approximation
      a. L'hypothèse du « champ moléculaire »
      b. Équation pour l'aimantation
      c. L'hamiltonien de Heisenberg dans l'approximation de champ moyen
      d. Fonction de partition et énergie libre
   2. Propriétés du système en champ nul
      a. Température critique
      b. Aimantation
      c. Énergie et capacité calorifique
   3. Propriétés du système en présence de champ extérieur
      a. Comportement au-dessus de T_c
      b. Comportement en dessous de T_c
   4. Les exposants critiques
   5. Comparaison avec les résultats expérimentaux

III. Le modèle d'Ising et ses solutions exactes
   1. Présentation du modèle d'Ising
   2. Solution du modèle d'Ising à une dimension
      a. Modèle d'Ising pour un cristal unidimensionnel
      b. Calcul de la fonction de partition et de l'énergie libre
      c. Aimantation
      d. Conclusion
   3. La solution à deux dimensions
      a. Fonction de partition exacte pour un réseau carré
      b. Existence d'une transition de phase
      c. Comparaison avec l'approximation de champ moyen
   4. Conclusion

IV. Indications succinctes sur les méthodes d'approximation élaborées

Certains matériaux, appelés ferromagnétiques, peuvent présenter une aimantation permanente en l'absence de champ magnétique extérieur[1]. Les plus courants sont le fer, le cobalt, le nickel et quelques alliages comprenant ces éléments[2].

Les recherches entreprises en mécanique statistique et en physique du solide pour comprendre et expliquer ce phénomène sont considérables. Il ne saurait donc être question ici d'autre chose que d'aborder le problème et de présenter les solutions les plus simples qui y ont été apportées.

1. On réserve le terme « ferromagnétique » aux solides pour lesquels, dans l'état fondamental, tous les moments magnétiques pointent dans la même direction. Lorsqu'il y a plusieurs ions magnétiques par cellule élémentaire du cristal, les uns peuvent pointer leur moment dans une direction, les autres dans la direction opposée, avec cependant une aimantation résultante non nulle; on parle alors de *ferrimagnétisme.*

2. Le ferromagnétisme est une propriété relativement rare dans l'ensemble des solides. Citons encore, parmi les éléments ferromagnétiques, le gadolinium (Gd) et le dysprosium (Dy). On trouve aussi des alliages ferromagnétiques dont aucun composant n'est ferromagnétique lorsqu'il est pur, par exemple Au_2MnAl. La plupart des ferromagnétiques sont des métaux ou des alliages métalliques; quelques-uns sont des composés ioniques : $CrBr_3$, EuO, EuS, etc.

I. LE PHÉNOMÈNE PHYSIQUE ET SON ORIGINE

1. LES RÉSULTATS EXPÉRIMENTAUX

a. Aimantation et température de Curie

Le fait d'expérience suivant est bien connu : un aimant attire un clou en fer, il est sans action sur un clou en cuivre. Dans les deux cas le champ magnétique de l'aimant induit une aimantation dans le clou, mais cette aimantation reste trop faible dans le cas du cuivre pour interagir efficacement avec l'aimant, alors qu'elle est considérable pour le fer. Les corps qui se comportent comme le fer sont dits ferromagnétiques.

Cependant, si on chauffe le clou en fer au-dessus de 770 °C, il n'est plus attiré par l'aimant, même si on l'approche très près (fig. 1) : il a perdu ses propriétés ferromagnétiques, devenant simplement paramagnétique. Il redevient ferromagnétique dès que sa température redescend au-dessous de 770 °C. Pour chaque corps ferromagnétique, il existe ainsi une *température de transition* T_c, appelée *température de Curie ou température critique* à laquelle il passe de l'état ferromagnétique (pour $T < T_c$) à l'état paramagnétique (pour $T > T_c$) : le matériau subit une *transition de phase* (de la « phase ferromagnétique » à la « phase paramagnétique ») à la température T_c.

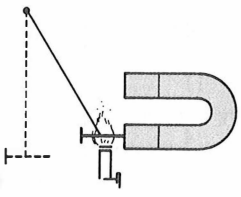


FIGURE 1
Expérience du clou chauffé. Le clou en fer suspendu par un fil est attiré par l'aimant. Lorsque sa température devient trop élevée, il revient à la verticale du point de suspension; lorsqu'il s'y est suffisamment refroidi, il est de nouveau attiré par l'aimant.

On peut étudier quantitativement le phénomène en mesurant l'aimantation M d'un barreau de fer en fonction du champ magnétique B_0 (uniforme et dirigé suivant l'axe du barreau) auquel il est soumis (fig. 2). En général l'aimantation du barreau est

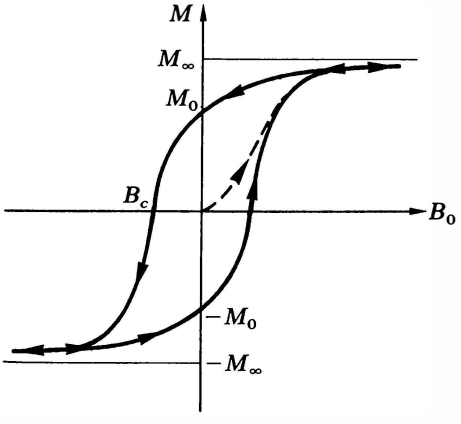


FIGURE 2
Aimantation d'un barreau de fer en fonction du champ appliqué B_0 pour une température inférieure à la température de Curie. La courbe tiretée est la *courbe de première aimantation.* Lorsqu'on augmente et diminue alternativement le champ appliqué de part et d'autre de $B_0 = 0$, l'aimantation décrit un *« cycle d'hystérésis »* (en trait plein).

initialement nulle pour $B_0 = 0$. Lorsqu'on fait croître le champ, M augmente, lentement au départ puis plus rapidement, pour finalement atteindre une valeur limite M_∞ appelée *aimantation à saturation.* On a ainsi fait décrire au barreau sa *« courbe de première aimantation »*. Si on diminue ensuite B_0, M décroît mais ne repasse pas par les mêmes valeurs. En particulier, elle ne s'annule pas lorsque B_0 redevient nul : la valeur de M qui subsiste alors est appelée *aimantation rémanente;* c'est seulement pour un champ dirigé en sens inverse, le *champ coercitif* B_c, que M s'annulera à nouveau. L'état du barreau ferromagnétique à un instant donné ne dépend donc pas seulement des conditions (champ B_0) qui lui sont imposées à cet instant, mais aussi de son « histoire » antérieure, c'est-à-dire de la façon dont il a été amené dans cet état. Ainsi, lorsqu'on fait varier le champ alternativement entre deux valeurs opposées, l'aimantation du barreau décrit une courbe fermée qu'on appelle un *cycle d'hystérésis.* Au-dessus de la température de Curie T_c, dans la phase paramagnétique, la boucle d'hystérésis disparaît et on retrouve le comportement paramagnétique usuel (complément III.A, § II).

Le tableau I donne les valeurs de T_c et de M_∞ pour quelques corps ferromagnétiques[3].

TABLEAU I

Température de Curie T_c et aimantation à saturation M_∞ pour quelques ferromagnétiques.

	T_c(K)	M_∞($\times 10^4$ A/m)
Fer (Fe)	1 043	14
Cobalt (Co)	1 388	11
Nickel (Ni)	627	4
Gadolinium (Gd)	293	16
Dysprosium (Dy)	85	24
Magnétite (Fe_3O_4)	858	4

b. La structure en domaines

Pour expliquer le phénomène d'hystérésis, P. Weiss émit en 1910 l'hypothèse (aujourd'hui confirmée par observation microscopique directe) qu'en dessous de la température de Curie un cristal ferromagnétique est structuré en *domaines :* chaque domaine, très petit à l'échelle macroscopique (mais comprenant un très grand nombre d'atomes) possède une aimantation non nulle; mais la direction de celle-ci varie d'un domaine à l'autre (fig. 3). L'aimantation d'un cristal ferromagnétique peut ainsi être nulle par compensation entre les différents domaines. L'apparition d'une aimantation globale non nulle sous l'action d'un champ extérieur se fait selon deux processus principaux : les domaines qui présentent une aimantation dans le sens du champ grossissent au détriment des autres (processus prépondérant aux faibles champs magnétiques), l'aimantation des domaines mal orientés pivote dans le sens du champ

3. On notera les valeurs considérables de M_∞ : quelques 10^4 A/m! Comme chacun a pu en faire l'expérience, les effets magnétiques d'un aimant permanent sont nettement plus forts que ceux des courants électriques usuels.

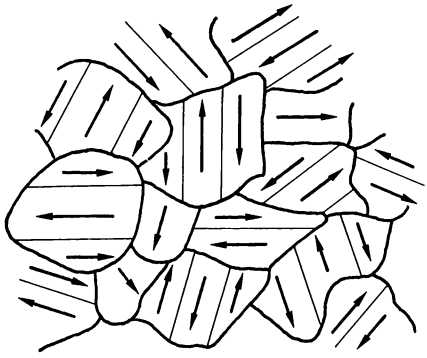


FIGURE 3
Structure schématique d'un matériau ferromagnétique globalement non magnétisé. Chaque grain cristallin est divisé en plusieurs domaines possédant chacun une aimantation (spontanée).

(aux forts champs). De plus, chaque domaine bien orienté voit son aimantation croître avec le champ.

Notre but n'est pas d'étudier cette structure en domaines et ses caractéristiques[4], mais de nous attacher à la compréhension du phénomène primaire du ferromagnétisme, l'aimantation d'un domaine particulier. C'est pourquoi, dans toute la suite, nous supposerons le *cristal suffisamment petit*[5] *pour être constitué d'un seul domaine.* Les courbes des figures 4 et 5 résument le comportement d'un tel cristal : en dessous de T_c, il présente une *aimantation spontanée* M_0 en l'absence de champ extérieur; cette aimantation est une fonction décroissante de la température, qui s'annule à la température critique[6].

4. Sur ce sujet, on pourra consulter les références 1 à 4.

5. La taille des domaines s'ajuste pour que l'énergie libre totale du cristal soit minimum. Cette minimisation résulte d'une compétition entre l'énergie d'interaction magnétique dipolaire, faible et à longue portée, et l'énergie d'échange (voir plus bas paragraphe I.4), forte et à courte portée. La première diminue lorsque le nombre des domaines augmente et tend donc à les multiplier; la seconde tend au contraire à étendre un domaine unique à l'ensemble du cristal.

6. La température critique T_c sépare les domaines d'existence de deux phases : l'une est ordonnée (la phase ferromagnétique), l'autre désordonnée (la phase paramagnétique), d'aimantation nulle. Plus la température est basse, plus la phase ferromagnétique est ordonnée, plus l'aimantation $M_0(T)$ est grande. $M_0(T)$ caractérise ainsi le degré d'ordre du matériau; on l'appelle le *«paramètre d'ordre»* du système.

La transition paramagnétique $\longleftrightarrow$ ferromagnétique s'accompagne d'un changement de symétrie. Le paramètre d'ordre est nul dans la phase la plus symétrique (paramagnétique), non nul dans la phase la moins symétrique (ferromagnétique), dans laquelle la direction de l'aimantation introduit un axe privilégié. Évidemment, l'ensemble des directions possibles pour l'aimantation, dans la phase ferromagnétique, possède la symétrie du cristal (c'est-à-dire celle de la phase paramagnétique); mais le système doit « choisir » une direction particulière et perd ainsi une partie de la symétrie sous-jacente : on dit que la transition de la phase paramagnétique à la phase ferromagnétique donne lieu à une *brisure spontanée de symétrie.*

D'une façon générale, les transitions qui s'accompagnent d'une brisure de symétrie et dont le paramètre d'ordre est une fonction continue de la température (c'est-à-dire s'annule à la température critique comme sur la figure 5) sont appelées *transitions du second ordre.* Lors d'une telle transition, on passe d'une phase à l'autre (ici du ferromagnétisme au paramagnétisme ou inversement) sans que les deux phases coexistent jamais; l'entropie des deux phases tend vers la même valeur lorsque la température tend vers T_c : il n'y a pas de chaleur latente de transformation (on notera les différences avec les transitions du premier ordre comme la transition liquide-vapeur : complément III.G, § III.4).

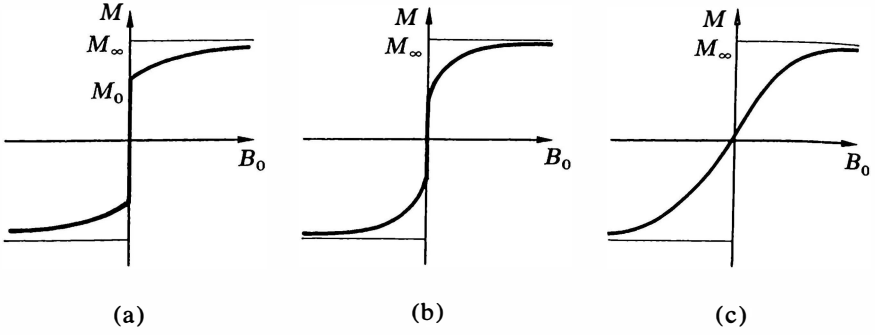


FIGURE 4
Courbes d'aimantation d'un domaine magnétique : a) au-dessous de la température de Curie ($T < T_c$); b) à la température de Curie ($T = T_c$); c) au-dessus de la température de Curie ($T > T_c$).

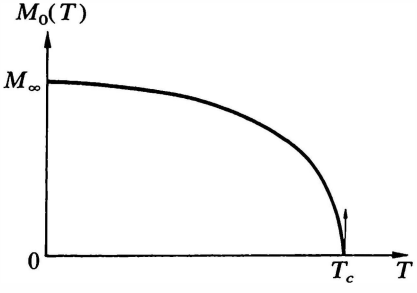


FIGURE 5
Variation avec la température de l'aimantation spontanée M_0 en champ nul.

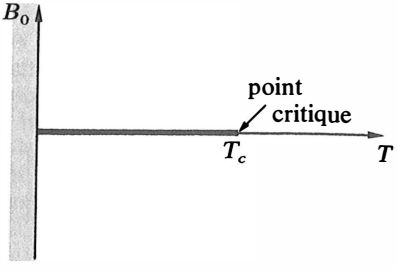


FIGURE 6
Demi-plan des paramètres extérieurs (T, B_0). L'aimantation $M(B_0, T)$ est une fonction définie et continue en tout point de ce demi-plan sauf sur le segment (0, T_c) de l'axe des températures.

L'état du cristal à l'équilibre est fixé par deux paramètres extérieurs : la température T et le champ magnétique appliqué B_0; son aimantation M est fonction de ces deux seuls paramètres[7]. Si nous considérons le demi-plan qu'ils définissent (fig. 6), avec T en abscisse ($T \geq 0$) et B_0 en ordonnée, les résultats présentés sur les figures 4 et 5 peuvent être résumés en disant que $M(B_0, T)$ est une fonction définie en tout point de ce demi-plan sauf sur le segment de l'axe des températures qui joint l'origine à la température critique T_c, puisque M est discontinue à la traversée de ce segment :

$$\lim_{B_0 \to 0_+} M(B_0, T) = - \lim_{B_0 \to 0_-} M(B_0, T) = M_0(T), \quad T < T_c \,. \tag{J.1}$$

Le point terminal de ce segment ($T = T_c$, $B_0 = 0$) s'appelle le *point critique.*

7. Rappelons que l'aimantation est le moment magnétique moyen par unité de volume. Elle ne dépend donc pas du volume V du cristal.

c. Le voisinage du point critique

De nombreuses propriétés macroscopiques des cristaux ferromagnétiques ont un comportement singulier au voisinage du point critique. Il semble que ces singularités puissent s'écrire comme des puissances fractionnaires de l'écart $|T - T_c|$ à la température critique. Par exemple, au voisinage de T_c, l'aimantation spontanée $M_0(T)$ est caractérisée par un *«exposant critique»*, traditionnellement appelé β, tel que[8]

$$M_0(T) \sim (T_c - T)^{\beta} \quad \text{pour} \quad 0 < (T_c - T)/T_c \ll 1 \,. \tag{J.2}$$

On définit ainsi tout un jeu d'exposants critiques dont les principaux concernent[9]

— la capacité calorifique en champ nul :

$$\begin{aligned} C_0(T) &\sim (T_c - T)^{-\alpha'} \quad \text{pour} \quad 0 < (T_c - T)/T_c \ll 1 \,, \\ &\sim (T - T_c)^{-\alpha} \quad \text{pour} \quad 0 < (T - T_c)/T_c \ll 1 \,, \end{aligned} \tag{J.3}$$

— la susceptibilité magnétique :

$$\begin{aligned} \chi(T) \mathrel{\widehat{=}} \lim_{B_0 \to 0} \frac{\partial M(B_0, T)}{\partial B_0} &\sim (T_c - T)^{-\gamma'} \quad \text{pour} \quad 0 < (T_c - T)/T_c \ll 1 \,, \\ &\sim (T - T_c)^{-\gamma} \quad \text{pour} \quad 0 < (T - T_c)/T_c \ll 1 \,. \end{aligned} \tag{J.4}$$

Expérimentalement, les valeurs mesurées pour les exposants critiques varient très peu d'un matériau à l'autre : β est toujours compris entre 0,34 et 0,38, γ et γ' sont tous deux égaux à 1,3-1,4 et α et α' sont de l'ordre de 0,1. Cette *universalité des exposants critiques*[10], *a priori* surprenante, indique qu'ils ne dépendent que des caractéristiques générales de la transition de phase et non de la structure microscopique propre à chaque matériau. On comprend ainsi leur importance et pourquoi les physiciens se sont attachés à les déterminer le plus précisément possible, tant expérimentalement que théoriquement.

2. LE PROBLÈME A RÉSOUDRE

On considère un cristal de volume V en équilibre avec un thermostat à la température T et soumis à un champ magnétique $\vec{B}_0$ (supposé uniforme). V, T, $\vec{B}_0$ constituent l'ensemble des paramètres extérieurs fixés, à partir desquels on cherche à déterminer les propriétés magnétiques du cristal : énergie magnétique moyenne, capacité calorifique associée, entropie... Nous nous intéresserons tout particulièrement à *l'aimantation* $\vec{M}$, c'est-à-dire au moment magnétique moyen $\vec{\mathcal{M}}$ par unité de volume :

$$\vec{M} \mathrel{\widehat{=}} \vec{\mathcal{M}} / V \,, \tag{J.5}$$

et nous chercherons *«l'équation d'état magnétique»* du cristal $\vec{M} = \vec{M}(T, \vec{B}_0)$. Nous supposons que $\vec{B}_0$ reste toujours dirigé suivant le même axe (l'axe Oz); $\vec{M}$ reste aussi

---

8. De façon plus précise,

$$\beta \mathrel{\widehat{=}} \lim_{T \to T_{c-}} \frac{\ln M_0(T)}{\ln (T_c - T)} \,.$$

Les autres exposants critiques ont des définitions analogues.

9. Signalons aussi, par exemple :

$$M(T_c, B_0) \sim B_0^{1/\delta} \quad \text{pour} \quad B_0 \longrightarrow 0 \,.$$

Pour une revue complète, voir référence 5.

10. Cette propriété dépasse même le cadre du ferromagnétisme. Il existe en effet toute une classe de transitions du second ordre qui ont les mêmes exposants critiques : on parle de *«classe d'universalité»*.

parallèle à cet axe et, par la suite, il suffira de déterminer la relation entre les valeurs algébriques M et B_0 de l'aimantation et du champ mesurées sur cet axe.

Pour un cristal paramagnétique parfait constitué de N spins 1/2 dont chacun porte un moment magnétique $\mu = g\mu_B/2$ (où μ_B est le magnéton de Bohr et g le facteur de Landé), l'équation d'état magnétique s'écrit (complément III.A, formule (A.21)) :

$$M = N\frac{g\mu_B}{2V}\,\text{th}\,\frac{g\mu_B B_0}{2kT} \quad \text{(paramagnétique parfait)}\,. \tag{J.6}$$

Dans ce système sans interactions, M s'annule avec le champ extérieur quelle que soit la température. Comme nous l'avons vu, il en va tout autrement pour un cristal ferromagnétique. Ce que nous cherchons ici c'est à expliquer, à partir d'un modèle microscopique, le comportement d'un cristal ferromagnétique monodomaine, tel qu'il est présenté sur les figures 4 et 5.

3. L'HAMILTONIEN DE HEISENBERG

Il nous faut tout d'abord un modèle microscopique physiquement sensé, et si possible simple, qui puisse rendre compte du phénomène.

C'est évidemment l'interaction entre les moments magnétiques atomiques, ou les moments cinétiques auxquels ils sont associés, qui tend à les aligner malgré l'agitation thermique pour former l'état ferromagnétique.

L'énergie d'interaction la plus simple qui puisse exister entre les spins $\vec{S}_i$ et $\vec{S}_j$ de deux atomes situés sur les sites (i) et (j) du cristal est de la forme

$$H_{ij} = -J_{ij}\,\vec{S}_i\,.\,\vec{S}_j\,. \tag{J.7}$$

Le « coefficient de couplage » J_{ij} doit être positif pour que l'interaction favorise l'alignement des spins même en l'absence de champ extérieur; il dépend de la distance entre les sites (i) et (j) et ne peut que décroître lorsque cette distance augmente.

Pour simplifier au maximum, nous considérerons un système constitué par N atomes placés aux nœuds d'un réseau cristallin et portant chacun un spin 1/2, et nous supposerons que J_{ij} décroît de façon suffisamment rapide quand la distance croît pour que nous n'ayons à tenir compte que de l'interaction entre voisins immédiats. A chaque spin $\vec{S}_i$ est associé un moment magnétique $\vec{\mu}_i = g\mu_B\vec{S}_i$. Placé dans un champ magnétique uniforme $\vec{B}_0$, le système a donc pour hamiltonien[11] :

$$H_H = -g\mu_B\vec{B}_0\,.\sum_{i=1}^{N}\vec{S}_i - J\sum_{(i,\,j)_v}\vec{S}_i\,.\,\vec{S}_j\,, \quad J>0\,; \tag{J.8}$$

la deuxième sommation porte sur toutes les *paires d'atomes* (i) *et* (j) *plus proches voisins;* si p est le nombre de voisins immédiats d'un atome particulier, elle comporte $Np/2$ termes.

Cet hamiltonien est connu sous le nom d'*hamiltonien de Heisenberg.* On vérifie aisément, puisque J est positif, que son état fondamental correspond au cas où tous les spins sont alignés dans la direction du champ extérieur.

4. ORIGINE PHYSIQUE DU FERROMAGNÉTISME

Avant d'étudier la mécanique statistique d'un système décrit par l'hamiltonien de Heisenberg, il nous reste à comprendre la cause de l'interaction entre les spins et à évaluer la constante de couplage J.

---

11. En toute rigueur, le champ $\vec{B}_0$ n'est pas le champ extérieur appliqué, mais le *champ magnétique local* « vu » par chaque spin. Le champ local prend donc en compte le champ magnétique créé au point considéré par tous les autres dipôles magnétiques. Il dépend de façon compliquée de la forme de l'échantillon. Dans les cas les plus simples, on l'obtient en multipliant le champ appliqué par un « facteur démagnétisant » (voir par exemple A. Sommerfeld, *Electrodynamics,* Academic Press, New York, paragraphe 13 H).

L'interaction magnétique dipolaire entre les moments magnétiques associés aux spins des atomes *est trop faible* pour rendre compte du phénomène. En effet, lorsque deux moments magnétiques μ sont séparés par une distance r, leur énergie d'interaction vaut, en ordre de grandeur[12],

$$E_{\text{mag}} \sim \frac{\mu_0}{4\pi} \times \frac{\mu^2}{r^3} . \tag{J.9}$$

Les moments magnétiques microscopiques sont de l'ordre du magnéton de Bohr ($\mu_B \widehat{=} e\hbar/2m_e \simeq 10^{-23}$ A.m^2), ce qui donne, pour deux atomes situés à environ 2 Å l'un de l'autre,

$$E_{\text{mag}} \sim 10^{-24} \text{ joule} . \tag{J.10}$$

L'interaction magnétique dipolaire entre deux atomes voisins ne dépasse donc pas 10^{-4} eV. Cette énergie ne pourrait « geler » le système dans l'état ferromagnétique que si elle était supérieure à l'énergie d'agitation thermique kT; il faudrait donc des températures inférieures à E_{mag}/k, c'est-à-dire à 1 K, valeur sans commune mesure avec les températures de Curie observées[13],[14].

Il existe en fait une interaction beaucoup plus forte entre les spins électroniques, d'origine quantique, qu'on appelle *l'interaction d'échange.* Elle est due à la *répulsion coulombienne* entre les deux électrons *combinée au principe d'exclusion de Pauli.*

On peut comprendre simplement le phénomène en considérant deux électrons seulement. La fonction d'onde des deux électrons peut s'écrire comme le produit[15]

$$\psi(1, 2) = \Phi(\vec{r}_1, \vec{r}_2)\chi_s(1, 2) \tag{J.11}$$

---

12. L'énergie d'interaction entre deux moments magnétiques $\vec{\mu}_1$ et $\vec{\mu}_2$ séparés par le vecteur $\vec{r}$ s'écrit

$$E_{\text{mag}} = \frac{\mu_0}{4\pi r^3}\left[\vec{\mu}_1 \cdot \vec{\mu}_2 - \frac{3}{r^2}(\vec{\mu}_1 \cdot \vec{r})(\vec{\mu}_2 \cdot \vec{r})\right] .$$

13. Ce qu'il faut en toute rigueur comparer à kT_c, c'est l'énergie d'interaction d'un moment magnétique particulier avec tous les autres; mais l'ordre de grandeur reste le même. Pour un cristal cubique, par exemple, on montre que, si tous les dipôles sont parallèles à l'un des axes du cristal, le champ magnétique créé sur un site particulier par l'ensemble des autres moments magnétiques vaut

$$\vec{B} = \frac{\mu_0}{3}\frac{\vec{\mu}}{a^3} ,$$

où a est la distance entre deux sites voisins. D'où une énergie d'interaction

$$|E_{\text{mag}}| = \vec{B} \cdot \vec{\mu} = \frac{\mu_0}{3a^3}\mu^2 .$$

14. L'interaction dipolaire magnétique joue cependant un rôle important dans la formation des domaines (cf. note 5).

15. Puisque l'hamiltonien du système ne dépend pas des spins, ses états propres peuvent s'écrire comme des produits tensoriels d'un état orbital par un état de spin. Les états de spin χ_s sont au nombre de 4 :

— un état singulet $(1/\sqrt{2})[|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle]$, de spin total nul, antisymétrique dans l'échange des deux particules (spins individuels « antiparallèles »),

— trois états triplets : $|\uparrow\uparrow\rangle$, $|\downarrow\downarrow\rangle$ et $(1/\sqrt{2})[|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle]$, de spin total égal à 1, symétriques dans l'échange des deux particules (spins individuels « parallèles »).

d'une fonction d'onde « orbitale » Φ dépendant de leurs positions $\vec{r}_1$ et $\vec{r}_2$ et d'une « fonction d'onde de spin » $\chi_s(1, 2)$. Les électrons étant des fermions, ψ doit être antisymétrique (principe de Pauli) :

$$\psi(1, 2) = -\psi(2, 1) . \tag{J.12}$$

Autrement dit, si les spins sont parallèles (χ_s symétrique), Φ doit être antisymétrique :

$$\text{spins parallèles} \implies \chi_s(1, 2) = \chi_s(2, 1)$$
$$\implies \Phi(\vec{r}_1, \vec{r}_2) = -\Phi(\vec{r}_2, \vec{r}_1) , \tag{J.13.a}$$

et inversement

$$\text{spins antiparallèles} \implies \chi_s(1, 2) = -\chi_s(2, 1)$$
$$\implies \Phi(\vec{r}_1, \vec{r}_2) = \Phi(\vec{r}_2, \vec{r}_1) . \tag{J.13.b}$$

Si les spins sont parallèles, la fonction d'onde s'annule donc lorsque les deux électrons sont au même endroit ($\vec{r}_1 = \vec{r}_2$) et elle est très petite lorsqu'ils sont voisins; dans le cas contraire, les deux électrons peuvent s'approcher l'un de l'autre sans être « gênés » par le principe d'exclusion. Ainsi deux spins parallèles seront rarement voisins; leur énergie d'interaction coulombienne (positive puisqu'ils ont même charge) sera donc plus faible que celle de deux spins antiparallèles. On peut donc s'attendre à ce que deux électrons de même spin aient une énergie plus faible que deux électrons de spins opposés. Cette différence d'énergie, d'origine électrostatique, est de l'ordre de l'eV. Elle est équivalente à une « force » entre les spins qu'on peut représenter par un hamiltonien effectif[16]

$$H_{\text{int}} = -J_{12} \vec{S}_1 . \vec{S}_2 . \tag{J.14}$$

---

16. Si on néglige dans un premier temps l'interaction coulombienne entre les deux électrons, l'hamiltonien est la somme de deux termes indépendants :

$$H_0 = H(\vec{p}_1, \vec{r}_1) + H(\vec{p}_2, \vec{r}_2) . \tag{1}$$

Soient $\phi_\alpha(\vec{r})$ les fonctions propres de H et ε_α les énergies correspondantes. Les états propres de H_0, antisymétriques dans l'échange des deux électrons, sont de la forme suivante (pour $\alpha \neq \beta$) :

— pour les états singulets (cf. note 15),

$$|\psi_S^{\alpha\beta}\rangle = \frac{1}{\sqrt{2}} [|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle] \otimes \frac{1}{\sqrt{2}} [\phi_\alpha(\vec{r}_1)\phi_\beta(\vec{r}_2) + \phi_\beta(\vec{r}_1)\phi_\alpha(\vec{r}_2)] ; \tag{2}$$

— pour les états triplets,

$$\left.\begin{aligned} |\psi_{T_1}^{\alpha\beta}\rangle &= |\uparrow\uparrow\rangle \\ |\psi_{T_2}^{\alpha\beta}\rangle &= \frac{1}{\sqrt{2}} [|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle] \\ |\psi_{T_3}^{\alpha\beta}\rangle &= |\downarrow\downarrow\rangle \end{aligned}\right\} \otimes \frac{1}{\sqrt{2}} [\phi_\alpha(\vec{r}_1)\phi_\beta(\vec{r}_2) - \phi_\beta(\vec{r}_1)\phi_\alpha(\vec{r}_2)] . \tag{3}$$

Ces quatre vecteurs propres de H_0 correspondent à la même valeur propre ($\varepsilon_\alpha + \varepsilon_\beta$). Si maintenant on tient compte de l'interaction coulombienne $e^2/4\pi\varepsilon_0 |\vec{r}_1 - \vec{r}_2|$ entre les deux électrons, l'état singulet se sépare des états triplets, avec une différence d'énergie qu'on peut évaluer par un calcul de perturbation au premier ordre :

$$J \hat{=} E_S^{\alpha\beta} - E_T^{\alpha\beta} = \langle \psi_S^{\alpha\beta} | \frac{e^2}{4\pi\varepsilon_0 |\vec{r}_1 - \vec{r}_2|} |\psi_S^{\alpha\beta}\rangle - \langle \psi_T^{\alpha\beta} | \frac{e^2}{4\pi\varepsilon_0 |\vec{r}_1 - \vec{r}_2|} |\psi_T^{\alpha\beta}\rangle$$
$$= 2 \int d^3r_1 \, d^3r_2 \left[\phi_\alpha^*(\vec{r}_1) \phi_\beta^*(\vec{r}_2) \frac{e^2}{4\pi\varepsilon_0 |\vec{r}_1 - \vec{r}_2|} \phi_\alpha(\vec{r}_2) \phi_\beta(\vec{r}_1) \right] . \tag{4}$$

.../...

J_{12} s'appelle *« l'intégrale d'échange »;* si chacun des deux électrons est localisé autour d'un noyau, J_{12} dépend de la distance entre les deux noyaux et décroît très rapidement lorsque cette distance augmente.

Le raisonnement précédent montre de façon simple comment une énergie d'interaction entre deux particules chargées peut être transformée par le principe de Pauli en une énergie d'interaction entre leurs spins. Mais la valeur de l'intégrale d'échange dans une situation réaliste, et même son signe, sont très difficiles à obtenir : on sait par exemple que les deux électrons d'une molécule d'hydrogène H_2 ont leurs spins antiparallèles dans l'état fondamental, contrairement à ce que suggère l'argument ci-dessus; c'est que l'attraction des noyaux sur les électrons, que nous avons ignorée plus haut, joue un rôle primordial dans la stabilité d'une molécule ou d'un solide. En fait, si on est sûr que « l'échange » est à la base du ferromagnétisme[17], on peut simplement espérer qu'une interaction de la forme (J.8) décrit raisonnablement le phénomène.

II. L'APPROXIMATION DE CHAMP MOYEN

Il n'a pas été possible (jusqu'à présent) de calculer exactement la fonction de partition d'un système décrit par l'hamiltonien de Heisenberg. Comme toujours, lorsqu'on a affaire à un système de particules en interaction, l'approximation la plus

---

Cette différence d'énergie, qui fait intervenir un élément de matrice purement coulombien, est donc de l'ordre de l'électron-volt. On l'appelle *« intégrale d'échange »* parce que les variables $\vec{r}_1$ et $\vec{r}_2$ s'échangent quand on passe d'un des produits $\phi_\alpha\phi_\beta$ à l'autre. Noter que J n'est différent de zéro que s'il existe une région de l'espace dans laquelle les fonctions $\phi_\alpha(\vec{r})$ et $\phi_\beta(\vec{r})$ sont simultanément non nulles; dans les situations où chacun des électrons est localisé autour d'un noyau, l'intégrale d'échange décroît donc très rapidement lorsque la distance entre les deux noyaux croît.

Par ailleurs, chaque opérateur de spin est tel que

$$\vec{S}_i^2 = \frac{1}{2}\left(\frac{1}{2}+1\right) = \frac{3}{4}\,. \tag{5}$$

Le spin total

$$\vec{S} = \vec{S}_1 + \vec{S}_2 \tag{6}$$

vérifie donc

$$\vec{S}^2 = \frac{3}{2} + 2\vec{S}_1.\vec{S}_2\,. \tag{7}$$

L'état singulet ($S=0$) est vecteur propre de $\vec{S}^2$ avec la valeur propre 0 et donc de $\vec{S}_1.\vec{S}_2$ avec la valeur propre $-3/4$; les trois états triplets ($S=1$) sont vecteurs propres de $\vec{S}^2$ avec pour valeur propre $S(S+1)=2$ et donc de $\vec{S}_1.\vec{S}_2$ avec pour valeur propre 1/4. Il est alors facile de vérifier que ces quatre états sont vecteurs propres de

$$H^{\text{spin}} = \frac{1}{4}(E_S + 3E_T) - (E_S - E_T)\vec{S}_1.\vec{S}_2 \tag{8}$$

avec les valeurs propres E_S pour l'état singulet, E_T pour les trois états triplets. L'interaction d'échange peut donc se représenter en ajoutant à H_0 un « hamiltonien effectif » qu'on peut écrire, à la constante $(E_S+3E_T)/4$ près,

$$H_{\text{eff}} = -(E_S - E_T)\vec{S}_1.\vec{S}_2 = -J\vec{S}_1.\vec{S}_2\,. \tag{9}$$

17. Il semble que ce sont les électrons des couches d ou f, incomplètement remplies dans les métaux de transition ou les terres rares, qui sont responsables du ferromagnétisme.

simple qui permet une description (au moins qualitative) du phénomène est l'approximation de champ moyen (cf. compléments III.G, III.H, III.K).

1. PRINCIPE DE L'APPROXIMATION

a. L'hypothèse du « champ moléculaire »

Considérons un site particulier du réseau, le site (i), et isolons, dans l'hamiltonien de Heisenberg, la partie H_i qui fait intervenir ce site :

$$H_i = -\vec{S}_i \cdot \left[g\mu_B \vec{B}_0 + J \sum_{j \in v(i)} \vec{S}_j \right] \tag{J.15}$$

(la sommation porte uniquement sur les sites (j) plus proches voisins du site (i)). Tout se passe, pour ce site (i), comme s'il était soumis à un champ magnétique effectif

$$\vec{B}^{(i)} \mathrel{\widehat{=}} \vec{B}_0 + \frac{J}{g\mu_B} \sum_{j \in v(i)} \vec{S}_j , \tag{J.16}$$

puisqu'alors

$$H_i = -g\mu_B \vec{S}_i \cdot \vec{B}^{(i)} . \tag{J.17}$$

Le second terme de (J.16), qui s'ajoute ainsi au champ $\vec{B}_0$, est le *« champ moléculaire »;* il traduit l'action sur le site (i) des autres sites du cristal.

Proposée par P. Weiss en 1907, l'approximation de champ moyen consiste à *négliger les fluctuations de ce champ moléculaire* et à le remplacer par sa valeur moyenne. Dans un cristal ferromagnétique, chaque spin a la même valeur moyenne, reliée à l'aimantation $\vec{M}$ par

$$\vec{M} = \frac{N}{V} g\mu_B \overline{\vec{S}_i} ; \tag{J.18}$$

le champ effectif moyen

$$\vec{B}_{\text{eff}} \mathrel{\widehat{=}} \overline{\vec{B}^{(i)}} \tag{J.19}$$

est donc le même en tout point du réseau, et de la forme

$$\vec{B}_{\text{eff}} = \vec{B}_0 + \lambda \vec{M} , \tag{J.20}$$

avec $$\lambda \mathrel{\widehat{=}} \frac{pJ}{(g\mu_B)^2} \times \frac{V}{N} \tag{J.21}$$

(p est le nombre des plus proches voisins d'un site donné). $\vec{M}$ étant parallèle à $\vec{B}_0$, $\vec{B}_{\text{eff}}$ est lui aussi uniforme et parallèle à cette direction.

Dans l'approximation de champ moyen, chacun des spins du cristal est soumis à la seule action du champ $\vec{B}_{\text{eff}}$, et nous sommes ramenés au problème de N *spins indépendants,* placés dans un champ magnétique uniforme (mais fonction de l'aimantation du cristal).

b. Équation pour l'aimantation

L'application de la formule (J.6), valable pour un système de spins 1/2 indépendants, donne immédiatement l'aimantation du système en fonction de B_{eff} :

$$M = N \frac{g\mu_B}{2V} \text{th} \left[\frac{g\mu_B}{2kT} B_{\text{eff}} \right] . \tag{J.22}$$

Mais B_{eff} est ici une fonction de M; cette relation va donc fournir en fait une *« équation d'autocohérence » déterminant M :* si on connaît M, on en déduit B_{eff} par (J.20);

connaissant B_{eff}, on en déduit M par (J.22); il faut que ces deux opérations soient cohérentes entre elles, c'est-à-dire que M doit vérifier l'équation

$$M = \frac{g\mu_B N}{2V} \text{ th} \left[\frac{g\mu_B}{2kT} (B_0 + \lambda M) \right] . \tag{J.23}$$

Il est commode d'analyser les solutions de cette équation en même temps que la stabilité des états correspondants. C'est pourquoi nous le ferons (aux § II.2 et § II.3) après avoir déterminé l'énergie libre du système.

c. L'hamiltonien de Heisenberg dans l'approximation de champ moyen

Dans l'approximation de champ moyen, l'hamiltonien du système est donc, à un terme constant près, l'hamiltonien de N spins indépendants placés dans le champ $\vec{B}_{\text{eff}}$:

$$H_{\text{c.m.}} = -g\mu_B \vec{B}_{\text{eff}} . \sum_{i=1}^{N} \vec{S}_i + K . \tag{J.24}$$

Le terme constant K peut être déterminé en comparant les valeurs moyennes des deux hamiltoniens (J.8) et (J.24). Le champ moyen, par principe, néglige les corrélations entre les spins des différents sites du cristal : la moyenne du produit $\vec{S}_i . \vec{S}_j$ est donc, dans cette approximation, égale au produit des moyennes de $\vec{S}_i$ et $\vec{S}_j$. Utilisant (J.19) et (J.16), on obtient

$$\overline{H}_{\text{c.m.}} = -g\mu_B \vec{B}_0 . N\overline{\vec{S}_i} - JpN\overline{\vec{S}_i}^2 + K \tag{J.25}$$

et

$$\overline{H}_H = -g\mu_B \vec{B}_0 . N\overline{\vec{S}_i} - J\frac{Np}{2} \overline{\vec{S}_i}^2 . \tag{J.26}$$

La valeur de K s'en déduit immédiatement et l'hamiltonien de Heisenberg approché s'écrit[18], si l'on remplace $\overline{\vec{S}_i}$ par sa valeur tirée de (J.18),

$$H_{\text{c.m.}} = -g\mu_B \vec{B}_{\text{eff}} . \sum_{i=1}^{N} \vec{S}_i + J\frac{Np}{2} \left(\frac{VM}{Ng\mu_B} \right)^2 . \tag{J.27}$$

Contrairement au cas de N spins véritablement indépendants, *le terme constant joue ici un rôle essentiel.* En effet K, comme d'ailleurs B_{eff}, dépend de l'aimantation M, c'est-à-dire de l'inconnue qu'on cherche à déterminer; on ne peut donc pas l'éliminer par un simple changement de l'origine des énergies, puisque cela reviendrait à prendre une origine des énergies différente pour chaque état d'aimantation du cristal.

Le principe de l'approximation de champ moyen consiste à calculer les propriétés du système (à partir de l'hamiltonien $H_{\text{c.m.}}$) comme si M était donnée. Mais M est en réalité une *variable interne qui doit s'ajuster pour que l'énergie libre soit minimum.* Nous allons voir qu'on retrouve ainsi l'équation d'autocohérence (J.23). Celle-ci permet ensuite de calculer les valeurs possibles de M et d'en déduire toutes les propriétés du système.

---

18. On trouve le même résultat si, notant que

$$\sum_{(i,j)_v} \ldots = \frac{1}{2} \sum_{i=1}^{N} \sum_{j \in v(i)} \ldots , \tag{1}$$

on effectue directement, dans l'hamiltonien de Heisenberg, la substitution

$$\vec{S}_i . \vec{S}_j \longrightarrow \vec{S}_i . \overline{\vec{S}_j} + \overline{\vec{S}_i} . \vec{S}_j - \overline{\vec{S}_i} . \overline{\vec{S}_j} \tag{2}$$

(on retranche le produit des valeurs moyennes pour conserver la valeur moyenne du produit $\vec{S}_i . \vec{S}_j$ dans la substitution).

d. Fonction de partition et énergie libre

Pour le système de N spins 1/2 « indépendants » décrit par l'hamiltonien (J.27), la fonction de partition se calcule sans difficulté[19] :

$$Z_{\text{c.m.}} = \left[2 \exp \left\{ -\frac{pJ}{2kT} \left(\frac{V}{N} \frac{M}{g\mu_B} \right)^2 \right\} \operatorname{ch} \left(\frac{g\mu_B}{2kT} B_{\text{eff}} \right) \right]^N . \tag{J.28}$$

On en déduit l'énergie libre

$$F_{\text{c.m.}} \mathrel{\widehat{=}} -kT \ln Z_{\text{c.m.}} = \frac{NJp}{2} \left(\frac{V}{N} \frac{M}{g\mu_B} \right)^2 - NkT \ln \left\{ 2 \operatorname{ch} \left[\frac{g\mu_B}{2kT} (B_0 + \lambda M) \right] \right\} . \tag{J.29}$$

La valeur de M à l'équilibre est celle qui rend $F_{\text{c.m.}}$ minimum. M est donc solution de

$$\partial F_{\text{c.m.}} / \partial M = 0 , \tag{J.30}$$

ce qui, comme il se doit, redonne l'équation d'autocohérence (J.23). Mais la connaissance de $F_{\text{c.m.}}$ en fonction de M va nous permettre en outre de décider lesquelles parmi les solutions de (J.23) correspondent à des états d'équilibre stables, c'est-à-dire à des *minimums* de l'énergie libre. Une fois déterminée une valeur de M acceptable, la véritable énergie libre du système s'obtient en remplaçant dans (J.29) M par cette valeur, fonction de B_0 et T.

2. PROPRIÉTÉS DU SYSTÈME EN CHAMP NUL

En l'absence de champ extérieur ($B_0 = 0$), l'équation (J.23) s'écrit, à l'aide de l'expression (J.21) de λ :

$$\frac{M}{M_\infty} = \operatorname{th} \frac{pJ}{4kT} \frac{M}{M_\infty} , \tag{J.31}$$

où $$M_\infty \mathrel{\widehat{=}} \frac{Ng\mu_B}{2V} \tag{J.32}$$

est l'aimantation maximum que peut prendre le cristal *(aimantation à saturation);* elle correspond à l'état parfaitement ordonné dans lequel tous les moments magnétiques pointent dans la même direction. L'équation transcendante (J.31) peut être commodément résolue graphiquement en traçant les deux courbes :

$$\frac{M}{M_\infty} = \operatorname{th} x , \tag{J.33}$$

$$\frac{M}{M_\infty} = \frac{4kT}{pJ} x , \tag{J.34}$$

et en cherchant leurs points d'intersection (fig. 7).

---

19. Le plus simple est d'écrire

$$H_{\text{c.m.}} = \sum_{i=1}^{N} \left[-g\mu_B \vec{B}_{\text{eff}} \cdot \vec{S}_i + \frac{1}{2} Jp \left(\frac{V}{N} \frac{M}{g\mu_B} \right)^2 \right] . \tag{1}$$

La fonction de partition se factorise donc :

$$Z_{\text{c.m.}} = (z)^N , \tag{2}$$

où z est la fonction de partition pour un seul spin. Celui-ci a seulement deux états, d'énergies respectives :

$$\varepsilon_\pm = \frac{1}{2} Jp \left(\frac{V}{N} \frac{M}{g\mu_B} \right)^2 \mp \frac{1}{2} g\mu_B B_{\text{eff}} . \tag{3}$$

D'où $$z = \mathrm{e}^{-\varepsilon_+/kT} + \mathrm{e}^{-\varepsilon_-/kT} , \tag{4}$$

et par suite le résultat (J.28).

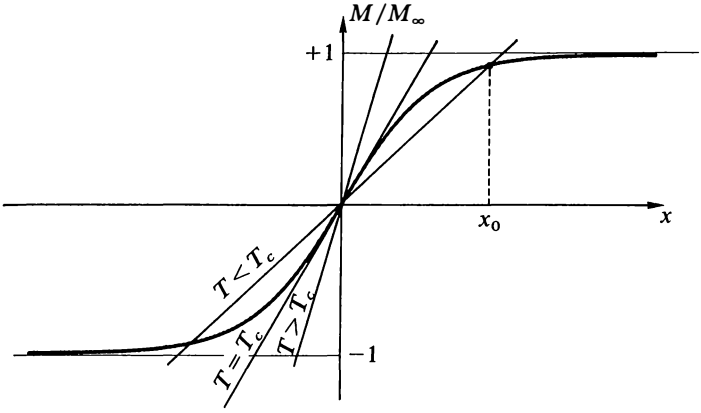


FIGURE 7

Résolution graphique de l'équation d'autocohérence en champ nul. Les points d'intersection entre la courbe $M/M_\infty = \text{th}\, x$ (en trait gras) et la droite de pente variable $M/M_\infty = xT/T_c$ (en trait léger, pour trois valeurs de T) déterminent les aimantations spontanées possibles dans l'approximation de champ moyen.

a. Température critique

Les deux courbes se coupent toujours à l'origine des coordonnées : $M = 0$ est toujours solution de l'équation d'autocohérence. Mais il existe deux autres intersections si la pente de la droite (J.34) est inférieure à celle de la tangente à la courbe (J.33) en $x = 0$. Autrement dit, si

$$T < T_c \cong \frac{pJ}{4k}, \tag{J.35}$$

il existe deux solutions symétriques $+M_0$ et $-M_0$, correspondant à une aimantation non nulle du cristal. En examinant la forme de l'énergie libre $F_{\text{c.m.}}$ en fonction de M (fig. 8), on se convainc aisément que pour $T < T_c$, $F_{\text{c.m.}}$ est minimum pour ces deux solutions symétriques, alors que l'aimantation nulle correspond à un maximum local. Pour $T > T_c$, il n'y a qu'une solution, $M = 0$, et $F_{\text{c.m.}}$ est minimum en ce point.

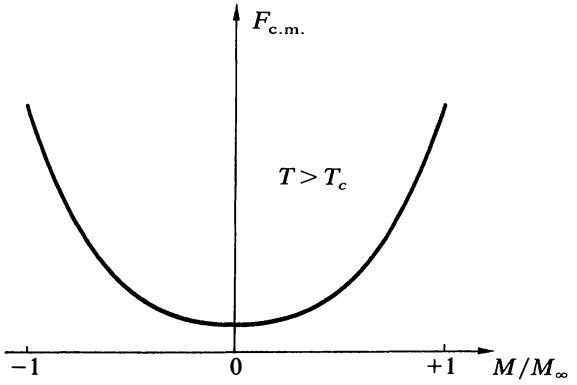
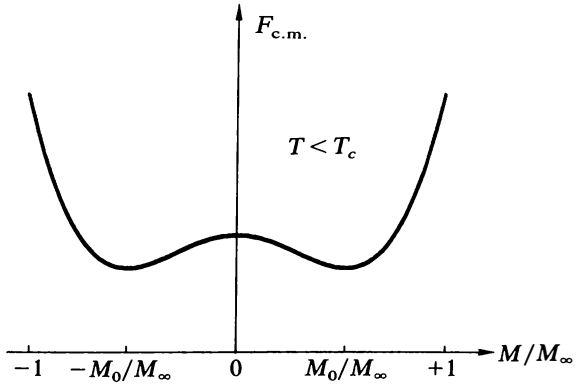


FIGURE 8

Énergie libre en champ nul, en fonction de M, dans l'approximation de champ moyen :

$$F_{\text{c.m.}}(M) = Nk\left[\frac{T_c}{2}\left(\frac{M}{M_\infty}\right)^2 - T \ln\left(2 \operatorname{ch} \frac{T_c}{T}\frac{M}{M_\infty}\right)\right].$$

Pour $T \geqslant T_c$, il n'y a qu'un minimum en $M = 0$ (fig. 8.a).



(b)

Pour $T < T_c$, la courbe présente trois extremums : un maximum local en $M = 0$ et deux minimums égaux pour des valeurs non nulles opposées M_0 et $-M_0$ (fig. 8.b).

La température T_c, qui sépare une région à aimantation non nulle ($T < T_c$) d'une région à aimantation nulle ($T > T_c$), est donc la *température de Curie* du cristal dans l'approximation de champ moyen.

b. Aimantation

En dessous de T_c, le système a plusieurs directions possibles, équivalentes, pour son aimantation. Pour atteindre l'équilibre, il doit cependant choisir l'une d'elles; l'équation (J.31) détermine en fait le module M_0 de cette aimantation d'équilibre. Il est facile de tracer la courbe $M_0(T)$ en faisant varier la pente (proportionnelle à T) de la droite (J.34) et en suivant sur la figure 7 son intersection avec la tangente hyperbolique du côté des x positifs.

A température nulle, les deux courbes se coupent pour x infini, et $M_0/M_\infty = 1$: le système se trouve dans son état fondamental, correspondant à l'aimantation M_∞. Quand la température augmente, la pente de la droite augmente et l'aimantation diminue; elle s'annule à $T = T_c$.

On peut donner une forme analytique approchée pour $M_0(T)$ dans la limite de basse température, et au voisinage de T_c. Aux basses températures ($T \ll T_c$), x_0 est très grand ($x_0 \gg 1$); le développement limité de th x conduit à[20]

$$\frac{M_0}{M_\infty} \simeq 1 - 2 \exp\left(-\frac{2T_c}{T}\right), \quad T \ll T_c . \tag{J.36}$$

20. Si $x \gg 1$,

$$\text{th } x \simeq 1 - 2e^{-2x}, \tag{1}$$

de sorte que

$$\frac{M}{M_\infty} = \frac{Tx}{T_c} \simeq 1 - 2e^{-2x} . \tag{2}$$

A l'ordre le plus bas,

$$x \simeq T_c/T , \tag{3}$$

ce qui, reporté dans le second membre de (2), donne le résultat (J.36).

Pour T voisin de T_c, x_0 est voisin de zéro ($x_0 \ll 1$); le développement de th x, à x petit cette fois, donne[21]

$$\frac{M_0}{M_\infty} \simeq \sqrt{3\left(1-\frac{T}{T_c}\right)}, \quad 0<\frac{T_c-T}{T_c}\ll 1 . \tag{J.37}$$

Dans la région intermédiaire, une résolution graphique, ou un calcul numérique, est nécessaire. L'allure de la courbe $M_0(T)$ est celle de la figure 5.

c. Énergie et capacité calorifique

L'énergie magnétique moyenne s'obtient à partir de la fonction de partition (J.28) prise pour $B_0 = 0$[22] :

$$\overline{E}_{\text{c.m.}}(T;\, B_0=0) = -\frac{\partial \ln Z_{\text{c.m.}}}{\partial\beta} = -\frac{NJp}{8}\left(\frac{M_0(T)}{M_\infty}\right)^2 . \tag{J.38}$$

Comme on pouvait s'y attendre, $\overline{E}_{\text{c.m.}}$ n'est rien d'autre que la valeur moyenne (J.26) de l'hamiltonien de Heisenberg (pour $B_0 = 0$) quand on y néglige les corrélations. $\overline{E}_{\text{c.m.}}$ est identiquement nulle au-dessus de T_c. Pour $T < T_c$, l'allure de la courbe $\overline{E}_{\text{c.m.}}(T;\, B_0=0)$ se déduit de celle pour $M_0(T)$ (fig. 9). Au voisinage de T_c, on obtient grâce à (J.37)

$$\begin{aligned} \overline{E}_{\text{c.m.}}(T;\, B_0=0) &\simeq -\frac{3}{8}NJp\left(1-\frac{T}{T_c}\right), \quad 0<\frac{T_c-T}{T_c}\ll 1 , \\ &= 0 \qquad\qquad , \quad T > T_c . \end{aligned} \tag{J.39}$$

---

21. Pour $x \ll 1$,

$$\text{th}\, x \simeq x - x^3/3 . \tag{1}$$

On a donc à résoudre l'équation approchée

$$\frac{M_0}{M_\infty} = \frac{Tx}{T_c} \simeq x - \frac{x^3}{3} . \tag{2}$$

Si on écarte la solution triviale $x = 0$ (qui correspond à un maximum de l'énergie libre), il reste

$$x \simeq \sqrt{3(1 - T/T_c)} , \tag{3}$$

d'où (J.37).

22. Pour dériver $\ln Z_{\text{c.m.}}$ par rapport à β, il faut en principe tenir compte du fait que M dépend de β; mais cette contribution est nulle, car la condition d'équilibre implique

$$\frac{\partial \ln Z_{\text{c.m.}}}{\partial M} = 0 . \tag{1}$$

On obtient donc simplement

$$\overline{E}_{\text{c.m.}}(T;\, B_0=0) = -N\left[-\frac{pJ}{2}\left(\frac{V}{N}\frac{M}{g\mu_B}\right)^2 + \frac{\lambda g\mu_B M}{2}\,\text{th}\left(\frac{g\mu_B}{2}\beta\lambda M\right)\right]. \tag{2}$$

D'où, en utilisant (J.23) et (J.21),

$$\overline{E}_{\text{c.m.}}(T;\, B_0=0) = -\frac{NpJ}{2}\left(\frac{V}{N}\frac{M}{g\mu_B}\right)^2. \tag{3}$$

On en tire (J.38) à l'aide de la définition (J.32) de M_∞.

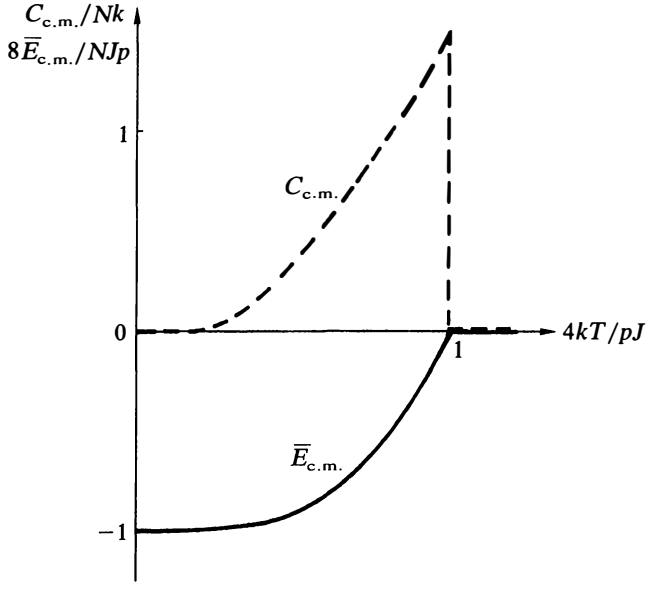


FIGURE 9
Énergie magnétique moyenne en champ nul (trait plein) et capacité calorifique associée (tirets) dans l'approximation de champ moyen.

La capacité calorifique correspondante

$$C_{c.m.} \hat{=} \frac{\partial \overline{E}_{c.m.}}{\partial T} = -\frac{NJp}{4M_\infty^2} M_0(T) \frac{dM_0(T)}{dT} \tag{J.40}$$

est, elle aussi, identiquement nulle pour $T > T_c$. En dessous de T_c, elle est positive (puisque $M_0(T)$ est une fonction décroissante de T). Son allure est donnée sur la figure 9. Au voisinage de T_c, elle est discontinue (cf. (J.39)) :

$$C_{c.m.}(T_c^-) - C_{c.m.}(T_c^+) = \frac{3}{8}\frac{NJp}{T_c} = \frac{3}{2} Nk \,. \tag{J.41}$$

3. PROPRIÉTÉS DU SYSTÈME EN PRÉSENCE DE CHAMP EXTÉRIEUR

Il nous faut maintenant résoudre l'équation complète (J.23) en gardant le champ magnétique B_0. Là encore, on est conduit à une résolution graphique; le plus simple est de laisser la tangente hyperbolique fixe et de faire varier la droite avec les paramètres, c'est-à-dire de poser

$$\frac{M}{M_\infty} = \text{th } x \,, \tag{J.42}$$

et

$$\frac{M}{M_\infty} = \frac{Tx}{T_c} - \frac{g\mu_B}{2kT_c} B_0 \,. \tag{J.43}$$

Les deux courbes sont tracées sur les figures 10 (pour $T > T_c$) et 11 (pour $T < T_c$).

T fixe la pente de la droite (J.43) et B_0 détermine son ordonnée à l'origine[23]. Le comportement du système étant évidemment symétrique dans le changement $B_0 \longrightarrow -B_0$, il suffit d'étudier le cas $B_0 > 0$.

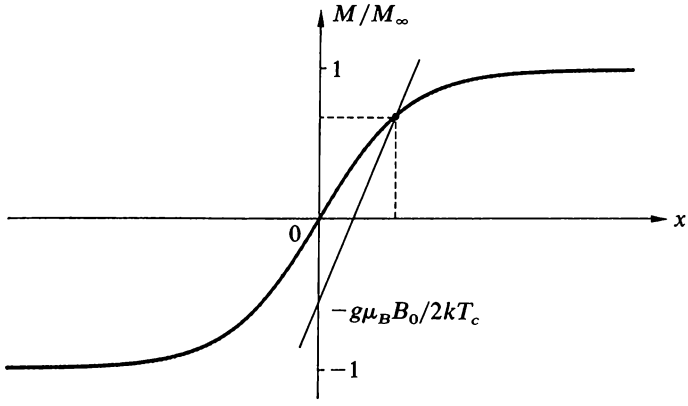


FIGURE 10

Résolution graphique de l'équation d'autocohérence en présence du champ B_0 pour $T > T_c$. Le point d'intersection de la courbe $M/M_\infty = \text{th}\, x$ et de la droite $M/M_\infty = Tx/T_c - g\mu_B B_0/2kT_c$ détermine l'aimantation du système.

a. Comportement au-dessus de T_c

Si $T > T_c$ (figure 10), la pente de la droite (J.43) est supérieure à la dérivée de la tangente hyperbolique en tout point. Il y a toujours un point d'intersection et un seul, correspondant à un minimum unique de l'énergie libre. En faisant varier (à température constante) l'ordonnée à l'origine de la droite, on voit aisément que M est une fonction continue de B_0, croissant de façon monotone de 0 à M_∞ comme sur la figure 4.c.

Quand B_0 tend vers zéro, le point d'intersection des deux courbes se rapproche de l'origine; un développement limité de th x est alors possible, et on trouve[24] :

$$\frac{M}{M_\infty} \simeq \frac{g\mu_B B_0}{2k(T - T_c)} \quad , \qquad T > T_c \, ; \, B_0 \ll \frac{2k}{g\mu_B}(T - T_c) \tag{J.44}$$

---

23. Cette ordonnée à l'origine est égale à

$$-\frac{g\mu_B B_0}{2kT_c} = \frac{-g\mu_B B_0}{2} \times \frac{1}{pJ/4} \, ; \tag{1}$$

c'est donc le rapport entre l'énergie d'un des spins dans le champ B_0 et la valeur maximum que peut prendre l'énergie d'interaction de ce spin avec ses p voisins.

On peut aussi l'écrire

$$-\frac{g\mu_B B_0}{2kT_c} = -B_0 \times \frac{1}{2kT_c/g\mu_B} \, ; \tag{2}$$

elle apparaît alors comme le rapport entre B_0 et la valeur maximum $pJ/2g\mu_B$ du champ moléculaire. En ordre de grandeur, ce champ maximum vaut T_c teslas : dans tous les cas pratiques, l'ordonnée à l'origine est petite devant l'unité.

24. Lorsque B_0 tend vers 0 à température fixée, la droite de la figure 10 coupe la courbe th x dans une région où celle-ci est pratiquement confondue avec sa tangente à l'origine. Il suffit donc de remplacer th x par x dans (J.42) et de reporter dans (J.43) pour obtenir (J.44).

On en déduit la susceptibilité magnétique au-dessus de la température critique :

$$\chi \cong \lim_{B_0 \to 0} \frac{\partial M}{\partial B_0} = \lim_{B_0 \to 0} \frac{M}{B_0} = \frac{N}{V} \left(\frac{g\mu_B}{2}\right)^2 \frac{1}{k(T - T_c)} , \quad T > T_c . \tag{J.45}$$

La susceptibilité magnétique pour $T > T_c$ se déduit donc simplement de la loi de Curie obtenue pour un paramagnétique parfait (complément III.A, formule (A.25)) en remplaçant la température T par $(T - T_c)$: c'est ce qu'on appelle la *loi de Curie-Weiss.*

b. Comportement en dessous de T_c

Deux cas sont à considérer suivant les valeurs de B_0 (fig. 11). Pour les fortes valeurs de B_0, il n'y a qu'un seul point d'intersection; l'énergie libre possède un seul minimum qui correspond à l'aimantation d'équilibre du système. Lorsque B_0 diminue, en dessous d'une valeur $B_0^0(T)$ apparaissent deux autres points d'intersection[25];

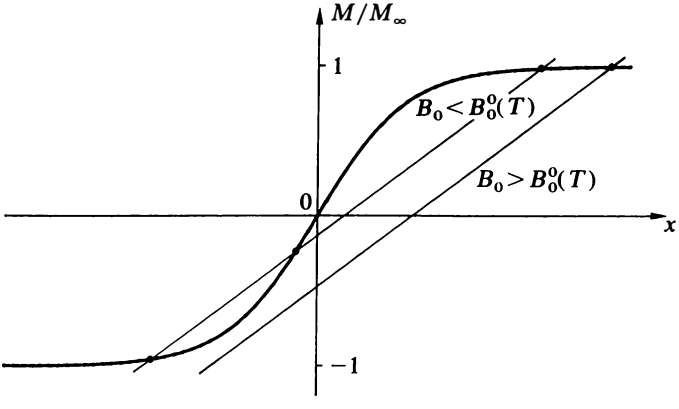


FIGURE 11

Résolution graphique de l'équation d'autocohérence en présence du champ B_0 pour $T < T_c$. Pour $B_0 > B_0^0(T)$, il y a une seule solution. Pour $B_0 < B_0^0(T)$, apparaissent deux autres points d'intersection; mais c'est encore l'aimantation de même signe que B_0 qui correspond à l'équilibre stable.

25. La valeur $B_0^0(T)$ du champ extérieur correspond à la situation où la courbe th x et la droite (J.43) sont tangentes. Elle se détermine le plus aisément par une méthode graphique (fig. 12).

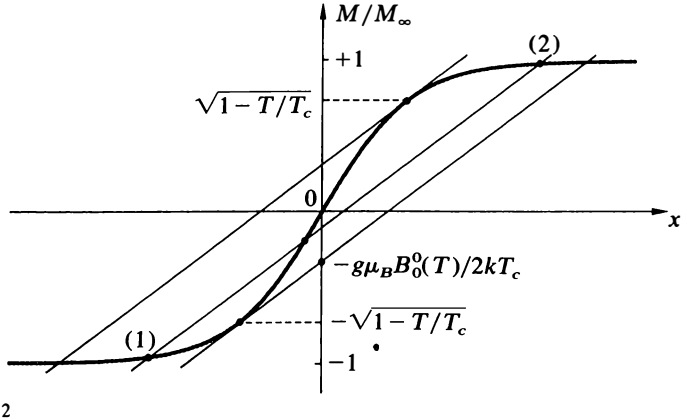


FIGURE 12

Détermination graphique de la valeur $B_0^0(T)$ du champ extérieur en dessous de laquelle la droite et la tangente hyperbolique se coupent en trois points.

.../...

l'énergie libre $F_{c.m.}(M)$ présente alors deux minimums séparés par un maximum (fig. 13). Tant que B_0 reste positif, le minimum correspondant à la valeur positive de l'aimantation est plus bas que celui qui correspond à la valeur négative[26]; c'est donc lui qui fixe l'aimantation d'équilibre du système. Cette valeur de l'aimantation ne s'annule pas avec B_0; lorsque B_0 s'annule, cependant, les deux minimums de $F_{c.m.}$ se

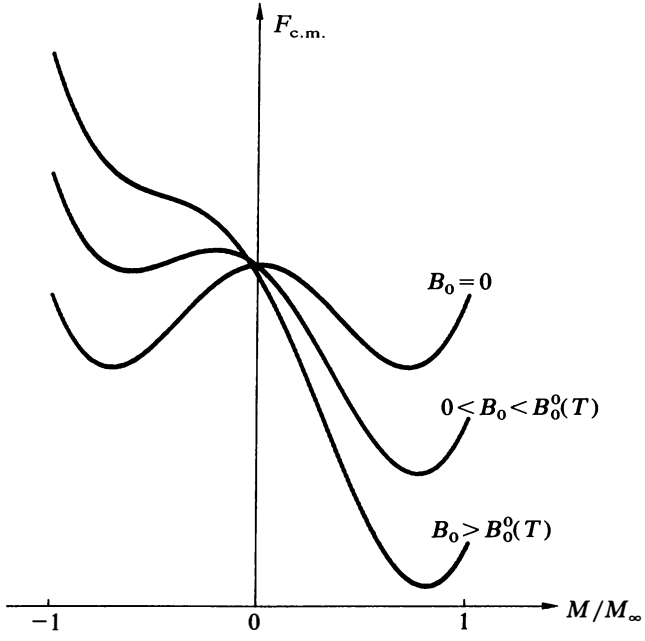


FIGURE 13
Énergie libre en fonction de l'aimantation, à température fixée inférieure à la température critique, pour différentes valeurs du champ B_0.

La dérivée de th x vaut $(1-\text{th}^2 x)$; en l'égalant à la pente T/T_c de la droite on obtient

$$\text{th } x = \pm\sqrt{1 - T/T_c}\,.$$

Il suffit donc de pointer les valeurs $\pm\sqrt{1-T/T_c}$ sur l'axe des ordonnées et de tracer les tangentes à th x aux points correspondants; ces droites coupent l'axe des ordonnées en $\pm(g\mu_B/2kT_c)\,B_0^0(T)$.

26. C'est physiquement évident, reste à s'en convaincre mathématiquement. En utilisant les définitions (J.32) et (J.35) de M_∞ et de T_c, on peut écrire l'énergie libre (J.29) sous la forme

$$F_{c.m.} = -NkT \ln 2 + \frac{Nk}{2}\left\{T_c\left(\frac{M}{M_\infty}\right)^2 - T \ln \text{ch}^2\left[\frac{g\mu_B}{2kT} B_0 + \frac{T_c}{T}\frac{M}{M_\infty}\right]\right\}. \qquad (1)$$

Supposons $B_0 > 0$ et soit $M_1 < 0$ et $M_2 > 0$ les deux valeurs de l'aimantation qui rendent $F_{c.m.}$ minimum. On vérifie directement sur la figure 12 que

$$\sqrt{1-T/T_c} < \left|\frac{M_1}{M_\infty}\right| < \left|\frac{M_2}{M_\infty}\right| < 1\,. \qquad (2)$$

Ces deux valeurs M_1 et M_2 vérifient les équations (J.42) et (J.43); on a donc

$$\text{ch}^2\left[\frac{g\mu_B B_0}{2kT} + \frac{T_c}{T}\frac{M_i}{M_\infty}\right] = \text{ch}^2 x = \frac{1}{1-\text{th}^2 x} = \frac{1}{1-(M_i/M_\infty)^2}\,; \quad i = 1, 2\,. \qquad (3)$$

.../...

produisent pour des valeurs opposées de l'aimantation, et ils deviennent égaux. Si maintenant B_0 change de sens ($B_0<0$), le comportement s'inverse : c'est alors l'état d'aimantation négative qui devient le plus favorable. On obtient ainsi, si on trace M en fonction de B_0 à température fixée, une courbe symétrique, discontinue en $B_0=0$, tout à fait semblable à celle de la figure 4.a. Lorsque T diminue, la pente de la droite (J.43) diminue, et on voit sur la figure 11 que la discontinuité s'accentue.

Remarquons que l'existence, pour les valeurs de $|B_0|$ inférieures à $B_0^\circ(T)$, d'un minimum relatif de $F_{c.m.}(M)$, conduit à la possibilité d'états métastables (cf. complément III.G, § III.4.c) : si le système a été préparé dans un état d'aimantation positive à l'aide d'un champ B_0 positif, il pourra rester « piégé » dans cet état lorsqu'on inversera le sens du champ, et cela tant que le minimum local de $F_{c.m.}(M)$ correspondant à une valeur positive de l'aimantation existera, c'est-à-dire tant que $|B_0|<B_0^\circ(T)$. Quand le champ extérieur, négatif, deviendra trop important en valeur absolue, le système « sautera » dans l'état d'aimantation négative correspondant au minimum absolu de $F_{c.m.}(M)$, c'est-à-dire dans son véritable état d'équilibre[27].

Reste à déterminer le comportement du système à faible B_0 pour en déduire sa susceptibilité magnétique. Lorsque B_0 tend vers zéro, l'aimantation tend vers la valeur $M_0(T)$ en champ nul, de sorte que

$$M(T;B_0)\simeq M_0(T)+\chi(T)B_0 \quad , \quad B_0\ll\frac{2kT_c}{g\mu_B} . \tag{J.46}$$

La susceptibilité $\chi(T)$ se calcule aisément en dérivant les deux membres de l'équation d'autocohérence (J.23)[28] :

$$\chi(T)\widehat{=}\lim_{B_0\to 0}\frac{\partial M}{\partial B_0}=\frac{M_\infty g\mu_B}{2kT_c}\,\frac{1-\left(\dfrac{M_0(T)}{M_\infty}\right)^2}{\dfrac{T}{T_c}-1+\left(\dfrac{M_0(T)}{M_\infty}\right)^2} \quad , \quad T<T_c . \tag{J.47}$$

L'énergie libre correspondante vaut

$$F_{c.m.}(M_i)=-NkT\ln 2+\frac{Nk}{2}\left\{T_c\left(\frac{M_i}{M_\infty}\right)^2+T\ln\left[1-\left(\frac{M_i}{M_\infty}\right)^2\right]\right\} ; \quad i=1,2 . \tag{4}$$

Considérée comme fonction de la variable $u=(M_i/M_\infty)^2$, $F_{c.m.}$ est décroissante pour

$$(1-T/T_c)<u<1, \tag{5}$$

de sorte que

$$F_{c.m.}(M_2)<F_{c.m.}(M_1) . \tag{6}$$

La position d'équilibre correspond donc à la valeur positive de l'aimantation.

27. L'hystérésis observée dans les matériaux magnétiques est plus complexe. Les interactions magnétiques dipolaires et la structure en domaines y jouent un rôle important.

28. La dérivation de (J.23) donne

$$\frac{1}{M_\infty}\frac{\partial M}{\partial B_0}=\left[1-\mathrm{th}^2\left(\frac{g\mu_B}{2kT}B_0+\frac{T_c}{T}\frac{M}{M_\infty}\right)\right]\left[\frac{g\mu_B}{2kT}+\frac{T_c}{T}\frac{1}{M_\infty}\frac{\partial M}{\partial B_0}\right] , \tag{1}$$

ou encore

$$\frac{1}{M_\infty}\frac{\partial M}{\partial B_0}=\left[1-\left(\frac{M}{M_\infty}\right)^2\right]\left[\frac{g\mu_B}{2kT}+\frac{T_c}{T}\frac{1}{M_\infty}\frac{\partial M}{\partial B_0}\right] . \tag{2}$$

On résout alors l'équation (2) en $(1/M_\infty)(\partial M/\partial B_0)$, puis on obtient χ en remplaçant M par sa valeur en champ nul $M_0(T)$.

Au voisinage de T_c, il suffit de remplacer $M_0(T)/M_\infty$ par son expression approchée (J.37) et M_∞ par sa définition (J.32) pour obtenir

$$\chi(T) \simeq \frac{N}{2V}\left(\frac{g\mu_B}{2}\right)^2 \frac{1}{k(T_c - T)} \quad , \quad 0 < \frac{T_c - T}{T_c} \ll 1 \ , \tag{J.48}$$

c'est-à-dire la même expression qu'au-dessus de T_c (formule (J.45)) à un facteur $(-1/2)$ près.

4. LES EXPOSANTS CRITIQUES

Les résultats précédents fournissent les valeurs des exposants critiques dans l'approximation de champ moyen. De la formule (J.37) on déduit l'exposant associé à l'aimantation (formule (J.2)) :

$$\beta_{\text{c.m.}} = 0{,}5 \ . \tag{J.49}$$

La capacité calorifique (J.40) est discontinue en T_c, mais elle tend vers une valeur finie de part et d'autre de T_c; conventionnellement, l'exposant associé à une discontinuité de ce type est pris égal à zéro[29] :

$$\alpha_{\text{c.m.}} = \alpha'_{\text{c.m.}} = 0 \ . \tag{J.50}$$

Enfin[30], la susceptibilité magnétique diverge de part et d'autre de T_c; de (J.45) et (J.48) on déduit

$$\gamma_{\text{c.m.}} = \gamma'_{\text{c.m.}} = 1 \ . \tag{J.51}$$

Les exposants critiques calculés dans l'approximation de champ moyen sont donc *universels :* ils ne dépendent ni de la force du couplage, ni du nombre p de plus proches voisins; ils sont les mêmes pour tous les matériaux (ils resteraient même inchangés si l'on étudiait un cristal dans un espace de dimension différente de 3).

5. COMPARAISON AVEC LES RÉSULTATS EXPÉRIMENTAUX

Comme dans d'autres circonstances (cf. compléments III.G, III.H et III.K), l'approximation de champ moyen reproduit bien, tout au moins qualitativement, le phénomène physique tel que nous l'avons décrit au premier paragraphe. Au vu de la courbe de la figure 14, on peut même dire que l'accord avec les résultats expérimentaux est tout à fait satisfaisant, surtout si on considère la rusticité de l'approximation.

---

29. Cette convention résulte en fait de la définition de l'exposant critique donnée à la note 8.

30. On peut aussi déterminer l'exposant δ (note 9). La relation (J.23) peut en effet s'écrire

$$\frac{g\mu_B}{2kT_c} B_0 = -\frac{M}{M_\infty} + \frac{T}{T_c} \operatorname{Argth} \frac{M}{M_\infty} \ . \tag{1}$$

Pour $T = T_c$ et à faible B_0, M/M_∞ est petit et on peut effectuer un développement limité :

$$\operatorname{Argth} \frac{M}{M_\infty} \simeq \frac{M}{M_\infty} + \frac{1}{3}\left(\frac{M}{M_\infty}\right)^3 \ . \tag{2}$$

On obtient ainsi

$$\frac{g\mu_B B_0}{2kT_c} = \frac{1}{3}\left(\frac{M(T_c, B_0)}{M_\infty}\right)^3 \ , \tag{3}$$

d'où l'on déduit

$$\delta = 3 \ . \tag{4}$$

Cependant, le voisinage immédiat du point critique est mal décrit : les exposants critiques ont des valeurs assez nettement différentes de celles qui sont observées (0,5 au lieu de 0,35 pour β; 1 au lieu de 1,3 pour γ; ...). Il aurait d'ailleurs été tout à fait étonnant qu'il en soit autrement : dans son principe, l'approximation néglige les fluctuations; or c'est justement dans la région du point critique que les fluctuations sont les plus importantes, de sorte que l'approximation ne peut être valable près de T_c[31].

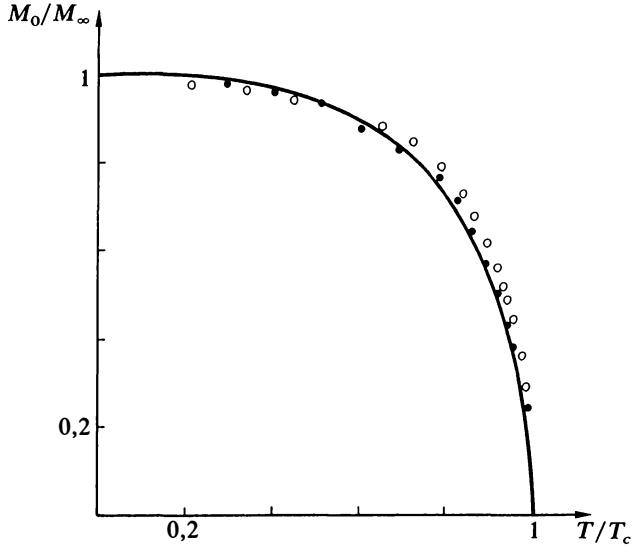


FIGURE 14
Courbe d'aimantation spontanée en fonction de la température dans l'approximation de champ moyen, comparée avec les points expérimentaux : ○ fer, • cobalt et nickel (d'après F. Tyler, *Phil. Mag.* 11, 596 (1931)).

III. LE MODÈLE D'ISING ET SES SOLUTIONS EXACTES

1. PRÉSENTATION DU MODÈLE D'ISING

La difficulté de l'hamiltonien de Heisenberg provient du fait que les trois composantes d'un opérateur de spin ne commutent pas entre elles. C'est pourquoi l'approximation d'Ising ne retient que les composantes dans la direction Oz du champ appliqué et décrit donc le système par l'hamiltonien

$$H_I = -g\mu_B B_0 \sum_{i=1}^{N} S_i^z - J \sum_{(i,j)_v} S_i^z S_j^z , \quad J>0 . \tag{J.52}$$

31. De toute façon, les fluctuations et les corrélations ne sont jamais vraiment négligeables : on peut calculer exactement, dans le modèle de Heisenberg, le comportement de l'aimantation dans la limite des basses températures (exercice VI.29); on trouve qu'elle s'écarte de la valeur M_∞ par un terme variant en $T^{3/2}$ et non en $\exp(-2T_c/T)$ comme nous l'avons obtenu dans l'approximation de champ moyen (formule (J.36)).

L'aimantation étant dirigée selon le champ appliqué, on peut arguer, pour justifier ce modèle, du fait que les valeurs moyennes des autres composantes des spins sont nulles[32]. Mais la véritable justification du modèle d'Ising réside dans la simplicité de son hamiltonien : tous les S_i^z commutent entre eux, et H_I est ainsi diagonal dans la représentation où les S_i^z le sont; les valeurs propres et les fonctions propres de H_I sont donc connues. C'est le modèle le plus simple qui conserve les ingrédients physiques indispensables à une compréhension minimale du phénomène étudié[33].

Dans l'approximation de champ moyen, le modèle d'Ising donne évidemment (cf. note 32) les mêmes résultats que le modèle de Heisenberg. Ce qui lui confère une importance théorique considérable, c'est qu'on en connaît des *solutions exactes*. En fait, malgré la simplicité du modèle, on ne connaît une telle solution, c'est-à-dire une expression analytique de la fonction de partition, que pour un système unidimensionnel ou bidimensionnel (et encore, dans ce dernier cas, seulement en l'absence de champ extérieur).

2. SOLUTION DU MODÈLE D'ISING À UNE DIMENSION

a. Modèle d'Ising pour un cristal unidimensionnel

Un réseau unidimensionnel est constitué par une chaîne linéaire de N atomes régulièrement espacés. Pour simplifier, nous supposerons cette chaîne fermée sur elle-même (fig. 15), ce qui a l'avantage de rendre tous les sites équivalents, sans changer (dans la limite $N \gg 1$) les propriétés du système. Chaque atome a un spin 1/2 qui interagit avec ses deux voisins et l'hamiltonien d'Ising s'écrit

$$H_I = -\sum_{i=1}^{N} [g\mu_B B_0 S_i^z + J S_i^z S_{i+1}^z] , \quad J > 0 , \tag{J.53}$$

où le site $(N+1)$ et le site (1) coïncident.

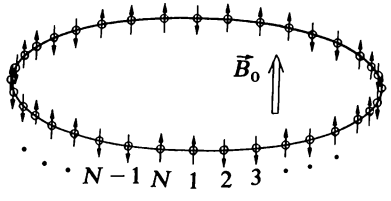


FIGURE 15
Chaîne linéaire de N spins fermée sur elle-même. Chaque spin interagit avec ses deux voisins. L'ensemble est placé dans le champ $\vec{B}_0$.

Un tel système possède 2^N états : chaque opérateur S_i^z peut en effet prendre deux valeurs, $+1/2$ et $-1/2$. La fonction de partition a donc pour expression

$$Z = \sum_{S_1^z = \pm\frac{1}{2}} \sum_{S_2^z = \pm\frac{1}{2}} \cdots \sum_{S_N^z = \pm\frac{1}{2}} \exp\left[\frac{1}{kT} \sum_{i=1}^{N} [g\mu_B B_0 S_i^z + J S_i^z S_{i+1}^z]\right] . \tag{J.54}$$

32. Ceci revient à dire qu'on traite ici les composantes de spin suivant les axes x et y en champ moyen, espérant faire mieux pour les composantes suivant Oz.

33. Il existe des cristaux dans lesquels l'interaction d'échange est très anisotrope (par exemple $FeCl_2$ et $FeCO_3$), et pour lesquels le modèle d'Ising est une bonne approximation. Il peut aussi s'appliquer tel quel à d'autres problèmes physiques ne faisant intervenir, sur chaque site, que deux états : c'est le cas des alliages binaires (cf. complément III.K).

b. Calcul de la fonction de partition et de l'énergie libre

Si l'on pose

$$\mathcal{C}(S_i^z, S_j^z) \widehat{=} \exp \frac{1}{kT} \left[\frac{g\mu_B B_0}{2} (S_i^z + S_j^z) + J S_i^z S_j^z \right], \tag{J.55}$$

on vérifie aisément que Z prend la forme

$$Z = \sum_{S_1^z = \pm\frac{1}{2}} \sum_{S_2^z = \pm\frac{1}{2}} \cdots \sum_{S_N^z = \pm\frac{1}{2}} \mathcal{C}(S_1^z, S_2^z)\, \mathcal{C}(S_2^z, S_3^z) \cdots \mathcal{C}(S_N^z, S_1^z) . \tag{J.56}$$

Les coefficients $\mathcal{C}(S_i^z, S_j^z)$ peuvent être considérés comme les éléments d'une matrice 2×2, *la même pour tous les couples de sites :*

$$\mathbb{T} \widehat{=} \begin{bmatrix} \mathcal{C}\left(+\frac{1}{2}, +\frac{1}{2}\right) & \mathcal{C}\left(+\frac{1}{2}, -\frac{1}{2}\right) \\ \mathcal{C}\left(-\frac{1}{2}, +\frac{1}{2}\right) & \mathcal{C}\left(-\frac{1}{2}, -\frac{1}{2}\right) \end{bmatrix} . \tag{J.57}$$

$\mathbb{T}$ est appelée *« matrice de transfert »* car, par définition même du produit des matrices,

$$\sum_{S_j^z = \pm\frac{1}{2}} \mathcal{C}(S_i^z, S_j^z)\, \mathcal{C}(S_j^z, S_k^z) = (\mathbb{T}^2)_{S_i S_k} . \tag{J.58}$$

Cette technique matricielle permet d'effectuer formellement, dans l'expression (J.56) de la fonction de partition, les sommations sur S_2^z, S_3^z, ..., S_N^z, ce qui donne le résultat simple

$$Z = \sum_{S_i^z = \pm\frac{1}{2}} (\mathbb{T}^N)_{S_i S_i} = \text{Trace}\, (\mathbb{T}^N) . \tag{J.59}$$

La matrice $\mathbb{T}$, réelle et symétrique, est diagonalisable et ses valeurs propres λ_1 et λ_2 sont réelles. La formule (J.59) s'écrit donc[34]

$$Z = \lambda_1^N + \lambda_2^N . \tag{J.60}$$

Soit λ_1 la plus grande des deux valeurs propres (en valeur absolue); on peut mettre Z sous la forme

$$Z = \lambda_1^N \left[1 + \left(\frac{\lambda_2}{\lambda_1} \right)^N \right], \quad \text{avec} \quad \left| \frac{\lambda_2}{\lambda_1} \right| < 1 , \tag{J.61}$$

ce qui, dans la limite $N \gg 1$, est équivalent à

$$Z = \lambda_1^N . \tag{J.62}$$

---

34. La trace d'une matrice est une caractéristique de l'opérateur associé; sa valeur est indépendante de la base sur laquelle la matrice est exprimée. Si on prend comme base celle des vecteurs propres de $\mathbb{T}$, alors

$$\mathbb{T} = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \quad \text{et} \quad \mathbb{T}^N = \begin{pmatrix} \lambda_1^N & 0 \\ 0 & \lambda_2^N \end{pmatrix} ; \tag{1}$$

d'où $\text{Trace}\, (\mathbb{T}^N) = \lambda_1^N + \lambda_2^N$. (2)

Reste alors à calculer cette plus grande valeur propre de la matrice $\mathbb{T}$; elle vaut[35]

$$\lambda_1 = \exp\frac{J}{4kT} \times \left\{ \operatorname{ch}\frac{g\mu_B B_0}{2kT} + \left(\operatorname{sh}^2 \frac{g\mu_B B_0}{2kT} + \exp -\frac{J}{kT} \right)^{1/2} \right\} . \tag{J.63}$$

On en déduit l'énergie libre dans le modèle d'Ising à une dimension :

$$\begin{aligned} F(B_0, T) &\widehat{=} -kT \ln Z \\ &= -NkT \left\{ \frac{J}{4kT} + \ln \left[\operatorname{ch}\frac{g\mu_B B_0}{2kT} + \left(\operatorname{sh}^2 \frac{g\mu_B B_0}{2kT} + \exp -\frac{J}{kT} \right)^{1/2} \right] \right\} . \end{aligned} \tag{J.64}$$

En l'absence de champ extérieur ($B_0 = 0$), F prend une forme beaucoup plus simple :

$$F(B_0 = 0, T) = -NkT \ln \left(2 \operatorname{ch}\frac{J}{4kT} \right) . \tag{J.65}$$

c. Aimantation

Le moment magnétique moyen $\mathcal{M}$ du système est donné, au facteur $g\mu_B$ près, par la valeur moyenne de la somme des S_i^z :

$$\begin{aligned} \mathcal{M}(B_0, T) = g\mu_B \overline{\sum_{i=1}^{N} S_i^z} = \frac{1}{Z} \sum_{S_1^z = \pm\frac{1}{2}} \cdots \sum_{S_N^z = \pm\frac{1}{2}} \left(\sum_{i=1}^{N} g\mu_B S_i^z \right) \\ \exp \frac{1}{kT} \left[\sum_{i=1}^{N} (g\mu_B B_0 S_i^z + J S_i^z S_{i+1}^z) \right] . \end{aligned} \tag{J.66}$$

Si l'on compare (J.66) avec l'expression (J.54) de Z, on voit que le moment magnétique s'écrit simplement

$$\mathcal{M}(B_0, T) = kT \frac{\partial \ln Z}{\partial B_0} = -\frac{\partial F(B_0, T)}{\partial B_0} . \tag{J.67}$$

L'aimantation, qui est ici le moment magnétique moyen par unité de longueur, s'obtient en divisant par la longueur L de la chaîne :

$$M(B_0, T) = -\frac{1}{L}\frac{\partial F}{\partial B_0} = \frac{N}{L}\frac{g\mu_B}{2} \frac{\operatorname{sh}\dfrac{g\mu_B B_0}{2kT}}{\left[\operatorname{sh}^2 \dfrac{g\mu_B B_0}{2kT} + \exp -\dfrac{J}{kT} \right]^{1/2}} . \tag{J.68}$$

---

35. λ_1 et λ_2 sont les solutions de l'équation

$$\left[\mathcal{C}\left(+\frac{1}{2}, +\frac{1}{2} \right) - \lambda \right] \left[\mathcal{C}\left(-\frac{1}{2}, -\frac{1}{2} \right) - \lambda \right] - \mathcal{C}\left(+\frac{1}{2}, -\frac{1}{2} \right) \mathcal{C}\left(-\frac{1}{2}, +\frac{1}{2} \right) = 0 \tag{1}$$

avec, d'après (J.55),

$$\mathcal{C}\left(+\frac{1}{2}, +\frac{1}{2} \right) = \exp \left[\frac{1}{kT} \left(\frac{g\mu_B B_0}{2} + \frac{J}{4} \right) \right] , \tag{2}$$

$$\mathcal{C}\left(-\frac{1}{2}, -\frac{1}{2} \right) = \exp \left[\frac{1}{kT} \left(-\frac{g\mu_B B_0}{2} + \frac{J}{4} \right) \right] , \tag{3}$$

$$\mathcal{C}\left(+\frac{1}{2}, -\frac{1}{2} \right) = \mathcal{C}\left(-\frac{1}{2}, +\frac{1}{2} \right) = \exp \left(-\frac{J}{4kT} \right) . \tag{4}$$

Les racines de l'équation (1) sont donc

$$\lambda = \exp \left(\frac{J}{4kT} \right) \times \left[\operatorname{ch}\frac{g\mu_B B_0}{2kT} \pm \left(\operatorname{sh}^2 \frac{g\mu_B B_0}{2kT} + \exp \left(-\frac{J}{kT} \right) \right)^{1/2} \right] . \tag{5}$$

La plus grande en valeur absolue λ_1 est celle qui correspond au signe + ; elle est visiblement positive.

d. Conclusion

Comme on peut le vérifier sur l'expression (J.64), l'énergie libre est une fonction continue, à dérivées continues, de B_0 et de T pour tout champ magnétique et toute température. Il en est de même de l'énergie magnétique moyenne et de la capacité calorifique. L'aimantation (J.68) est, elle aussi, une fonction analytique; elle s'annule pour $B_0 = 0$ quelle que soit la température, de sorte que la courbe représentative de M en fonction de B_0 a toujours l'allure représentée sur la figure 4.c; la susceptibilité magnétique reste finie à toute température non nulle. *Le modèle d'Ising unidimensionnel ne présente donc pas de transition de phase.*

Ces résultats exacts sont radicalement différents de ceux qu'on obtient dans l'approximation de champ moyen : celle-ci prédit en effet une transition de phase (les calculs sont analogues à ceux du paragraphe II), qui n'existe pas dans la solution exacte. A une dimension, les fluctuations et les corrélations ne sont donc jamais négligeables et *l'approximation de champ moyen conduit à des résultats totalement erronés.*

3. LA SOLUTION À DEUX DIMENSIONS

Les résultats précédents sont quelque peu décevants, puisqu'ils ne présentent pas de singularités, caractéristiques essentielles d'une transition de phase. Il n'en est heureusement pas de même pour les systèmes bidimensionnels.

C'est en 1944 que Lars Onsager publia la première détermination analytique exacte de l'énergie libre d'un système d'Ising à deux dimensions en champ nul, et c'est en 1952 que C.N. Yang donnait le calcul de l'aimantation spontanée de ce système en fonction de la température.

Depuis, plusieurs versions de ces calculs ont été publiées; aucune n'est simple, aucune n'est courte. Malgré les très nombreuses simplifications qu'il a fallu faire pour arriver à un modèle soluble (le modèle d'Ising), la solution reste un véritable exploit mathématique et il n'a pas été possible, jusqu'à présent, de résoudre exactement le problème en présence de champ extérieur. Nous ne donnerons ici que les résultats, et pour le seul réseau carré[36].

a. Fonction de partition exacte pour un réseau carré

Dans le cas où chaque atome possède 4 premiers voisins équivalents, la fonction de partition exacte s'écrit (si $\vec{B}_0 = \vec{0}$)

$$Z = \left[2 \operatorname{ch} \frac{J}{2kT} . e^{I}\right]^{N}, \tag{J.69}$$

avec $$I \widehat{=} \frac{1}{2\pi} \int_0^{\pi} d\phi \ln \left[\frac{1 + (1 - x^2 \sin^2 \phi)^{1/2}}{2}\right], \tag{J.70}$$

où $$x \widehat{=} 2 \frac{\operatorname{sh} J/2kT}{\operatorname{ch}^2 J/2kT} . \tag{J.71}$$

---

36. On peut se faire une idée de la difficulté à la seule lecture des dates :

1925 — Modèle d'Ising et solution à une dimension.

1928 — Modèle de Heisenberg.

1944 — Calcul de l'énergie libre pour le modèle d'Ising à deux dimensions par L. Onsager (*Phys. Rev. 65,* 117 (1944)).

1949 — L. Onsager annonce le résultat pour l'aimantation dans le modèle d'Ising à deux dimensions à la conférence de Florence.

1952 — C.N. Yang publie le calcul de cette aimantation (*Phys. Rev. 85,* 808 (1952)).

Pour la démonstration, on pourra consulter les articles originaux, mais nous conseillons plutôt les références 7, 8 et 9.

De l'expression de Z, on déduit les propriétés du système. Par exemple, on trouve pour l'énergie magnétique moyenne

$$\bar{E} = -\frac{\partial \ln Z}{\partial \beta} = -\frac{NJ}{2} \operatorname{th} \frac{J}{2kT} - \frac{NJ}{2} \frac{\operatorname{sh}^2 J/2kT - 1}{\operatorname{sh} J/kT} \left[\frac{2}{\pi} K(x) - 1 \right], \tag{J.72}$$

où

$$K(x) \widehat{=} \int_0^{\pi/2} \mathrm{d}\phi \, \frac{1}{(1 - x^2 \sin^2 \phi)^{1/2}} \tag{J.73}$$

est «l'intégrale elliptique complète de première espèce».

b. Existence d'une transition de phase

L'existence d'une transition de phase, si transition il y a, doit se manifester par un comportement singulier des fonctions thermodynamiques ou de leurs dérivées. Ces fonctions dépendent de la température explicitement, mais aussi implicitement par l'intermédiaire du paramètre x (formule (J.71)). Ce paramètre est toujours positif; il s'annule à température nulle et à température infinie; il passe par un maximum égal à 1 pour la valeur T_c de la température telle que[37]

$$\operatorname{sh} \frac{J}{2kT_c} = 1 . \tag{J.74}$$

Or l'intégrale elliptique (J.73) est définie quel que soit $x < 1$, mais diverge pour $x = 1$. Au voisinage de cette valeur[38],

$$K(x) = -\frac{1}{2} \ln(1 - x^2) + \ln 4 + \dots . \tag{J.75}$$

Pour T voisine de T_c, $K(x)$ diverge donc[39] :

$$K(x) \sim -\ln \left| \frac{T - T_c}{T_c} \right| . \tag{J.76}$$

---

37. Si on pose

$$x(t) \widehat{=} 2 \operatorname{sh} t / \operatorname{ch}^2 t , \tag{1}$$

il vient

$$\frac{\mathrm{d}x}{\mathrm{d}t} = \frac{2(1 - \operatorname{sh}^2 t)}{\operatorname{ch}^3 t} , \tag{2}$$

qui s'annule pour la valeur t_c telle que

$$\operatorname{sh} t_c = 1 . \tag{3}$$

Pour cette valeur,

$$\operatorname{ch}^2 t_c = 1 + \operatorname{sh}^2 t_c = 2 , \tag{4}$$

et donc

$$x(t_c) = 1 . \tag{5}$$

38. I.S. Gradshteyn et I.M. Ryzhik, *Table of Integrals, Series, and Products* (Academic Press 1980), paragraphe 8.11.

39. Au voisinage de t_c (note 37),

$$x(t) = x(t_c) + (t - t_c) \left(\frac{\mathrm{d}x}{\mathrm{d}t} \right)_{t=t_c} + \frac{1}{2} (t - t_c)^2 \left(\frac{\mathrm{d}^2 x}{\mathrm{d}t^2} \right)_{t=t_c} + \dots . \tag{1}$$

Or

$$x(t_c) = 1, \quad \left(\frac{\mathrm{d}x}{\mathrm{d}t} \right)_{t=t_c} = 0 \quad \text{et} \quad \left(\frac{\mathrm{d}^2 x}{\mathrm{d}t^2} \right)_{t=t_c} = -2 . \tag{2}$$

On a donc

$$1 - x^2 \simeq 2(t - t_c)^2 , \tag{3}$$

d'où l'on déduit, en remplaçant t par $J/2kT$ et en reportant dans (J.75), la relation (J.76).

La fonction $\overline{E}$ reste cependant continue, puisqu'au voisinage de T_c le deuxième terme de (J.72) se comporte comme[40] $(T-T_c)\ln|T-T_c|$; mais sa dérivée, la capacité calorifique, diverge comme $\ln|T-T_c|$:

$$C_V \sim -\frac{NJ^2}{2k\pi T_c^2}\ln\left|\frac{T-T_c}{T_c}\right| , \quad \left|\frac{T-T_c}{T_c}\right| \ll 1 . \tag{J.77}$$

Cette singularité est caractéristique d'une *transition de phase à la température* T_c. Ce résultat est d'ailleurs confirmé par le calcul de l'aimantation en champ nul de Yang, qui donne

$$M_0(T) = \begin{cases} \dfrac{Ng\mu_B}{2s}\left[1-\dfrac{1}{\mathrm{sh}^4\, J/2kT}\right]^{1/8} & \text{pour} \quad T<T_c \\ 0 & \text{pour} \quad T>T_c \end{cases} \tag{J.78}$$

(s est la surface du cristal). M_0 décroît de la valeur maximum $M_\infty = Ng\mu_B/2s$, atteinte pour $T \to 0$, à la valeur zéro obtenue pour $T \geqslant T_c$. Le système passe ainsi de l'état ferromagnétique à l'état paramagnétique à la température T_c définie par (J.74); la température de Curie du cristal bidimensionnel carré vaut donc

$$T_c = 0{,}567\,296\ \ldots\ J/k . \tag{J.79}$$

Les exposants critiques s'obtiennent par simple inspection des relations (J.78) et (J.77) : la divergence de l'aimantation est caractérisée par[41]

$$\beta = 1/8 , \tag{J.80}$$

et on associe à la singularité logarithmique de la capacité calorifique les valeurs[42]

$$\alpha = \alpha' = 0 . \tag{J.81}$$

---

40. La partie singulière de (J.72) s'écrit

$$\overline{E}_S = -\frac{NJ}{\pi}\frac{\mathrm{sh}^2(J/2kT)-1}{\mathrm{sh}(J/kT)}K(x) . \tag{1}$$

Or, au voisinage de t_c,

$$\frac{\mathrm{sh}^2(J/2kT)-1}{\mathrm{sh}(J/kT)} = \frac{\mathrm{sh}^2\, t-1}{\mathrm{sh}\, 2t} \sim (t-t_c) = \frac{J}{2kT_c^2}(T_c-T) . \tag{2}$$

On a donc

$$\overline{E}_S \simeq \frac{NJ^2}{2\pi kT_c^2}(T_c-T)\ln\left|\frac{T-T_c}{T_c}\right| , \tag{3}$$

ce qui conduit à une capacité calorifique divergente de la forme (J.77).

41. Au voisinage de T_c, (J.78) s'écrit en effet

$$M_0(T) \simeq \frac{Ng\mu_B}{2s}\left[\frac{4J}{\sqrt{2}kT_c^2}(T_c-T)\right]^{1/8}$$

42. Le résultat (J.81) découle directement de la définition de la note 8. Cette définition ne permet pas de distinguer entre une discontinuité simple (comme en champ moyen) et une singularité logarithmique, qui correspondent pourtant à des comportements très différents.

c. Comparaison avec l'approximation de champ moyen

On peut comparer les courbes représentant les expressions exactes de l'énergie moyenne $\bar{E}(T)$ et de la capacité calorifique $C(T)$ avec celles que l'on obtient dans l'approximation de champ moyen (fig. 16 et 17). Les résultats sont assez différents :

— L'approximation de champ moyen surestime largement la température de Curie (J/k au lieu de 0,567 ... J/k).

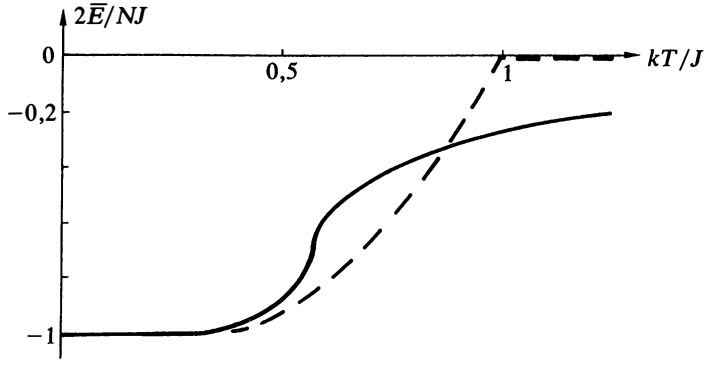


FIGURE 16
Énergie magnétique moyenne pour le modèle d'Ising d'un réseau carré en fonction de la température : valeur exacte (trait plein), approximation de champ moyen (tirets).

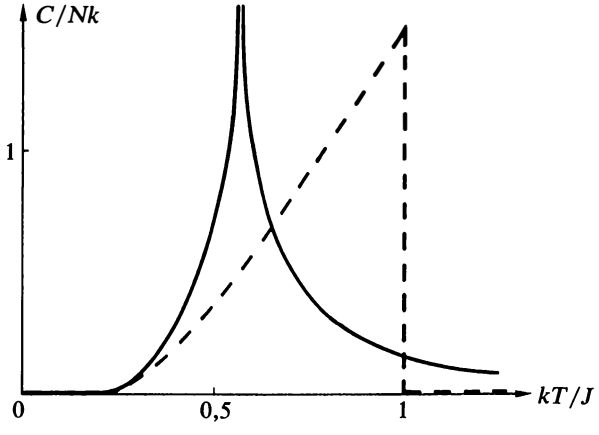


FIGURE 17
Capacité calorifique pour le modèle d'Ising d'un réseau carré : valeur exacte (trait plein) et approximation de champ moyen (tirets).

— Au-dessus de T_c, l'énergie magnétique vaut zéro en champ moyen alors qu'elle reste non nulle dans le résultat exact; c'est que le champ moyen postule une distribution des spins totalement aléatoire au-dessus de T_c; le résultat exact montre qu'il n'en est rien : même si l'échantillon dans son ensemble a une aimantation totale nulle, il subsiste localement, sur de courtes distances, une tendance à l'alignement des spins qui implique une énergie moyenne non nulle.

— Les comportements critiques sont très différents (les coefficients α sont les mêmes, mais la valeur zéro recouvre des singularités de type très différent dans les deux cas : discontinuité en champ moyen, divergence logarithmique pour le résultat exact).

4. CONCLUSION

Avant de terminer ce paragraphe, insistons encore sur l'importance théorique considérable de la solution du modèle d'Ising à deux dimensions : *un calcul exact* de la fonction de partition d'un système de particules en interaction mutuelle *prévoit l'existence d'une transition de phase* sans la postuler *a priori* (comme on le fait dans l'« approximation » de champ moyen). Les méthodes de la mécanique statistique sont donc aptes à décrire de tels phénomènes, et c'est seulement pour des raisons de complication mathématique qu'on arrive rarement à un résultat sans devoir faire appel à des approximations (pas toujours bien contrôlées).

En outre, le fait de disposer de solutions exactes dans certaines situations particulières permet, en y confrontant les résultats, d'évaluer les divers types d'approximations auxquelles on peut avoir recours. On peut ainsi mieux comprendre ces approximations et leurs limites de validité pour pouvoir les appliquer ensuite de façon plus pertinente aux cas que l'on ne sait pas traiter autrement.

Enfin, les solutions du modèle d'Ising mettent en évidence *l'importance des dimensions de l'espace* dans le problème des transitions de phase, puisque les comportements à une et deux dimensions sont radicalement différents.

IV. INDICATIONS SUCCINCTES SUR LES MÉTHODES D'APPROXIMATION ÉLABORÉES

Il n'existe actuellement aucun calcul exact de la fonction de partition pour aucun modèle physiquement raisonnable de particules en interaction mutuelle à trois dimensions. On en est réduit à des approximations et des calculs numériques souvent très élaborés. Nous nous contenterons de passer en revue rapidement les principales méthodes.

Citons d'abord les *méthodes dérivées du champ moyen :* le comportement du système est calculé à partir des propriétés d'un amas constitué par un très petit nombre d'atomes voisins, l'interaction avec le reste du cristal étant traitée approximativement (cf. complément III.K, § III)[43].

On peut également faire appel à des *méthodes purement numériques :* on extrapole le comportement, étudié à l'aide d'ordinateurs, de systèmes grands à l'échelle microscopique mais pas encore macroscopiques à cause des limitations techniques du calcul (les systèmes que l'on peut traiter complètement comportent au mieux quelques centaines d'atomes).

Dans les *méthodes de perturbation,* on développe les grandeurs physiques en série de puissances d'un paramètre petit devant 1 (on utilise par exemple des développements de haute température, où ce paramètre est en $1/T$); la difficulté est d'obtenir suffisamment de termes pour pouvoir en déduire par extrapolation le comportement de la série. Pour le modèle d'Ising à trois dimensions, on a pu calculer jusqu'à 40 termes[44].

---

43. Citons, parmi les plus classiques, les méthodes de H. Bethe (*Proc. Roy. Soc. London A 150,* 552 (1935)), de R.H. Fowler et E.A. Guggenheim (*Proc. Roy. Soc. London A 174,* 189 (1940)), et de R. Kikuchi (*Phys. Rev. 81,* 988 (1951)).

44. Pour plus de détails sur ces méthodes, voir les références 10 et 11.

Les deux premières techniques prédisent des comportements corrects et peuvent même être assez précises, sauf au voisinage immédiat du point critique. Elles négligent en effet, ou traitent approximativement, les corrélations entre un grand nombre de particules ou entre deux particules très éloignées, alors qu'au voisinage de T_c la portée des corrélations devient infinie. La troisième méthode est alors plus fiable, mais elle nécessite pour cela un travail considérable.

La méthode la plus récente, et celle qui ouvre les plus larges perspectives pour l'étude du comportement critique, est celle du *« groupe de renormalisation »* proposée par K. Wilson[45]. Elle consiste à calculer la fonction de partition par étapes successives; à chaque étape, une resommation partielle (par « blocs de spins ») remplace le système initial par un système constitué d'un nombre de « particules » plus faible décrit par un hamiltonien « renormalisé ». Si l'on est capable de caractériser correctement le passage d'une étape à la suivante, on peut en principe obtenir le comportement exact du système[46]. En fait, il est impossible de résoudre le problème sans approximations; mais il est remarquable que ces approximations, même grossières, conduisent directement pour les exposants critiques à des valeurs proches de la réalité. Dès le début des années 70, la méthode du groupe de renormalisation et sa formulation en théorie des champs permettaient de comprendre l'existence des « classes d'universalité » (cf. note 10) et des « lois d'échelle » (qui conduisent à des relations entre exposants critiques[47]) observées expérimentalement dans les comportements critiques des transitions du second ordre[48]. Plus récemment, des calculs faisant appel à des techniques mathématiques très sophistiquées ont permis, par cette méthode, d'obtenir une prédiction quantitative précise des exposants critiques, en très bon accord avec les valeurs expérimentales[49].

---

45. On trouvera une introduction élémentaire dans

— K.G. Wilson, « Les phénomènes de physique et les échelles de longueur », *Pour la science 24* (octobre 1979);

— H. Maris, L. Kadonoff, *Am. J. Physics 46,* 652 (1978);

et une revue plus complète dans

— K.G. Wilson et J.B. Kogut, *Phys. Rep. C12,* 75 (1974);

— M.E. Fisher, *Reviews of Modern Physics 46,* 597 (1974);

— (voir aussi les références 10 et 11).

46. Pour mieux comprendre, nous conseillons l'exercice III.38 qui applique la méthode au cas simple du modèle d'Ising à une dimension.

47. On peut ainsi montrer qu'au voisinage du point critique, la partie singulière

$$F_S\left[(T-T_c)/T_c,\ B\right]$$

de l'énergie libre vérifie une loi d'échelle : quel que soit λ, il existe p et q tels que

$$F_S\left\{\lambda^p\left[(T-T_c)/T_c\right],\ \lambda^q B\right\}=\lambda F_S\left[(T-T_c)/T_c,\ B\right]. \qquad (1)$$

On en déduit facilement (exercice III.37) la relation

$$\alpha+2\beta+\gamma=2\,. \qquad (2)$$

48. Une conséquence intéressante (et inattendue) de la théorie est que les exposants critiques calculés dans l'approximation de champ moyen deviennent exacts si le nombre de dimensions de l'espace est supérieur à 4.

49. Les estimations les plus précises sont actuellement celles de J.C. Le Guillou et J. Zinn-Justin, *Phys. Rev. B21,* 3976 (1980) et *Journal de Physique 48,* 19 (1987).

Il est intéressant de remarquer que les valeurs calculées par ces auteurs sont affectées d'incertitudes plus faibles que les données expérimentales actuelles.

Les tableaux II et III donnent quelques-uns des résultats obtenus pour des systèmes à 2 et 3 dimensions.

TABLEAU II

Comparaison, dans divers modèles, entre la température critique $T_c^{c.m.} = pJ/4k$ donnée par l'approximation de champ moyen et celle que l'on obtient par des méthodes plus élaborées.

Modèle	Dimension d	Nombre p de plus proches voisins	$T_c/T_c^{c.m.}$
Ising réseau carré	2	4	0,567 (solution exacte)
Ising réseau triangulaire	2	6	0,607 (solution exacte)
Ising réseau cubique simple	3	6	0,752
Ising réseau cubique à faces centrées	3	12	0,816
Heisenberg réseau cubique simple	3	6	0,562

TABLEAU III

Principaux exposants critiques : définition et valeurs dans quelques modèles simples.

Définition	Champ moyen	Ising ($d=2$) (exact)	Ising ($d=3$)	Heisenberg ($d=3$)
$M_0 \propto (T_c - T)^{\beta}$	$\beta = 1/2$	$\beta = 1/8$	$\beta = 0{,}327$	$\beta = 0{,}365$
$C \propto \begin{matrix}(T - T_c)^{-\alpha} \\ (T_c - T)^{-\alpha'}\end{matrix}$	$\alpha = \alpha' = 0$ (discontinuité)	$\alpha = \alpha' = 0$ (divergence logarithmique)	$\alpha = \alpha' = 0{,}107$	$\alpha = \alpha' = -0{,}115$
$\chi \propto \begin{matrix}(T - T_c)^{-\gamma} \\ (T_c - T)^{-\gamma'}\end{matrix}$	$\gamma = \gamma' = 1$	$\gamma = \gamma' = 7/4$	$\gamma' = \gamma = 1{,}239$	$\gamma' = \gamma = 1{,}386$

Références

Sur le ferromagnétisme :

1. Pour un exposé élémentaire, Feynman, Leighton, Sands, *Le cours de physique de Feynman,* Tome 2, chapitres 36 et 37, Inter-Éditions, Paris (1979).
2. N. Ashcroft, D. Mermin, *Solid State Physics,* Holt-Saunders, New York, chapitres 32-33 (1981).

3. C. Kittel, *Introduction à la physique de l'état solide,* Dunod, Paris, chapitre 16 (1972).

4. D. Mattis, *The Theory of Magnetism,* Harper and Row, New York (1965).

Sur les phénomènes critiques en général :

5. M.E. Fisher, *Reports on Progress in Physics 30,* 615 (1967).

6. N. Boccara, *La physique des transitions,* Que Sais-je 1389, PUF, Paris (1970); *Symétries brisées,* Hermann, Paris (1976).

Sur le modèle d'Ising et la solution à deux dimensions :

7. L. Landau, E. Lifchitz, *Physique statistique,* Éditions Mir, Moscou (1967), paragraphe 141.

8. K. Huang, *Statistical Mechanics,* John Wiley, New York (1966), chapitres 16 et 17.

9. R. Baxter, *Exactly Solved Models in Statistical Mechanics,* Academic Press, New York (1980).

Sur les méthodes d'approximation élaborées :

10. *Phase Transitions and Critical Phenomena* (C. Domb et M.S. Green eds), Academic Press, New York (1976), volume 6.

11. *Phase Transitions — Cargèse 1980* (M. Lévy, J.C. Le Guillou, J. Zinn-Justin eds), Plenum, New York (1982).

COMPLÉMENT III.K

Transition ordre-désordre dans les alliages binaires

I. La transition ordre-désordre
   1. Un exemple : l'alliage cuivre-zinc
   2. Description du modèle. Énergie de configuration
   3. Définition du paramètre d'ordre
   4. Origine physique de la transition
      a. Effet coopératif
      b. Analogie avec le modèle d'Ising du ferromagnétisme

II. L'approximation de Bragg-Williams
   1. Principe de l'approximation
   2. Détermination du paramètre d'ordre
      a. Calcul de l'énergie E en fonction de s
      b. Fonction de partition et énergie libre pour s fixé
      c. Équation donnant le paramètre d'ordre à la température T
   3. Propriétés du système
      a. Température critique
      b. Capacité calorifique

III. L'approximation de Bethe-Peierls
   1. Principe de l'approximation
   2. Équation d'autocohérence
   3. Calcul du paramètre d'ordre et de l'énergie
      a. Le paramètre d'ordre moyen $\bar{s}$
      b. Le paramètre $\bar{\sigma}$ et l'énergie moyenne
   4. Propriétés du système
      a. Température critique
      b. Comportement du système pour $T > T_c$
      c. Comportement du système pour $T < T_c$
      d. Capacité calorifique
   5. Conclusion

Les solides métalliques purs présentent pour la plupart une structure cristalline simple : cubique centrée, cubique à faces centrées ou hexagonale compacte. Lorsqu'on parvient à incorporer à un métal A une proportion appréciable d'atomes d'un élément B différent, on réalise un *alliage* binaire.

Il existe plusieurs types distincts d'alliages. Par exemple, si les atomes B sont beaucoup plus petits que les atomes A, ils peuvent se glisser dans les interstices du réseau cristallin du métal A, ce qui donne une *« solution solide d'insertion »* : les aciers, solutions de carbone dans le fer, sont de ce type. Dans d'autres cas, les atomes B que l'on ajoute au cristal peuvent remplacer les atomes A en certains nœuds du réseau cristallin; on parle alors de *« solution solide de substitution »*. Ce type d'alliage est obtenu lorsque les atomes A et B sont de taille comparable : par exemple, le cuivre et le zinc peuvent former de telles solutions solides quelles que soient leurs proportions relatives, depuis du cuivre pratiquement pur jusqu'à du zinc pratiquement pur.

Nous allons nous intéresser dans ce complément aux solutions solides de substitution. Le plus souvent, les atomes A et B sont distribués au hasard sur les nœuds du réseau cristallin commun, la probabilité pour qu'un site donné soit occupé par un atome A étant simplement égale à la proportion $N_A/(N_A + N_B)$ des atomes A dans l'alliage; on dit alors qu'ils constituent une solution solide *désordonnée*. Mais il arrive que, pour certains alliages, apparaisse à basse température une *« phase ordonnée »* dans laquelle les atomes sont répartis de façon plus régulière, chaque espèce (A ou B) se trouvant plutôt sur certains sites du réseau que sur les autres. Il existe alors une température critique T_c qui sépare deux domaines de température : l'alliage se trouve dans la phase ordonnée si $T < T_c$, dans la phase désordonnée si $T > T_c$.

Nous n'étudierons ici que le cas où les composants A et B interviennent dans une proportion simple, mais il peut exister une transition ordre-désordre pour une proportion N_A/N_B quelconque (voir exercice III.39).

I. LA TRANSITION ORDRE-DÉSORDRE

1. UN EXEMPLE : L'ALLIAGE CUIVRE-ZINC

Considérons, pour fixer les idées, l'alliage équiatomique ($N_A = N_B$) cuivre-zinc. Dans cet alliage, les atomes Cu et Zn occupent les sites d'un réseau cubique centré.

A température nulle, le cristal est parfaitement ordonné, c'est-à-dire que les atomes des deux espèces sont disposés de façon parfaitement régulière (fig. 1) : on peut distinguer deux sous-réseaux cubiques (α) et (β), imbriqués de telle sorte que les sommets de l'un coïncident avec les centres des mailles cubiques de l'autre; les atomes de cuivre occupent par exemple les sites α et les atomes de zinc les sites β; ainsi, un atome de cuivre est situé au centre d'un cube dont les sommets portent chacun un atome de zinc, et vice versa.

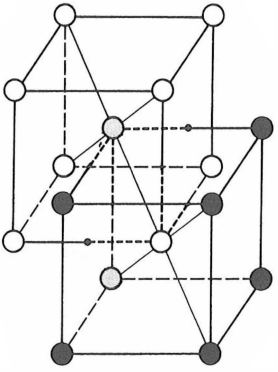


FIGURE 1
Alliage ordonné CuZn. A température nulle, les atomes de cuivre (en noir) sont tous situés aux nœuds d'un réseau cubique (sites α) d'arête a. Les atomes de zinc occupent le centre des cubes formés par les atomes de cuivre; ils constituent ainsi un deuxième réseau cubique (sites β), de même taille, mais décalé par rapport au premier.

A température T non nulle, cet ordre rigoureux s'atténue sous l'effet de l'agitation thermique : deux atomes d'espèce différente échangent parfois leurs places, de sorte que certains sites α sont occupés par un atome de zinc et certains sites β par un atome de cuivre. Mais, pour des températures pas trop élevées, la probabilité pour que ce soit un atome Cu qui se trouve sur un site α choisi au hasard (ou un atome Zn sur un site β) est supérieure à 1/2, de sorte qu'on a encore affaire à la phase ordonnée.

Lorsque T dépasse la température critique T_c (qui vaut 742 K dans l'exemple considéré), l'alliage devient désordonné : bien qu'en nombre toujours égal, les atomes Cu et Zn sont distribués de façon totalement aléatoire sur les sites α et β du réseau, c'est-à-dire que la présence, sur un site déterminé choisi au hasard, d'un atome Cu a même probabilité que celle d'un atome Zn.

La transition ordre-désordre s'observe directement en analysant la diffraction de rayons X ou de neutrons par un échantillon de l'alliage considéré[1]. On mesure

1. Voir par exemple C. Kittel, *Introduction à la Physique de l'État Solide,* Dunod, Paris (1972), chapitre 2 (pour l'étude générale de la diffraction) et D. Ter Haar, *Elements of Statistical Mechanics,* Holt Rinehart Winston, New York (1960), p. 251 (pour son application à la transition ordre-désordre).

également, au passage de la température critique, des variations très marquées de certaines grandeurs physiques telles que la capacité calorifique (fig. 2) ou la résistivité électrique.

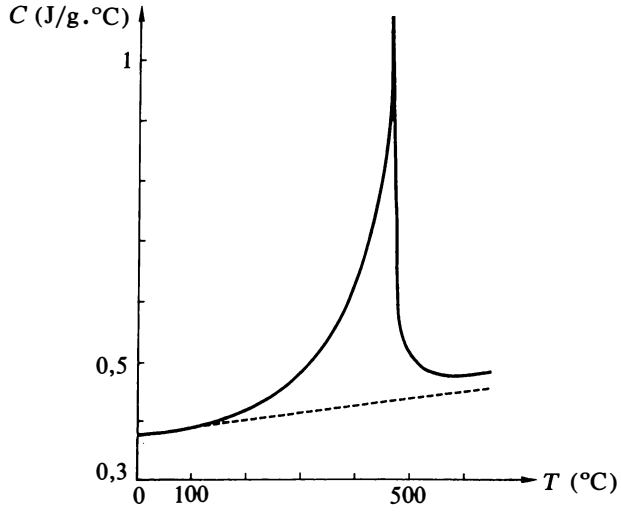


FIGURE 2
Variation de la chaleur spécifique avec la température pour un alliage CuZn. La ligne pointillée représente la chaleur spécifique de l'alliage évaluée à partir de celles du cuivre et du zinc purs.

2. DESCRIPTION DU MODÈLE. ÉNERGIE DE CONFIGURATION

Nous considérons un alliage binaire équiatomique AB comprenant $N/2$ atomes de type A et $N/2$ atomes de type B. Ces atomes peuvent se placer sur deux sous-réseaux équivalents, comme dans le cas de l'alliage CuZn. Chaque site α est entouré de z plus proches voisins ($z=8$ dans le cas de l'alliage CuZn) appartenant tous au réseau (β) et vice versa. La cohésion du cristal est assurée par les interactions entre atomes; celles-ci sont à courte portée, et nous ne tiendrons compte que des *interactions entre plus proches voisins.*

Nous ne nous intéressons qu'à l'agencement des atomes dans le cristal, c'est-à-dire à sa «configuration»; nous ne considérerons pas les effets liés aux vibrations des atomes (cf. complément III.E) ni à leur structure interne. L'énergie du cristal sera alors uniquement son *énergie de configuration*, déterminée par les nombres N_{AA}, N_{AB}, N_{BB} de liaisons A—A, A—B et B—B respectivement. Si ε_{AA}, ε_{AB}, ε_{BB} sont les énergies d'interaction entre les couples de plus proches voisins correspondants, l'énergie du cristal s'écrit dans ce modèle

$$E = N_{AA}\,\varepsilon_{AA} + N_{AB}\,\varepsilon_{AB} + N_{BB}\,\varepsilon_{BB}\ . \tag{K.1}$$

Les nombres N_{AA}, N_{AB}, N_{BB} ne sont *pas indépendants.* Le nombre total de liaisons entre plus proches voisins étant égal à $zN/2$, on a en premier lieu

$$N_{AA} + N_{AB} + N_{BB} = zN/2\ . \tag{K.2}$$

D'autre part, dans un alliage équiatomique, les $N/2$ atomes A participent à $zN/2$ liaisons de type A—A ou A—B, de sorte que

$$2N_{AA} + N_{AB} = zN/2\ . \tag{K.3.a}$$

Il en est de même pour les atomes B :

$$2N_{BB} + N_{AB} = zN/2 \ . \tag{K.3.b}$$

D'où les relations

$$N_{AA} = N_{BB} = zN/4 - N_{AB}/2 \ . \tag{K.4}$$

L'énergie (K.1) prend alors la forme

$$E = \frac{zN}{4}(\varepsilon_{AA} + \varepsilon_{BB}) + N_{AB}\left(\varepsilon_{AB} - \frac{\varepsilon_{AA} + \varepsilon_{BB}}{2}\right) . \tag{K.5}$$

Nous voulons qu'à $T = 0$ le cristal soit parfaitement ordonné, avec tous les atomes A sur des sites α et tous les atomes B sur des sites β. L'énergie doit donc être minimale dans cette situation, où $N_{AA} = N_{BB} = 0$ et $N_{AB} = zN/2$. Ceci suppose que[2]

$$v \mathrel{\widehat{=}} \frac{1}{2}(\varepsilon_{AA} + \varepsilon_{BB}) - \varepsilon_{AB} > 0 \ . \tag{K.6}$$

3. DÉFINITION DU PARAMÈTRE D'ORDRE

Appelons N_A^α le nombre d'atomes A placés sur des sites α et définissons le paramètre s $(-1 \leqslant s \leqslant 1)$ par la relation

$$N_A^\alpha \mathrel{\widehat{=}} \frac{N}{4}(1+s) \ . \tag{K.7.a}$$

Le nombre N_A^β d'atomes A situés sur des sites β est alors

$$N_A^\beta = \frac{N}{2} - N_A^\alpha = \frac{N}{4}(1-s) \ . \tag{K.7.b}$$

Les sites α qui ne sont pas occupés par des atomes A le sont par des atomes B :

$$N_B^\alpha = \frac{N}{2} - N_A^\alpha = \frac{N}{4}(1-s) \ , \tag{K.7.c}$$

et de même

$$N_B^\beta = \frac{N}{4}(1+s) \ . \tag{K.7.d}$$

On voit facilement que $s = +1$ correspond à l'ordre parfait dans lequel tous les atomes A sont sur des sites α. Le cas $s = -1$ correspond aussi à un ordre total dans lequel les réseaux (α) et (β) sont simplement échangés. Il suffira en fait d'appeler (α) le réseau sur lequel les atomes A sont majoritaires et de ne considérer que le cas $s \geqslant 0$. Si $s = 0$, les atomes A sont répartis également sur les sites α et β, de même que les atomes B. *La valeur de s caractérise donc l'ordre dans le cristal : on l'appelle «paramètre d'ordre».*

Comme les interactions dans le cristal sont à courte portée, l'ordre global du cristal ne peut qu'être la conséquence d'un *«ordre local» :* un atome donné a tendance à s'entourer d'atomes de type différent, quelle que soit par ailleurs sa position sur un site α ou β. Cet ordre local se manifeste dans le nombre de liaisons A—B. On le caractérise par le paramètre σ (compris entre -1 et $+1$) :

$$\frac{1+\sigma}{2} \mathrel{\widehat{=}} \frac{N_{AB}}{N_{AA} + N_{AB} + N_{BB}} = \frac{2N_{AB}}{zN} \ . \tag{K.8}$$

2. Si, au contraire, v est négatif, les atomes A d'une part, et B d'autre part, ont tendance à se regrouper et à former des agrégats de métal pur (voir exercice III.40).

Si $\sigma = 1$ l'ordre est parfait : il n'y a que des liaisons A—B. Si les atomes sont disposés de façon totalement aléatoire, $N_{AA} = N_{BB} = N_{AB}/2$ et $\sigma = 0$. Le cas $\sigma = -1$ correspond à l'absence totale de liaisons A—B; il y a encore un ordre parfait mais les atomes A et B sont complètement séparés (voir note 2).

Chacun à sa manière, σ et s caractérisent l'ordre dans l'alliage. Mais alors que s décrit un ordre à « grande distance », en référant aux positions des atomes A et B relativement aux deux sous-réseaux (α) et (β), σ décrit un ordre « local » en référant à l'environnement d'un atome indépendamment de son appartenance à (α) ou (β).

Remarquons que l'énergie d'un état microscopique de l'alliage ne dépend que de la valeur du paramètre σ dans cet état. En effet, (K.5) peut s'écrire

$$E = E_0 - \frac{zN}{4} \sigma v \,, \tag{K.9}$$

où nous avons posé

$$E_0 = \frac{zN}{8} (\varepsilon_{AA} + \varepsilon_{BB} + 2\varepsilon_{AB}) \,. \tag{K.10}$$

Il existe en général, pour une valeur donnée de σ, un grand nombre d'états microscopiques distincts, qui ne correspondent pas tous à la même valeur de s. Le dénombrement de ces états et la sommation nécessaire au calcul de la fonction de partition n'a été jusqu'à présent possible (pour un cristal à trois dimensions) qu'à l'aide d'approximations. Nous montrerons aux paragraphes II et III comment les plus simples d'entre elles permettent d'obtenir les valeurs d'équilibre des paramètres σ et s et d'en déduire les propriétés du système.

4. ORIGINE PHYSIQUE DE LA TRANSITION

a. Effet coopératif

Lorsque l'ordre est parfait, un atome donné est partout entouré d'atomes d'espèce différente; ceci correspond à la configuration d'énergie minimale, réalisée à température nulle. Lorsqu'à température non nulle un atome A s'établit sur un site β, il est toujours énergétiquement plus favorable qu'il soit entouré d'atomes B; ceux-ci doivent alors se trouver sur des sites α. Le déplacement d'un atome d'un site sur l'autre entraîne donc le déplacement en cascade d'autres atomes. Autrement dit, plus le désordre est grand, plus il a tendance à augmenter. Cet effet d'avalanche conduit à la disparition abrupte, à une certaine température T_c, de l'ordre global dans le cristal. T_c est appelée « température critique ».

Au-dessus de T_c, il y a autant d'atomes A sur des sites α que sur des sites β. Un ordre local subsiste cependant au-delà de T_c; il diminue lorsque T augmente mais ne s'annule qu'à température infinie.

b. Analogie avec le modèle d'Ising du ferromagnétisme

Le ferromagnétisme est un autre exemple de transition ordre-désordre. D'ailleurs le modèle que nous venons de décrire est en tout point semblable au modèle d'Ising pour des spins 1/2 (cf. complément III.J, § III). Dans les deux cas, il s'agit de comprendre le comportement d'un système de particules en interaction placées sur des sites cristallins; chaque site a deux états possibles et seules les forces entre plus proches voisins sont prises en compte.

Montrons que l'on peut transformer l'expression (K.9) de l'énergie pour lui donner une forme analogue à celle d'un système de N spins d'Ising en interaction. Affectons à chaque site i ($i = 1, 2 \ldots N$) un coefficient S_i qui prend la valeur $+1$ s'il s'agit d'un site α occupé par un atome A ou d'un site β occupé par un atome B, -1 s'il s'agit d'un site α occupé par un atome B ou d'un site β occupé par un atome A.

On calcule alors facilement la somme suivante, prise sur tous les couples de plus proches voisins :

$$\sum_{(i,\,j)_v} S_i S_j = -N_{AA} - N_{BB} + N_{AB}$$
$$= \frac{zN}{2}\,\sigma\,. \qquad \text{(K.11)}$$

On a donc

$$E = E_0 - \frac{v}{2} \sum_{(i,\,j)_v} S_i S_j\,, \qquad \text{(K.12)}$$

c'est-à-dire, à la constante E_0 près, la même expression[3] que pour l'hamiltonien d'Ising en l'absence de champ magnétique (formule (J.52) du complément III.J).

Les résultats que nous avons obtenus au complément III.J sont directement transposables; on connaît donc des solutions exactes du problème pour des alliages à une et deux dimensions. On peut en particulier affirmer qu'il n'existe pas de phase ordonnée dans un cristal à une dimension, car la solution exacte du modèle d'Ising ne présente, dans ce cas, aucune transition de phase. Inversement, les méthodes que nous allons développer ici peuvent s'appliquer au ferromagnétisme : l'approximation de Bragg-Williams (§ II) redonne les résultats du champ moyen (III.J, § II); l'approximation de Bethe-Peierls (§ III), un peu plus élaborée, se transpose elle aussi sans difficulté.

II. L'APPROXIMATION DE BRAGG-WILLIAMS

1. PRINCIPE DE L'APPROXIMATION

De nombreux travaux ont été consacrés à la transition ordre-désordre. La première théorie satisfaisante a été proposée en 1934 par W.L. Bragg et J.L. Williams. Elle est fondée sur l'approximation qui consiste à *négliger les corrélations entre sites :* on suppose que la probabilité qu'a un site donné d'être occupé par un atome A ou B est indépendante de l'environnement particulier de ce site et dépend seulement des propriétés moyennes du cristal dans son ensemble. *L'approximation de Bragg-Williams est donc une approximation de champ moyen.*

2. DÉTERMINATION DU PARAMÈTRE D'ORDRE

Le système étant à la température T, le paramètre d'ordre s est une variable interne libre de fluctuer. Pour déterminer sa valeur la plus probable s_m, nous allons d'abord calculer l'énergie libre partielle $F(T, s)$ dans le cas où le paramètre s est fixé; nous écrirons ensuite que s_m est la valeur de s qui rend F minimale à T donnée.

a. Calcul de l'énergie *E* en fonction de *s*

Montrons tout d'abord que, dans cette approximation, lorsque s est fixé, σ et donc l'énergie E le sont aussi.

---

3. Le paramètre d'ordre moyen $\bar{s}$ joue ici le même rôle que l'aimantation pour la transition ferro-paramagnétique. En effet, d'après (K.7),

$$s = \frac{1}{N}\,(N_A^\alpha + N_B^\beta - N_A^\beta - N_B^\alpha) = \frac{1}{N} \sum_{i=1}^{N} S_i\,.$$

La probabilité P_A^α pour qu'un site α soit occupé par un atome A vaut

$$P_A^\alpha = \frac{N_A^\alpha}{N/2} = \frac{1+s}{2} ; \tag{K.13.a}$$

de même

$$P_A^\beta = \frac{1-s}{2} , \tag{K.13.b}$$

$$P_B^\alpha = \frac{1-s}{2} , \tag{K.13.c}$$

et

$$P_B^\beta = \frac{1+s}{2} . \tag{K.13.d}$$

Dans le modèle de Bragg-Williams, les probabilités d'occupation de deux sites, même voisins, sont statistiquement indépendantes; on a donc[4]

$$N_{AA} = N_{BB} = \frac{zN}{8}(1-s^2) , \tag{K.14}$$

$$N_{AB} = \frac{zN}{4}(1+s^2) . \tag{K.15}$$

Ainsi, dans l'approximation de Bragg-Williams, σ et s sont reliés (cf. (K.8)) par

$$\sigma = s^2 ; \tag{K.16}$$

l'énergie a donc pour expression

$$E = E_0 - \frac{zN}{4} v s^2 . \tag{K.17}$$

b. Fonction de partition et énergie libre pour s fixé

La fonction de partition du système pour s fixé s'écrit

$$Z(T, s) = \sum_{(\mathscr{C}) \in \mathcal{L}(s)} \mathrm{e}^{-E(s)/kT} , \tag{K.18}$$

où $\mathcal{L}(s)$ est l'ensemble des états microscopiques pour lesquels le paramètre d'ordre prend la valeur s. Or, d'après (K.17), tous ces états ont même énergie; on a donc simplement

$$Z(T, s) = \Omega(s) \mathrm{e}^{-E(s)/kT} , \tag{K.19}$$

où $\Omega(s)$ est le nombre d'états microscopiques du système correspondant à la valeur s

---

4. Le nombre total de liaisons peut être obtenu en multipliant le nombre $N/2$ de sites α par le nombre z de sites β qui entourent chaque site α. Pour trouver N_{AA}, il suffit alors de multiplier ce nombre par la probabilité P_A^α qu'un site α soit occupé par un atome A et par celle P_A^β qu'un site β soit aussi occupé par un atome A :

$$N_{AA} = \frac{N}{2} z \times \frac{1}{2}(1+s) \times \frac{1}{2}(1-s) = \frac{zN}{8}(1-s^2) . \tag{1}$$

Pour obtenir N_{AB}, il faut ajouter la probabilité d'avoir A en α et B en β à celle d'avoir B en α et A en β :

$$N_{AB} = \frac{N}{2} z \left[\frac{1}{4}(1+s)^2 + \frac{1}{4}(1-s)^2\right] = \frac{N}{4} z(1+s^2) . \tag{2}$$

On peut aussi utiliser la relation (K.4).

choisie. $\Omega(s)$ est égal au nombre de façons de répartir N_A^α atomes sur $N/2$ sites α et N_A^β atomes sur $N/2$ sites β :

$$\Omega(s)=\frac{(N/2)!}{\left[\frac{N}{4}(1+s)\right]!\left[\frac{N}{4}(1-s)\right]!}\cdot\frac{(N/2)!}{\left[\frac{N}{4}(1-s)\right]!\left[\frac{N}{4}(1+s)\right]!}\,. \tag{K.20}$$

On en déduit l'énergie libre partielle correspondant à cette valeur de s :

$$F(T,\,s)=-kT\,\ln\,\Omega(s)+E(s)=E_0-\frac{zN}{4}\,vs^2$$
$$+NkT\left[\frac{1}{2}(1+s)\ln\frac{1}{2}(1+s)+\frac{1}{2}(1-s)\ln\frac{1}{2}(1-s)\right] \tag{K.21}$$

(on a utilisé comme toujours la formule de Stirling).

c. Équation donnant le paramètre d'ordre à la température *T*

La valeur la plus probable s_m de s à la température T s'obtient en minimisant F :

$$\frac{\partial F}{\partial s}(T,\,s_m)=-\frac{zN}{2}\,vs_m+NkT\left[\frac{1}{2}\ln\frac{1}{2}(1+s_m)-\frac{1}{2}\ln\frac{1}{2}(1-s_m)\right]=0\,. \tag{K.22}$$

Elle est donc solution de l'équation

$$s_m=\frac{kT}{zv}\ln\frac{1+s_m}{1-s_m}\,, \tag{K.23}$$

que l'on transforme facilement en

$$s_m=\text{th}\,\frac{zvs_m}{2kT}\,. \tag{K.24}$$

L'équation (K.24) peut être résolue en traçant les deux courbes

$$s=\text{th}\,x \tag{K.25.a}$$

et

$$s=\frac{2kT}{zv}\,x \tag{K.25.b}$$

sur un même graphique (fig. 3), et en cherchant leurs points d'intersection.

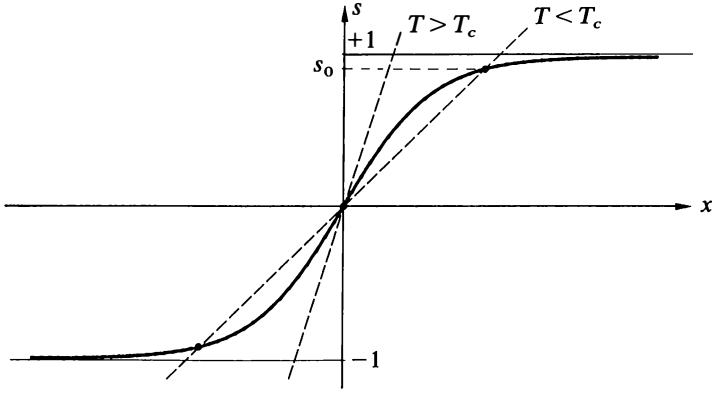


FIGURE 3

Résolution graphique de l'équation (K.24). Les points d'intersection de la courbe $s=\text{th}\,x$ (en trait plein) avec les droites $s=2kTx/zv=Tx/T_c^{BW}$ donnent les valeurs possibles s_m du paramètre d'ordre le plus probable.

3. PROPRIÉTÉS DU SYSTÈME

Comme on devait s'y attendre (cf. § I.4.b), l'équation (K.24) donnant le paramètre d'ordre est tout à fait semblable à celle qui donne l'aimantation dans le modèle de champ moyen pour le ferromagnétisme. Nous pouvons donc transposer ici les résultats obtenus au paragraphe II.2 du complément III.J.

a. Température critique

Les deux courbes (K.25) se coupent toujours à l'origine des coordonnées : $s_m = 0$ est toujours une solution de (K.24). Mais, pour les températures T inférieures à la température critique

$$T_c^{BW} \widehat{=} \frac{zv}{2k} , \tag{K.26}$$

il existe deux autres solutions $\pm s_0 \neq 0$. La solution $s_m = 0$ n'est stable[5] que pour les températures supérieures à T_c^{BW}. Pour $T < T_c^{BW}$, ce sont les solutions $\pm s_0$ qui correspondent à un minimum de F. Nous ne gardons ici que la solution $s_0 > 0$ puisque le changement de s en $-s$ correspond simplement à l'échange des sites α et β (voir paragraphe I.3).

A $T = 0$, $s_0 = +1$: l'alliage est totalement ordonné. Lorsque la température augmente, le paramètre d'ordre le plus probable diminue; il s'annule pour $T = T_c^{BW}$. Enfin, au-dessus de T_c^{BW}, s_m est identiquement nul. La température de transition T_c^{BW} sépare donc une région où l'alliage est ordonné ($s \neq 0$) d'un domaine de température où les atomes sont répartis aléatoirement sur l'ensemble des sites du cristal.

L'analogie avec le ferromagnétisme permet d'utiliser la formule (J.37) du complément III.J pour préciser le comportement du paramètre d'ordre lorsque T tend vers la température critique :

$$s_0 \simeq \sqrt{3(1 - T/T_c^{BW})} \quad \text{pour} \quad 0 < \frac{T_c^{BW} - T}{T_c^{BW}} \ll 1 . \tag{K.27}$$

L'allure de s_m en fonction de la température est représentée sur la figure 4.

b. Capacité calorifique

Maintenant que l'on a calculé la valeur d'équilibre de s à la température T, il est facile de déterminer la capacité calorifique

$$C \widehat{=} \frac{\partial \overline{E}}{\partial T} . \tag{K.28}$$

---

5. Développons F autour de $s = 0$:

$$F(s, T) = E_0 - NkT \ln 2 - \frac{1}{4} zNvs^2 + NkT \left[\frac{1}{2} s^2 + \frac{1}{12} s^4 + \ldots \right] \tag{1}$$

$$= E_0 - NkT \ln 2 + \frac{1}{2} N \left(kT - \frac{zv}{2} \right) s^2 + \frac{NkT}{12} s^4 + \ldots \tag{2}$$

Le coefficient de s^2 est positif pour

$$T > \frac{zv}{2k} \widehat{=} T_c^{BW} ; \tag{3}$$

l'énergie libre est alors minimum en $s = 0$. Elle est maximum en ce point dans le cas contraire $T < T_c^{BW}$.

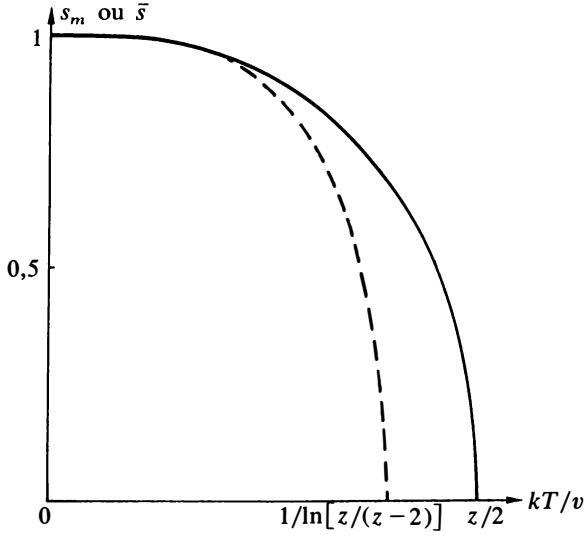


FIGURE 4
Variation du paramètre d'ordre avec la température T dans l'approximation de Bragg-Williams (trait plein) et dans celle de Bethe-Peierls (trait tireté).

L'énergie moyenne (cf. (K.17))

$$\bar{E} = E_0 - \frac{zN}{4} v\overline{s^2} \tag{K.29}$$

dépend de T à travers la valeur moyenne $\overline{s^2}$ du carré du paramètre d'ordre. Comme les fluctuations sont négligeables, $\overline{s^2}$ coïncide pratiquement avec le carré de s_m, solution de l'équation (K.24); on a donc

$$C = -\frac{zN}{4} v \frac{\partial s_m^2}{\partial T} . \tag{K.30}$$

Au-dessus de T_c^{BW}, s_m est identiquement nul et $C = 0$. A très basse température, on trouve également que C tend vers zéro. Au voisinage de la température critique (mais pour $T < T_c^{BW}$), la formule (K.27) donne

$$C \simeq \frac{3}{2} Nk \quad \text{pour} \quad T = T_{c^-}^{BW} . \tag{K.31}$$

La capacité calorifique C est donc discontinue en T_c (fig. 5) :

$$C(T_{c^-}^{BW}) - C(T_{c^+}^{BW}) = \frac{3}{2} Nk . \tag{K.32}$$

Le modèle de Bragg-Williams reproduit qualitativement la transition ordre-désordre telle que nous l'avons décrite au début de ce complément. Si l'on compare ses prédictions avec les résultats expérimentaux, on constate qu'il donne correctement l'allule générale des phénomènes, mais que l'accord n'est pas très bon sur le plan quantitatif. Ce n'est pas surprenant, car ce modèle assez rudimentaire souffre d'un défaut important : il ne prend en compte que les probabilités pour qu'un atome de type A ou B occupe un site α ou β et ignore donc toute corrélation entre les positions des atomes; or les configurations dans lesquelles, localement, un atome A est entouré

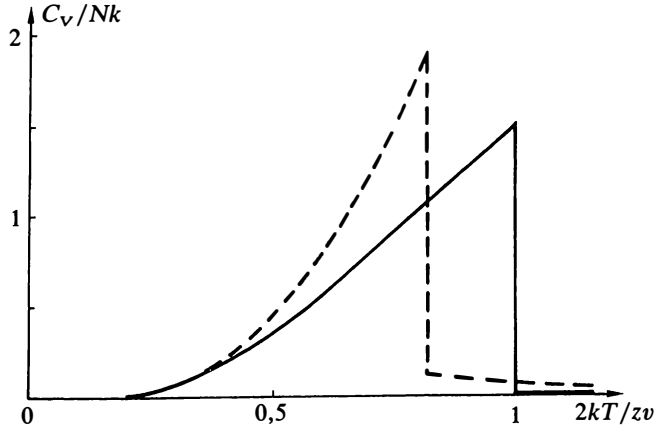


FIGURE 5
Chaleur spécifique d'un alliage dans les modèles de Bragg-Williams (trait plein) et de Bethe-Peierls (tirets).

préférentiellement d'atomes B et vice versa sont énergétiquement plus favorables, donc plus probables.

Dans le paragraphe suivant, nous allons décrire un modèle un peu plus élaboré, qui tient partiellement compte des corrélations à courte distance dans la répartition des atomes.

III. L'APPROXIMATION DE BETHE-PEIERLS

1. PRINCIPE DE L'APPROXIMATION

Jusqu'à présent, nous avons tout simplement négligé les corrélations locales entre atomes, puisque nous avons supposé que toutes les configurations correspondant à un paramètre d'ordre s donné étaient également probables. Dans le modèle de Bethe-Peierls, on isole dans le système un petit sous-ensemble composé d'un atome central et de ses z plus proches voisins (fig. 6). Les z atomes de la couronne interagissent d'une part avec l'atome central, mais aussi avec les atomes extérieurs. On traite ici exactement les interactions entre l'atome central et ceux de la couronne, mais on utilise une approximation de champ moyen pour les interactions des z atomes de la couronne avec le reste du cristal.

Commençons par choisir dans le cristal l'un des deux sous-réseaux imbriqués, par exemple le sous-réseau (α). Prenons ensuite un site particulier de ce sous-réseau. Supposons que ce soit un atome A qui occupe le site considéré, et qu'il soit entouré de n atomes A et donc de $(z-n)$ atomes B. L'énergie de ce petit amas de $z+1$ atomes peut s'écrire

$$\varepsilon_{A,\,nA} = n\varepsilon_{AA} + (z-n)\,\varepsilon_{AB} + nu_A + (z-n)\,u_B \;; \tag{K.33}$$

les deux premiers termes de (K.33) décrivent l'interaction de l'atome central avec ses z voisins; les termes nu_A et $(z-n)\,u_B$ prennent en compte les interactions des atomes A et B de la couronne avec le reste du cristal. Si, au contraire, c'est un atome B qui occupe le site central et qu'il est entouré de n atomes A et de $(z-n)$ atomes B, on écrit l'énergie du petit amas correspondant

$$\varepsilon_{B,\,nA} = n\varepsilon_{AB} + (z-n)\,\varepsilon_{BB} + nu_A + (z-n)\,u_B \;. \tag{K.34}$$

Dans les expressions (K.33) et (K.34), les interactions des z atomes de la couronne avec l'extérieur de l'amas sont traitées de façon moyenne : l'énergie correspondante est simplement prise égale à u_A pour chaque atome A et à u_B pour chaque atome B, quel que soit l'environnement immédiat du sous-système microscopique que l'on a individualisé. L'approximation de Bethe-Peierls est donc encore une méthode de champ moyen, mais appliquée à des groupes d'atomes et non plus à des atomes individuels.

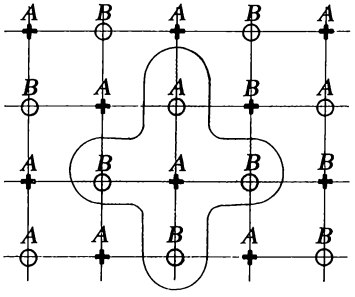


FIGURE 6
Modèle de Bethe-Peierls pour un cristal à deux dimensions. Chaque atome a $z = 4$ plus proches voisins. Les sites α sont représentés par des croix, les sites β par des points. L'amas d'atomes considéré comporte ici un atome A central entouré d'une couronne constituée de $n = 1$ atome A et $z - n = 3$ atomes B.

2. ÉQUATION D'AUTOCOHÉRENCE

Les énergies u_A et u_B sont pour l'instant des inconnues. Elles ne sont cependant pas quelconques : elles doivent vérifier une « condition d'autocohérence » qui les relie aux énergies ε_{AA}, ε_{BB} et ε_{AB} des liaisons entre proches voisins. En réalité, nous allons le voir, c'est seulement la différence

$$u \widehat{=} u_A - u_B \tag{K.35}$$

qui figure dans l'équation d'autocohérence.

Caleulons tout d'abord la probabilité $P(A, nA)$ pour qu'un site α donné soit occupé par un atome A et qu'il soit entouré de n atomes A. Puisque le système est à la température T, on a

$$P(A, nA) = \frac{1}{Q} C_z^n \, e^{-\beta \varepsilon_{A,nA}} , \tag{K.36.a}$$

où C_z^n est le nombre de façons de distribuer les n atomes A sur les z sites de la couronne et Q un facteur de normalisation. On obtient de même la probabilité $P(B, nA)$ pour que le site central soit occupé par un atome B et entouré de n atomes A :

$$P(B, nA) = \frac{1}{Q} C_z^n \, e^{-\beta \varepsilon_{B,nA}} . \tag{K.36.b}$$

Posons

$$y \widehat{=} e^{-\beta u} \tag{K.37}$$

et

$$x \widehat{=} e^{-\beta v} , \tag{K.38}$$

et supposons, pour ne pas alourdir inutilement les calculs,

$$\varepsilon_{AA} = \varepsilon_{BB} , \tag{K.39}$$

ce qui entraîne (cf. (K.6))

$$v = \varepsilon_{AA} - \varepsilon_{AB} = \varepsilon_{BB} - \varepsilon_{AB} . \tag{K.40}$$

Il vient alors

$$P(A, nA)=\frac{1}{q}\, C_z^n x^n y^n \tag{K.41.a}$$

et $$P(B, nA)=\frac{1}{q}\, C_z^n x^{z-n} y^n\,, \tag{K.41.b}$$

avec $$q \widehat{=} Q\, e^{\beta z(\varepsilon_{AB}+u_B)}\,. \tag{K.42}$$

La probabilité $\mathcal{P}(A)$ de trouver un atome A sur le site central α quel que soit son environnement s'en déduit aussitôt :

$$\mathcal{P}(A)=\sum_{n=0}^{z} P(A, nA)=\frac{1}{q}(1+xy)^z\,. \tag{K.43}$$

De même, la probabilité $\mathcal{P}(B)$ pour que ce soit un atome B qui occupe le site central s'écrit

$$\mathcal{P}(B)=\sum_{n=0}^{z} P(B, nA)=\frac{1}{q}(x+y)^z\,. \tag{K.44}$$

La somme de ces deux probabilités complémentaires vaut évidemment 1, ce qui permet de déterminer le facteur de normalisation q :

$$q=(1+xy)^z+(x+y)^z\,. \tag{K.45}$$

Dans l'alliage AB, il y a autant d'atomes A situés sur des sites β que d'atomes B sur des sites α. Aucun site ne jouant un rôle particulier, la probabilité $\mathcal{P}(B)$ de trouver un atome B sur un site central α doit être égale à la probabilité $\mathcal{P}'(A)$ de trouver un atome A en un site β *donné* de la couronne entourant le site α :

$$\mathcal{P}(B)=\mathcal{P}'(A)\,. \tag{K.46}$$

Or $$\mathcal{P}'(A)=\sum_{n=0}^{z}\frac{n}{z}\,P(A, nA)+\sum_{n=0}^{z}\frac{n}{z}\,P(B, nA)\,. \tag{K.47}$$

Si l'on utilise les expressions (K.41) et (K.44), il vient donc

$$\frac{1}{q}(x+y)^z=\frac{1}{q}\sum_{n=0}^{z}\frac{n}{z}\,C_z^n x^n y^n+\frac{1}{q}\sum_{n=0}^{z}\frac{n}{z}\,C_z^n x^{z-n} y^n\,. \tag{K.48}$$

La relation ci-dessus est une « équation d'autocohérence » qui définit implicitement y (c'est-à-dire en fait u) comme une fonction de v et de T. On met facilement[6] cette équation sous la forme

$$y=\left(\frac{x+y}{1+xy}\right)^{z-1} \tag{K.49}$$

---

6. En effet,

$$\sum_{n=0}^{z}\frac{n}{z}\,C_z^n (xy)^n=\frac{xy}{z}\,\frac{\partial}{\partial(xy)}\sum_{n=0}^{z} C_z^n (xy)^n=\frac{xy}{z}\,\frac{\partial}{\partial(xy)}(1+xy)^z=xy(1+xy)^{z-1}\,. \tag{1}$$

De même

$$\sum_{n=0}^{z}\frac{n}{z}\,C_z^n x^{z-n} y^n=y(x+y)^{z-1}\,. \tag{2}$$

L'équation (K.48) s'écrit alors

$$(x+y)^z=xy(1+xy)^{z-1}+y(x+y)^{z-1}\,, \tag{3}$$

d'où (K.49).

3. CALCUL DU PARAMÈTRE D'ORDRE ET DE L'ÉNERGIE

a. Le paramètre d'ordre moyen $\bar{s}$

Reportons-nous à la définition (K.7.a) de s :

$$\frac{N}{4}(1+s)\widehat{=}N_A^\alpha . \tag{K.50}$$

Nous savons calculer le nombre moyen d'atomes A sur les sites α :

$$\overline{N_A^\alpha}=\frac{N}{2}\mathcal{P}(A) . \tag{K.51}$$

Donc, d'après (K.43),

$$1+\bar{s}=\frac{2}{q}(1+xy)^z , \tag{K.52}$$

ce qui donne, si l'on remplace q par son expression (K.45)[7],

$$\bar{s}=\frac{1-y^{z/(z-1)}}{1+y^{z/(z-1)}} . \tag{K.53}$$

b. Le paramètre $\bar{\sigma}$ et l'énergie moyenne

Nous avons introduit, pour chaque site α, la probabilité $P(A, nA)$ qu'il soit occupé par un atome A entouré de n atomes A. Cette configuration comportant n liaisons A—A, le nombre moyen de ces liaisons dans le cristal vaut

$$\overline{N_{AA}}=\frac{N}{2}\sum_{n=0}^{z} nP(A, nA) , \tag{K.54}$$

c'est-à-dire (cf. note 6)

$$\overline{N_{AA}}=\frac{z}{q}\frac{N}{2}xy(1+xy)^{z-1} . \tag{K.55}$$

D'autre part (cf. (K.4) et (K.8)),

$$\overline{N_{AA}}=\frac{zN}{4}-\frac{1}{2}\overline{N_{AB}}=\frac{zN}{8}(1-\bar{\sigma}) . \tag{K.56}$$

On en tire aussitôt, en utilisant (K.45) et (K.49), le paramètre moyen $\bar{\sigma}$:

$$\bar{\sigma}=1-\frac{4xy}{1+xy}\frac{1}{1+y^{z/(z-1)}} . \tag{K.57}$$

Il suffit ensuite de reporter cette formule dans (K.9) pour obtenir l'énergie moyenne de l'alliage.

---

7. Écrivons (K.52) sous la forme

$$\frac{2}{1+\bar{s}}=\frac{q}{(1+xy)^z}=1+\left(\frac{x+y}{1+xy}\right)^z . \tag{1}$$

On reconnaît au second membre l'expression (K.49) de y :

$$\frac{2}{1+\bar{s}}=1+y^{z/(z-1)} . \tag{2}$$

Il ne reste plus qu'à inverser cette relation pour obtenir $\bar{s}$.

4. PROPRIÉTÉS DU SYSTÈME

a. Température critique

Pour expliciter les valeurs de $\bar{s}$ et $\bar{\sigma}$, il faut d'abord résoudre l'équation (K.49) qui donne y à la température T, puis reporter la solution dans (K.53) et (K.57). On peut tout de suite remarquer que

(i) $y=1$ est toujours solution et correspond à $\bar{s}=0$;

(ii) si y est solution, $1/y$ l'est aussi; mais changer y en $1/y$ change seulement $\bar{s}$ en $-\bar{s}$;

(iii) les valeurs $\bar{s}=\pm 1$ et $\bar{\sigma}=1$ seront obtenues si $y=0$ ou $y=\infty$ est solution.

Résolvons graphiquement l'équation d'autocohérence (K.49). Suivant la valeur du paramètre x (voir figure 7), les deux courbes

$$\xi = y \tag{K.58.a}$$

et

$$\xi = \left(\frac{x+y}{1+xy}\right)^{z-1} \tag{K.58.b}$$

possèdent un ou trois points d'intersection. Si la pente γ de la tangente à la courbe (K.58.b) au point $y=1$ est inférieure à 1, la seule solution est $y=1$. Au contraire, si γ est supérieure à 1, il y a trois solutions :

$$y=1, \quad y=y_0 \quad \text{et} \quad y=1/y_0.$$

On trouve facilement, en dérivant (K.58.b),

$$\gamma = (z-1)\frac{1-x}{1+x}. \tag{K.59}$$

La valeur de x qui donne à γ la valeur 1, et qui sépare donc les domaines où (K.49) possède une ou trois solutions, est égale à

$$x_c = \frac{z-2}{z}. \tag{K.60}$$

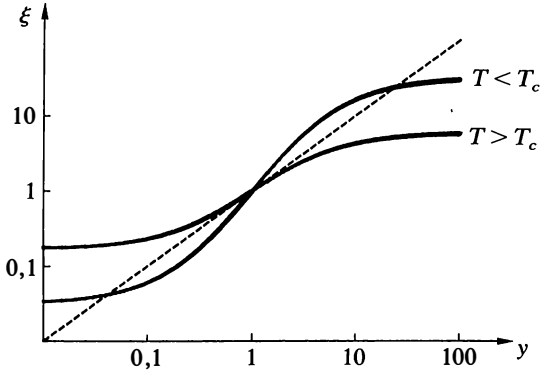


FIGURE 7

Résolution graphique de l'équation d'autocohérence pour deux valeurs de x. On cherche les points d'intersection de la courbe $\xi = \left(\frac{x+y}{1+xy}\right)^{z-1}$ avec la droite $\xi = y$ (noter la double échelle logarithmique). La solution $y=1$ existe toujours, alors que les solutions y_0 et $1/y_0$ n'existent que si la pente γ de la courbe $\xi(y)$ au point $y=1$ est supérieure à 1, c'est-à-dire que si $x<x_c$ (ou $T<T_c$).

On définit la température critique T_c^{BP} par

$$\exp[-v/kT_c^{BP}] \widehat{=} x_c \,, \tag{K.61}$$

ce qui donne

$$kT_c^{BP} = \frac{v}{\ln[z/(z-2)]} \,. \tag{K.62}$$

On retrouve donc, comme dans le modèle de Bragg-Williams, l'existence d'une température critique au-dessus de laquelle le paramètre d'ordre moyen $\bar{s}$ est identiquement nul (puisque $y=1$ est alors la seule solution de l'équation d'autocohérence).

b. Comportement du système pour $T > T_c$

Le paramètre d'ordre $\bar{s}$ étant nul, la phase de haute température est désordonnée à l'échelle de l'ensemble du cristal. En revanche, le paramètre $\bar{\sigma}$ qui caractérise l'ordre local ne s'annule pas : il vaut, d'après (K.57),

$$\bar{\sigma} = 1 - \frac{2x}{1+x} \,, \tag{K.63}$$

c'est-à-dire

$$\bar{\sigma} = \frac{1-x}{1+x} = \text{th}\,\frac{\beta v}{2} \quad \text{pour} \quad T > T_c^{BP} \,. \tag{K.64}$$

$\bar{\sigma}$ tend donc vers zéro à haute température; mais, contrairement à $\bar{s}$, il ne s'annule pas dès que la température dépasse T_c : *un ordre à courte distance subsiste donc aux températures supérieures à* T_c.

c. Comportement du système pour $T < T_c$

Pour $T < T_c^{BP}$, l'équation (K.49) possède trois solutions. Seules nous intéressent celles qui correspondent à un minimum de l'énergie libre. On peut montrer que, comme dans l'approximation de Bragg-Williams, la solution $y=1$ (c'est-à-dire $\bar{s}=0$) ne correspond pas à un état d'équilibre stable du système. Il subsiste alors deux solutions acceptables, y_0 et $1/y_0$. Nous ne conservons que la solution

$$y_0 < 1 \,, \tag{K.65}$$

à laquelle correspond un paramètre d'ordre (cf. (K.53))

$$\bar{s} > 0 \,. \tag{K.66}$$

La variation de $\bar{s}$ avec la température est représentée sur la figure 4.

On vérifie facilement[8] que, lorsque $T \to 0$, $y_0 \to 0$; $\bar{s}$ et $\bar{\sigma}$ tendent alors vers 1 : l'alliage est parfaitement ordonné.

d. Capacité calorifique

Parmi toutes les grandeurs thermodynamiques, considérons plus particulièrement la capacité calorifique C. On la calcule à partir de la valeur moyenne de l'énergie (K.9) :

$$C \widehat{=} \frac{\partial \bar{E}}{\partial T} = -\frac{zN}{4} v \frac{\mathrm{d}\bar{\sigma}}{\mathrm{d}T} \,. \tag{K.67}$$

---

8. Lorsque $T \to 0$, $x \widehat{=} \mathrm{e}^{-\beta v} \to 0$. On voit sur la figure 7 que l'intervalle de variation de la courbe

$$\xi(y) = \left(\frac{x+y}{1+xy}\right)^{z-1}$$

s'étend alors de 0 à l'∞; y_0 tend donc vers 0.

α. Capacité calorifique de la phase désordonnée

Pour $T > T_c^{BP}$, $\bar{\sigma}$ est donné par (K.64); il vient immédiatement

$$\frac{d\bar{\sigma}}{dT} = -\frac{1}{kT^2}\frac{d\bar{\sigma}}{d\beta} = -\frac{v}{2kT^2}\frac{1}{\mathrm{ch}^2(v/2kT)}, \tag{K.68}$$

et donc

$$C^{BP} = \frac{zN}{8}\frac{v^2}{kT^2}\frac{1}{\mathrm{ch}^2(v/2kT)} \quad \text{pour} \quad T > T_c^{BP}. \tag{K.69}$$

Pour $T = T_c^{BP}$, la capacité calorifique vaut (cf. (K.60) et (K.62))

$$C^{BP}(T_{c^+}^{BP}) = \frac{Nk}{8}\left[\ln\frac{z}{z-2}\right]^2 \frac{z^2(z-2)}{(z-1)^2}. \tag{K.70}$$

Elle décroît lorsque T croît, pour tendre vers zéro à haute température (fig. 5).

β. Capacité calorifique de la phase ordonnée

Calculer le paramètre $\bar{\sigma}$ est ici beaucoup plus difficile, car on ne sait pas résoudre analytiquement l'équation (K.49). Mais un calcul approché est possible pour des températures voisines de la température critique. La valeur y_0 de y est alors proche de 1 (cf. figure 7). En développant $\bar{\sigma}$, donné par (K.57), autour de $y = 1$, on obtient pour la chaleur spécifique l'expression approchée suivante[9] :

$$\begin{aligned} C^{BP} &= -\frac{zN}{4} v \frac{d\bar{\sigma}}{dT} \\ &\simeq \frac{zN}{4} v^2 \left(\frac{z}{z-1}\right)^2 \frac{3z/2-1}{kT^2} e^{-v/kT}, \qquad 0 < \frac{T_c^{BP} - T}{T_c^{BP}} \ll 1. \end{aligned} \tag{K.71}$$

---

9. Lorsque T s'approche de T_c^{BP} par valeurs inférieures, on peut écrire

$$y = 1 + \delta, \quad \text{avec} \quad \delta \ll 1. \tag{1}$$

Développons alors (K.49), prise sous la forme

$$x = \frac{y^{1/(z-1)} - y}{1 - y^{z/(z-1)}}, \tag{2}$$

jusqu'au 3e ordre en δ au numérateur et au dénominateur; nous obtenons (après avoir éliminé la solution évidente $\delta = 0$)

$$x = \frac{z-2}{z}\left[1 - \frac{\delta^2}{6(z-1)} + \ldots\right]. \tag{3}$$

Remplaçons maintenant x et y par leurs expressions (1) et (3) dans (K.57) : le développement au deuxième ordre en δ donne

$$\bar{\sigma} = 1 - \frac{z-2}{z-1}\left[1 - \frac{\delta^2}{4(z-1)^2}\left(z^2 - \frac{2}{3}z\right) + \ldots\right]. \tag{4}$$

En éliminant δ entre (3) et (4), on aboutit à une expression approchée de $\bar{\sigma}$ en fonction de x, valable au voisinage de T_c^{BP} :

$$\bar{\sigma} \simeq 1 - \frac{z-2}{z-1}\left[1 - \frac{(3z/2-1)z}{z-1}\left(1 - \frac{zx}{z-2}\right)\right]. \tag{5}$$

On en déduit

$$\frac{d\bar{\sigma}}{dx} \simeq -\frac{z^2(3z/2-1)}{(z-1)^2}. \tag{6}$$

.../...

Compte tenu de la définition (K.62) de T_c^{BP}, l'expression précédente donne, lorsque T tend vers T_c^{BP} par valeurs inférieures,

$$C^{BP}(T_{c^-}^{BP}) = \frac{Nk}{4}(z-2)\left(\frac{3}{2}z-1\right)\left(\frac{z}{z-1}\right)^2\left[\ln\frac{z}{z-2}\right]^2 . \qquad \text{(K.72)}$$

La capacité calorifique est donc *discontinue* en $T = T_c^{BP}$. Le rapport $C^{BP}(T_{c^-})/C^{BP}(T_{c^+})$ s'exprime de façon simple :

$$\frac{C^{BP}(T_{c^-})}{C^{BP}(T_{c^+})} = 3z-2 . \qquad \text{(K.73)}$$

La courbe est représentée sur la figure 5.

5. CONCLUSION

Les résultats que nous venons d'obtenir dans le cadre du modèle de Bethe-Peierls représentent une amélioration sensible par rapport à ceux du modèle de Bragg-Williams. La principale différence provient du fait qu'un ordre local subsiste ici au-delà de la température critique T_c, de sorte que la capacité calorifique ne s'annule pas brutalement en $T = T_c$. Cette prédiction est effectivement conforme aux résultats expérimentaux.

La valeur de la température critique est également modifiée par rapport au modèle de Bragg-Williams : les formules (K.26) et (K.62) dépendent toutes deux de z, mais sont différentes. On note qu'elles donnent des valeurs voisines si z est grand devant 1.

Cependant, la comparaison des résultats résumés sur la figure 5 avec les données expérimentales de la figure 2 montre que le modèle de Bethe-Peierls ne reproduit lui-même qu'assez grossièrement le comportement des systèmes réels au voisinage de la température critique. On peut améliorer encore la description en traitant exactement les interactions à l'intérieur d'amas comportant un atome central, ses plus proches voisins et ses seconds voisins[10]; l'amélioration reste cependant limitée.

On peut aussi s'appuyer sur le fait que l'on connaît la solution exacte du modèle d'Ising *à deux dimensions* (complément III.J, § III.3) pour comparer à ce résultat exact les formules approchées que donnent les deux modèles que nous avons décrits (fig. 8). On note que la température critique T_c obtenue dans l'approximation de Bethe-Peierls, plus basse que celle du modèle de Bragg-Williams, est plus proche de la valeur correcte[11]. Mais le comportement au voisinage immédiat de T_c est toujours mal reproduit dans ces modèles de champ moyen (cf. complément III.J, § III.3.c).

De façon plus générale, comme la transition ferro-paramagnétique et la transition ordre-désordre peuvent être analysées à l'aide de formalismes analogues (§ I.4.b), les indications que nous avons données au complément III.J (§ IV) sur les difficultés de cette analyse et les méthodes utilisées concernent également la théorie des alliages.

On a donc finalement, en utilisant (K.38),

$$\frac{d\bar{\sigma}}{dT} = \frac{d\bar{\sigma}}{dx}\cdot\frac{dx}{dT} = -\frac{v}{kT^2}\,e^{-v/kT}\,\frac{z^2(3z/2-1)}{(z-1)^2} \quad \text{pour} \quad T = T_{c^-}^{BP} . \qquad (7)$$

10. H.A. Bethe et H.H. Wills, *Proc. Roy. Soc. 150A*, 552 (1935).

11. On peut aussi remarquer que pour $z = 2$, c'est-à-dire dans un modèle à une dimension, l'approximation de Bethe-Peierls aboutit à une température critique nulle (cf. (K.62)). C'est effectivement ce que montre la solution exacte du modèle d'Ising à une dimension.

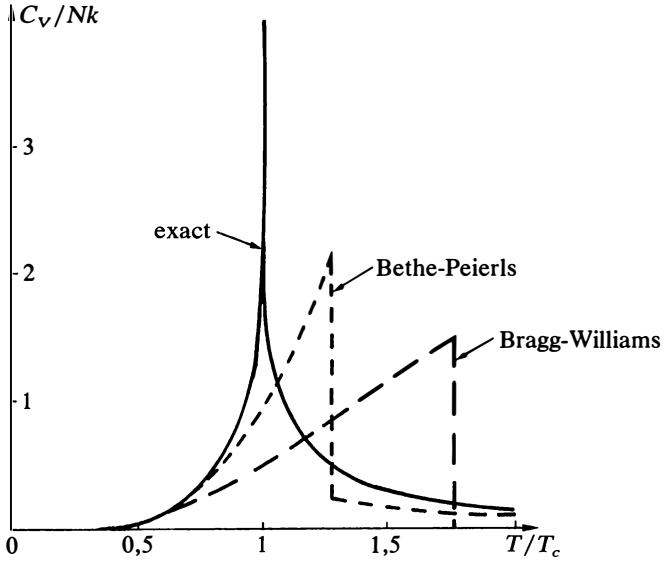


FIGURE 8
Capacité calorifique du réseau carré à 2 dimensions ($z = 4$) représenté sur la figure 6. On compare ici le résultat exact tiré du modèle d'Ising (trait plein) avec les valeurs que donnent l'approximation de Bragg-Williams et celle de Bethe-Peierls.

COMPLÉMENT III.L

Chaleur et travail. Le « second principe » de la thermodynamique

I. Variation d'entropie dans une transformation avec échange de chaleur
   1. Variation d'entropie d'un thermostat fournissant de la chaleur
   2. Variation d'entropie d'un système échangeant de la chaleur

II. Le « second principe » de la thermodynamique

III. Interprétation statistique de la chaleur et du travail
   1. Les deux expressions de l'accroissement d'énergie
   2. Interprétation statistique de la chaleur
   3. Chaleur et travail échangés dans une transformation réversible

Nous allons achever ici la démonstration des « principes » de la thermodynamique, entreprise au complément II.D, et donner l'interprétation statistique des notions de chaleur et de travail.

Nous limitons notre étude aux *systèmes thermodynamiques :* les fluctuations sont négligeables et les descriptions microcanonique et canonique équivalentes (voir chapitre III, § III.C)[1].

I. VARIATION D'ENTROPIE DANS UNE TRANSFORMATION AVEC ÉCHANGE DE CHALEUR

1. VARIATION D'ENTROPIE D'UN THERMOSTAT FOURNISSANT DE LA CHALEUR

Supposons que le système $\mathcal{S}$ que nous voulons étudier soit en contact avec un thermostat $\mathcal{T}$ de température T^{ex}, l'ensemble $\mathcal{S} \cup \mathcal{T}$ étant isolé. Évaluons pour commencer la variation d'entropie *du thermostat* lorsqu'il fournit au système $\mathcal{S}$ la quantité de chaleur Q (nous continuons à compter positivement les quantités de chaleur reçues par le système $\mathcal{S}$).

La réponse découle directement de la définition du thermostat. En effet, lorsque l'énergie du système $\mathcal{S}$ est E, celle du thermostat est $E_{tot} - E$, où E_{tot} est l'énergie (fixée) du système global $\mathcal{S} \cup \mathcal{T}$. Or E est par hypothèse très petite devant E_{tot}, de sorte qu'on peut développer $S^*_{\mathcal{T}}(E_{tot} - E)$, entropie microcanonique du thermostat, près de $E = 0$; par définition du thermostat (III, § I.A.1), ce développement s'arrête au premier ordre, puisque la température T^{ex} est invariable :

$$S^*_{\mathcal{T}}(E_{tot} - E) = S^*_{\mathcal{T}}(E_{tot}) - \frac{E}{T^{ex}} . \qquad (L.1)$$

Si maintenant E s'accroît de Q, il en résulte une variation de $S^*_{\mathcal{T}}$ égale à

$$\Delta S^*_{\mathcal{T}} = -\frac{Q}{T^{ex}} . \qquad (L.2)$$

1. Ceci permet en particulier de généraliser la définition des transformations thermodynamiques donnée au complément II.D (§ II.1) : les états initial et final d'une transformation sont nécessairement des états d'équilibre macroscopique du système étudié; mais celui-ci peut tout aussi bien être considéré comme en contact avec un thermostat (équilibre à température fixée) qu'isolé (équilibre à énergie fixée). De même, les diverses grandeurs (énergie, entropie, ...) sont ici des *grandeurs thermodynamiques*, et nous ne distinguons plus désormais entre leurs déterminations microcanonique et canonique.

Ce résultat reste valable *même si le système $\mathcal{S}$ n'est pas en équilibre avec le thermostat* (cf. III, § II.D.1) : ce dernier est par définition constamment à l'équilibre microcanonique, sa température étant invariable, quel que soit l'état du système $\mathcal{S}$.

Si l'on compare l'égalité (L.2) à la formule (D.59) du complément II.D qui donne la variation d'entropie d'un système dans une transformation réversible (en tenant compte du fait que la chaleur reçue par le thermostat est $-Q$), on conclut que *les échanges de chaleur auxquels participe un thermostat lui apparaissent toujours comme s'ils étaient réversibles*[2],[3].

2. VARIATION D'ENTROPIE D'UN SYSTÈME ÉCHANGEANT DE LA CHALEUR

Considérons maintenant un système $\mathcal{S}$ subissant une transformation A au cours de laquelle il reçoit de l'extérieur le travail W et la chaleur Q, celle-ci lui étant fournie par un thermostat $\mathcal{T}$ de température T^{ex} (nous supposons pour simplifier que cette transformation s'effectue à *nombre de particules constant*).

Définissons le système $\mathcal{U}$ comme la *réunion de $\mathcal{S}$ et $\mathcal{T}$*. Ce système $\mathcal{U}$ n'est en général pas isolé : il reçoit le travail W. Il est cependant isolé thermiquement (puisque l'échange de chaleur entre $\mathcal{S}$ et $\mathcal{T}$ fait maintenant partie des processus internes à $\mathcal{U}$) : *pour $\mathcal{U}$, la transformation A est adiabatique.* On peut donc lui appliquer le résultat fondamental démontré dans le complément II.D (formule (D.63)) : *son entropie ne peut que croître* au cours de la transformation A :

$$\Delta S_{\mathcal{U}} \geqslant 0 , \tag{L.3}$$

l'égalité ayant lieu si et seulement si la transformation A est réversible[4]. L'entropie du système $\mathcal{U}$ est la somme de l'entropie S du système $\mathcal{S}$ et de celle du thermostat $\mathcal{T}$. En utilisant pour celui-ci le résultat (L.2), on met la relation (L.3) sous la forme

$$\Delta S \geqslant \frac{Q}{T^{ex}} , \tag{L.4}$$

l'*égalité* se produisant pour une *transformation réversible,* et seulement dans ce cas.

---

2. Ils ne sont *pas réversibles* si le système $\mathcal{S}$ et le thermostat $\mathcal{T}$ ne sont pas en équilibre thermique. La formule (L.2) reste pourtant valable dans ce cas : *du point de vue du thermostat,* tout se passe comme si l'échange de chaleur était réversible. Par contre, la formule (D.59) du complément II.D n'est applicable au système $\mathcal{S}$ que si la transformation est réversible, c'est-à-dire si $\mathcal{S}$ est à l'équilibre, sa température étant égale à celle du thermostat.

3. Le résultat (L.2) se généralise aisément au cas où le système $\mathcal{S}$ échange des particules, de la chaleur et même du volume avec un *réservoir* $\mathcal{R}$, beaucoup plus grand que lui, dont la température T^{ex}, le potentiel chimique μ^{ex} et la pression p^{ex} sont invariables, le système global $\mathcal{S} \cup \mathcal{R}$ étant isolé (cf. chapitre V et complément III.D). La formule (L.1) est alors remplacée par

$$S^*_{\mathcal{R}}(E_{tot} - E,\ N_{tot} - N,\ V_{tot} - V) = S^*_{\mathcal{R}}(E_{tot},\ N_{tot},\ V_{tot}) - \frac{E}{T^{ex}} + \frac{\mu^{ex}}{T^{ex}} N - \frac{p^{ex}}{T^{ex}} V . \tag{1}$$

Mais, si le volume V du système s'accroît de ΔV, son nombre N de particules de ΔN et son énergie E de ΔE, on a (formule (D.51) du complément II.D)

$$\Delta E = Q - p^{ex}\Delta V + \mu^{ex}\Delta N , \tag{2}$$

de sorte que l'accroissement d'entropie du réservoir est encore

$$\Delta S^*_{\mathcal{R}} = -\frac{Q}{T^{ex}} . \tag{3}$$

4. C'est à vrai dire la transformation (adiabatique) *du grand système* $\mathcal{U}$ qui doit être réversible. Mais elle ne peut l'être que si $\mathcal{U}$ est à chaque instant à l'équilibre (condition pour que la transformation soit quasistatique), ce qui suppose que les échanges de chaleur entre $\mathcal{S}$ et le thermostat $\mathcal{T}$ sont réversibles (égalité des températures). C'est donc vraiment la réversibilité de la transformation A initiale qui assure l'égalité dans (L.3).

Le raisonnement se généralise aisément au cas où le système $\mathcal{S}$ échange de la chaleur avec plusieurs thermostats : le thermostat $\mathcal{C}_\alpha$, de température T_α^{ex}, fournit au système $\mathcal{S}$ la quantité de chaleur Q_α, avec $\alpha = 1, 2, \ldots, q$. Le système $\mathcal{U}$ réunit dans ce cas $\mathcal{S}$ et l'ensemble des q thermostats $\mathcal{C}_\alpha$, et la relation (L.3) donne

$$\Delta S \geqslant \sum_{\alpha=1}^{q} \frac{Q_\alpha}{T_\alpha^{ex}} . \tag{L.5}$$

Il est facile enfin de passer à la limite où la température extérieure T^{ex} varie de façon continue au cours de la transformation $1 \longrightarrow 2$ considérée[5]. Le résultat (L.5) s'écrit alors

$$\boxed{\Delta S \geqslant \int_{(1)}^{(2)} \frac{đQ}{T^{ex}}} \; ; \tag{L.6}$$

dans le cas d'une transformation réversible, l'égalité des deux membres reproduit la formule (D.60) du complément II.D :

$$\boxed{\Delta S^{rév} = \int_{(1)}^{(2)} \frac{đQ^{rév}}{T}} \; , \tag{L.7}$$

où T est aussi bien la température du système que la température extérieure T^{ex}, puisqu'elles sont alors égales.

II. LE « SECOND PRINCIPE » DE LA THERMODYNAMIQUE

La relation (L.5) ou (L.6) ci-dessus, dite *inégalité de Clausius,* et l'égalité (L.7) qui en est un cas limite, constituent l'*expression générale du «second principe»* de la thermodynamique. Les formules

$$\Delta S^{ad} \geqslant 0 \, , \tag{L.8.a}$$

$$\Delta S^{ad,rév} = 0 \, , \tag{L.8.b}$$

que nous avions démontrées au complément II.D pour les transformations adiabatiques, apparaissent maintenant comme des cas particuliers de l'inégalité de Clausius et de l'égalité qu'elle devient si la transformation est réversible. *Le «second principe» découle donc de la mécanique statistique,* c'est-à-dire en dernier ressort du postulat fondamental que nous avons posé au chapitre II.

Il ne saurait être question d'aborder ici les développements considérables qu'inspire en thermodynamique le «second principe», ni ses applications très variées. Nous renvoyons pour cela à un cours de thermodynamique proprement dite, nous contentant de dégager quelques points importants.

---

5. Pour qu'une transformation $A_{1\to 2}$ puisse être analysée de cette façon, il faut que, à chacune de ses étapes, on puisse définir la *température* T^{ex} *de la source extérieure* qui fournit la chaleur $đQ$ au système. Les autres aspects de l'échange, tels que la nature et la taille de la source extérieure, étant sans importance pour ce qui concerne le système $\mathcal{S}$ et la transformation qu'il subit, nous pouvons imaginer que, à chaque stade de la transformation, c'est un *thermostat* de température T^{ex} qui fournit la chaleur $đQ$ au système. Comme T^{ex} varie en général au cours de la transformation, il faut supposer que l'on dispose d'une batterie de thermostats de températures différentes, avec lesquels on met le système en contact successivement.

Au-delà du « premier principe » qui établit leur *équivalence* (conservation de l'énergie), le « second principe » met en évidence une *profonde dissymétrie entre travail et chaleur :* le travail reçu par un système dans des conditions réversibles ne provoque aucune variation de son entropie, alors que la chaleur reçue, même de façon réversible, change nécessairement son entropie. Nous préciserons au paragraphe III l'origine physique, microscopique, de cette dissymétrie. Rappelons que, historiquement, c'est elle qui a motivé le développement de la thermodynamique, à partir de l'étude empirique des moteurs thermiques, destinés à transformer de la chaleur en travail mécanique.

Comme l'inégalité de Clausius devient *égalité seulement pour les transformations réversibles,* celles-ci apparaissent comme un *cas limite idéal.* Les transformations réalisables en pratique sont toujours irréversibles; il faut prendre des précautions particulières pour s'approcher de la limite réversible : transformation suffisamment lente pour être quasistatique, et ajustement du « milieu extérieur » pour que les déséquilibres nécessaires aux échanges puissent être négligés.

Avant la découverte de la mécanique statistique, la *notion d'entropie* était fondée en thermodynamique sur le second principe : à partir d'énoncés initialement qualitatifs (« il est impossible de réaliser un processus qui... »), les travaux de Clausius, notamment, ont conduit à la définition d'une nouvelle « fonction d'état », baptisée « entropie ». En mécanique statistique, l'entropie est une grandeur de portée beaucoup plus générale : elle mesure le « manque d'information » associé à la distribution de probabilités caractérisant un état macroscopique; l'entropie statistique existe donc même pour des états hors d'équilibre macroscopique, ou pour des systèmes n'ayant pas une taille suffisante pour avoir atteint la limite thermodynamique.

En outre, le second principe ne déterminait en thermodynamique que des *différences d'entropie* entre les divers états d'équilibre d'un même système (différences calculables par la formule (L.7)). On faisait ensuite appel à un *« troisième principe »* pour fixer la valeur de l'entropie de façon absolue : ce « troisième principe », ou *« principe de Nernst »,* stipule que *l'entropie de tout système tend vers zéro avec la température.* En mécanique statistique, l'entropie est entièrement déterminée lorsqu'on a choisi la constante multiplicative qui figure dans son expression; sa limite de température nulle est donc parfaitement fixée. On trouve facilement[6] que

$$S \underset{T \to 0}{\longrightarrow} k \ln g_0 , \tag{L.9}$$

où g_0 est le degré de dégénérescence du niveau fondamental E_0 du système. Donc, *le « troisième principe » n'est vérifié que si le niveau fondamental du système est non*

---

6. Écrivons la fonction de partition sous la forme

$$Z = \mathrm{e}^{-E_0/kT} \sum_{(\ell)} \mathrm{e}^{-(E_\ell - E_0)/kT} , \tag{1}$$

où E_0 est le niveau fondamental du système. Lorsque T tend vers zéro, la contribution $\exp\left[-(E_\ell - E_0)/kT\right]$ d'un état (ℓ) dont l'énergie est différente de E_0 (c'est-à-dire supérieure) tend exponentiellement vers zéro. Si g_0 est le nombre d'états distincts ayant l'énergie E_0, la fonction de partition devient donc

$$Z \underset{T \to 0}{\sim} g_0 \, \mathrm{e}^{-E_0/kT} . \tag{2}$$

On en déduit l'énergie libre

$$F \mathrel{\widehat{=}} -kT \ln Z \underset{T \to 0}{\sim} E_0 - kT \ln g_0 , \tag{3}$$

et finalement l'entropie :

$$S = -\frac{\partial F}{\partial T} \underset{T \to 0}{\longrightarrow} k \ln g_0 . \tag{4}$$

dégénéré ($g_0 = 1$). C'est le cas semble-t-il pour la plupart des systèmes. Mais l'analyse de leur comportement à très basse température est délicate : au fur et à mesure que T décroît, il faut prendre en compte dans l'hamiltonien des termes de plus en plus petits, puisqu'ils deviennent supérieurs à l'énergie caractéristique kT, qui tend vers zéro. De toutes façons, pour un système macroscopique, la limite (L.9) est toujours extrêmement petite par rapport aux valeurs de l'entropie pour des températures non nulles.

Dans le cas de *transformations infinitésimales* (à *nombre de particules constant,* pour simplifier), le premier et le second principes peuvent être résumés de la façon suivante. L'accroissement d'énergie du système est donné par (complément II.D, § II.2)

$$\mathrm{d}E = đQ + đW = đQ + \sum_i X_i^{ex}\, \mathrm{d}x_i^{ex} \,. \tag{L.10}$$

Cette relation, qui définit en thermodynamique la fonction d'état « énergie du système », découle du premier principe. Le second principe définit quant à lui la fonction d'état « entropie », dont la différentielle s'écrit (cf. complément II.D, formule (D.56))

$$\mathrm{d}S = \frac{1}{T}\,\mathrm{d}E + \sum_i \frac{\overline{X}_i}{T}\,\mathrm{d}x_i \,, \tag{L.11}$$

ce que l'on peut mettre sous la forme

$$\mathrm{d}E = T\,\mathrm{d}S - \sum_i \overline{X}_i\,\mathrm{d}x_i \,. \tag{L.12}$$

Les deux expressions (L.10) et (L.12) sont *simultanément valables pour une transformation infinitésimale quelconque :* c'est le travail $đW$ des forces extérieures qui intervient dans $\mathrm{d}E$ avec la chaleur reçue $đQ$ (formule (L.10)), et c'est l'opposé du travail des forces intérieures lorsqu'on introduit la différentielle de l'entropie (formule (L.12)). Si les conditions de *réversibilité « mécanique »*

$$X_i^{ex} = \varepsilon_i \overline{X}_i \quad (\varepsilon_i = \pm 1) \tag{L.13}$$

sont satisfaites, on peut *identifier terme à terme* les deux expressions :

$$\left.\begin{aligned} đQ &= T\,\mathrm{d}S && \text{(L.14.a)} \\ đW &= -\sum_i \overline{X}_i\,\mathrm{d}x_i && \text{(L.14.b)} \end{aligned}\right\} \text{ si réversibilité mécanique.}$$

(en effet, (L.14.b) est une conséquence directe de (L.13), et (L.14.a) en découle par différence). Si l'on a réversibilité mécanique, mais que la température T du système est différente de la température extérieure T^{ex}, la transformation n'est pas véritablement réversible; l'inégalité de Clausius (L.6), qui s'écrit alors

$$\mathrm{d}S = \frac{đQ}{T} \geqslant \frac{đQ}{T^{ex}} \quad \text{si réversibilité mécanique} \,, \tag{L.15}$$

indique que le signe de $đQ$ est celui de $(T^{ex} - T)$, c'est-à-dire que la chaleur va de l'extérieur au système si $T^{ex} > T$, et en sens inverse si $T > T^{ex}$; comme toujours, cette inégalité devient égalité si et seulement si la transformation est entièrement réversible (réversibilité thermique $T^{ex} = T$ en plus de la réversibilité « mécanique » (L.13))[7].

7. Si les conditions de réversibilité « mécanique » ne sont pas satisfaites, il n'est plus vrai que la chaleur passe nécessairement dans le sens des températures décroissantes. En effet, l'inégalité de Clausius est toujours applicable sous la forme

$$\mathrm{d}S \geqslant \frac{đQ}{T^{ex}} \,; \tag{1}$$

.../...

III. INTERPRÉTATION STATISTIQUE DE LA CHALEUR ET DU TRAVAIL

1. LES DEUX EXPRESSIONS DE L'ACCROISSEMENT D'ÉNERGIE

Considérons un *système thermodynamique* à l'équilibre, et écrivons son énergie E sous la forme

$$E = \sum_{(\ell)} E_\ell P_\ell \,, \tag{L.16}$$

l'énergie étant identifiée avec sa valeur moyenne puisque les fluctuations sont négligeables. Sa variation au cours d'une transformation infinitésimale (que nous supposons à *nombre de particules constant*), peut être exprimée sous la forme

$$\mathrm{d}E = \sum_{(\ell)} E_\ell \, \mathrm{d}P_\ell + \sum_{(\ell)} P_\ell \, \mathrm{d}E_\ell \,, \tag{L.17}$$

où $\mathrm{d}E_\ell$ est le déplacement de l'énergie E_ℓ de l'état microscopique (ℓ) provoqué par la transformation, et $\mathrm{d}P_\ell$ la variation de la «population» P_ℓ de cet état (ℓ).

Si, au cours de cette transformation infinitésimale, le système reçoit le travail $đW$ et la chaleur $đQ$, son énergie s'accroît de

$$\mathrm{d}E = đW + đQ \,. \tag{L.18}$$

Nous allons comparer les formules (L.17) et (L.18) : la première découle de la mécanique statistique, la seconde exprime le «premier principe» de la thermodynamique.

2. INTERPRÉTATION STATISTIQUE DE LA CHALEUR

Examinons d'abord le cas d'une transformation dans laquelle le système reçoit *seulement de la chaleur :*

$$đW = 0 \,. \tag{L.19}$$

Dans ce cas, aucun des paramètres extérieurs x ne varie en général[8] dans la transformation, de sorte que

$$\mathrm{d}E_\ell = 0 \quad \text{pour tout } (\ell) \,. \tag{L.20}$$

On en déduit

$$đQ = \sum_{(\ell)} E_\ell \, \mathrm{d}P_\ell \quad \text{si} \quad đW = 0 \,. \tag{L.21}$$

Ainsi, *lorsqu'un système reçoit seulement de la chaleur, son énergie varie par modification des populations de ses états microscopiques :* si la quantité de chaleur reçue est positive, la probabilité des états d'énergie élevée augmente, au détriment de celle des états de plus basse énergie.

---

mais, si l'on y remplace $\mathrm{d}S$ par son expression générale (cf. complément II.D, formule (D.58)), elle s'écrit

$$\left(\frac{1}{T} - \frac{1}{T^{\mathrm{ex}}}\right) đQ + \sum_i \frac{1}{T} \left(\overline{X}_i - \varepsilon_i X_i^{\mathrm{ex}}\right) \mathrm{d}x_i \geqslant 0 \,. \tag{2}$$

Cette inégalité peut très bien être vérifiée sans que le premier terme soit nécessairement positif.

8. Dans certaines transformations pourtant (c'est le cas pour la *«détente de Joule»* décrite au chapitre II, § II.B.1), un paramètre extérieur x du système peut varier sans que cela se traduise par un échange de travail : x^{ex}, le paramètre analogue pour le milieu extérieur, reste constant. Nous excluons ici de telles transformations, qui ne vérifient pas la condition (L.20).

Si l'on tient compte du fait que le système doit être à l'équilibre dans les états initial et final de la transformation, on constate[9] que, comme on s'y attendait, *la température du système a varié :* elle a augmenté si la quantité de chaleur reçue est positive.

On comprend alors la *nature physique profonde de la chaleur,* qui apparaissait, avant la découverte de la mécanique statistique, d'abord comme un « fluide » mystérieux puis comme une forme tout aussi mystérieuse d'énergie échangée entre des systèmes : si l'énergie d'un système macroscopique peut être modifiée sans qu'on lui fournisse du travail, c'est qu'on peut modifier la distribution statistique de ses états microscopiques (ℓ) sans changer ces états ni leur énergie E_ℓ.

3. CHALEUR ET TRAVAIL ÉCHANGÉS DANS UNE TRANSFORMATION RÉVERSIBLE

Les choses se compliquent lorsque le système reçoit du travail. Ceci signifie en général[10] qu'un ou plusieurs des paramètres extérieurs x_i varie(nt) dans la

9. Choisissons, pour décrire les états d'équilibre macroscopiques du système, la distribution canonique

$$P_\ell = \frac{1}{Z}\, e^{-\beta E_\ell}\,, \tag{1}$$

où la température apparaît directement. Calculons alors

$$\frac{\partial P_\ell}{\partial T} = -\frac{1}{kT^2}\,\frac{\partial P_\ell}{\partial \beta}\,; \tag{2}$$

il vient

$$\frac{\partial P_\ell}{\partial \beta} = -E_\ell P_\ell - \frac{1}{Z^2}\,\frac{\partial Z}{\partial \beta}\, e^{-\beta E_\ell} = (\overline{E} - E_\ell)\, P_\ell\,. \tag{3}$$

Ainsi, un accroissement dT de la température (les paramètres extérieurs x restant constants) induit une variation des populations P_ℓ donnée par

$$dP_\ell = \frac{dT}{kT^2}\,(E_\ell - \overline{E})\, P_\ell\,. \tag{4}$$

Ce résultat correspond exactement à la variation de population produite par un apport de chaleur (formule (L.21)) : augmentation de la probabilité pour les états dont l'énergie E_ℓ est supérieure à l'énergie initiale $E = \overline{E}$ du système, diminution pour ceux dont l'énergie E_ℓ est inférieure à E.

10. Dans certaines transformations, on fournit au système du travail sans modifier les paramètres extérieurs autres que l'énergie. C'est le cas par exemple dans la célèbre expérience de Joule, schématisée sur la figure 1, qui démontre l'équivalence entre la chaleur et le travail (« premier principe »).

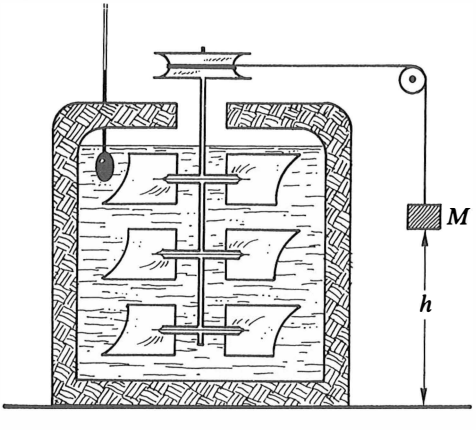


FIGURE 1
Schéma de l'expérience de Joule (1842). Le système est constitué par un liquide enfermé dans un récipient calorifugé de volume invariable. On lui fournit le travail effectué par la force de pesanteur s'exerçant sur une masse *M* qui descend d'une hauteur *h*. Ce travail *Mgh* est transmis à des pales pouvant tourner dans le liquide. Il y est transformé en chaleur par les frottements de type visqueux freinant le mouvement des pales, et provoque une augmentation de température du liquide.

transformation. Supposons pour simplifier qu'un seul de ces paramètres, x, est modifié, de sorte que

$$đW = X^{ex}\, dx^{ex} = -\varepsilon X^{ex}\, dx \qquad (\varepsilon = \pm 1)\,. \tag{L.22}$$

Les niveaux d'énergie E_ℓ du système sont alors déplacés :

$$dE_\ell = \frac{\partial E_\ell}{\partial x}\, dx = -X_\ell\, dx \tag{L.23}$$

(cf. complément II.D, formule (D.2)).

Le second terme de l'expression (L.17) s'écrit alors

$$\sum_{(\ell)} P_\ell\, dE_\ell = -\sum_{(\ell)} P_\ell X_\ell\, dx = -\overline{X}\, dx\,, \tag{L.24}$$

où $\overline{X}$ est la force moyenne conjuguée du paramètre x (complément II.D, formule (D.6)); on reconnaît là, au signe près, le travail infinitésimal de la force intérieure au système. Quant au premier terme de (L.17), on l'identifie facilement comme[11]

$$\sum_{(\ell)} E_\ell\, dP_\ell = T\, dS\,. \tag{L.25}$$

On retrouve donc exactement, explicitée en termes statistiques, la situation que nous avons analysée à la fin du paragraphe II (comparaison des formules (L.10) et (L.12))[12]; les conclusions sont les mêmes.

Dans le cas général, il n'y a pas de relation simple entre $đW$ *et* $đQ$ *d'une part et chacun des deux termes de l'expression statistique de* dE *d'autre part.* Ainsi, lorsqu'on fournit du travail à un système de façon irréversible, on modifie à la fois ses niveaux d'énergie E_ℓ et les populations P_ℓ de ses états microscopiques. Cependant,

11. D'après la définition de l'entropie statistique, son accroissement s'écrit

$$dS = -k\sum_{(\ell)} dP_\ell\,.(\ln P_\ell + 1) = -k\sum_{(\ell)} dP_\ell\,.\ln P_\ell\,, \tag{1}$$

car
$$\sum_{(\ell)} dP_\ell = 0 \tag{2}$$

par conservation de la normalisation des probabilités P_ℓ. Si nous prenons encore la distribution canonique pour évaluer dS, nous obtenons

$$dS = -k\sum_{(\ell)} dP_\ell\left[-\frac{E_\ell}{kT} - \ln Z\right] = \sum_{(\ell)} \frac{1}{T} E_\ell\, dP_\ell \tag{3}$$

(où nous avons à nouveau utilisé la condition (2)). La relation (L.25) s'en déduit aussitôt.

12. Notons cependant que les expressions (L.17) et (L.18) seraient encore valables pour l'accroissement d'énergie du système dans une étape intermédiaire infinitésimale d'une transformation quelconque, alors que la formule (L.25) suppose la transformation quasistatique : le calcul de la note 11, démontrant la relation (L.25), s'appuie sur le fait que les probabilités P_ℓ constituent une distribution d'équilibre.

dans le cas particulier d'une *transformation réversible,* les expressions (L.17) et (L.18) de l'accroissement d'énergie se correspondent terme à terme[13] :

$$đW^{\text{rév}} = \sum_{(\ell)} P_\ell \, dE_\ell^{\text{rév}} , \tag{L.26.a}$$

$$đQ^{\text{rév}} = \sum_{(\ell)} E_\ell \, dP_\ell^{\text{rév}} . \tag{L.26.b}$$

Ainsi, l'interprétation statistique du travail et de la chaleur reçus par un système est simple dans le cas d'une transformation réversible (formules (L.26)) ou dans le cas d'une transformation sans échange de travail (formule (L.21)). Cette interprétation montre clairement l'origine de la *dissymétrie entre chaleur et travail,* en même temps qu'elle élucide la *nature physique de la chaleur.*

13. Il suffit en réalité (voir paragraphe II) que la transformation soit réversible du point de vue « mécanique », c'est-à-dire que les « forces » intérieure et extérieure s'équilibrent; il n'est pas indispensable que les températures intérieure et extérieure soient égales.

Exercices III

GAZ PARFAITS

III.1. On considère une mole de gaz parfait monoatomique enfermée dans un récipient de volume V à la température $T = 300$ K. Calculer numériquement son énergie moyenne, ainsi que la probabilité pour que l'énergie du gaz diffère de plus de 10^{-4} % de cette valeur moyenne.

III.2. En mécanique relativiste, l'énergie ε d'une particule de masse m est reliée à sa quantité de mouvement $\vec{p}$ par la relation

$$\varepsilon^2 = c^2 \vec{p}^{\,2} + m^2 c^4 .$$

1) Quelle est la densité d'états $\rho(\varepsilon)$ d'une telle particule enfermée dans un récipient de volume V (cf. exercice I.23)?

2) Donner l'expression de la fonction de partition Z d'un gaz parfait constitué de N particules relativistes de masse m.

3) Montrer que, à haute température, Z peut être calculée en supposant $\varepsilon \gg mc^2$ (gaz ultrarelativiste). Préciser le domaine de validité de l'approximation.

4) En déduire les propriétés thermodynamiques d'un gaz parfait ultrarelativiste : énergie moyenne, capacité calorifique, équation d'état, ...

III.3. Un récipient de volume V est séparé en deux compartiments (1) et (2), de volumes respectifs V_1 et V_2, par une cloison amovible. Chacun des compartiments contient un gaz parfait monoatomique : N_1 molécules de masse m_1 dans (1), N_2 molécules de masse m_2 dans (2). L'ensemble est en équilibre thermique avec un thermostat à la température T.

1) Quelle est l'énergie libre F_i de l'ensemble des deux gaz séparés par la cloison?

2) On retire la cloison et on attend que le nouvel état d'équilibre s'établisse. On suppose que les deux gaz forment un mélange idéal, c'est-à-dire que les interactions entre leurs molécules sont négligeables.

a) Calculer la fonction de partition Z_f et l'énergie libre F_f du mélange des deux gaz. Comparer F_f et F_i.

b) Calculer l'énergie moyenne du système. Le thermostat a-t-il reçu ou fourni de la chaleur au cours de l'opération de mélange?

c) Déterminer l'équation d'état du mélange et montrer qu'il vérifie la loi de Dalton : la pression totale est la somme des pressions partielles des deux gaz (la pression partielle d'un gaz est celle qu'il exercerait s'il occupait seul le volume du récipient).

d) Calculer l'entropie du système et utiliser le résultat pour retrouver l'expression de l'entropie de mélange (complément II.C, § II.2.b).

3) Comment doit-on modifier les résultats de la question 2 lorsque les deux parties du récipient sont initialement remplies par des molécules de même nature?

III.4. Un gaz constitué de N molécules HCl à la température T dans un récipient de volume V est placé dans un champ électrique uniforme $\vec{\mathcal{E}}$. Chaque molécule possède un moment dipolaire électrique $p = 0{,}36 \cdot 10^{-29}$ C.m qui s'oriente dans le champ $\vec{\mathcal{E}}$.

1) En admettant que l'orientation d'un dipôle est indépendante de celle des autres et que cette orientation peut être traitée par la mécanique classique, déterminer la probabilité pour que le moment dipolaire électrique d'une molécule pointe dans un angle solide infinitésimal $d\Omega$ faisant l'angle θ avec la direction du champ $\vec{\mathcal{E}}$ (voir exercice III.27, question 3).

2) En déduire la polarisation $\vec{P}$, c'est-à-dire le moment dipolaire électrique moyen par unité de volume du gaz. Calculer la permittivité relative ε_r $\left(\widehat{=} \lim_{\mathcal{E} \to 0} (1 + P/\varepsilon_0 \mathcal{E})\right)$ du gaz HCl dans les conditions normales de température et de pression. Comment varie-t-elle, à pression donnée, avec la température?

3) Calculer l'énergie moyenne du gaz à la température T dans le champ $\vec{\mathcal{E}}$, ainsi que sa capacité calorifique à volume constant.

III.5. Retrouver les propriétés du gaz parfait à partir de l'ensemble T-p (complément III.D).

III.6. Calculer, pour $T = 300$ K puis $T = 3\,000$ K, la population des divers états (ν) de vibration de la molécule Br_2, puis de Cl_2, enfin de N_2 (on utilisera les données du complément III.B, tableau II).

Montrer que la probabilité pour que la molécule étudiée ne se trouve pas dans son état fondamental de vibration s'exprime très simplement en fonction de T_v/T. Calculer numériquement cette probabilité d'excitation pour les trois molécules précédentes à $T = 300$ K puis $T = 3\,000$ K.

III.7. Calculer, pour $T = 300$ K puis $T = 30$ K, la population des divers niveaux de rotation de la molécule H_2. Comparer les résultats avec ceux que l'on obtient pour HCl et pour N_2 (on utilisera les données du complément III.B, tableau II, sans oublier de tenir compte du facteur de dégénérescence $(2j + 1)$).

III.8. On étudie un gaz constitué de N molécules monoatomiques identiques et indépendantes, enfermé dans un récipient de volume V et maintenu à la température T par un thermostat. On traitera ce gaz dans l'approximation de Maxwell-Boltzmann.

Le cortège électronique de chacun des atomes du gaz possède, dans son état fondamental, un moment cinétique orbital total L et un spin total S; leur couplage est responsable d'une *«structure fine»* du niveau fondamental de l'atome. On prendra le cas où $L = 1$ et $S = 1/2$; l'atome possède alors deux niveaux proches : le plus bas a l'énergie $-\varepsilon_0$ (avec $\varepsilon_0 > 0$) et le moment cinétique total $J = 1/2$, de sorte qu'il lui correspond deux états distincts; le suivant a l'énergie $-\varepsilon_0 + \Delta$ (où Δ est de l'ordre de 10^{-4} eV) et le moment cinétique $J = 3/2$, il est donc quatre fois dégénéré. Le niveau d'énergie immédiatement supérieur, de valeur $-\varepsilon_1$, se situe plusieurs eV au-dessus des deux précédents (on prendra $\varepsilon_0 - \varepsilon_1 \simeq 5$ eV).

1) Expliquer pourquoi, si la température ne dépasse pas quelques milliers de kelvins, on peut se limiter à ne tenir compte que des deux niveaux les plus bas, $-\varepsilon_0$ et $-\varepsilon_0 + \Delta$.

2) Donner l'expression de la fonction de partition relative à une molécule. On notera z_e et z_t les contributions respectives des états électroniques et des états de translation à cette fonction de partition. En déduire la fonction de partition Z et l'énergie libre F du gaz.

3) Calculer l'énergie moyenne $\overline{E}$ et l'entropie S du gaz en fonction de N, V et T.

4) *a*) Que deviennent les expressions de Z, $\overline{E}$ et S dans le domaine où $kT \ll \Delta$, puis dans celui où $kT \gg \Delta$? Quelle est, dans ces deux domaines, l'expression de la chaleur spécifique c_V?

b) Peut-on comprendre simplement la différence entre les formules obtenues dans ces deux cas limites?

c) Étant donné l'ordre de grandeur de Δ, ces deux cas sont-ils tous deux réalistes?

III.9. On veut analyser par des calculs simples la façon dont la chaleur spécifique de rotation $c_V^{(r)}$ d'un gaz diatomique hétéropolaire tend vers zéro à basse température et vers R à haute température (fig. 4 du complément III.B).

1) Dans la région où $T \ll T_r$, on peut ne garder que les deux premiers termes de la somme donnant la fonction de partition ζ_r (formule (B.46) du complément III.B). Montrer que, à cette approximation, l'énergie libre rotationnelle du gaz s'écrit

$$F_r \simeq -3NkT \exp[-2T_r/T] .$$

Calculer l'énergie moyenne et l'entropie correspondantes, puis la chaleur spécifique à volume constant.

2) Dans la région de haute température, on veut évaluer les corrections à l'approximation consistant à remplacer la somme discrète donnant ζ_r par une intégrale (formule (B.47) du complément III.B). Ceci peut se faire grâce à la *formule d'Euler-MacLaurin :*

$$\sum_{n=a}^{b} f(n) = \int_a^b f(n)\,\mathrm{d}n + \frac{1}{2}[f(b)+f(a)] + \sum_{q=1}^{\infty} (-1)^q \frac{B_q}{(2q)!}\left[f^{(2q-1)}(a) - f^{(2q-1)}(b)\right],$$

où $f^{(k)}$ est la dérivée k-ième de la fonction f et les coefficients B_q sont les nombres de Bernoulli ($B_1 = 1/6$, $B_2 = 1/30$, $B_3 = 1/42$, ...).

Montrer que le développement de ζ_r à haute température s'écrit

$$\zeta_r = \frac{T}{T_r}\left\{1 + \frac{1}{3}\frac{T_r}{T} + \frac{1}{15}\left(\frac{T_r}{T}\right)^2 + \ldots\right\} .$$

En calculant le premier terme correctif en T_r/T à la chaleur spécifique $c_V^{(r)}$, montrer que celle-ci tend vers sa limite de haute température par valeurs supérieures (cf. figure 4 du complément III.B). Calculer à la même approximation l'entropie rotationnelle du gaz.

III.10. On veut évaluer la correction apportée à la fonction de partition d'une molécule diatomique par le couplage vibration-rotation et l'anharmonicité des vibrations (§ II.4.b du complément III.B).

L'énergie η_{vr} de vibration et de rotation de la molécule peut être écrite (C.D.L., p. 880)

$$\eta_{vr} = \left(\nu + \frac{1}{2}\right)\hbar\omega + j(j+1)\frac{\hbar^2}{2I} - a_1[j(j+1)]^2\frac{\hbar^2}{2I} - a_2\left(\nu+\frac{1}{2}\right)j(j+1)\frac{\hbar^2}{2I} - c\left(\nu+\frac{1}{2}\right)^2\hbar\omega ,$$

l'origine étant toujours prise au minimum du potentiel $u(\rho)$: a_1, a_2 et c sont des constantes sans dimension, positives, et très petites devant 1 (a_1 caractérise la variation du moment d'inertie sous l'effet de la « force centrifuge » de rotation, a_2 le couplage rotation-vibration et c l'effet d'anharmonicité; ces paramètres peuvent être déterminés expérimentalement par l'analyse spectroscopique du rayonnement émis par les molécules). On cherche la fonction de partition corrigée sous la forme approchée

$$\zeta_{vr} = \zeta_v \cdot \zeta_r \left[1 + a_1 f_1(T) + a_2 f_2(T) + c g(T)\right],$$

où ζ_v et ζ_r sont données par les formules (B.27) et (B.48) du complément III.B.

1) Calculer ζ_{vr} au premier ordre en a_1, a_2 et c. Montrer que f_1, f_2 et g peuvent se déduire des valeurs moyennes de l'énergie et de son carré pour la rotation et pour la vibration.

2) Évaluer numériquement la correction correspondante à l'énergie moyenne et à la chaleur spécifique de l'oxygène à la température ordinaire; on utilisera les données du complément III.B (tableau II) et les valeurs suivantes des paramètres : $a_1 \simeq 3 \times 10^{-6}$, $a_2 \simeq 10^{-2}$, $c \simeq 8 \times 10^{-3}$.

THÉORIE CINÉTIQUE

III.11. Calculer la valeur moyenne et l'écart quadratique moyen de la vitesse relative de deux molécules d'un gaz à l'équilibre à la température T. Comparer ces résultats avec ceux que l'on obtient pour la vitesse d'une seule molécule.

III.12. On considère des atomes (ou molécules) qui émettent, lorsqu'ils sont au repos dans le référentiel du laboratoire, une raie lumineuse pratiquement monochromatique, de fréquence ν_0. Des atomes de ce type, à l'état gazeux, sont placés dans un four chauffé à la température T; une fenêtre ménagée dans la paroi du four permet l'observation du rayonnement émis dans la direction Ox.

1) Lorsque l'atome qui émet est animé par rapport au référentiel du laboratoire d'une vitesse $\vec{v}$, de composante v_x suivant Ox, la fréquence ν perçue par l'observateur est déplacée par effet Doppler. Exprimer ν en fonction de ν_0 et v_x.

2) L'intensité lumineuse $I(\nu)\mathrm{d}\nu$ reçue dans un intervalle de fréquence (ν, $\nu+\mathrm{d}\nu$) est proportionnelle au nombre d'atomes qui ont, à $\mathrm{d}v_x$ près, la vitesse correspondant au déplacement Doppler $\nu - \nu_0$. Comment $I(\nu)$ varie-t-il avec ν? Comment la mesure de $I(\nu)$ permet-elle de déterminer la température du four?

III.13. Effusion

1) Un gaz supposé parfait, constitué de N atomes de masse m, est placé dans un récipient de volume V et maintenu à la température T. Calculer le nombre d'atomes qui viennent, pendant un intervalle de temps $\mathrm{d}t$, heurter un élément de surface $\mathrm{d}s$ de la paroi avec une vitesse de module v (à $\mathrm{d}v$ près).

2) On découpe dans la paroi un trou de surface $\mathrm{d}s$, assez petit pour que le comportement des atomes à l'intérieur du récipient reste décrit par la distribution de Maxwell. On appelle « effusion » ce phénomène de fuite lente du gaz à travers un trou de petite dimension. Calculer le nombre d'atomes qui quittent le récipient pendant un intervalle de temps $\mathrm{d}t$, ainsi que la quantité de mouvement qu'ils emportent. A l'aide d'un diaphragme, on sélectionne les atomes dont la vitesse est normale à la paroi. Écrire la distribution statistique des vitesses à l'intérieur de ce jet.

3) Applications

a) Deux récipients, maintenus à des températures différentes, contiennent des gaz constitués des mêmes atomes, sous des pressions différentes. Ils sont séparés par une paroi adiabatique percée d'un trou assez petit pour que l'on puisse admettre que l'équilibre règne en permanence dans chacun des récipients. Écrire, en fonction des températures et des pressions des deux compartiments, la condition pour que les nombres d'atomes qu'ils contiennent restent constants dans le temps.

b) Deux gaz différents, formés de particules de masses m_1 et m_2, sont mélangés dans un récipient dont la paroi est percée d'un petit trou. Montrer que la proportion des deux types de particules n'est pas la même dans le récipient et dans le gaz qui s'en échappe par effusion. On enrichit ainsi l'uranium naturel en son isotope « léger » 235, utilisé dans les centrales nucléaires et les piles : l'hexafluorure d'uranium UF_6 étant gazeux, on en provoque l'effusion. Évaluer l'enrichissement que l'on peut ainsi obtenir. Combien de fois faut-il procéder à un enrichissement par effusion pour obtenir 2,5 % d'uranium 235 (la composition isotopique de l'uranium naturel est d'environ 0,7 % de 235 pour 99,3 % de 238)?

FLUCTUATIONS

III.14. Expérience de Kappler. Kappler détermina expérimentalement (*Annalen der Physik 11,* 233 (1931)) la constante de Boltzmann en étudiant les fluctuations d'un petit miroir suspendu verticalement à un fil de torsion et placé dans une enceinte maintenue à température constante. Il enregistra, sur un long intervalle de temps, les fluctuations de l'angle θ que fait le miroir avec sa position d'équilibre et en déduisit un angle carré moyen $\overline{\theta^2} = 4{,}178.10^{-6}$ (en radian^2). Sachant que l'enceinte était maintenue à 278,1 K et que la constante de rappel du fil de torsion valait $9{,}428.10^{-9}$ g.cm^2.s^{-2}, quelle valeur Kappler trouva-t-il pour la constante de Boltzmann?

III.15. Un système en équilibre avec un thermostat à la température T est constitué de N particules indépendantes. Soit ε l'énergie d'une particule du système. Calculer son écart quadratique moyen $\Delta\varepsilon$ en fonction de T, N et de la capacité calorifique à volume constant C_V du système. Comment varie $\Delta\varepsilon/\bar{\varepsilon}$ avec N? Combien vaut ce rapport pour un gaz parfait monoatomique? Comparer avec $\Delta E/\bar{E}$ pour l'énergie totale E des N particules.

III.16. Deux systèmes macroscopiques $\mathcal{S}$ et $\mathcal{S}'$ sont en équilibre thermique.

1) On suppose d'abord que $\mathcal{S}'$ est suffisamment gros pour constituer un thermostat vis-à-vis de $\mathcal{S}$.

a) Montrer que l'écart quadratique moyen ΔE sur l'énergie de $\mathcal{S}$ peut s'écrire

$$(\Delta E)^2 = kT^2 C_V\,, \qquad (1)$$

où C_V est la capacité calorifique à volume constant de $\mathcal{S}$. Comment varie $\Delta E/\bar{E}$ avec la taille du système $\mathcal{S}$?

b) Calculer $\overline{(E-\bar{E})^3}$. Comment le rapport $[\overline{(E-\bar{E})^3}]^{1/3}/\bar{E}$ varie-t-il avec la taille de $\mathcal{S}$?

2) On se place maintenant dans le cas où $\mathcal{S}'$ n'est pas nécessairement beaucoup plus gros que $\mathcal{S}$. L'énergie totale E_{tot} du système global réunissant $\mathcal{S}$ et $\mathcal{S}'$ est fixée.

a) Quelle est la probabilité $P(E)$ pour que S ait l'énergie E?

b) Comment trouve-t-on la valeur la plus probable E_m de E?

c) En développant $\ln P(E)$ autour de $E = E_m$, montrer que l'écart quadratique moyen ΔE a ici pour expression

$$(\Delta E)^2 = \frac{kT^2}{1/C_V + 1/C'_V} . \tag{2}$$

Que représente T dans cette formule?

d) Retrouver à partir de là le résultat (1) et son domaine de validité.

III.17. Statistique d'une corde vibrante en présence d'un thermostat. Une corde vibrante, de masse M, de longueur au repos L_0, est tendue entre deux points fixes A et B distants de L ($>L_0$). La droite AB, position de repos de la corde (on note alors F la tension de la corde) est prise comme axe des x, avec origine en A. Les déformations de la corde sont repérées sur un axe Ay perpendiculaire à Ax (on suppose pour simplifier que les mouvements de la corde sont « polarisés » dans le plan xAy) : $y(x, t)$ est le déplacement du point d'abscisse x à l'instant t. On a évidemment $y(0, t) = y(L, t) \equiv 0$ et on suppose que les déformations restent petites en tout point (plus exactement, on supposera dans tout le problème $\partial y/\partial x \ll 1$).

1) Lorsque la corde est écartée de sa position de repos, sa longueur augmente de ℓ ($\ll L$), et sa tension de $f (= k\ell \ll F)$.

a) Montrer que son énergie potentielle élastique augmente alors de $F\ell$.

b) Montrer que, lorsque le mouvement de la corde est décrit par la fonction $y(x, t)$, son énergie $E\{y(x, t)\}$ (cinétique et potentielle), mesurée à partir de sa position de repos, vaut (dans la limite des faibles déplacements)

$$E\{y(x, t)\} = \frac{1}{2}\int_0^L \mathrm{d}x \left[\frac{M}{L}\left(\frac{\partial y}{\partial t}\right)^2 + F\left(\frac{\partial y}{\partial x}\right)^2\right] .$$

2) On veut étudier l'équilibre statistique de la corde en présence d'un thermostat à la température T. Pour cela on développe $y(x, t)$ en série de Fourier sous la forme

$$y(x, t) = \sum_{n=1}^{N} A_n(t) \sin \frac{n\pi x}{L} .$$

Si la corde était vraiment un milieu continu, N serait égal à l'infini (ce qui poserait problème pour le calcul de l'énergie moyenne ou la capacité calorifique); c'est la nature discontinue, au niveau microscopique, du matériau qui impose à n une valeur maximum N (cf. complément III.E, § II.1.c et note 8).

a) Montrer que

$$E\{y(x, t)\} = \sum_{n=1}^{N} \left[\frac{M}{4}\dot{A}_n^2 + \frac{n^2\pi^2 F}{4L} A_n^2\right] .$$

b) Les A_n et $\dot{A}_n$ constituent ainsi un système de $2N$ coordonnées généralisées définissant l'état de la corde à l'instant t. On admettra qu'elles peuvent varier de façon continue. Quelle est la probabilité pour que, lorsque la corde est en équilibre à la température T, ces coordonnées prennent les valeurs $A_1, A_2, \ldots, A_N, \dot{A}_1, \dot{A}_2, \ldots, \dot{A}_N$ à $\mathrm{d}A_1, \mathrm{d}A_2, \ldots, \mathrm{d}A_N, \mathrm{d}\dot{A}_1, \ldots, \mathrm{d}\dot{A}_N$ près?

c) En déduire l'énergie moyenne de la corde, ainsi que sa capacité calorifique à la température T et sous la tension F.

3) On se propose maintenant de déterminer l'amplitude des mouvements de la corde provoqués par l'agitation thermique.

a) Déterminer la valeur moyenne de A_n et du produit $A_n A_{n'}$ pour n et n' fixés (on distinguera les cas $n = n'$ et $n \neq n'$).

b) En déduire que la position moyenne de la corde est le segment de droite AB $(\bar{y}(x, t) = 0)$.

c) Exprimer le carré moyen $\overline{y^2}(x, t)$ sous forme d'une somme sur n. Montrer que l'écart quadratique moyen de $y(x, t)$ est maximum au point d'abscisse $x = L/2$ et calculer cet écart quadratique moyen maximum $\left(\text{on rappelle que } \sum_{n=1}^{\infty} n^{-2} = \pi^2/8\right)$.

ÉQUILIBRES ET DÉTERMINATION DE VARIABLES INTERNES

III.18. On considère un cristal semiconducteur extrinsèque, de volume V, maintenu à la température T par un thermostat. Ce cristal contient N pièges localisés, identiques, et N électrons qui peuvent être soit piégés, leur énergie étant alors égale à $-\Delta$ $(\Delta > 0)$, soit libres.

1) On considère que le nombre d'électrons libres est fixé. Écrire la fonction de partition du système, qui contient alors n électrons libres et $N - n$ électrons piégés.

2) En fait, le nombre n d'électrons libres est une variable interne. Écrire la fonction de partition Z du système sous forme d'une somme sur les valeurs de n. Quelle est la probabilité pour qu'il y ait n_p électrons piégés? En déduire la valeur la plus probable n_p^0 de ce nombre. Étudier les cas limites $T \to 0$ et $T \to \infty$.

3) Quelle est la fonction de partition d'un électron? Pourquoi ne peut-on pas en déduire simplement la fonction de partition Z des N électrons? Montrer que si n_p^0 est très inférieur à N, cette difficulté disparaît. Retrouver dans ce cas les résultats du 2.

III.19. On veut étudier *l'équilibre liquide-vapeur* dans le cadre d'un modèle microscopique très simple.

1) *Modèle pour la phase gazeuse*

La vapeur est considérée comme un gaz parfait classique constitué de molécules polyatomiques : l'énergie ε de chacune d'elles s'écrit

$$\varepsilon = \frac{\hbar^2 \vec{k}^2}{2m} + \varepsilon'_\tau ,$$

où m est la masse de la molécule et ε'_τ représente l'énergie associée à son état interne (τ) (rotation, vibration, ...).

Écrire la fonction de partition et l'énergie libre d'un gaz constitué de N_g telles molécules enfermées dans un volume V_g à la température T. On posera

$$z_i = \sum_{(\tau)} e^{-\varepsilon'_\tau / kT} .$$

2) *Modèle pour la phase liquide*

Dans le liquide, les interactions entre molécules ne peuvent plus être négligées. On va les schématiser en traitant ici aussi les molécules comme indépendantes, mais en considérant que chacune d'elles possède, en plus de l'énergie ε écrite ci-dessus, une énergie potentielle constante $-\varepsilon_0$ (avec $\varepsilon_0 > 0$). Le liquide est en outre pratiquement incompressible; on va supposer que son volume vaut toujours $N_\ell v_0$, où N_ℓ est le nombre de molécules du liquide et v_0 un volume élémentaire constant.

Calculer dans ce modèle la fonction de partition et l'énergie libre d'un liquide constitué de N_ℓ molécules à la température T.

3) *Équation de la courbe de vaporisation*

a) Écrire la fonction de partition $Z(T, V, N_g, N_\ell)$ et l'énergie libre d'un système liquide-vapeur dont on aurait fixé le nombre N_g de molécules de la phase vapeur et le nombre N_ℓ de molécules de la phase liquide, ainsi que la température T et le volume total V.

b) En réalité, N_g et N_ℓ sont des variables internes, dont la somme N est seule fixée.

On suppose les fluctuations négligeables (limite thermodynamique) et l'on traitera dans la suite les variables internes comme si elles étaient fixées à leur valeur la plus probable (cf. exercice II.19).

Trouver l'équation donnant la valeur la plus probable de N_g et N_ℓ.

c) Montrer que cette équation peut être considérée comme donnant la pression p du système en fonction de la température T. Cette relation caractérise la *« courbe de vaporisation »* dans le plan (T, p).

d) Résoudre l'équation en supposant $pv_0/kT \ll 1$. Justifier cette hypothèse.

4) *Chaleur latente et relation de Clapeyron*

a) Calculer la différence d'entropie $S_g - S_\ell$ entre une mole de vapeur et une mole de liquide à l'équilibre, en fonction de la température T. En déduire la *« chaleur latente »* molaire de vaporisation

$$L \widehat{=} T(S_g - S_\ell).$$

b) Soit de même V_g le volume d'une mole de vapeur et V_ℓ celui d'une mole de liquide, toujours à l'équilibre. Montrer, dans le cadre de l'approximation introduite en 3.*d*, que la *« relation de Clapeyron »*

$$L = T(V_g - V_\ell)\frac{\mathrm{d}p}{\mathrm{d}T}$$

est vérifiée dans ce modèle.

5) *Évaluation de ε_0 et discussion du modèle*

On prend l'exemple de l'eau, dont on connaît les principales caractéristiques. On se place à la température d'ébullition T_B sous la pression atmosphérique.

a) Évaluer numériquement ε_0/kT_B à partir de la chaleur latente $L = 40{,}7$ kJ/mole.

b) Estimer d'autre part ε_0/kT_B à partir de l'équation de la courbe de vaporisation.

c) Comparer ces deux valeurs numériques et discuter la validité du modèle. En particulier, l'expérience montre que la chaleur latente est une fonction décroissante de la température; peut-on comprendre simplement l'origine physique de cette dépendance?

III.20. Le but de ce problème est d'étudier la *sublimation* d'un solide cristallin à partir d'un modèle microscopique simplifié.

Un gaz monoatomique et un solide cristallin constitué des mêmes atomes, de masse m et de spin nul, coexistent à l'équilibre dans une enceinte de volume V maintenue à la température T par un thermostat. On négligera le volume du cristal devant celui du gaz.

La vapeur est assimilée à un *gaz parfait* classique.

Dans le solide, les atomes sont fixés aux nœuds d'un réseau cristallin. Ils sont supposés *indépendants* les uns des autres, le mouvement de chacun d'eux étant celui d'un *oscillateur harmonique à 3 dimensions* (modèle d'Einstein) : l'état d'un atome est caractérisé par 3 entiers positifs ou nuls n_x, n_y, n_z, l'énergie correspondante étant

$$\varepsilon_{n_x, n_y, n_z} = -\varepsilon_0 + \left(n_x + \frac{1}{2} + n_y + \frac{1}{2} + n_z + \frac{1}{2}\right)\hbar\omega ,$$

où ε_0 et ω sont des constantes positives caractéristiques du cristal, et supposées connues.

1) *Condition d'équilibre entre le solide et la vapeur*

a) Écrire la fonction de partition $Z(T, V, N_g, N_s)$ du système global lorsque N_s atomes sont à l'état solide et N_g à l'état gazeux.

b) En réalité, seul le nombre total N d'atomes du système est fixé; N_g et N_s sont des variables internes libres de fluctuer. Quelle est la probabilité pour que N_g atomes se trouvent à l'état gazeux?

c) En déduire la valeur la plus probable N_{gm} du nombre d'atomes de la vapeur.

d) A température T donnée, comment N_{gm} varie-t-il avec V? Montrer que, pour N et T fixés, l'équilibre solide-vapeur n'est possible que si le volume V de l'enceinte est inférieur à une certaine valeur V_0 que l'on déterminera. Que se passe-t-il si l'on impose au système un volume V supérieur à V_0?

Dans la suite, on suppose qu'on reste toujours dans le domaine $V < V_0$.

2) *Pression du système à l'équilibre*

a) Écrire l'expression exacte de la fonction de partition $Z(T, V, N)$ du système solide-vapeur comme une somme sur toutes les valeurs possibles de N_g. En déduire l'expression de la pression p du système; montrer qu'elle est proportionnelle au nombre moyen $\overline{N}_g$ d'atomes du gaz.

b) Sachant que, pour un grand système, $\overline{N}_g$ et N_{gm} coïncident pratiquement, écrire la pression p en fonction des paramètres extérieurs; montrer qu'elle ne dépend que de la température.

3) *Chaleur latente de sublimation*

On admettra que l'entropie S du système à l'équilibre est en pratique la somme des entropies respectives d'une phase vapeur constituée de N_{gm} atomes et d'une phase solide constituée de $(N - N_{gm})$ atomes.

a) Montrer que S peut se mettre sous la forme

$$S = Ns_1(T) + N_{gm}s_2(T)$$

et donner l'expression de $s_1(T)$ et $s_2(T)$.

b) N et T étant fixés, on augmente le volume de l'enceinte de manière à faire passer N_A atomes de la phase solide à la phase vapeur (N_A est le nombre d'Avogadro). Calculer l'augmentation ΔV de volume nécessaire pour obtenir ce transfert.

c) Calculer l'accroissement ΔS de l'entropie du système dans cette opération. En déduire la *« chaleur latente de sublimation »* (molaire) L_s, définie par

$$L_s \widehat{=} T\Delta S .$$

d) Vérifier sur les résultats des questions précédentes la *« relation de Clapeyron »*

$$L_s = T\Delta V \frac{\mathrm{d}p}{\mathrm{d}T} .$$

III.21. Équilibre de dissociation des molécules de chlore. Dans une enceinte de volume V, initialement à la température ordinaire, on introduit N_0 molécules Cl_2. On chauffe ensuite jusqu'à une température T suffisante pour que puisse se produire la réaction réversible de dissociation-recombinaison :

$$Cl_2 \rightleftarrows 2\,Cl\ .$$

Coexistent alors dans l'enceinte un gaz monoatomique et un gaz diatomique; on suppose qu'ils sont parfaits, que l'approximation de Maxwell-Boltzmann leur est applicable, et qu'ils ne sont couplés entre eux que par la réaction ci-dessus.

1) *Étude du gaz de chlore monoatomique*

a) Écrire la fonction de partition $Z_a(T, V, N_a)$ d'un gaz parfait constitué de N_a atomes Cl de masse m.

b) En déduire l'énergie libre, l'énergie moyenne et la pression de ce gaz de chlore monoatomique.

2) *Étude du gaz de chlore diatomique*

On considère maintenant un gaz parfait de N_d molécules Cl_2, de masse $m_d = 2m$. Pour caractériser l'état d'une molécule, il faut se donner :

(i) l'impulsion $\vec{p}$ de son centre de masse, qui fixe son état de translation;

(ii) l'état de vibration des deux atomes l'un par rapport à l'autre;

(iii) l'état de rotation de la molécule autour de son centre de masse.

Le degré de liberté de vibration est assimilé à un oscillateur harmonique à une dimension : l'état de vibration est caractérisé par un entier n positif ou nul, l'énergie correspondante valant

$$\varepsilon_n^{(v)} = -u_0 + \left(n + \frac{1}{2}\right)\hbar\omega_0\ ,$$

où ω_0 et u_0 sont des constantes positives connues. Les degrés de liberté de rotation sont traités par la mécanique classique (exercice III.27).

Écrire la fonction de partition $Z_d(T, V, N_d)$ du système constitué par les N_d molécules Cl_2. En déduire l'énergie moyenne et la pression de ce gaz de chlore diatomique.

3) *Étude de l'équilibre*

a) En utilisant les résultats des questions 1 et 2, donner l'expression de la probabilité $P(T, V, N_0; N)$ pour que N molécules Cl_2 parmi les N_0 soient dissociées à la température T.

b) Calculer la valeur la plus probable du nombre N de molécules dissociées à la température T.

c) Montrer que, pour cette valeur la plus probable de N, les pressions partielles (la pression partielle d'un gaz dans un mélange est celle qu'il exercerait s'il était seul à occuper le volume disponible) p_d et p_a du gaz diatomique et du gaz monoatomique vérifient la relation

$$\frac{p_a^2}{p_d} = K(T)\ ,$$

où $K(T)$ est une fonction de la température seule, que l'on explicitera.

d) Indiquer l'allure des variations de $K(T)$. En utilisant les données du complément III.B (tableau II), préciser où se situe en pratique le domaine de températures dans lequel la dissociation est appréciable.

III.22. La molécule d'hydrogène H_2 et celle de deutérium D_2 ont des températures rotationnelles T_r suffisamment élevées (et des points de liquéfaction suffisamment bas) pour que les effets quantiques décrits au paragraphe II.3.c du complément III.B soient observables. Le spin i du noyau d'hydrogène est 1/2, celui du noyau de deutérium 1. Les molécules dont les spins nucléaires sont dans un état symétrique sont appelées « ortho-hydrogène » ou « ortho-deutérium », celles où l'état de spin est antisymétrique « para-hydrogène » ou « para-deutérium ».

1) A partir des résultats du paragraphe II.3.c du complément III.B, écrire, pour H_2 et D_2, la fonction de partition associée aux degrés de liberté de rotation et aux degrés de liberté de spin des noyaux.

Montrer que, si T devient petite devant T_r, la somme sur les j pairs tend vers 1 et celle sur les j impairs tend exponentiellement vers zéro. En déduire que, dans ce domaine de température, le gaz se comporte comme s'il était monoatomique (gel des degrés de liberté de rotation). Calculer à cette approximation son énergie libre, son entropie et son potentiel chimique. Préciser numériquement jusqu'à quelles valeurs de T/T_r on peut considérer ces résultats comme réalistes.

2) Montrer que le rapport entre le nombre de molécules ortho et le nombre de molécules para est donné par

$$\frac{N_o(H_2)}{N_p(H_2)} = \frac{3\zeta_{\text{impair}}(H_2)}{\zeta_{\text{pair}}(H_2)} \qquad \text{pour l'hydrogène,}$$

et

$$\frac{N_o(D_2)}{N_p(D_2)} = \frac{2\zeta_{\text{pair}}(D_2)}{\zeta_{\text{impair}}(D_2)} \qquad \text{pour le deutérium.}$$

Trouver les limites de haute et basse température de ces rapports, et tracer l'allure de leurs variations en fonction de la température.

3) Nous avons jusqu'ici supposé qu'il y avait équilibre entre les variétés ortho et para. En pratique, cet équilibre est très long à s'établir (car le couplage entre spins nucléaires est très faible), sauf en présence de catalyseurs adaptés. On peut donc observer, à toute température, des états métastables dans lesquels la proportion de molécules ortho et de molécules para est fixée à une valeur différente de celle qui a été déterminée à la question 2.

Considérons en particulier le comportement à basse température d'un gaz d'ortho-hydrogène pur ou de para-deutérium pur. En conservant dans ζ_r seulement le premier terme de la somme, calculer l'énergie libre du gaz. En déduire son énergie et sa chaleur spécifique, ainsi que son entropie et son potentiel chimique.

4) En calculant numériquement (sur calculette ou ordinateur) les sommes ζ_{pair} et ζ_{impair} en fonction de T/T_r, préciser les variations de la chaleur spécifique de l'ortho et du para-hydrogène, du mélange ortho-para à l'équilibre (en présence d'un catalyseur) et du mélange ortho-para en proportions 3 : 1 (mélange métastable obtenu en laissant s'établir l'équilibre à température ordinaire avant de refroidir à la température T choisie). Les courbes correspondantes se trouvent dans *D.* McQuarrie, *« Statistical Mechanics »,* Harper and Row (1976), p. 106. Des calculs analogues sont évidemment possibles pour le deutérium.

III.23. La « loi d'action de masse » pour les équilibres chimiques. Considérons une réaction chimique réversible, que nous écrirons sous la forme générale

$$\nu_1 A_1 + \nu_2 A_2 + \ldots + \nu_q A_q \rightleftarrows \nu'_1 A'_1 + \nu'_2 A'_2 + \ldots + \nu'_{q'} A'_{q'} : \qquad (1)$$

les A_i ($i = 1, 2, \ldots, q$) et $A'_{i'}$ ($i' = 1, 2, \ldots, q'$) sont les symboles chimiques des molécules prenant part à la réaction et les ν_i et $\nu'_{i'}$ des nombres *entiers* positifs, appelés les « coefficients stœchiométriques ». On suppose que la réaction se produit en phase gazeuse, dans un récipient de volume V fixé et à la température T, et que les réactifs

se comportent comme des *gaz parfaits classiques* n'interagissant pas les uns avec les autres (mélange «idéal»), sauf à travers la réaction (1).

1) Écrire la fonction de partition Z du système des $(q+q')$ gaz parfaits indépendants lorsqu'on fixe les nombres $N_1, N_2, \ldots, N_q$ et $N'_1, N'_2, \ldots, N'_{q'}$ des divers types de molécules. On prendra la fonction de partition z_j d'une molécule A_j sous la forme

$$z_j = V f_j(T) , \tag{2}$$

où $f_j(T)$ est une fonction connue de la température seule.

2) En réalité, les nombres N_i $(i=1, 2, \ldots, q)$ et $N'_{i'}$ $(i'=1, 2, \ldots, q')$ sont des *variables internes* qui s'ajustent pour amener le mélange à l'équilibre chimique.

a) Montrer que les accroissements de ces $(q+q')$ variables internes sont liés, de sorte qu'il reste une seule variable interne indépendante, que nous noterons n :

$$\frac{dN'_1}{\nu'_1} = \frac{dN'_2}{\nu'_2} = \ldots = \frac{dN'_{q'}}{\nu'_{q'}} = -\frac{dN_1}{\nu_1} = -\frac{dN_2}{\nu_2} = \ldots = -\frac{dN_q}{\nu_q} \widehat{=} dn . \tag{3}$$

b) Montrer que les *valeurs d'équilibre* (valeurs les plus probables) des N_i et $N'_{i'}$ sont données par la condition

$$\sum_{i=1}^{q} \nu_i \mu_i = \sum_{i'=1}^{q'} \nu'_{i'} \mu'_{i'} , \tag{4}$$

où μ_i et $\mu'_{i'}$, fonctions des variables T, V, N_j $(j=1, 2, \ldots, q)$ et $N'_{j'}$ $(j'=1, 2, \ldots, q')$, sont les potentiels chimiques des molécules A_i et $A'_{i'}$.

3) La *«loi d'action de masse»* se déduit de la condition d'équilibre (4); on peut en donner diverses formes équivalentes.

a) On définit les fugacités

$$\varphi_j \widehat{=} e^{\mu_j / kT} . \tag{5}$$

Montrer que la condition (4) peut s'écrire

$$\frac{\varphi'^{\nu'_1}_1 \times \varphi'^{\nu'_2}_2 \times \ldots \times \varphi'^{\nu'_{q'}}_{q'}}{\varphi^{\nu_1}_1 \times \varphi^{\nu_2}_2 \times \ldots \times \varphi^{\nu_q}_q} = 1 . \tag{6}$$

b) Montrer que, si les réactifs sont des *gaz parfaits indépendants,* la loi d'action de masse (6) prend la forme

$$\frac{N'^{\nu'_1}_1 \times N'^{\nu'_2}_2 \times \ldots \times N'^{\nu'_{q'}}_{q'}}{N^{\nu_1}_1 \times N^{\nu_2}_2 \times \ldots \times N^{\nu_q}_q} = V^{\Delta\nu} K(T) , \tag{7}$$

où

$$\Delta\nu \widehat{=} \sum_{i'=1}^{q'} \nu'_{i'} - \sum_{i=1}^{q} \nu_i . \tag{8}$$

La «constante d'équilibre» K dépend de la température seule.

c) Toujours dans le cas de gaz parfaits indépendants, montrer que la loi d'action de masse peut aussi s'exprimer en termes des «pressions partielles» :

$$\frac{p'^{\nu'_1}_1 \times p'^{\nu'_2}_2 \times \ldots \times p'^{\nu'_{q'}}_{q'}}{p^{\nu_1}_1 \times p^{\nu_2}_2 \times \ldots \times p^{\nu_q}_q} = \widetilde{K}(T) , \tag{9}$$

où

$$p_j \widehat{=} \frac{N_j kT}{V} . \tag{10}$$

La «constante d'équilibre» $\widetilde{K}$ dépend seulement de la température.

d) On peut également définir la *concentration molaire* du réactif A_j par

$$c_j \widehat{=} \frac{N_j}{N}, \tag{11}$$

où $$N \widehat{=} \sum_{i=1}^{q} N_i + \sum_{i'=1}^{q'} N'_{i'} \tag{12}$$

est le nombre total de molécules du système (qui n'est pas fixé à l'avance, mais varie avec les N_i et $N'_{i'}$). En déduire une troisième forme de la loi d'action de masse :

$$\frac{c_1'^{\nu'_1} \times c_2'^{\nu'_2} \times \ldots \times c_{q'}'^{\nu'_{q'}}}{c_1^{\nu_1} \times c_2^{\nu_2} \times \ldots \times c_q^{\nu_q}} = p^{-\Delta\nu} \widetilde{K}(T), \tag{13}$$

où p est la pression globale du système.

4) On augmente la pression (ou on diminue le volume) à température constante. Ceci déplace l'équilibre, c'est-à-dire augmente (ou diminue) la quantité des réactifs qui figurent au second membre de la réaction (1), au détriment (ou en faveur) de ceux du premier membre. Montrer que le sens de ce déplacement dépend du signe de $\Delta\nu$ (c'est la « loi de Le Châtelier »).

Application : on veut synthétiser de l'ammoniac par la réaction

$$3H_2 + N_2 \rightleftarrows 2NH_3 \tag{14}$$

en phase gazeuse; a-t-on intérêt à opérer sous forte ou faible pression?

5) *a*) Exprimer (pour des gaz parfaits indépendants), à l'aide des fonctions $f_j(T)$ (formule (2)), l'énergie E d'un système constitué de ν_1 molécules A_1, ν_2 molécules A_2, ... et ν_q molécules A_q dans un récipient de volume V à la température T; exprimer de même l'énergie E' d'un système constitué de ν'_1 molécules A'_1, ν'_2 molécules A'_2, ... et $\nu'_{q'}$ molécules $A'_{q'}$ dans le même volume et à la même température. La différence

$$Q_V \widehat{=} E' - E \tag{15}$$

est appelée *« chaleur de réaction à volume constant »*.

b) Démontrer la relation

$$\frac{\mathrm{d}}{\mathrm{d}T} \ln K(T) = \frac{1}{kT^2} Q_V. \tag{16}$$

c) On augmente la température, à volume constant. Montrer que le sens dans lequel se déplace l'équilibre dépend du signe de Q_V (autre forme de la loi de Le Châtelier). Quand on veut préparer un composé chimique, dans quels cas a-t-on intérêt à opérer à température élevée?

6) On suppose maintenant que la réaction (1) se produit, non plus en phase gazeuse, mais *en solution* dans un solvant neutre (c'est-à-dire ne participant pas à la réaction). Si la solution est *suffisamment diluée,* c'est-à-dire si les concentrations molaires c_j (définies comme en (11) mais en prenant pour N le nombre de molécules du solvant) des divers réactifs restent suffisamment petites devant 1, leurs potentiels chimiques μ_j sont de la forme (cf. complément III.F, formule (F.47.b))

$$\mu_j = \psi_j(T, p) + kT \ln c_j. \tag{17}$$

Montrer qu'on peut à nouveau, dans ce cas, établir une loi d'action de masse portant sur les concentrations molaires.

III.24. Un liquide, à la température ordinaire, contient en solution des particules étrangères ayant la particularité de posséder un spin 1/2 et donc un moment magnétique $\vec{\mu}_0$, alors que les molécules du solvant ne sont pas (para)magnétiques.

On place cette solution dans l'entrefer d'un électroaimant créant un champ magnétique fortement inhomogène : sa plus grande composante, B_z, est une fonction linéairement croissante de z, prenant la valeur B_1 au fond du récipient ($z=0$) et la valeur $B_2 > B_1$ à la surface du liquide ($z=h$).

Quelle est la probabilité $w(z)\,dz$ qu'a chacune des particules magnétiques de se trouver à l'altitude z à dz près (dans l'un ou l'autre de ses deux états de spin possibles)? En déduire le rapport $n(h)/n(0)$, où $n(z)$ est la densité volumique d'impureté à l'altitude z.

Application numérique : pour $\mu_0 \simeq 10^{-23}$ J/T, $B_1 = 0$ et $B_2 = 5$ teslas, combien vaut $n(h)/n(0)$ à la température ordinaire? Est-il possible d'utiliser ce dispositif pour purifier le liquide?

III.25. Conditions de séparation d'un mélange dilué en deux phases (liquide et gaz). Cet exercice fait suite au complément III.F : il démontre les résultats énoncés au paragraphe III.2.c et résumés sur la figure 2; les notations sont les mêmes.

On fixe une fois pour toutes la pression à une valeur p_0; à cette pression, les phases liquide et gazeuse du solvant pur sont en équilibre à la température T_0. On choisit une température T voisine de T_0 et on va analyser le comportement du mélange lorsqu'on fait varier sa concentration *globale* c (en la maintenant très petite devant 1). Dans ces conditions (T et p fixées), l'état d'équilibre du système est celui qui correspond au *minimum de son enthalpie libre* G (complément III.D, § I.4). On posera pour simplifier

$$g(T, p_0, c) \widehat{=} \frac{1}{N} G(T, p_0, N, c)\,, \tag{1}$$

où N est le nombre de molécules du solvant.

1) Si le mélange est entièrement liquide, la fonction g s'écrit

$$g(T, p_0, c) = g_\ell(T, p_0, c)\,; \tag{2}$$

elle s'écrit

$$g(T, p_0, c) = g_g(T, p_0, c) \tag{3}$$

si le mélange est gazeux homogène. Dessiner l'allure des fonctions g_ℓ et g_g de c pour T et p_0 fixées (et $c \ll 1$).

2) Expliquer pourquoi, si ces deux courbes ne se coupent pas, le mélange reste homogène pour toute concentration $c(\ll 1)$. Montrer que l'abscisse du point d'intersection des deux courbes (s'il existe) est donnée par

$$c = \frac{\mu_{0\ell}(T, p_0) - \mu_{0g}(T, p_0)}{\psi_g(T, p_0) - \psi_\ell(T, p_0)}\,. \tag{4}$$

En utilisant le fait que l'entropie du solvant pur est plus grande en phase gazeuse qu'en phase liquide, montrer que le numérateur de l'expression (4) est positif pour $T > T_0$ et négatif pour $T < T_0$. En déduire que

(i) si $\psi_g(T_0, p_0) < \psi_\ell(T_0, p_0)$, le mélange reste gazeux homogène, pour toute concentration $c(\ll 1)$, dans la région $T > T_0$;

(ii) si $\psi_g(T_0, p_0) > \psi_\ell(T_0, p_0)$, il reste homogène liquide, quelle que soit $c(\ll 1)$, dans la région $T < T_0$.

3) On suppose maintenant T choisie de telle sorte que les deux courbes se coupent pour une valeur acceptable de c. Tracer leur tangente commune. Montrer que les abscisses des deux points de contact ne sont autres que c_ℓ et c_g, concentrations des deux phases en équilibre à la température T (et la pression p_0). Discuter l'état d'équilibre stable du système : si la concentration globale c est comprise entre c_ℓ et c_g, le mélange est séparé en deux phases; si c est extérieure à cet intervalle, il est homogène. Retrouver ainsi les résultats résumés sur la figure 2 du complément III.F.

SYSTÈMES CLASSIQUES

III.26. On considère un *oscillateur harmonique classique* à 3 dimensions, isotrope, et on veut discuter sur cet exemple les conditions d'application du *théorème d'équipartition de l'énergie.*

1) On choisit d'abord les *coordonnées cartésiennes.* La fonction de Hamilton du système s'écrit alors

$$\mathcal{H}(x, y, z; p_x, p_y, p_z) = \frac{1}{2m}(p_x^2 + p_y^2 + p_z^2) + \frac{1}{2} m\omega^2 (x^2 + y^2 + z^2) .$$

Montrer que chacun des six termes du second membre vérifie les conditions permettant de lui appliquer le théorème d'équipartition de l'énergie. En déduire que la valeur moyenne de l'énergie de l'oscillateur à la température T est $3kT$.

2) En *coordonnées sphériques,* la fonction de Hamilton prend la forme

$$\mathcal{H}(r, \theta, \varphi; p_r, p_\theta, p_\varphi) = \frac{p_r^2}{2m} + \frac{1}{2mr^2}\left(p_\theta^2 + \frac{1}{\sin^2\theta} p_\varphi^2\right) + \frac{1}{2} m\omega^2 r^2 .$$

a) Montrer que chacun des trois termes provenant de l'énergie cinétique vérifie encore les hypothèses du théorème. En déduire que la valeur moyenne de l'énergie cinétique vaut bien $3kT/2$ comme à la question 1.

b) Pourquoi la valeur moyenne du terme quadratique $m\omega^2 r^2/2$ n'est-elle pas donnée par le théorème d'équipartition de l'énergie?

3) On veut retrouver en coordonnées sphériques la valeur moyenne $3kT/2$ de l'énergie potentielle (*attention :* les intégrations sur l'espace des phases se font avec l'élément d'intégration $dr\, d\theta\, d\varphi\, dp_r\, dp_\theta\, dp_\varphi$). On écrit la fonction de Hamilton sous la forme

$$\mathcal{H} = \frac{1}{2m} p_r^2 + \frac{1}{2mr^2} \vec{L}^2 + \frac{1}{2} m\omega^2 r^2 .$$

a) Montrer, en intégrant par parties sur la variable r, que la valeur moyenne de $m\omega^2 r^2/2 = ar^2$ est reliée à celle de $\vec{L}^2/2mr^2 = b/r^2$.

b) Déduire du résultat obtenu que la valeur moyenne de $m\omega^2 r^2/2$ est effectivement égale à $3kT/2$.

III.27. Fonction de partition rotationnelle d'une molécule diatomique : calcul classique. On considère une molécule diatomique. Son orientation par rapport à un trièdre trirectangle de directions fixes attaché au centre de masse G est repérée par les angles polaires θ et φ (voir figure 3 du complément III.B). La rotation de la molécule peut être étudiée dans le cadre de la mécanique classique (§ II.3.a du complément III.B). Comme cette rotation s'effectue librement (pas d'énergie potentielle), la fonction de Lagrange $\mathcal{L}(\theta, \varphi; \dot{\theta}, \dot{\varphi})$ se réduit à l'énergie cinétique de rotation et s'écrit donc

$$\mathcal{L}(\theta, \varphi; \dot{\theta}, \dot{\varphi}) = \varepsilon^{(r)}(\theta, \varphi; \dot{\theta}, \dot{\varphi}) = \frac{1}{2} I(\dot{\theta}^2 + \dot{\varphi}^2 \sin^2\theta) ,$$

où I est le moment d'inertie de la molécule par rapport à G.

1) Calculer les moments conjugués p_θ et p_φ des variables θ et φ (formule (9) de l'appendice III). Exprimer l'énergie de rotation $\varepsilon^{(r)}$ en fonction des variables canoniques θ, φ, p_θ, p_φ, ce qui donne ici la fonction de Hamilton :

$$\mathcal{H}(\theta, \varphi; p_\theta, p_\varphi) = \varepsilon^{(r)}(\theta, \varphi; p_\theta, p_\varphi) .$$

2) Calculer, à partir de son expression en mécanique classique, la fonction de partition $\zeta_r(\beta)$ associée aux états de rotation de la molécule, et retrouver ainsi le résultat (B.48) du complément III.B.

3) Supposons maintenant que s'ajoute à l'énergie cinétique de rotation de la molécule une énergie potentielle $u(\theta, \varphi)$. Montrer que la probabilité $w(\theta, \varphi)\mathrm{d}\theta\ \mathrm{d}\varphi$ pour que la molécule pointe dans la direction (θ, φ), à $(\mathrm{d}\theta, \mathrm{d}\varphi)$ près, est proportionnelle à

$$\mathrm{e}^{-u(\theta, \varphi)/kT} \sin\theta\ \mathrm{d}\theta\ \mathrm{d}\varphi\ .$$

III.28. 1) On veut calculer, en mécanique classique, la fonction de partition rotationnelle d'une molécule polyatomique non linéaire, à partir des données de la note 22 du complément III.B. Cette fonction de partition s'écrit

$$\zeta_r = \frac{1}{\sigma}\frac{1}{(2\pi\hbar)^3}\int_0^{2\pi}\mathrm{d}\psi\int_0^{2\pi}\mathrm{d}\varphi\int_0^{\pi}\mathrm{d}\theta\int_{-\infty}^{+\infty}\mathrm{d}p_\psi\int_{-\infty}^{+\infty}\mathrm{d}p_\varphi\int_{-\infty}^{+\infty}\mathrm{d}p_\theta$$
$$\exp\left[-\beta\mathcal{H}(\theta, \varphi, \psi; p_\theta, p_\varphi, p_\psi)\right],$$

où $\mathcal{H}$ est donné par les formules (2) et (3) de la note citée.

On effectue d'abord l'intégration sur les trois moments conjugués, en utilisant le changement de variables défini par la formule (3). Montrer, en calculant le jacobien de cette transformation, que

$$\mathrm{d}p_\theta\ \mathrm{d}p_\varphi\ \mathrm{d}p_\psi = \frac{1}{\sin\theta}\,\mathrm{d}p_1\ \mathrm{d}p_2\ \mathrm{d}p_3\ .$$

Calculer ensuite les trois intégrales sur p_1, p_2, p_3, qui sont indépendantes.

Intégrer enfin le résultat obtenu sur θ, φ, ψ, ce qui donne la formule (B.64) du complément III.B.

2) Montrer que, si la fonction de Hamilton comporte en outre un terme d'énergie potentielle $u(\theta, \varphi, \psi)$, la probabilité $w(\theta, \varphi, \psi)\mathrm{d}\theta\ \mathrm{d}\varphi\ \mathrm{d}\psi$ pour que l'orientation de la molécule soit repérée par (θ, φ, ψ), à $(\mathrm{d}\theta, \mathrm{d}\varphi, \mathrm{d}\psi)$ près, est proportionnelle à

$$\mathrm{e}^{-u(\theta, \varphi, \psi)/kT} \sin\theta\ \mathrm{d}\theta\ \mathrm{d}\varphi\ \mathrm{d}\psi\ .$$

III.29. Distribution canonique et mouvement macroscopique. La démonstration que nous avons donnée de la distribution canonique suppose implicitement que le thermostat avec lequel le système échange de l'énergie est macroscopiquement au repos. On se propose ici d'établir cette distribution dans le cas où le thermostat est en mouvement macroscopique.

1) On considère un système $\mathcal{S}$ en contact thermique avec un thermostat $\mathcal{T}$. Le système global formé par la réunion de $\mathcal{S}$ et $\mathcal{T}$ est supposé *totalement* isolé et à l'équilibre statistique. Dans cette question, on suppose que le système global est simplement animé d'un mouvement de translation uniforme (sans rotation); l'énergie E_0 et la quantité de mouvement $\vec{p}_0$ de l'ensemble sont donc fixées :

$$E + E_{\mathcal{T}} = E_0\ ,$$
$$\vec{p}\ + \vec{p}_{\mathcal{T}} = \vec{p}_0\ .$$

La masse $M_{\mathcal{T}}$ du thermostat est bien supérieure à la masse M du système $\mathcal{S}$ et on négligera systématiquement dans toute la suite les termes non significatifs d'ordre $M/M_{\mathcal{T}}$. Enfin on rappelle (complément II.E, § I.1.c) que le nombre Ω d'états accessibles à un système d'énergie E et de quantité de mouvement $\vec{p}$ données ne dépend que de son « énergie interne » $E - \vec{p}^{\,2}/2M$.

a) Écrire le nombre d'états accessibles au système global $\mathcal{S} \cup \mathcal{T}$ lorsque l'énergie de $\mathcal{S}$ est E et sa quantité de mouvement $\vec{p}$. En déduire que les valeurs les plus probables de E et de $\vec{p}$ correspondent à l'égalité des températures microcanoniques de $\mathcal{S}$ et $\mathcal{T}$ et à l'égalité des vitesses de leurs centres de masse.

b) Montrer que la probabilité P_ℓ de trouver $\mathcal{S}$ dans un état microscopique (ℓ), d'énergie E_ℓ et de quantité de mouvement $\vec{p}_\ell$, quel que soit l'état de $\mathcal{C}$, est proportionnelle à

$$\Omega_{\mathcal{C}}\left(E_0-\frac{\vec{p}_0^{\,2}}{2M_{\mathcal{C}}}-E_\ell+\vec{v}_0\cdot\vec{p}_\ell\right),$$

où $\vec{v}_0=\vec{p}_0/M_{\mathcal{C}}$ est la vitesse d'ensemble du système global (ou, ce qui revient au même, du thermostat). En déduire que

$$P_\ell \propto \exp-[E_\ell-\vec{v}_0\cdot\vec{p}_\ell]/kT,$$

en précisant la signification de T.

2) On suppose maintenant que le centre de masse de $\mathcal{S}\cup\mathcal{C}$ est immobile $(\vec{p}_0=\vec{0})$, mais que le système global est animé d'un mouvement de rotation d'ensemble caractérisé par un moment cinétique $\vec{\mathcal{M}}_0$ fixé. On a alors

$$E+E_{\mathcal{C}}=E_0,$$
$$\vec{\mathcal{M}}+\vec{\mathcal{M}}_{\mathcal{C}}=\vec{\mathcal{M}}_0.$$

On rappelle (complément II.E, formule (E.37)) que les composantes du vecteur vitesse angulaire $\vec{\omega}$ d'un système d'entropie microcanonique $S^*(E, \vec{\mathcal{M}})$ sont données par

$$\frac{\omega_i}{T^*}=-\frac{\partial S^*(E,\vec{\mathcal{M}})}{\partial \mathcal{M}_i}, \quad i=x, y, z.$$

En adaptant à ce cas les raisonnements de 1, montrer

a) que la vitesse angulaire $\vec{\omega}_m$ la plus probable de $\mathcal{S}$ est égale à la vitesse angulaire $\vec{\omega}_0$ du thermostat.

b) que la probabilité P_ℓ de trouver $\mathcal{S}$ dans un état microscopique (ℓ), d'énergie E_ℓ et de moment cinétique $\vec{\mathcal{M}}_\ell$, est proportionnelle à

$$\exp-[E_\ell-\vec{\omega}_0\cdot\vec{\mathcal{M}}_\ell]/kT.$$

3) Application au cas d'un gaz parfait

a) Le gaz est en translation uniforme, son centre de masse ayant une vitesse $\vec{v}_0$ constante. Montrer que, à l'équilibre à la température T, la probabilité pour qu'une molécule du gaz soit animée de la vitesse $\vec{v}$ (à d^3v près) est proportionnelle à

$$\mathrm{d}^3v \exp\left[-m(\vec{v}-\vec{v}_0)^2/2kT\right].$$

b) Le gaz est animé d'un mouvement de rotation d'ensemble, de vitesse angulaire constante $\vec{\omega}_0$, autour d'un axe passant par son centre de masse O supposé immobile. Montrer que, à l'équilibre à la température T, la probabilité de trouver une molécule du gaz au point $\vec{r}\,(=\overrightarrow{OM})$ à d^3r près avec une vitesse $\vec{v}$ à d^3v près est proportionnelle à

$$\mathrm{d}^3r\,\mathrm{d}^3v \exp\left\{-m\left[(\vec{v}-\vec{\omega}_0\wedge\vec{r})^2-(\vec{\omega}_0\wedge\vec{r})^2\right]/2kT\right\}.$$

Interpréter chacun des deux termes de l'exponentielle et comparer avec la question 2 de l'exercice V.10.

SYSTÈMES DE PARTICULES EN INTERACTION

III.30. Un réseau unidimensionnel est constitué par une chaîne linéaire, fermée sur elle-même, de N atomes régulièrement espacés. Chaque atome possède un spin 1/2 qui interagit seulement avec ses deux plus proches voisins suivant l'hamiltonien (cf. complément III.J, § III.2)

$$H=-J\sum_{i=1}^{N} S_i^z S_{i+1}^z, \quad J>0.$$

Dans l'état fondamental, tous les spins pointent dans la même direction et l'énergie vaut $E_0 = -NJ/4$.

Considérons alors une situation dans laquelle n spins consécutifs ont été retournés.

1) Combien d'états du système réalisent une telle situation? Quelle est l'énergie de chacun de ces états?

2) Calculer l'énergie libre du système à la température T pour une valeur donnée de n.

3) Montrer que, dans la limite $N \to \infty$, la situation dans laquelle tous les spins pointent dans la même direction ne correspond jamais à un équilibre stable du système. En déduire qu'il n'existe pas de transition ferromagnétique pour un tel système unidimensionnel.

III.31. On considère le système constitué par N atomes assujettis à se déplacer sur une droite, dans un récipient unidimensionnel de longueur L. Chaque atome, de diamètre a, se comporte comme une « sphère dure » impénétrable aux autres particules. En dehors de cette restriction, qui les empêche en particulier de se croiser, les atomes se meuvent librement à l'intérieur du récipient, leur mouvement étant décrit par la mécanique classique.

1) Calculer la fonction de partition du système à la température T. On pourra résoudre d'abord le problème dans la limite $a = 0$, et passer ensuite au cas $a \neq 0$ en utilisant le changement de variables $x'_i = x_i - \left(i - \frac{1}{2}\right) a$, où x_i est le centre de l'atome (i).

2) En déduire l'équation d'état et la capacité calorifique du système.

III.32. Dans un gaz réel, l'énergie potentielle totale U est en première approximation la somme

$$U = \frac{1}{2} \sum_{i \neq j} u\left(\vec{r}_i - \vec{r}_j\right) \qquad i, j = 1, 2, 3, \ldots, N$$

des énergies d'interaction $u\left(\vec{r}_i - \vec{r}_j\right)$ des molécules prises deux à deux. Pour de nombreuses substances, u ne dépend que de la distance $r_{ij} = \left|\vec{r}_i - \vec{r}_j\right|$ entre les deux molécules et peut s'écrire sous la forme

$$u(r) = \varepsilon\, v(r/\sigma) ,$$

où v est une fonction universelle, indépendante de la nature du gaz; seules les deux constantes ε et σ varient d'une substance à l'autre.

1) Montrer que la fonction de partition d'un tel gaz, constitué de N molécules enfermées dans un récipient de volume V, est de la forme

$$Z(T, V, N) = Z_p(T, V, N)\, f\left(\frac{kT}{\varepsilon}, \frac{V}{\sigma^3}, N\right) ,$$

où Z_p est la fonction de partition qu'aurait le gaz si les molécules n'interagissaient pas et f une fonction elle aussi universelle.

2) En utilisant la propriété d'extensivité de l'énergie libre, montrer que $N^{-1} \ln f$ ne dépend que des deux variables intensives T et V/N.

3) En déduire que l'équation d'état d'un tel gaz est de la forme

$$\frac{pV}{NkT} = h\left(\frac{kT}{\varepsilon}, \frac{V}{N\sigma^3}\right) .$$

4) Montrer que les coordonnées T_c, p_c, V_c du point critique vérifient

$$\frac{p_c V_c}{N k T_c} = a \,,$$

où a est un nombre, le même pour toutes les substances considérées (cf. tableau I du complément III.G).

III.33. Potentiel dans une solution d'électrolyte fort. On considère une solution d'électrolyte fort contenant deux types d'ions de charges opposées $+e$ et $-e$, en nombre égal (n_0 ions de chaque type par unité de volume). On introduit dans la solution une plaque métallique plane portée au potentiel V_0. On se propose de déterminer comment le potentiel électrique $V(x)$ dans la solution varie avec la distance x à la plaque (on choisira l'origine des potentiels de sorte que $V(\infty)=0$).

1) Écrire l'équation de Poisson qui relie $V(x)$ aux densités de charge $\rho_+(x)$ et $\rho_-(x)$ des deux types d'ions.

2) L'énergie potentielle d'un ion de charge q placé dans le potentiel $V(x)$ étant $qV(x)$, montrer que la distribution de Boltzmann à la température T de la solution permet d'exprimer $\rho_+(x)$ et $\rho_-(x)$ en fonction de $V(x)$.

3) En combinant les résultats des questions 1 et 2, écrire l'équation différentielle que doit satisfaire $V(x)$. Résoudre cette équation dans le domaine(*) où $eV_0 \ll kT$. Tracer les courbes représentant $V(x)$, $\rho_+(x)$ et $\rho_-(x)$ à l'équilibre et montrer que la solution est perturbée seulement au voisinage immédiat de l'électrode, sur une distance que l'on déterminera et que l'on évaluera numériquement. Cette région perturbée s'appelle la « double couche diffuse » de Gouy et Chapman.

III.34. On considère un cristal constitué de N atomes portant chacun un spin 1/2 et interagissant suivant l'hamiltonien

$$H = -\frac{J}{2} \sum_{i,\,j} S_i^z S_j^z \,, \quad J>0 \,, \quad i,\, j = 1,\, 2,\, 3,\, \ldots,\, N \,,$$

où S_i^z est la composante suivant Oz du spin situé sur le site (i). Dans ce modèle d'Ising (complément III.J, § III) généralisé aux interactions à longue portée, les spins interagissent donc deux à deux indépendamment de la distance qui les sépare.

1) Montrer que l'énergie d'un état microscopique du système ne dépend que des nombres de spins n_+ et n_- pointant dans les deux directions de l'axe Oz.

2) Le cristal étant en équilibre thermique à la température T, quelle est la probabilité $P(n_+)$ pour que n_+ parmi les N spins pointent dans le sens positif de l'axe Oz?

3) Quelle est la valeur la plus probable de n_+? Montrer que le système subit une transition de phase d'un état ferromagnétique (aimantation non nulle) à un état paramagnétique (aimantation nulle), à une température T_c que l'on calculera.

4) Étudier les propriétés du système lorsqu'il est plongé dans un champ magnétique uniforme $\vec{B}_0$ parallèle à Oz.

(*) L'équation pour $V(x)$ peut être résolue exactement sans faire appel à l'approximation des hautes températures; on obtient alors le résultat de Gouy-Chapman (cf. J. Bockris et A. Reddy, *Modern Electrochemistry*, Plenum (1977), § 7.4.3 et § 7.4.4). Mais l'amélioration obtenue est illusoire (cf. complément III.H, note 6).

III.35. Étude d'un cristal antiferromagnétique. On considère un cristal cubique centré. A chaque site du réseau (sommet ou centre d'un cube) se trouve un atome portant un moment magnétique $\vec{\mu}$ dont la projection sur un axe Oz ne peut prendre que deux valeurs, $+\mu$ et $-\mu$. On suppose que l'interaction entre les moments magnétiques est limitée aux plus proches voisins : un atome appartenant au sous-réseau (a) formé par les sommets des cubes interagit seulement avec les 8 atomes voisins appartenant au sous-réseau (b) formé par les centres des cubes; de même, les atomes du sous-réseau (b) n'interagissent qu'avec leurs 8 voisins du sous-réseau (a). On suppose que, en présence d'un champ magnétique extérieur uniforme $\vec{B}_0$ (parallèle à Oz), l'hamiltonien du système est de la forme

$$H = -\vec{B}_0 \cdot \Big(\sum_{i \in (a)} \vec{\mu}_i + \sum_{j \in (b)} \vec{\mu}_j \Big) + J \sum_{\substack{i \in (a),\, j \in (b) \\ i,\, j \text{ p.v.}}} \vec{\mu}_i \cdot \vec{\mu}_j \,,$$

où J est une constante *positive.*

1) Le système étant en équilibre thermique à la température T, on suppose que la valeur moyenne du moment magnétique est la même pour tous les atomes d'un sous-réseau donné :

$$\overline{\vec{\mu}_{i \in (a)}} = \vec{\mu}_a \,, \qquad \overline{\vec{\mu}_{j \in (b)}} = \vec{\mu}_b \,,$$

$\vec{\mu}_a$ et $\vec{\mu}_b$ étant tous deux parallèles à Oz.

L'approximation de champ moyen consiste à remplacer dans H le terme $\vec{\mu}_i \cdot \vec{\mu}_j$ par $\vec{\mu}_i \cdot \vec{\mu}_b + \vec{\mu}_a \cdot \vec{\mu}_j - \vec{\mu}_a \cdot \vec{\mu}_b$ (complément III.J, note 18). Montrer que, dans cette approximation, l'hamiltonien du système peut s'écrire

$$H_{\text{c.m.}} = -\sum_{i \in (a)} \vec{\mu}_i \cdot \vec{B}_a^{\,\text{eff}} - \sum_{j \in (b)} \vec{\mu}_j \cdot \vec{B}_b^{\,\text{eff}} + H_0 \,,$$

où $\vec{B}_a^{\,\text{eff}}$ et $\vec{B}_b^{\,\text{eff}}$ sont des champs effectifs et H_0 une constante que l'on déterminera.

En traitant $\vec{B}_a^{\,\text{eff}}$ et $\vec{B}_b^{\,\text{eff}}$ comme des champs extérieurs fixés, calculer $\vec{\mu}_a$ et $\vec{\mu}_b$. En déduire le système d'équations autocohérent que doivent satisfaire μ_a et μ_b.

2) On se propose de résoudre ce système en l'absence de champ extérieur ($\vec{B}_0 = \vec{0}$).

a) Vérifier que ses solutions sont telles que

$$\mu_a = -\mu_b \mathrel{\widehat{=}} x \,,$$

et que x n'est différent de zéro que pour $T < T_N$, où T_N (appelée «température de Néel») est une constante que l'on déterminera.

b) Que vaut l'aimantation M du système pour $T < T_N$ et pour $T > T_N$?

c) Que vaut l'énergie du système pour $T < T_N$ et $T > T_N$? Calculer la capacité calorifique et étudier plus particulièrement son comportement au voisinage de la température de Néel.

3) On étudie maintenant l'effet d'un champ extérieur $\vec{B}_0$. Déterminer, dans l'approximation de champ moyen, la susceptibilité magnétique du système

$$\chi \mathrel{\widehat{=}} \lim_{B_0 \to 0} \frac{\partial M}{\partial B_0}$$

au voisinage de T_N, en distinguant les deux cas $T < T_N$ et $T > T_N$. Donner l'allure de la courbe $\chi(T)$.

III.36. Résultats exacts sur le modèle d'Ising. On considère un système de N spins 1/2 placés aux nœuds d'un réseau cristallin de volume V et plongés dans un champ magnétique extérieur B. Chaque spin interagit avec ses p plus proches voisins suivant l'hamiltonien d'Ising (complément III.J, § III.1)

$$H_I = -g\mu_B B \sum_{i=1}^{N} S_i^z - J \sum_{(i,j)_v} S_i^z S_j^z , \quad J>0 .$$

A $T=0$, même lorsque B tend vers zéro, l'aimantation prend sa valeur maximum. Dans la limite inverse $T \to \infty$, l'interaction entre les spins devient négligeable et le système se comporte comme un paramagnétique parfait de susceptibilité $\chi_0 \simeq N(g\mu_B)^2/4VkT$. On se propose ici de *calculer exactement* les premières corrections dans les deux domaines des très basses et des très hautes températures.

1) On étudie d'abord le domaine des basses températures ($T \ll T_c$).

a) Que vaut l'énergie E_0 de l'état fondamental du système? Cet état est-il dégénéré? Quel est, dans cet état, le moment magnétique M_0 du système?

b) Montrer que le premier niveau excité a pour énergie $E_1 = E_0 + g\mu_B B + pJ/2$. Déterminer sa dégénérescence ainsi que le moment magnétique associé M_1. Montrer que les autres niveaux excités sont séparés du premier par un écart énergétique toujours non nul (même pour $B=0$) et que le moment magnétique associé aux états correspondants est strictement inférieur à M_1.

c) Calculer, à la température T, la probabilité P_0 pour que le moment magnétique du système soit M_0 et la probabilité P_1 pour qu'il soit M_1. Montrer que, dans le domaine $T \ll T_c$, on peut négliger les probabilités des autres valeurs du moment magnétique. Calculer, dans ce domaine, le moment magnétique moyen, pour $B \to 0$, et en déduire que l'aimantation vaut alors

$$M(T) \simeq \frac{Ng\mu_B}{2V}\left(1-2e^{-pJ/2kT}\right) , \quad T \ll T_c .$$

Comparer avec le résultat obtenu dans le même domaine par l'approximation de champ moyen(*).

2) Considérons maintenant le domaine des hautes températures ($T \gg T_c$).

a) Montrer que la susceptibilité magnétique du système peut s'écrire

$$\chi(T) \widehat{=} \lim_{B \to 0} \frac{\partial M}{\partial B} = \frac{1}{VkT}(g\mu_B)^2 \overline{\left(\sum_i S_i^z\right)^2}_{B=0} ,$$

la moyenne canonique de $\left(\sum_i S_i^z\right)^2$ étant prise pour $B=0$. Pour obtenir la correction cherchée à χ_0, il nous suffit donc de calculer cette moyenne jusqu'à l'ordre $1/T$, en négligeant les ordres supérieurs.

b) Étudions tout d'abord la fonction de partition du système. Montrer qu'elle s'écrit

$$Z = \sum_{S_1^z=\pm\frac{1}{2}} \sum_{S_2^z=\pm\frac{1}{2}} \cdots \sum_{S_N^z=\pm\frac{1}{2}} \left(1+\frac{J}{kT}\sum_{(i,j)_v} S_i^z S_j^z\right) + \mathcal{O}\left(\frac{1}{T^2}\right) ,$$

c'est-à-dire

$$Z = 2^N + \mathcal{O}(1/T^2).$$

(*) La conclusion est tout à fait différente pour l'hamiltonien de Heisenberg (cf. complément III.J, note 31 et exercice VI.29).

c) Dans l'expression de $\overline{\left(\sum_i S_i^z\right)^2}$ intervient au numérateur la quantité

$$Q = \sum_{S_1^z=\pm\frac{1}{2}} \cdots \sum_{S_N=\pm\frac{1}{2}} \left(\sum_\ell S_\ell^z\right)^2 \left(1+\frac{J}{kT}\sum_{(i,\,j)_v} S_i^z S_j^z\right) + \mathcal{O}\left(\frac{1}{T^2}\right).$$

Montrer que

$$Q = 2^N \frac{N}{4}\left[1+\frac{pJ}{4kT}\right] + \mathcal{O}\left(\frac{1}{T^2}\right).$$

d) En déduire l'expression de $\chi(T)$ jusqu'au deuxième ordre en $1/T$ inclus. Comparer avec le résultat de l'approximation de champ moyen(*).

III.37. Soit $F(T, B)$ l'énergie libre par unité de volume d'un corps ferromagnétique à la température T, en présence du champ magnétique extérieur B.

1) Exprimer, à partir de F et de ses dérivées, la capacité calorifique à volume constant $C_V(T, B=0)$ et l'aimantation $M(T, B=0)$ en champ nul, ainsi que la susceptibilité magnétique $\chi \widehat{=} \lim_{B\to 0} \partial M/\partial B$.

2) Au voisinage de la température critique T_c, ces grandeurs ont un comportement singulier caractérisé par les « exposants critiques » α, β et γ :

$$C_V \propto |T-T_c|^{-\alpha}, \quad M \propto (T_c-T)^\beta, \quad \chi \propto |T-T_c|^{-\gamma}.$$

On suppose par ailleurs que, au voisinage de T_c, la partie singulière de F vérifie une « loi d'échelle » : quel que soit λ, on a

$$F\left(\frac{T-T_c}{T_c}, B\right) = \frac{1}{\lambda} F\left(\lambda^p \frac{T-T_c}{T_c}, \lambda^q B\right),$$

où p et q sont deux constantes. Exprimer α, β et γ en fonction de p et q (on pourra choisir λ tel que $\lambda^p(T-T_c)/T_c = 1$). En déduire la relation

$$\alpha + 2\beta + \gamma = 2.$$

Confronter ce résultat aux données du tableau III du complément III.J.

III.38. Application simple du « groupe de renormalisation »(**). La technique la plus puissante pour l'étude des systèmes présentant une transition de phase du second ordre est sans aucun doute la méthode du « groupe de renormalisation ». En général elle fournit des résultats concrets seulement au voisinage du point critique, où elle permet une excellente détermination des exposants critiques (cf. complément III.J, § IV). Nous nous proposons ici de l'appliquer au modèle d'Ising à une dimension (qui ne présente pas de transition de phase) dont elle donne la fonction de partition à toute température. On considère donc un système de N spins 1/2, régulièrement espacés sur une chaîne fermée sur elle-même (figure 15, complément III.J), et interagissant suivant l'hamiltonien

$$H = -J\sum_{i=1}^{N} S_i^z S_{i+1}^z, \quad J>0, \; N+1 \equiv 1.$$

(*) La méthode peut évidemment être étendue aux termes suivants du développement, mais les décomptes deviennent vite fastidieux.

(**) Cet exercice est inspiré de l'article de H.J. Maris et L.P. Kadanoff (*Am. J. Phys. 46,* 652 (1978)) auquel nous renvoyons pour l'application (plus intéressante, mais un peu plus élaborée) au système bidimensionnel.

1) Montrer que la fonction de partition de ce système peut s'écrire

$$Z = \sum_{\sigma_1=\pm1} \sum_{\sigma_2=\pm1} \cdots \sum_{\sigma_N=\pm1} e^{K(\sigma_1\sigma_2+\sigma_2\sigma_3)}\, e^{K(\sigma_3\sigma_4+\sigma_4\sigma_5)} \cdots ,$$

avec $K \widehat{=} J/4kT$ et $\sigma_i \widehat{=} 2S_i^z$,

ou encore, après sommation sur les σ d'indice pair,

$$Z = \sum_{\sigma_1=\pm1} \sum_{\sigma_3=\pm1} \cdots (e^{K(\sigma_1+\sigma_3)} + e^{-K(\sigma_1+\sigma_3)}) \times (e^{K(\sigma_3+\sigma_5)} + e^{-K(\sigma_3+\sigma_5)}) \times \cdots .$$

2) Montrer que, pour les valeurs permises à σ_1 et σ_3, on a toujours

$$e^{K(\sigma_1+\sigma_3)} + e^{-K(\sigma_1+\sigma_3)} = f(K)\, e^{K'\sigma_1\sigma_3} , \tag{1}$$

avec $K' \widehat{=} \frac{1}{2} \ln[\operatorname{ch} 2K]$, (2)

$$f(K) \widehat{=} 2 \operatorname{ch}^{\frac{1}{2}} 2K . \tag{3}$$

En déduire que

$$Z(N, K) = f(K)^{N/2} Z(N/2, K') : \tag{4}$$

le calcul de $Z(N, K)$ est ainsi ramené à celui de la fonction de partition pour un nombre de spins moitié et pour une force de couplage plus faible (ou une température plus élevée).

3) L'extensivité de l'énergie libre, pour $N \gg 1$, impose que

$$\ln Z(N, K) = N\zeta(K) . \tag{5}$$

Montrer que $\zeta(K)$ vérifie l'équation

$$\zeta(K) = \frac{1}{2} \ln f(K) + \frac{1}{2} \zeta(K') = \frac{1}{2}\left[\ln 2 + K' + \zeta(K')\right] , \tag{6}$$

et que celle-ci est satisfaite par la solution exacte du problème (complément III.J, formule (J.65)).

4) Si on ne connaît pas cette solution exacte, l'équation (6) permet de résoudre le problème de façon approchée. A très haute température en effet (ou à très faible valeur du couplage), les spins se comportent comme s'ils étaient pratiquement indépendants; autrement dit $Z(N, K \ll 1) \simeq 2^N$. Partant donc d'une température très élevée ($K' \ll 1$) où l'on connaît une bonne approximation de la fonction de partition, on déduit de (6) la fonction de partition pour une température plus basse (correspondant à la valeur K déduite de K' à l'aide de (2)); on itère le procédé (en prenant pour K' et $\zeta(K')$ les valeurs que l'on vient de trouver) et l'on peut ainsi calculer la fonction de partition à toute température.

En partant par exemple de $K' = 10^{-2}$ ou 10^{-3}, calculer numériquement quelques valeurs de $\zeta(K)$ et comparer avec celles que donne la solution exacte.

III.39. Transition ordre-désordre dans un alliage. On se propose de généraliser l'approximation de Bragg-Williams à l'étude de la transition ordre-désordre dans le cas d'un alliage binaire de concentration quelconque. Soit $x \widehat{=} N_A/N$ la concentration du composant A et $(1-x) = N_B/N$ celle de B (on prendra d'abord $x \leqslant 1/2$). Les notations sont celles du complément III.K.

Le réseau cristallin sur lequel sont placés les N atomes peut être décomposé en deux sous-réseaux identiques (α) et (β) imbriqués. Il y a $N/2$ sites α et $N/2$ sites β et chaque atome possède z plus proches voisins. Lorsque le système est ordonné, les atomes A sont préférentiellement situés sur les sites α et les atomes B sur les sites β.

1) On définit le paramètre d'ordre s par la relation

$$N_A^\alpha \mathrel{\widehat{=}} \frac{Nx}{2}(1+s) .$$

Écrire les expressions de N_A^β, N_B^α et N_B^β, et celles de N_{AA}, N_{AB} et N_{BB} en fonction de s.

Montrer que $s=1$ correspond à l'ordre maximum compatible avec la valeur de x choisie et $s=0$ au désordre maximum.

2) En adaptant la méthode exposée au complément III.K (§ II), déterminer, en fonction des énergies d'interaction entre plus proches voisins, la relation à laquelle doit satisfaire s lorsque le système est en équilibre à la température T.

3) Montrer que, pour une concentration x donnée, il existe une phase ordonnée lorsque la température T est inférieure à une température critique $T_c(x)$. Calculer $T_c(x)$ en fonction de T_c^0, température critique pour $x=1/2$.

Tracer la courbe $T_c(x)/T_c^0$ en fonction de x; la compléter pour $1/2 \leqslant x \leqslant 1$. Cette courbe sépare le plan (x, T) en deux régions : l'une où le système est ordonné, l'autre où il est désordonné. Identifier ces deux régions sur le diagramme.

III.40. Démixtion dans un alliage. Un alliage binaire contient Nx atomes A et $N(1-x)$ atomes B occupant les sites d'un réseau cristallin; chacun d'eux possède z plus proches voisins. On admet que seules sont importantes les interactions entre plus proches voisins et on note ε_{AA}, ε_{BB} et ε_{AB} les énergies correspondantes.

On suppose ici (contrairement à ce qui a été fait dans l'exercice précédent et au complément III.K) que les énergies d'interaction entre paires sont telles que

$$v \mathrel{\widehat{=}} \frac{1}{2}(\varepsilon_{AA}+\varepsilon_{BB})-\varepsilon_{AB} < 0 ,$$

de sorte que les atomes de même espèce ont tendance à se regrouper, au moins à basse température.

On se propose de montrer que, *au-dessus d'une certaine température critique* T_c, le système est dans un *état désordonné homogène,* et que, au-dessous de T_c, il est susceptible de se séparer en *deux phases distinctes* ayant des concentrations différentes (on dit alors qu'il y a « démixtion »).

1) On considère d'abord l'état homogène. On suppose valable l'approximation de champ moyen : la probabilité de trouver un atome A (ou B) sur un site donné ne dépend pas de la nature des atomes qui l'entourent.

a) Calculer le nombre moyen de paires de type AA, AB et BB. Que vaut l'énergie du système?

b) Calculer son entropie.

c) Montrer que l'énergie libre du système peut s'écrire

$$F_H(T, x) = E_0(x) - Nzvx(1-x) + NkT\left[x \ln x + (1-x)\ln(1-x)\right] ,$$

avec $$E_0(x) \mathrel{\widehat{=}} \frac{Nz}{2}\left[x\varepsilon_{AA} + (1-x)\varepsilon_{BB}\right] .$$

2) On suppose maintenant que le système, dont la composition globale est x_0, se sépare en deux phases de concentrations respectives x_1 et x_2. Soit $N\alpha$ le nombre d'atomes entrant dans la phase de composition x_1.

a) Quelle est la relation entre α, x_1, x_2 et x_0?

b) Montrer que l'énergie libre $F_D(T, x_0, x_1, x_2)$ du système formé par la juxtaposition des deux phases peut s'écrire

$$F_D(T, x_0, x_1, x_2) = \alpha F_H(T, x_1) + (1-\alpha) F_H(T, x_2) .$$

3) On veut maintenant rechercher l'état d'équilibre, à la température T, d'un système de concentration globale x_0 et déterminer notamment s'il reste homogène ou s'il se sépare en deux phases.

a) Quel est l'état d'équilibre du système si

$$\Delta F \widehat{=} F_D(T, x_0, x_1, x_2) - F_H(T, x_0)$$

est positif pour toutes les valeurs de x_1 et x_2?

b) On pose

$$\widehat{F}_H(T, x) \widehat{=} F_H(T, x) - E_0(x) ,$$

et $$\widehat{F}_D(T, x_0, x_1, x_2) \widehat{=} F_D(T, x_0, x_1, x_2) - E_0(x_0) .$$

On a donc

$$\Delta F = \widehat{F}_D(T, x_0, x_1, x_2) - \widehat{F}_H(T, x_0) .$$

Exprimer $\widehat{F}_D(T, x_0, x_1, x_2)$ en fonction de $\widehat{F}_H(T, x_1)$ et de $\widehat{F}_H(T, x_2)$. Vérifier que $\widehat{F}_D(T, x_0, x_1, x_2)$ est égale à l'ordonnée du point d'abscisse x_0 sur la droite passant par les points $(x_1, \widehat{F}_H(T, x_1))$ et $(x_2, \widehat{F}_H(T, x_2))$.

c) Étudier la variation de $\widehat{F}_H(T, x)$ avec x. Montrer que, si T est supérieure à une température critique T_c que l'on calculera, la courbe $\widehat{F}_H(T, x)$ ne possède qu'un minimum en $x = 1/2$. Montrer que pour $T < T_c$ la courbe possède trois extremums : un maximum en $x = 1/2$ et deux minimums dont on notera x_m et x'_m les positions (il peut être commode d'analyser la dérivée seconde de $\widehat{F}_H$). Tracer les courbes $\widehat{F}_H(T, x)$ en fonction de x pour différentes valeurs de la température T.

d) Montrer que, pour $T > T_c$, l'état d'équilibre est l'état homogène quel que soit x_0.

e) Montrer que, pour $T < T_c$, l'état stable est encore l'état homogène pour les faibles concentrations en A ou B tandis que, pour $x_m < x_0 < x'_m$, le système se sépare en deux phases de compositions respectives x_m et x'_m.

Donner l'allure du diagramme de phase (diagramme où $x_m(T)$ et $x'_m(T)$ sont représentées en portant T en *ordonnée* et la concentration en abscisse) et interpréter les différentes régions de ce diagramme.

DIVERS

III.41. Équation d'état d'un solide. Comme nous l'avons montré au complément III.E, l'étude d'un solide cristallin isolant de N atomes se ramène à celle d'un système de $3N$ oscillateurs harmoniques indépendants. Chaque oscillateur est caractérisé par un vecteur d'onde $\vec{k}$, une polarisation λ et une pulsation $\omega_\lambda(\vec{k})$. Nous noterons E_0 l'énergie du cristal dans son état fondamental.

1) Donner l'expression de l'énergie libre d'un tel cristal en équilibre à la température T.

2) E_0 est une fonction du volume V du cristal, de même que les fréquences propres de vibration(*).

(*) Dans un cristal parfaitement harmonique, les pulsations propres $\omega_\lambda(\vec{k})$ ne dépendent pas de V (on pourra le montrer facilement sur l'exemple du cristal unidimensionnel (complément III.E, § II.1)). C'est l'anharmonicité des vibrations qui conduit à une dépendance en V des $\omega_\lambda(\vec{k})$.

A une bonne approximation on peut écrire

$$\frac{\partial \ln \omega_\lambda(\vec{k})}{\partial \ln V} = -\gamma ,$$

où la «constante de Gruneisen» γ est pratiquement indépendante de $\vec{k}$ et de λ. En déduire que la pression du cristal s'écrit alors

$$p = -\frac{dE_0}{dV} + \gamma \frac{\overline{E} - E_0}{V} ,$$

où $\overline{E}$ est l'énergie moyenne du cristal à la température T. Cette relation est connue sous le nom d'«équation d'état de Mie-Gruneisen».

3) Montrer que le coefficient volumique de dilatation isobare

$$\alpha \widehat{=} \frac{1}{V}\left(\frac{\partial V}{\partial T}\right)_p ,$$

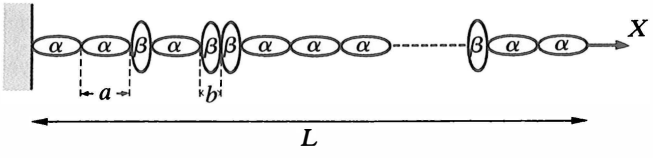
le module de compression

$$B \widehat{=} -V\left(\frac{\partial p}{\partial V}\right)_T$$

et la capacité calorifique par unité de volume C_V du cristal vérifient la relation

$$B\alpha = \gamma C_V .$$

III.42. On considère une chaîne unidimensionnelle constituée par $N(\gg 1)$ molécules identiques. Chaque molécule peut se trouver dans deux états, α et β, qu'on supposera de même énergie ε (on pourra prendre $\varepsilon = 0$). Une molécule dans l'état α apporte une contribution a à la longueur de la chaîne; quand elle est dans l'état β, cette contribution est b $(b < a)$ (voir figure)(*). La chaîne, en contact avec un thermostat à la température T, est soumise à une force de tension X.



1) Calculer la fonction de partition généralisée T-X (complément III.D, § II.1) de ce système. En déduire la longueur moyenne $\overline{L}$ de la chaîne à l'équilibre, en fonction de X et T.

2) Donner l'allure de la courbe $\overline{L} = f(X/T)$. On étudiera plus particulièrement le domaine $aX/kT \ll 1$ et on montrera que, dans ce domaine, la chaîne vérifie la loi de Hooke

$$X = \chi(T)[\overline{L} - L_0] .$$

Déterminer L_0 et $\chi(T)$.

3) Calculer l'énergie libre de la chaîne. Donner son expression en fonction de T et $\overline{L}$ dans le domaine $aX/kT \ll 1$.

(*) Ce modèle très simple schématise le comportement des chaînes de kératine présentes dans la laine.

III.43. La molécule d'éthane C_2H_6 est constituée de deux groupements CH_3 pyramidaux dont la liaison C—C joint les sommets. Pour repérer l'orientation relative de ces deux radicaux, on choisit dans chacun d'eux une liaison C—H particulière, et on donne l'angle α entre les deux plans C—C—H ainsi déterminés. Comme les atomes H se repoussent, l'énergie potentielle de la molécule est maximale lorsque les hydrogènes des deux radicaux se font exactement face, c'est-à-dire pour $\alpha = 0$ (ou $2\pi/3$, ou $4\pi/3$), et minimale lorsqu'ils s'imbriquent « en quinconce », c'est-à-dire pour $\alpha = \pi/3$ (ou π, ou $5\pi/3$). Soit Δ la différence d'énergie entre les maximums et les minimums ($\Delta \simeq 0{,}1\,\text{eV}$).

1) A basse température ($kT \ll \Delta$), la molécule oscille autour d'une de ses positions d'équilibre, disons $\alpha = \pi/3$; ce mouvement est l'un de ses modes normaux de vibration (torsion de la molécule autour de l'axe C—C).

Comment se comporte dans ce domaine la chaleur spécifique correspondante? Vers quelle valeur limite tendrait-elle à haute température si le comportement de la molécule restait vibratoire?

2) A haute température ($kT \gg \Delta$), les deux radicaux CH_3 peuvent tourner de façon pratiquement libre autour de la liaison C—C, de sorte que le degré de liberté considéré devient purement rotationnel. Quelle est la valeur limite de la chaleur spécifique correspondante à haute température?

3) Dans quel domaine de température se situe en pratique la transition entre les deux régimes? Esquisser l'allure des variations de la chaleur spécifique depuis les basses jusqu'aux hautes températures.

III.44. Les énergies d'excitation électroniques d'un atome d'hydrogène sont de la forme $\varepsilon_n = -Ry/n^2$ ($n = 1, 2, 3 \ldots$), où Ry est la constante de Rydberg; chaque niveau est $2n^2$ fois dégénéré. Montrer que ces états électroniques, si on les prend sans précaution, donnent une contribution infinie à la fonction de partition de l'atome, mais que le paradoxe se résout si l'on tient compte de la taille finie du récipient qui contient l'atome. On rappelle que le « rayon » de l'atome dans le niveau n est $r_n = n^2 a_0$. Montrer que, aux températures usuelles (ou inférieures), les excitations électroniques donnent une contribution tout à fait négligeable à la fonction de partition, et que l'atome reste pratiquement « gelé » dans son niveau électronique fondamental $n = 0$.

III.45. Soit un système dont les énergies E_ℓ associées aux états microscopiques (ℓ) dépendent d'un paramètre extérieur x. La force conjuguée X du paramètre x est une variable interne définie par la donnée de ses valeurs possibles

$$X_\ell \widehat{=} -\frac{\partial E_\ell}{\partial x} .$$

1) Montrer que, lorsque le système est en équilibre à la température T, la valeur moyenne $\overline{X}$ de la force X est donnée par

$$\overline{X} = kT \frac{\partial}{\partial x} \ln Z(T, x) .$$

2) Utiliser cette relation pour démontrer le résultat annoncé à la note 14 du complément III.D. *Attention :* dans cette note, X désigne la force *extérieure* imposée au système.

III.46. On considère un système *macroscopique* $\mathcal{S}$ en équilibre avec un thermostat $\mathcal{C}$. On note N le nombre de particules du système, $N_{\mathcal{C}}$ celui du thermostat. $\mathcal{C}$ est beaucoup plus gros que $\mathcal{S}$:

$$\frac{N}{N_{\mathcal{C}}} \ll 1 , \qquad (1)$$

mais pas suffisamment pour que $N^2/N_{\mathcal{C}}$ soit petit devant 1; on supposera quand même

$$\frac{N^3}{N_{\mathcal{C}}^2} \ll 1 \,. \tag{2}$$

On imagine d'abord, pour concrétiser les raisonnements, que le thermostat $\mathcal{C}$ est un gaz parfait monoatomique; on connaît alors la dépendance en énergie de sa densité d'états :

$$\rho_{\mathcal{C}}(E_{\mathcal{C}}) = A_{\mathcal{C}}(E_{\mathcal{C}})^{3N_{\mathcal{C}}/2} \,. \tag{3}$$

1) Montrer que la température microcanonique du thermostat est donnée par

$$T^*_{\mathcal{C}} = \frac{2}{3}\frac{E_{\mathcal{C}}}{N_{\mathcal{C}}k} \,. \tag{4}$$

2) Lorsqu'on développe l'entropie microcanonique du thermostat pour en déduire la probabilité canonique P^c_ℓ du système (formule III (I.14)), on ne peut *a priori* négliger que les termes petits devant 1, puisqu'il s'agira ensuite de prendre l'exponentielle du développement (cf. complément I.C, note 6). Montrer qu'il faut alors écrire

$$\frac{1}{k} S^*_{\mathcal{C}}(E_{\text{tot}} - E_\ell) \simeq \frac{1}{k} S^*_{\mathcal{C}}(E_{\text{tot}}) - \beta \mathrm{E}_\ell - \frac{1}{2}\gamma E_\ell^2 \,, \tag{5}$$

avec

$$\beta = \frac{3}{2}\frac{N_{\mathcal{C}}}{E_{\text{tot}}} \,, \qquad \gamma = \frac{3}{2}\frac{N_{\mathcal{C}}}{E_{\text{tot}}^2} = \frac{2}{3}\frac{\beta^2}{N_{\mathcal{C}}} \,. \tag{6}$$

La probabilité pour que le système $\mathcal{S}$ se trouve dans l'état (ℓ) vaut donc

$$P^c_\ell = \frac{1}{Z}\, \mathrm{e}^{-\beta E_\ell - \gamma E_\ell^2/2} \,. \tag{7}$$

3) Le système $\mathcal{S}$ étant macroscopique, sa fonction de partition Z peut être calculée en intégrant sur l'énergie (cf. formule III (II.2)) :

$$Z = \int_{E_0}^{+\infty} \mathrm{d}E \rho(E)\, \mathrm{e}^{-\beta E - \gamma E^2/2} \,. \tag{8}$$

En outre, cette intégrale peut être évaluée dans le cadre de l'approximation gaussienne (III, § I.C.1.b.β et § III.B.1).

Montrer que la valeur la plus probable E_m de l'énergie du système est donnée par l'équation

$$\frac{1}{k}\frac{\partial S^*}{\partial E}(E_m) - \beta - \gamma E_m = 0 \,, \tag{9}$$

où $S^*(E, V, N)$ est l'entropie microcanonique de $\mathcal{S}$ (pour faciliter les raisonnements, on pourra supposer dans un premier temps que $\mathcal{S}$ est aussi un gaz parfait monoatomique). En déduire que E_m est de l'ordre de N/β, avec un terme correctif en $N^2/N_{\mathcal{C}}\beta$ négligeable devant le précédent.

4) Montrer (cf. III, § III.B.1) que la fonction de partition s'écrit à une bonne approximation

$$Z \simeq \sqrt{2\pi(\Delta E)^2}\, \exp\left\{\frac{1}{k} S^*(E_m) - \beta E_m - \frac{1}{2}\gamma E_m^2\right\} \,, \tag{10}$$

avec

$$\frac{1}{(\Delta E)^2} = -\frac{1}{k}\frac{\partial^2 S^*}{\partial E^2}(E_m) + \gamma \,. \tag{11}$$

Peut-on négliger γ dans l'expression de ΔE? dans l'exponentielle?

5) Calculer l'énergie libre du système. Diffère-t-elle sensiblement de celle que l'on obtient lorsqu'on ne tient pas compte du troisième terme dans le développement (5)?

6) Généraliser la conclusion aux autres propriétés du système $\mathcal{S}$, puis au cas où ni le système ni le thermostat ne sont des gaz parfaits monoatomiques.

IV

I. L'ÉQUATION MAITRESSE

A. Transitions entre états microscopiques
   1. Origine physique des transitions
   2. Description quantitative : probabilités de transition

B. Système isolé. Le postulat d'évolution
   1. Démonstration heuristique de l'équation maîtresse
   2. Énoncé du postulat
   3. Discussion
      a. Est-il possible de démontrer le postulat d'évolution?
      b. L'équation maîtresse n'est pas invariante par renversement du temps

C. Équation maîtresse pour un système en contact avec un thermostat
   1. Établissement de l'équation
      a. Équation maîtresse pour l'ensemble système + thermostat
      b. Équation maîtresse pour le système seul
   2. Propriétés des coefficients

II. PRINCIPALES IMPLICATIONS DE L'ÉQUATION MAITRESSE

A. Propriétés de l'équilibre : la relation de « bilan détaillé »
   1. Solutions d'équilibre de l'équation maîtresse
   2. Système isolé : la distribution microcanonique
      a. La distribution microcanonique est une solution d'équilibre
      b. Elle vérifie en outre la relation de « bilan détaillé »
      c. C'est la seule solution d'équilibre
   3. Système en contact avec un thermostat : la distribution canonique
   4. Origine physique de la relation de « bilan détaillé »

B. Propriétés de l'évolution spontanée
   1. Évolution vers l'équilibre
      a. Propriété fondamentale de l'équation maîtresse
      b. La notion de temps de relaxation
   2. Système isolé : le théorème H
      a. Énoncé et démonstration du théorème
      b. Corollaire : l'équilibre comme maximum de l'entropie
   3. Système en contact avec un thermostat
      a. Version modifiée du théorème H
      b. Corollaire : l'équilibre comme minimum de l'énergie libre
   4. Irréversibilité de l'évolution spontanée
      a. Équation maîtresse et irréversibilité
      b. Analyse physique d'un exemple simple
      c. D'où vient l'irréversibilité?

CHAPITRE IV

Évolution vers l'équilibre et irréversibilité

Le postulat fondamental que nous avons énoncé au chapitre II concerne seulement l'état d'équilibre d'un système macroscopique isolé, état dans lequel les probabilités P_ℓ des divers états microscopiques (ℓ) sont indépendantes du temps. La mécanique statistique permet également d'étudier les systèmes hors d'équilibre, c'est-à-dire les *distributions de probabilités* $\{P_\ell(t)\}$ *dépendant du temps.* Ce sont les bases de cette étude que nous allons esquisser dans le présent chapitre : dans le reste du livre, nous nous intéressons presque exclusivement aux états d'équilibre; il est cependant important de connaître les lois fondamentales qui régissent l'évolution des systèmes hors d'équilibre, ne serait-ce que parce qu'elles permettent une compréhension plus profonde des distributions d'équilibre elles-mêmes.

La mécanique statistique hors d'équilibre est fondée sur un postulat général que nous appellerons *«postulat d'évolution»*. Son énoncé consiste à poser l'*«équation maîtresse»* donnant l'évolution dans le temps des probabilités $P_\ell(t)$[1]. Les états d'équilibre correspondent à des solutions particulières de l'équation maîtresse (probabilités P_ℓ indépendantes du temps), de sorte que le postulat microcanonique (chapitre II) apparaît ici comme une conséquence du postulat d'évolution.

I. L'ÉQUATION MAITRESSE

A. Transitions entre états microscopiques

Avant d'énoncer le postulat d'évolution, introduisons la notion fondamentale de probabilité de transition entre états microscopiques.

1. ORIGINE PHYSIQUE DES TRANSITIONS

Comme nous l'avons déjà indiqué, il est commode de choisir pour base, dans l'espace des états quantiques du système étudié, les états propres (ℓ) de l'hamiltonien H du système, et de considérer ces états (ℓ) comme les états microscopiques possibles du système; son état macroscopique est alors un mélange statistique d'états (ℓ), caractérisé par l'ensemble des probabilités P_ℓ pour que le système se trouve dans chacun des états microscopiques (ℓ) possibles.

Pour un système isolé, l'hamiltonien H est indépendant du temps et chacun de ses états propres (ℓ) est stationnaire, c'est-à-dire que ses propriétés restent invariables au cours du temps. Si donc nous connaissions parfaitement l'hamiltonien du système, les états (ℓ) seraient stationnaires, de sorte que si le système se trouvait, à un instant donné, dans l'un de ces états (ℓ), il y demeurerait indéfiniment, et ses propriétés seraient strictement indépendantes du temps.

Pour un système de taille macroscopique, la situation est radicalement différente : *le système passe constamment d'un état* (ℓ) *à un autre,* même s'il est isolé, même s'il est à l'équilibre du point de vue macroscopique. La raison en est que, *pour un grand système, l'hamiltonien n'est qu'imparfaitement défini.* Il est en effet

1. Les notions et raisonnements contenus dans ce chapitre restent à un niveau élémentaire. Certains résultats généraux sur les propriétés de l'équation maîtresse, admis sans démonstration dans le chapitre, sont discutés plus en détail dans les compléments.

impossible de répertorier de façon exhaustive et suffisamment précise tous les phénomènes qui se produisent au sein d'un système macroscopique ou qui peuvent l'influencer. Ainsi, lorsque nous délimitons un système macroscopique dit « isolé » (pour pouvoir restreindre l'analyse à ce système sans avoir à considérer en même temps le reste de l'univers), son isolement n'est jamais parfait : il existe toujours des forces, ayant leur origine à l'extérieur du système, qui sont certes trop faibles pour être prises en compte explicitement dans l'hamiltonien, mais qui induisent pourtant des transitions entre les états propres (ℓ) de cet hamiltonien. Des exemples de tels effets sont analysés dans le complément IV.B.

2. DESCRIPTION QUANTITATIVE : PROBABILITÉS DE TRANSITION

Comme toujours en mécanique statistique, c'est en termes probabilistes que l'on décrit les transitions entre états microscopiques : nous noterons

$$\mathcal{T}_{m\ell}(t_0,\ t_0+\delta t) \tag{I.1}$$

la probabilité pour que le système, partant de l'état (ℓ) au temps t_0, se trouve dans l'état (m) au temps $t_0+\delta t$; $\mathcal{T}_{m\ell}(t_0,\ t_0+\delta t)$ est appelée *probabilité de transition* de l'état (ℓ) vers l'état (m) entre les instants t_0 et $t_0+\delta t$. Nous préciserons dans le complément IV.D comment la mécanique quantique permet de relier ces probabilités de transition entre états microscopiques aux termes supplémentaires non contrôlables qu'ajoutent à l'hamiltonien du système les effets rappelés au paragraphe 1.

D'après sa définition même, la probabilité de transition $\mathcal{T}_{m\ell}$ entre deux états (ℓ) et (m) différents est nulle pour $\delta t = 0$. Mais l'important est de connaître l'ordre de grandeur de l'intervalle de temps δt nécessaire pour produire des probabilités de transition appréciables, et de le comparer à la durée d'observation τ_{exp} caractéristique des expériences que l'on veut effectuer sur le système. Un raisonnement d'ordres de grandeur sur les effets responsables des transitions montre facilement (complément IV.B) que *les temps correspondants sont très courts devant la durée d'observation* τ_{exp}; par conséquent, les transitions entre états microscopiques constituent un phénomène très rapide à l'échelle de temps qui nous intéresse.

Nous savons de plus qu'un système, lorsqu'il se trouve à l'équilibre macroscopique, a « perdu la mémoire » de la façon dont il y est parvenu, et du temps qui s'est écoulé depuis lors. Autrement dit, il n'existe pas dans ce cas d'origine des temps objective. Il est donc naturel de penser que *les probabilités de transition* $\mathcal{T}_{m\ell}(t_0,\ t_0+\delta t)$ *dépendent seulement de l'intervalle de temps* δt sur lequel on les évalue, et pas de l'instant initial t_0[2].

Soit alors un intervalle de temps $\mathrm{d}t$ suffisamment court pour pouvoir être considéré comme infinitésimal à l'échelle macroscopique, mais pourtant long devant le temps caractéristique des transitions entre états microscopiques. La probabilité de transition $\mathcal{T}_{m\ell}(t_0,\ t_0+\mathrm{d}t)$, nulle pour $\mathrm{d}t=0$ et indépendante de t_0, est de la forme

$$\mathcal{T}_{m\ell}(t_0,\ t_0+\mathrm{d}t) = a_{m\ell}\ \mathrm{d}t \quad \text{pour} \quad (m)\neq(\ell)\ , \tag{I.2}$$

où $a_{m\ell}$, dite *probabilité de transition par unité de temps* de l'état (ℓ) vers l'état (m), est *indépendante du temps* et bien sûr positive pour (m) différent de (ℓ). Ce sont ces probabilités par unité de temps $a_{m\ell}$ qui vont jouer le rôle central dans la suite de ce chapitre[3].

---

2. Cette propriété reste vraie lorsque le système n'est pas à l'équilibre macroscopique : les probabilités de transition entre *états microscopiques* sont les mêmes quel que soit l'*état macroscopique* du système.

3. On peut également définir la probabilité $\mathcal{T}_{\ell\ell}(t_0,\ t_0+\mathrm{d}t)$ pour que le système, se trouvant à l'instant t_0 dans l'état (ℓ), s'y trouve encore à l'instant $t_0+\mathrm{d}t$. Cette probabilité tend vers 1 lorsque $\mathrm{d}t \to 0$, de sorte qu'elle est de la forme

$$\mathcal{T}_{\ell\ell}(t_0,\ t_0+\mathrm{d}t) = 1 + a_{\ell\ell}\ \mathrm{d}t\ . \tag{1}$$

.../...

B. Système isolé. Le postulat d'évolution

Considérons un système macroscopique *isolé,* c'est-à-dire dont les échanges d'énergie, de particules, de volume, ... avec le reste de l'univers sont négligeables à l'échelle macroscopique. Nous désignons comme toujours par (ℓ) ses états microscopiques possibles, états propres de son hamiltonien H. A un instant t donné, son état macroscopique est caractérisé par l'ensemble des probabilités $P_\ell(t)$ pour qu'il se trouve dans chacun des états (ℓ) possibles. Nous noterons $a^*_{m\ell}$ la probabilité de transition par unité de temps de l'état (ℓ) vers l'état (m), l'étoile indiquant par convention que le système est isolé.

1. DÉMONSTRATION HEURISTIQUE DE L'ÉQUATION MAITRESSE

Choisissons un état (ℓ) déterminé et évaluons l'*accroissement* $\mathrm{d}P_\ell$ *de la probabilité* $P_\ell(t)$ correspondante, entre les instants t et $t+\mathrm{d}t$. Cet accroissement est la différence de deux termes : le premier est égal à la probabilité pour que, partant à l'instant t d'un état (m) quelconque, le système aboutisse dans l'état (ℓ) à l'instant $t+\mathrm{d}t$ (*« peuplement » de l'état* (ℓ)); le second, qui vient en déduction du précédent, est égal à la probabilité pour que le système quitte l'état (ℓ) pendant le temps $\mathrm{d}t$, vers l'un quelconque des états (m) (*« dépeuplement » de l'état* (ℓ)).

Ces deux termes sont faciles à calculer. En effet, la probabilité pour que le système se trouve dans l'état (m) à l'instant t puis dans l'état (ℓ) à l'instant $t+\mathrm{d}t$ est le produit

$$a^*_{\ell m}\ \mathrm{d}t\,.\,P_m(t) \tag{I.3}$$

de la probabilité $P_m(t)$ qu'il soit dans l'état (m) à l'instant t par la probabilité $a^*_{\ell m}\ \mathrm{d}t$ pour que, se trouvant dans l'état (m), il passe dans l'état (ℓ) au cours de l'intervalle $\mathrm{d}t$. Le terme de peuplement de l'état (ℓ) entre t et $t+\mathrm{d}t$ s'obtient donc en sommant l'expression (I.3) sur tous les états (m) possibles différents de (ℓ). De même, la probabilité pour que le système se trouve dans l'état (ℓ) à l'instant t puis dans l'état (m) à l'instant $t+\mathrm{d}t$ s'écrit

$$a^*_{m\ell}\ \mathrm{d}t\,.\,P_\ell(t)\ ; \tag{I.4}$$

sommée sur $(m)\neq(\ell)$, cette expression donne le dépeuplement total de l'état (ℓ) entre t et $t+\mathrm{d}t$. Finalement[4],

$$\mathrm{d}P_\ell = \sum_{(m)} a^*_{\ell m} P_m(t)\,\mathrm{d}t - \sum_{(m)} a^*_{m\ell} P_\ell(t)\,\mathrm{d}t\ . \tag{I.5}$$

2. ÉNONCÉ DU POSTULAT

Nous poserons comme *postulat d'évolution* le résultat que nous venons de déduire d'un bilan probabiliste simple : l'évolution dans le temps des probabilités

---

Si le système est dans l'état (ℓ) à l'instant t_0, la somme des probabilités pour que, durant l'intervalle de temps $\mathrm{d}t$, il passe dans un état (m) différent quelconque ou qu'il reste dans l'état (ℓ) doit être égale à 1. On en déduit

$$a_{\ell\ell} = -\sum_{(m)\neq(\ell)} a_{m\ell}\ ; \tag{2}$$

le coefficient $a_{\ell\ell}$ est évidemment négatif.

4. D'après le raisonnement, chacune des deux sommations de l'égalité (I.5) devrait porter sur l'ensemble des états (m) possibles, à l'exclusion de l'état (ℓ) lui-même. Mais les deux termes correspondant à $(m)=(\ell)$ se compensent évidemment, de sorte qu'on peut tout aussi bien les inclure (voir note 3).

d'occupation $P_\ell(t)$ des états microscopiques est donnée par l'*équation maîtresse*

$$\boxed{\frac{dP_\ell}{dt} = \sum_{(m)} \left[a^*_{\ell m} P_m(t) - a^*_{m\ell} P_\ell(t) \right]} \quad . \tag{I.6}$$

Nous compléterons ce postulat par l'hypothèse que les coefficients de l'équation maîtresse, qui sont, pour $(\ell) \neq (m)$, *positifs et indépendants du temps* puisque ce sont des probabilités de transition par unité de temps (§ A.2), sont en outre *symétriques*[5] :

$$\boxed{a^*_{m\ell} = a^*_{\ell m}} \tag{I.7.a}$$

et respectent enfin la *conservation de l'énergie* (c'est-à-dire qu'ils ne relient que des états (ℓ) et (m) de même énergie) :

$$\boxed{a^*_{\ell m} = 0 \quad \text{si} \quad |E_\ell - E_m| > \delta E} \quad , \tag{I.7.b}$$

où δE est comme précédemment (II, § I.B.1) l'incertitude affectant la donnée de l'énergie au niveau macroscopique.

3. DISCUSSION

a. Est-il possible de démontrer le postulat d'évolution?

On pourrait penser que les arguments du paragraphe 1 constituent une démonstration de l'équation maîtresse. Nous les avons effectivement présentés comme tels, et leur validité est d'ailleurs plus générale : toute une classe de processus aléatoires, dits *«processus markoviens stationnaires»*, sont régis par une équation maîtresse de la forme (I.6) (voir complément IV.D, § I). Cependant, la situation qui nous intéresse ici est moins simple : la mécanique quantique fait intervenir en premier lieu des *amplitudes de probabilité,* et non pas directement les probabilités elles-mêmes. C'est ce qui explique que l'équation maîtresse (I.6) ne découle pas directement et simplement des lois de la mécanique quantique.

Il est pourtant possible, moyennant certaines hypothèses supplémentaires, de démontrer à partir de la mécanique quantique l'équation maîtresse (I.6) et les propriétés (I.7) de ses coefficients dans le cas où le système est isolé (cf. complément IV.D). L'ensemble du « postulat d'évolution » que nous avons posé ici peut donc être considéré plutôt comme un résultat découlant de lois plus fondamentales que comme un véritable postulat. Cependant, les hypothèses supplémentaires sans lesquelles la démonstration ne serait pas possible ont une signification et une portée qu'il n'est pas facile de cerner de façon parfaitement claire. C'est la raison qui nous fait présenter ici les relations (I.6) et (I.7) comme constituant un postulat; bien entendu, on peut adopter le point de vue inverse (ces relations sont démontrées à partir de la mécanique quantique) sans rien changer aux développements qui suivent.

b. L'équation maîtresse n'est pas invariante par renversement du temps

L'équation maîtresse constitue en fait un *système d'équations différentielles* : la dérivée dP_ℓ/dt de la probabilité d'occupation d'un état (ℓ) particulier est donnée en fonction de $P_\ell(t)$ elle-même et de toutes les autres probabilités d'occupation $P_m(t)$ des états (m) différents de (ℓ); il y a une telle équation pour chacun des états (ℓ)

5. L'origine physique de cette propriété de symétrie sera précisée plus loin (§ II.A.4).

possibles. Ce système est susceptible de fournir les diverses probabilités P_ℓ en fonction du temps t si l'on connaît les coefficients $a^*_{\ell m}$ (c'est-à-dire les probabilités de transition par unité de temps entre les divers états microscopiques possibles) et les conditions initiales (c'est-à-dire la distribution $\{P_\ell(t_0)\}$ à un instant donné t_0).

Une caractéristique essentielle de ce système différentiel est que les équations sont du *premier ordre par rapport au temps*. Cette propriété a pour conséquence que *l'équation maîtresse n'est pas invariante par renversement du temps*. Supposons en effet que, ayant choisi une origine des temps $t=0$, nous effectuions dans l'équation (I.6) le changement de variable

$$t' = -t \,. \tag{I.8}$$

Ceci revient à inverser le sens du temps : pendant que t croît de 0 à $+\infty$, t' décroît de 0 à $-\infty$. Le changement de variable (I.8) transforme l'équation (I.6) en

$$-\frac{\mathrm{d}}{\mathrm{d}t'} P_\ell(-t') = \sum_{(m)} \left[a^*_{\ell m} P_m(-t') - a^*_{m\ell} P_\ell(-t') \right] . \tag{I.9}$$

A cause du changement de signe du premier membre, dû au fait qu'y figure une dérivée première par rapport au temps, l'équation (I.9) est différente de l'équation initiale (I.6) : plus précisément, si l'on pose

$$P'_\ell(t') \widehat{=} P_\ell(-t') \,, \tag{I.10}$$

l'ensemble des $P'_\ell(t')$ ne constitue pas une distribution possible puisque l'équation tirée de (I.9) qu'ils vérifient diffère de l'équation maîtresse.

Ce comportement est différent de celui des *équations fondamentales* régissant le mouvement du système *au niveau microscopique* qui sont, elles, *invariantes par renversement du temps*. Si par exemple le système peut être décrit par la mécanique classique, ces équations microscopiques découlent de l'équation fondamentale de la dynamique newtonienne

$$m \frac{\mathrm{d}^2 \vec{r}}{\mathrm{d}t^2} = \vec{F}(\vec{r}) \,, \tag{I.11}$$

que nous écrivons pour simplifier dans le cas d'une seule particule de masse m : $\vec{r}$ repère la position de cette particule, $\vec{F}$ désigne la force qui s'exerce sur elle, que nous supposons indépendante du temps. L'équation (I.11) étant du deuxième ordre par rapport au temps, le changement de variable (I.8) ne modifie pas ici le signe du premier membre :

$$m \frac{\mathrm{d}^2}{\mathrm{d}t'^2} \vec{r}(-t') = \vec{F}(\vec{r}) \,. \tag{I.12}$$

Posons alors

$$\vec{r}\,'(t') \widehat{=} \vec{r}(-t') \,. \tag{I.13}$$

La fonction $\vec{r}\,'(t')$ est donnée par une équation identique à l'équation fondamentale de la dynamique :

$$m \frac{\mathrm{d}^2}{\mathrm{d}t'^2} \vec{r}\,'(t') = \vec{F}(\vec{r}\,') \,. \tag{I.14}$$

Ainsi, à toute solution $\vec{r}(t)$ de l'équation initiale (I.11), qui décrit un mouvement possible de la particule, on peut associer une autre fonction $\vec{r}\,'(t')$, qui vérifie la même équation fondamentale et décrit donc aussi un mouvement possible du système; par rapport au mouvement précédent caractérisé par $\vec{r}(t)$, la trajectoire est la même mais elle est parcourue en sens inverse (vitesse opposée en chaque point)[6].

6. La figure 1 (§ II.B.4.b ci-après) donne un exemple simple et concret de mouvement ainsi inversé.

La question est un peu moins simple en mécanique quantique : l'équation fondamentale (équation de Schrödinger) est du premier ordre par rapport au temps, mais elle s'applique à une fonction d'onde à valeurs complexes, ce qui la rend elle aussi invariante par renversement du temps[7].

On se trouve donc confronté à un *problème fondamental* : les équations (classiques ou quantiques) qui régissent le mouvement du système à l'échelle microscopique sont invariantes par renversement du temps[8]; pourtant, l'évolution macroscopique du système, qui s'effectue selon l'équation maîtresse (I.6), ne possède

---

7. Dans le cas le plus simple (particule sans spin de masse m plongée dans un potentiel scalaire $U(\vec{r})$), l'équation de Schrödinger s'écrit

$$i\hbar \frac{\partial}{\partial t} \psi(\vec{r}, t) = H\psi(\vec{r}, t) , \tag{1}$$

où H est un opérateur *réel :*

$$H = -\frac{\hbar^2}{2m} \Delta + U(\vec{r}) . \tag{2}$$

Le changement de variable $t' = -t$ transforme l'équation (1) en

$$-i\hbar \frac{\partial}{\partial t'} \psi(\vec{r}, -t') = H\psi(\vec{r}, -t') , \tag{3}$$

qui diffère de l'équation initiale (1) par le signe du premier membre. Mais prenons le complexe conjugué des deux membres de l'équation (3) :

$$+i\hbar \frac{\partial}{\partial t'} \psi^*(\vec{r}, -t') = H\psi^*(\vec{r}, -t') , \tag{4}$$

où ψ^* désigne le complexe conjugué de ψ. Posons alors

$$\psi'(\vec{r}, t') \widehat{=} \psi^*(\vec{r}, -t') . \tag{5}$$

La nouvelle fonction d'onde $\psi'(r, t')$ vérifie une équation identique à l'équation de Schrödinger initiale :

$$i\hbar \frac{\partial}{\partial t'} \psi'(\vec{r}, t') = H\psi'(\vec{r}, t') . \tag{6}$$

Ainsi, à toute fonction d'onde $\psi(\vec{r}, t)$ solution de l'équation de Schrödinger (1) et représentant donc un mouvement possible de la particule, on peut associer une autre fonction d'onde $\psi'(\vec{r}, t')$ donnée par (5); cette nouvelle fonction d'onde $\psi'(\vec{r}, t')$ est solution de l'équation de Schrödinger (cf. (6)) et représente donc un autre mouvement possible de la particule. L'équation de Schrödinger est donc elle aussi invariante par renversement du temps.

8. On a découvert en 1964 en physique des particules, dans la désintégration des mésons K neutres, un effet provenant d'une interaction microscopique qui n'est *pas invariante* par renversement du temps. Cette interaction est extrêmement faible (son intensité est environ 500 fois moindre que celle des interactions faibles responsables du phénomène principal de désintégration des mésons K, mais le système K^0—$\overline{K}^0$ se comporte comme un interféromètre permettant de détecter des effets très fins), et son existence est totalement négligeable à l'échelle qui nous intéresse ici. On pense par contre qu'elle a joué un rôle important dans l'évolution de l'Univers primordial. En effet, selon les théories de « grande unification », quand les énergies mises en jeu dépassaient 10^{15} GeV (c'est-à-dire à moins de 10^{-36} s du « Big-Bang » initial), les interactions faibles, électromagnétiques et fortes avaient la même intensité et la violation de l'invariance par renversement du temps se manifestait de façon appréciable. Elle entre en particulier comme une composante essentielle dans l'explication (réf. 3, p. 28) de la dissymétrie observée dans l'Univers entre la matière et l'antimatière (l'Univers contient des protons, des neutrons et des électrons, mais pas ou très peu d'antiprotons, d'antineutrons et de positrons).

pas cette propriété. Nous analyserons plus concrètement ce problème au paragraphe II.B.4. Soulignons-en dès maintenant l'importance physique primordiale : la non invariance de l'équation maîtresse par renversement du temps explique que *l'évolution du système à l'échelle macroscopique est irréversible.*

C. Équation maîtresse pour un système en contact avec un thermostat

Ayant formulé le postulat d'évolution pour un système isolé, nous allons maintenant en déduire l'équation maîtresse régissant l'évolution d'un système en contact avec un thermostat.

1. ÉTABLISSEMENT DE L'ÉQUATION

Soit $\mathcal{S}$ le système considéré, dont nous désignons toujours par (ℓ) les états microscopiques; il est en contact avec un thermostat $\mathcal{T}$, de température T, dont nous notons (L) les états microscopiques.

a. Équation maîtresse pour l'ensemble système + thermostat

Le principe du raisonnement est le même qu'au chapitre III (§ I.B.1) : le système $\mathcal{S}$ lui-même n'est pas isolé, mais le système global $\mathcal{S} \cup \mathcal{T}$ l'est, et l'on peut donc lui appliquer le postulat d'évolution énoncé au paragraphe B.2.

A cause de la faiblesse de l'interaction assurant le contact thermique (II, § III.A.1.b), un état microscopique du système global $\mathcal{S} \cup \mathcal{T}$ est caractérisé par la donnée du couple (ℓ, L) formé par un état microscopique (ℓ) du système $\mathcal{S}$ et un état microscopique (L) du thermostat $\mathcal{T}$. L'énergie d'un tel état (ℓ, L) vaut simplement

$$E_{\ell, L} = E_\ell + E_L \,, \tag{I.15}$$

où E_ℓ est l'énergie de l'état (ℓ) du système et E_L celle de l'état (L) du thermostat. Appliqué au système global isolé $\mathcal{S} \cup \mathcal{T}$, le postulat d'évolution s'écrit

$$\frac{\mathrm{d}}{\mathrm{d}t} P_{\ell, L} = \sum_{(m),(M)} [a^*_{\ell, L; m, M} P_{m, M} - a^*_{m, M; \ell, L} P_{\ell, L}] \,, \tag{I.16}$$

avec
$$a^*_{\ell, L; m, M} = a^*_{m, M; \ell, L} \,, \tag{I.17.a}$$

$$a^*_{\ell, L; m, M} = 0 \quad \text{si} \quad |E_{\ell, L} - E_{m, M}| > \delta E \,. \tag{I.17.b}$$

b. Équation maîtresse pour le système seul

Seule nous intéresse en réalité la distribution statistique des états microscopiques du système $\mathcal{S}$. La probabilité pour que $\mathcal{S}$ se trouve dans l'état (ℓ), quel que soit par ailleurs l'état du thermostat, s'obtient simplement :

$$P_\ell = \sum_{(L)} P_{\ell, L} \,. \tag{I.18}$$

Or nous avons vu au chapitre III (§ II.D.1) que, par définition du thermostat, la probabilité conjointe $P_{\ell, L}$ se factorise sous la forme

$$P_{\ell, L} = P_\ell \,.\, P^*_L(E_{\mathrm{tot}} - E_\ell) \,, \tag{I.19}$$

où E_{tot} est l'énergie totale (fixée à δE près) du système global et $P^*_L(E_{\mathcal{T}})$ la *distribution microcanonique pour le thermostat.*

Il est dès lors facile de déduire de l'équation maîtresse initiale (I.16) l'équation d'évolution de la probabilité $P_\ell(t)$: d'après (I.18),

$$\frac{\mathrm{d}P_\ell}{\mathrm{d}t} = \sum_{(L)} \frac{\mathrm{d}}{\mathrm{d}t} P_{\ell, L} \,; \tag{I.20}$$

il suffit d'expliciter le second membre à l'aide de (I.16) et d'utiliser la forme (I.19) de $P_{\ell,L}$ pour trouver[9]

$$\boxed{\frac{dP_\ell}{dt}=\sum_{(m)}\left[a^T_{\ell m}P_m(t)-a^T_{m\ell}P_\ell(t)\right]}\ , \tag{I.21}$$

où l'on a posé

$$a^T_{m\ell}\widehat{=}\sum_{(L)}\sum_{(M)}a^*_{m,M;\ell,L}P^*_L(E_{\text{tot}}-E_\ell)\ . \tag{I.22}$$

2. PROPRIÉTÉS DES COEFFICIENTS

L'équation d'évolution (I.21) des probabilités d'occupation P_ℓ pour un système en contact avec un thermostat est une *équation maîtresse* semblable à celle qui a été postulée pour un système isolé : ses coefficients $a^T_{\ell m}$ sont *indépendants du temps* et *positifs* (pour $(\ell)\neq(m)$) d'après leur définition (I.22); ils peuvent donc être interprétés comme des probabilités de transition par unité de temps (§ A.2). Toutefois, leurs propriétés diffèrent de celles des $a^*_{\ell m}$.

D'une part, *ils relient des états* (ℓ) *et* (m) *d'énergie différente* : dans l'expression (I.22), les divers termes conservent tous l'énergie totale, c'est-à-dire qu'ils vérifient tous

$$E_m+E_M=E_\ell+E_L \quad \text{à } \delta E \text{ près ;} \tag{I.23}$$

mais cela autorise des énergies E_ℓ et E_m différentes pour le système $\mathcal{S}$ lui-même : le thermostat fournit ou encaisse les quantités de chaleur nécessaires à la conservation globale de l'énergie.

---

9. Compte tenu de (I.16), la formule (I.20) donne

$$\frac{dP_\ell}{dt}=\sum_{(L)}\sum_{(m)}\sum_{(M)}\left[a^*_{\ell,L;m,M}P_{m,M}-a^*_{m,M;\ell,L}P_{\ell,L}\right]\ . \tag{1}$$

Remplaçons au second membre les probabilités conjointes par leur expression factorisée (formule (I.19)) :

$$\frac{dP_\ell}{dt}=\sum_{(L)}\sum_{(m)}\sum_{(M)}\left[a^*_{\ell,L;m,M}\cdot P_m\cdot P^*_M(E_{\text{tot}}-E_m)-a^*_{m,M;\ell,L}\cdot P_\ell\cdot P^*_L(E_{\text{tot}}-E_\ell)\right]\ . \tag{2}$$

Dans le premier terme du second membre, P_m sort de la somme sur (L) et (M), ainsi que P_ℓ dans le deuxième terme :

$$\frac{dP_\ell}{dt}=\sum_{(m)}\left\{P_m\sum_{(L)}\sum_{(M)}a^*_{\ell,L;m,M}P^*_M(E_{\text{tot}}-E_m)-P_\ell\sum_{(L)}\sum_{(M)}a^*_{m,M;\ell,L}P^*_L(E_{\text{tot}}-E_\ell)\right\}\ . \tag{3}$$

Le facteur multipliant P_m possède alors deux indices libres, ℓ et m, et l'on peut poser

$$a^T_{\ell m}\widehat{=}\sum_{(L),(M)}a^*_{\ell,L;m,M}P^*_M(E_{\text{tot}}-E_m)\ ; \tag{4.a}$$

de même, le facteur multipliant P_ℓ dans le deuxième terme peut être noté

$$a^T_{m\ell}\widehat{=}\sum_{(L),(M)}a^*_{m,M;\ell,L}P^*_L(E_{\text{tot}}-E_\ell)\ . \tag{4.b}$$

Il est facile de s'assurer, en échangeant les rôles des indices muets L et M, que les formules (4.a) et (4.b) définissent bien le même ensemble de coefficients $a^T_{\ell m}$. Moyennant cette définition, l'équation (3) s'écrit aussitôt sous la forme (I.21).

D'autre part, on démontre facilement, en explicitant les P^*_L, que les coefficients $a^T_{\ell m}$ vérifient[10]

$$\boxed{a^T_{\ell m} e^{-E_m/kT} = a^T_{m\ell} e^{-E_\ell/kT}} \quad . \tag{I.24}$$

Cette relation remplace, pour les $a^T_{\ell m}$, la propriété de symétrie (I.7.a) des $a^*_{\ell m}$.

II. PRINCIPALES IMPLICATIONS DE L'ÉQUATION MAITRESSE

A. Propriétés de l'équilibre : la relation de « bilan détaillé »

Que le système étudié soit isolé ou en contact avec un thermostat, l'équation maîtresse qui régit son évolution permet de déterminer ses états d'équilibre.

---

10. On sait (II, § I.C.1) que la distribution $P^*_L(E_\mathcal{C})$ a pour expression

$$P^*_L(E_\mathcal{C}) = \exp\left[-\frac{1}{k} S^*_\mathcal{C}(E_\mathcal{C})\right] \quad \text{si} \quad E_\mathcal{C} \leqslant E_L \leqslant E_\mathcal{C} + \delta E$$

$$= 0 \qquad\qquad \text{sinon,} \tag{1}$$

où $S^*_\mathcal{C}(E_\mathcal{C})$ est l'entropie microcanonique du thermostat. D'autre part, la température T du thermostat étant par définition invariable, on a (formule III (I.14))

$$S^*_\mathcal{C}(E_{\text{tot}} - E_\ell) = S^*_\mathcal{C}(E_{\text{tot}}) - \frac{1}{T} E_\ell \, . \tag{2}$$

On en déduit aussitôt la forme du second facteur de la formule (I.19) :

$$P^*_L(E_{\text{tot}} - E_\ell) = A \exp\left[+\frac{E_\ell}{kT}\right] \quad \text{si} \quad E_{\text{tot}} - E_\ell \leqslant E_L \leqslant E_{\text{tot}} - E_\ell + \delta E$$

$$= 0 \qquad\qquad \text{sinon,} \tag{3}$$

où A est une constante de normalisation indépendante de (ℓ) et (L).

Les formules (4) de la note 9 s'écrivent alors

$$a^T_{\ell m} = \sum_{(L),(M)} a^*_{\ell, L; m, M} \cdot A e^{E_m/kT} \, , \tag{4.a}$$

$$a^T_{m\ell} = \sum_{(L),(M)} a^*_{m, M; \ell, L} \cdot A e^{E_\ell/kT} \, , \tag{4.b}$$

où les sommes portent sur les états accessibles au thermostat, compte tenu de l'énergie E_ℓ ou E_m du système. Elles impliquent donc

$$a^T_{\ell m} e^{-E_m/kT} = A \sum_{(L),(M)} a^*_{\ell, L; m, M} \, , \tag{5.a}$$

$$a^T_{m\ell} e^{-E_\ell/kT} = A \sum_{(L),(M)} a^*_{m, M; \ell, L} \, . \tag{5.b}$$

L'égalité de ces deux expressions découle immédiatement de la propriété de symétrie des coefficients $a^*_{\ell, L; m, M}$ (formule (I.17.a)).

Si le raisonnement et le résultat sont aussi simples c'est que, dans l'expression (I.22) des $a^T_{m\ell}$, la probabilité $P^*_L(E_{\text{tot}} - E_\ell)$ sort en réalité de la double somme sur les états du thermostat. Cette simplification provient du fait que le thermostat reste constamment à l'équilibre microcanonique dans lequel, pour une énergie donnée, ses divers états accessibles ont même probabilité.

1. SOLUTIONS D'ÉQUILIBRE DE L'ÉQUATION MAITRESSE

Considérons l'équation maîtresse

$$\frac{dP_\ell}{dt} = \sum_{(m)} \left[a_{\ell m} P_m(t) - a_{m\ell} P_\ell(t) \right], \tag{II.1}$$

dont les coefficients $a_{\ell m}$ sont indépendants du temps et positifs ou nuls pour $(\ell) \neq (m)$; ces coefficients sont symétriques si le système est isolé (formule (I.7.a)) ou vérifient la relation (I.24) si le système est en contact avec un thermostat.

Dans ce paragraphe A, nous nous restreignons aux solutions de l'équation maîtresse décrivant des *états d'équilibre* du système. Un tel état est caractérisé par une *distribution de probabilités indépendante du temps* $\{P_\ell^e\}$:

$$\frac{dP_\ell^e}{dt} = 0 \quad \text{quel que soit } (\ell). \tag{II.2}$$

Les solutions d'équilibre $\{P_\ell^e\}$ de l'équation maîtresse (II.1) sont donc données par les *équations d'équilibre* suivantes :

$$\sum_{(m)} [a_{\ell m} P_m^e - a_{m\ell} P_\ell^e] = 0 \quad \text{quel que soit } (\ell). \tag{II.3}$$

Ces équations s'interprètent de façon très simple : à l'équilibre, la probabilité P_ℓ^e d'un état (ℓ) donné est indépendante du temps non pas parce que le système reste constamment dans le même état microscopique, mais parce que le « peuplement » total de l'état (ℓ) pendant l'intervalle de temps dt, qui vaut

$$\sum_{(m)\neq(\ell)} a_{\ell m} P_m^e \, dt, \tag{II.4.a}$$

est exactement compensé par son « dépeuplement » total pendant le même intervalle de temps,

$$\sum_{(m)\neq(\ell)} a_{m\ell} P_\ell^e \, dt. \tag{II.4.b}$$

Pour trouver les distributions d'équilibre $\{P_\ell^e\}$ possibles, il faut donc en principe résoudre le système d'équations (II.3), qui comprend une équation pour chaque état (ℓ), c'est-à-dire en général une infinité dénombrable d'équations. C'est un système linéaire et homogène. Il ne faut cependant pas perdre de vue qu'une solution de ce système ne nous intéresse que si elle vérifie les conditions

$$P_\ell^e \geqslant 0 \quad \text{pour tout } (\ell), \tag{II.5.a}$$

$$\sum_{(\ell)} P_\ell^e = 1, \tag{II.5.b}$$

indispensables pour que $\{P_\ell^e\}$ soit une distribution de probabilités.

Nous allons montrer que l'équation maîtresse admet une seule solution indépendante du temps : la distribution microcanonique pour un système isolé, la distribution canonique pour un système en contact avec un thermostat.

2. SYSTÈME ISOLÉ : LA DISTRIBUTION MICROCANONIQUE

a. La distribution microcanonique est une solution d'équilibre

Considérons un système isolé, dont l'énergie est fixée et égale à E (à δE près). La distribution microcanonique correspondante[11]

$$P_\ell^*(E) = \frac{1}{\Omega(E)} \quad \text{si} \quad E \leqslant E_\ell \leqslant E + \delta E$$
$$= 0 \qquad \text{pour les autres états } (\ell) \tag{II.6}$$

11. Nous écartons d'emblée les états microscopiques qui ne vérifient pas les autres contraintes extérieures.

est solution des équations d'équilibre adaptées à ce cas :

$$\sum_{(m)} [a^*_{\ell m} P^*_m - a^*_{m\ell} P^*_\ell] = 0 \quad \text{pour tout } (\ell). \tag{II.7}$$

Distinguons en effet, comme dans (II.6), les états accessibles du système (II, § I.B.1.a) des autres états microscopiques. Choisissons d'abord *un état* (ℓ) *qui n'est pas accessible.* Sa probabilité P^*_ℓ est nulle d'après (II.6), ce qui annule directement le second terme figurant dans le crochet de la formule (II.7). Dans la somme sur (m) qui reste alors (premier terme de (II.7)), les seules contributions non nulles pourraient venir des états (m) accessibles, qui sont les seuls à avoir une probabilité P^*_m différente de zéro. Mais les coefficients $a^*_{\ell m}$ correspondants sont nuls d'après la propriété (I.7.b) (conservation de l'énergie). L'égalité (II.7) est donc vérifiée pour tout (ℓ) non accessible.

Prenons maintenant *un état* (ℓ) *accessible.* Toujours à cause de la propriété (I.7.b), la somme sur (m) se réduit aux seuls états (m) accessibles. Mais alors tous les P^*_m sont égaux entre eux d'après (II.6), et P^*_ℓ leur est aussi égal. La condition (II.7) s'écrit par conséquent

$$\frac{1}{\Omega} \sum_{(m)} [a^*_{\ell m} - a^*_{m\ell}] = 0 \quad \text{pour tout } (\ell) \text{ accessible.} \tag{II.8}$$

Elle est vérifiée ici aussi, compte tenu de la symétrie des coefficients $a^*_{\ell m}$ (formule (I.7.a)).

b. Elle vérifie en outre la relation de « bilan détaillé »

Il est intéressant d'examiner de plus près la manière dont la distribution microcanonique vérifie les équations d'équilibre.

Nous pouvons nous limiter aux états accessibles, puisque ce sont les seuls à contribuer de façon non triviale dans les égalités (II.7). On constate alors que, pour chaque état (ℓ), la somme sur (m) figurant au premier membre de (II.7) s'annule parce que *chacun des termes est séparément nul* :

$$a^*_{\ell m} P^*_m = a^*_{m\ell} P^*_\ell \quad \text{pour tout } (\ell) \text{ et tout } (m), \tag{II.9}$$

puisque P^*_m et P^*_ℓ sont égales et les coefficients $a^*_{\ell m}$ symétriques. Insistons sur le fait que *les équations d'équilibre* (II.3) *demandent seulement l'annulation,* pour un état (ℓ) donné, *d'une somme* sur tous les autres états (m); la distribution microcanonique vérifie donc cette condition de façon très particulière et inattendue.

L'égalité (II.9) porte le nom de *relation du « bilan détaillé »,* et s'interprète comme suit. Nous avons indiqué plus haut que l'équilibre statistique résulte de la compensation, durant chaque intervalle de temps $\mathrm{d}t$, entre le peuplement total d'un état (ℓ) donné à partir de tous les autres états (m) et son dépeuplement total vers ces autres états (m). Ce bilan s'équilibre en fait, dans la distribution microcanonique, de manière plus « détaillée » : la relation (II.9) exprime que la probabilité $a^*_{\ell m} P^*_m \, \mathrm{d}t$ pour que le système se trouve dans l'état (m) à l'instant t puis dans l'état (ℓ) à l'instant $t + \mathrm{d}t$ est égale à la probabilité pour qu'il se trouve dans l'état (ℓ) à l'instant t puis dans l'état (m) à l'instant $t + \mathrm{d}t$; ainsi, *le peuplement de l'état* (ℓ) *à partir d'un état* (m) *déterminé est exactement compensé par son dépeuplement vers ce même état* (m).

Cette interprétation de la relation du bilan détaillé peut être exprimée dans le langage des ensembles statistiques. L'*ensemble microcanonique* (II, § I.B.2) comporte à chaque instant autant de systèmes qui se trouvent dans l'état (ℓ) que dans l'état (m) (ces deux états étant supposés accessibles). Tous les systèmes de l'ensemble changent continuellement d'état microscopique (voir paragraphe I.A.1), mais, pendant un intervalle de temps quelconque, il y en a en moyenne *autant qui passent de l'état* (ℓ) *à l'état* (m) *que de l'état* (m) *à l'état* (ℓ).

c. C'est la seule solution d'équilibre

La distribution microcanonique est une solution d'équilibre de l'équation maîtresse pour un système isolé. On peut montrer que cette solution d'équilibre est *unique,* pour des valeurs données des paramètres extérieurs, en particulier de l'énergie. L'argument décisif en faveur de l'unicité de la solution d'équilibre est fourni par le *« théorème H »,* que nous démontrerons plus loin (§ B.2).

Lorsqu'on examine le problème plus en détail (complément IV.C), il apparaît pourtant une *condition d'unicité* : il faut que les probabilités de transition $a^*_{\ell m}$ fassent communiquer entre eux tous les états permis par la valeur de l'énergie et des autres paramètres extérieurs. Si par exemple ces états *a priori* permis se divisent en deux sous-ensembles séparés, tels que le système ne puisse pas passer d'un état de l'un des sous-ensembles à un état de l'autre, on aura deux solutions d'équilibre possibles. Mais une telle situation ne se produit en pratique que si une loi de conservation interdit le passage entre les deux sous-ensembles, ce qui revient à dire qu'ils correspondent à deux valeurs différentes d'un paramètre extérieur. Quoi qu'il en soit, la propriété des probabilités de transition qui garantit l'unicité de la solution d'équilibre est la même qui assure la validité du principe ergodique (I, § IV.B.3). Nous la supposerons vérifiée dans la suite.

3. SYSTÈME EN CONTACT AVEC UN THERMOSTAT : LA DISTRIBUTION CANONIQUE

Lorsque le système étudié est en contact avec un thermostat de température T, les coefficients $a^T_{\ell m}$ des équations d'équilibre (II.3) vérifient la relation de pseudo-symétrie (I.24). Dans ce cas, la distribution canonique

$$P^c_\ell = \frac{1}{Z}\, e^{-E_\ell/kT} \tag{II.10}$$

est solution de ces équations d'équilibre. En effet, la relation (I.24) peut être écrite

$$a^T_{\ell m} P^c_m = a^T_{m\ell} P^c_\ell \,, \tag{II.11}$$

ce qui assure de nouveau que, dans l'équation (II.3) relative à un état (ℓ) particulier, chacun des termes de la somme du premier membre est séparément nul. Autrement dit, *la distribution canonique est une solution d'équilibre de l'équation maîtresse* correspondant au cas où le système est en contact avec un thermostat, mais elle satisfait en plus la *relation du bilan détaillé* (II.11).

Celle-ci s'interprète de la même façon qu'au paragraphe 2.b : la probabilité pour que le système se trouve dans l'état (ℓ) à l'instant t puis dans l'état (m) à l'instant $t + dt$ est égale à la probabilité pour qu'il se trouve dans l'état (m) à l'instant t puis dans l'état (ℓ) à l'instant $t + dt$. Dans l'*ensemble canonique,* tous les états microscopiques ne sont pas représentés avec la même fréquence : le nombre de systèmes de l'ensemble qui sont, à chaque instant, dans l'état (ℓ) est proportionnel à $\exp[-E_\ell/kT]$. La constance de ce nombre dans le temps est assurée de façon « détaillée » : pendant un intervalle de temps donné, il y a en moyenne *autant de systèmes qui quittent l'état (ℓ) pour l'état (m) que de systèmes qui effectuent la transition inverse*[12].

---

12. Il est clair que la relation de bilan détaillé doit être vérifiée dans le cas présent si elle l'est pour un système isolé. Dans l'ensemble isolé système + thermostat, en effet, cette relation est valable pour tout couple de transitions inverses

$$(\ell, L) \leftrightarrows (m, M) \,; \tag{1}$$

elle le sera donc également pour les transitions correspondantes

$$(\ell) \leftrightarrows (m) \tag{2}$$

à l'intérieur du système, lorsqu'on aura sommé sur les états (L) et (M) du thermostat.

Ici aussi, la distribution canonique est la *seule solution d'équilibre* de l'équation maîtresse lorsque la température est fixée[13].

4. ORIGINE PHYSIQUE DE LA RELATION DE « BILAN DÉTAILLÉ »

Contrairement aux équations du mouvement microscopiques, l'équation maîtresse n'est pas invariante par renversement du temps (§ I.B.3.b). On peut cependant montrer que *l'invariance des équations microscopiques par renversement du temps se manifeste par la symétrie des coefficients* $a^*_{\ell m}$ *de l'équation maîtresse* relative à un système isolé. Or il est clair d'après les raisonnements des paragraphes 2 et 3 que la relation de bilan détaillé provient de la symétrie des $a^*_{\ell m}$. Ainsi, *le bilan détaillé a son origine dans l'invariance par renversement du temps*[14].

Ceci laisse bien entendu ouverte la question fondamentale posée par la perte de l'invariance dans le renversement du temps lors du passage des équations du mouvement microscopique à l'équation maîtresse (voir paragraphe B.4 et complément IV.D).

B. Propriétés de l'évolution spontanée

1. ÉVOLUTION VERS L'ÉQUILIBRE

Analysons les principales propriétés de l'évolution d'un système macroscopique vers son état d'équilibre. Reprenons pour cela l'équation maîtresse complète (II.1).

a. Propriété fondamentale de l'équation maîtresse

Nous avons déjà souligné que l'équation maîtresse est du premier ordre par rapport au temps. Cela implique en particulier qu'il suffit de se donner la distribution de probabilité $\{P_\ell(t_0)\}$ à un instant initial t_0 quelconque pour déterminer de façon unique la solution $\{P_\ell(t)\}$ correspondante de l'équation maîtresse.

On démontre alors (cf. complément IV.C) le résultat fondamental suivant : *la solution de l'équation maîtresse correspondant à une distribution initiale quelconque tend, lorsque* $t \to \infty$, *vers la solution d'équilibre*[15] :

$$P_\ell(t) \underset{t \to +\infty}{\longrightarrow} P^e_\ell \,. \tag{II.12}$$

Cette propriété fonde théoriquement le principe que nous avons constamment utilisé jusqu'ici, souvent de façon implicite : abandonné à lui-même, *un système macroscopique évolue spontanément vers son état d'équilibre* (macroscopique), qu'il finit par atteindre si l'on attend suffisamment longtemps.

---

13. Pour chaque énergie totale E, la solution d'équilibre est unique en ce qui concerne le système global $\mathcal{S} \cup \mathcal{C}$ (§ 2.c). A partir de là, le raisonnement du chapitre III (§ I.B.1) donne de manière univoque la distribution canonique comme distribution d'équilibre du système $\mathcal{S}$.

14. Dans l'évolution de l'Univers primordial sont intervenues, pense-t-on, des interactions microscopiques non invariantes dans le renversement du temps (cf. note 8). Dans ce cas, les coefficients $a^*_{\ell m}$ de l'équation maîtresse ne sont pas symétriques et la relation du bilan détaillé n'est plus vérifiée. On peut cependant montrer (réf. 3, p. 31) qu'il y a quand même évolution vers l'équilibre, et que celui-ci est caractérisé par les mêmes distributions qu'ici.

15. Ce résultat est valable pour un système isolé, auquel cas la distribution d'équilibre est la distribution microcanonique $\{P^*_\ell\}$, ou pour un système en contact avec un thermostat, auquel cas $\{P^e_\ell\}$ est la distribution canonique $\{P^c_\ell\}$. Comme dans les notes 12 et 13, il suffit que la propriété soit vraie pour tout système isolé pour qu'on puisse l'étendre de façon simple aux systèmes en contact avec un thermostat.

b. La notion de temps de relaxation

α. Définition

La propriété fondamentale que nous venons d'énoncer concerne la distribution asymptotique du système, atteinte au bout d'un temps « suffisamment long ». Au niveau pratique, il ne suffit pas de savoir qu'un système donné va évoluer vers tel état d'équilibre connu, il est essentiel de connaître également *l'échelle de temps* qui régit cette évolution. On la caractérise par ce qu'on appelle le *temps de relaxation* τ_r. Supposons que, à l'instant t_0, le système se trouve dans un état macroscopique hors d'équilibre caractérisé par une distribution de probabilité $\{P_\ell(t_0)\}$. Le temps de relaxation τ_r est l'intervalle de temps, donné en ordre de grandeur, tel que la formule mathématique (II.12) signifie en pratique[16]

$$P_\ell(t) = P_\ell^e \quad \text{pour} \quad (t - t_0) \gg \tau_r \,. \tag{II.13}$$

Les propriétés que nous avons analysées jusqu'ici s'appuyaient uniquement sur des caractéristiques générales des coefficients $a_{\ell m}$ de l'équation maîtresse : symétrie, conservation de l'énergie. Le temps de relaxation, au contraire, dépend des *valeurs* de ces coefficients. Or l'origine même des transitions entre états microscopiques (§ I.A.1) interdit le plus souvent un calcul précis des probabilités de transition $a_{\ell m}$. Pour évaluer le temps de relaxation, on fait appel généralement à des *modèles physiques simplifiés,* décrivant le système de façon approximative à l'échelle macroscopique; on peut également considérer τ_r comme un *paramètre phénoménologique* dont la valeur est déterminée empiriquement.

β. Discussion physique

Le temps de relaxation dépend d'un grand nombre de facteurs variés : nature du système, température, état initial (hors d'équilibre)... Il peut être fortement influencé par la présence de corps étrangers (impuretés ou poussières), par la nature des parois du récipient contenant le système, la manière dont on le manipule concrètement[17]... De façon générale, pour des systèmes homogènes de même nature, le temps de relaxation augmente rapidement avec leur taille; en effet l'équilibre s'établit d'abord localement, dans des sous-systèmes de petite taille; à ce stade intermédiaire, les propriétés du système varient d'un point à l'autre, c'est-à-dire d'un sous-système à l'autre; puis les échanges et interactions entre sous-systèmes voisins aboutissent finalement à l'équilibre global du système; celui-ci est donc atteint d'autant plus lentement que le système est plus grand.

L'ordre de grandeur des temps de relaxation varie dans des proportions énormes suivant le système considéré et sa situation. Certains sont extrêmement courts : si par exemple on applique brutalement un champ électrique à un échantillon de métal, les électrons de conduction se redistribuent à sa surface pour annuler le champ total à l'intérieur de l'échantillon (équilibre électrostatique d'un conducteur); pour les métaux courants, le temps de relaxation correspondant est inférieur à 10^{-12} s. Certains autres

16. Nous verrons dans le complément IV.C que l'approche de l'équilibre se fait dans la plupart des cas selon la fonction exponentielle. On peut alors définir le temps de relaxation τ_r de manière plus précise par la formule

$$\left| P_\ell(t) - P_\ell^e \right| \underset{t \to +\infty}{\sim} c_\ell \, \mathrm{e}^{-(t-t_0)/\tau_r} \,,$$

où τ_r est indépendant de l'état (ℓ) choisi et c_ℓ, qui dépend en général de (ℓ), est indépendant du temps.

17. Lorsque par exemple on plonge un morceau de sucre dans un verre d'eau, l'état d'équilibre (eau sucrée de façon homogène) est atteint beaucoup plus rapidement si l'eau est chaude que si elle est froide, et encore plus rapidement si l'on agite le mélange.

sont extrêmement longs : si par exemple on fabrique un verre de silice par fusion puis trempe (c'est-à-dire refroidissement brusque), sa transformation en quartz (sa forme cristalline, qui est son véritable état d'équilibre à température ordinaire) n'aura même pas commencé de façon appréciable après plusieurs centaines d'années. Tous les intermédiaires entre ces deux extrêmes se rencontrent dans la pratique[18].

L'important est de situer le temps de relaxation τ_r par rapport au temps d'observation τ_{exp} caractéristique des expériences que l'on se propose d'effectuer sur le système. Dans les cas où

$$\tau_{exp} \gg \tau_r \,, \tag{II.14}$$

la mécanique statistique d'équilibre, dont nous avons jeté les bases aux chapitres II et III, suffit à décrire les propriétés du système, puisque celui-ci aura pratiquement toujours atteint son état d'équilibre au tout début de l'expérience envisagée. En fait, on peut le plus souvent traiter de façon analogue la situation inverse

$$\tau_{exp} \ll \tau_r \,. \tag{II.15}$$

En effet, la considération de l'état d'équilibre, qui n'a pas dans ce cas le temps de s'établir pendant que nous étudions le système, est sans intérêt. Le système se trouve dans un état macroscopique qui paraît indépendant du temps puisqu'il n'évolue pas de façon appréciable à l'échelle de temps où nous le considérons, et que l'on appelle *«état métastable»*[19]. On traite alors cet état métastable comme un véritable état d'équilibre : pour étudier le verre de silice que nous avons pris plus haut pour exemple, on fera comme si les états microscopiques de type cristallin (quartz) étaient rendus inaccessibles par une contrainte extérieure, et l'on appliquera la mécanique statistique d'équilibre en tenant compte uniquement des états de type vitreux.

---

18. Il arrive même assez souvent que, dans un même système, différentes propriétés « relaxent » à des vitesses différentes, c'est-à-dire que certaines atteignent leur valeur d'équilibre avant d'autres : par exemple, la dissolution complète d'un morceau de sucre dans un verre d'eau (c'est-à-dire l'homogénéisation de la concentration en sucre) peut prendre plus de temps que l'établissement de l'équilibre thermique (homogénéisation de la température dans le verre). On peut dans de tels cas définir un temps de relaxation pour chacune des propriétés du système; c'est bien sûr le plus long d'entre eux qui régit l'approche vers l'équilibre final, et qui s'identifie donc au temps de relaxation global τ_r. Mais le système peut ainsi passer parfois, avant d'atteindre son véritable état d'équilibre, par des états d'*équilibre partiel, ou incomplet.* Un exemple intéressant d'une situation de ce type est décrit dans le complément II.B.

19. La silice vitreuse présente toutes les caractéristiques extérieures d'un véritable état d'équilibre. Certains autres états métastables sont au contraire fragiles : on peut par exemple maintenir de l'eau à l'état liquide au-dessous de son point de solidification (c'est-à-dire au-dessous de 0° Celsius à la pression atmosphérique); mais ceci demande quelques précautions (absence de poussières, parois aussi lisses que possible...) et l'état métastable obtenu est immédiatement détruit par un choc ou une vibration. Le phénomène est fondamentalement le même dans les deux cas : le système « ignore » certains états microscopiques accessibles (ici, les états cristallins : quartz ou glace); son état macroscopique correspond à un maximum d'entropie (s'il est isolé), mais il existe un autre maximum, où la valeur de l'entropie est supérieure, lorsqu'on prend en compte tous les états accessibles; c'est ce maximum absolu qui correspond à l'état d'équilibre stable. La différence entre les deux types d'états métastables provient simplement des ordres de grandeur : dans le cas de la silice, la « vallée » séparant les deux maximums de S^* est tellement profonde que le système « ne s'aperçoit jamais » de l'existence du véritable état d'équilibre; pour l'eau surfondue, au contraire, les fluctuations de densité liées à la présence de corps étrangers ou produites par des perturbations extérieures suffisent à faire basculer le système de l'état métastable à l'état d'équilibre stable. Tous les intermédiaires entre ces deux cas extrêmes sont évidemment possibles.

Le complément III.G permet de comprendre, dans le cadre d'un modèle simple, l'origine physique de ces états métastables pour la transition liquide-gaz.

Les cas les plus délicats sont ceux où temps de relaxation et temps d'observation sont du même ordre de grandeur, car l'étude théorique du système demande alors une connaissance plus détaillée des coefficients de l'équation maîtresse[20], et nous avons déjà souligné combien une telle connaissance est difficile à obtenir[21].

2. SYSTÈME ISOLÉ : LE THÉORÈME *H*

L'évolution spontanée d'un système macroscopique abandonné à lui-même dans un état hors d'équilibre se produit donc dans le sens qui l'amène à l'équilibre. Le « théorème *H* » de Boltzmann[22] permet, dans le cas d'un système isolé, de caractériser de façon simple cette évolution à sens unique.

a. Énoncé et démonstration du théorème

Pour un système hors d'équilibre, l'entropie statistique, définie comme toujours par

$$S \widehat{=} -k \sum_{(\ell)} P_\ell \ln P_\ell \,, \tag{II.16}$$

dépend du temps comme la distribution de probabilité $\{P_\ell(t)\}$. L'équation maîtresse et la symétrie de ses coefficients pour un système isolé permettent de démontrer que *l'entropie est une fonction non décroissante du temps :*

$$\boxed{\frac{dS}{dt} \geqslant 0 \quad \text{pour un système isolé.}} \tag{II.17}$$

La démonstration peut être décomposée en trois étapes. Commençons par calculer la dérivée par rapport au temps de l'expression (II.16) de l'entropie, en y reportant la dérivée des probabilités P_ℓ donnée par l'*équation maîtresse*; on obtient ainsi[23]

$$\frac{dS}{dt} = k \sum_{(\ell),(m)} [a^*_{m\ell} P_\ell - a^*_{\ell m} P_m] \ln P_\ell \,. \tag{II.18}$$

---

20. On est souvent amené, dans les situations de ce type, à rechercher des équations d'évolution approchées, moins générales que l'équation maîtresse mais plus faciles à traiter.

21. On apprécie mieux, à la lumière de cette discussion, la beauté et l'efficacité de la mécanique statistique : les lois qu'on en tire dans les deux premiers cas (II.14) et (II.15) sont indépendantes des détails de chaque situation particulière, et fournissent ainsi une description à la fois très générale et concrète d'une vaste classe de systèmes macroscopiques de nature variée. Dans le troisième cas, ce sont les difficultés d'explicitation des caractéristiques du problème qui limitent les possibilités de la mécanique statistique.

22. Boltzmann notait *H* une grandeur proportionnelle à l'entropie; d'où le nom du théorème.

23. En effet,

$$\frac{dS}{dt} = -k \left[\sum_{(\ell)} \frac{dP_\ell}{dt} \ln P_\ell + \sum_{(\ell)} \frac{dP_\ell}{dt} \right]. \tag{1}$$

Le second terme est nul par suite de la condition de normalisation :

$$\sum_{(\ell)} P_\ell = 1 \implies \sum_{(\ell)} \frac{dP_\ell}{dt} = 0 \,. \tag{2}$$

Reportons dans le premier l'information fournie par l'équation maîtresse; il vient

$$\frac{dS}{dt} = -k \sum_{(\ell)} \sum_{(m)} [a^*_{\ell m} P_m - a^*_{m\ell} P_\ell] \ln P_\ell \,. \tag{3}$$

En second lieu, utilisons la *symétrie des coefficients* de l'équation maîtresse,

$$a^*_{m\ell} = a^*_{\ell m} , \tag{II.19}$$

valable pour un *système isolé,* pour montrer que la formule (II.18) peut être écrite sous l'une des deux formes équivalentes suivantes[24] :

$$\frac{dS}{dt} = k \sum_{(\ell),(m)} a^*_{m\ell}(P_\ell - P_m) \ln P_\ell \; ; \tag{II.20.a}$$

$$\frac{dS}{dt} = k \sum_{(\ell),(m)} a^*_{m\ell}(P_m - P_\ell) \ln P_m \; . \tag{II.20.b}$$

La troisième et dernière étape consiste à additionner membre à membre les deux égalités (II.20), ce qui donne finalement

$$\frac{dS}{dt} = \frac{k}{2} \sum_{(\ell),(m)} a^*_{m\ell}(P_\ell - P_m)(\ln P_\ell - \ln P_m) \; . \tag{II.21}$$

Dans cette expression finale, seuls sont différents de zéro les termes $(\ell) \neq (m)$, pour lesquels les coefficients $a^*_{m\ell}$ sont positifs ou nuls. Il en est de même du produit des deux derniers facteurs, car

$$P_\ell \geqslant P_m \iff \ln P_\ell \geqslant \ln P_m \; . \tag{II.22}$$

La conclusion (II.17) découle donc immédiatement de la formule (II.21).

Soulignons que *cette démonstration* du théorème *H utilise seulement l'équation maîtresse et la symétrie de ses coefficients*[25].

b. Corollaire : l'équilibre comme maximum de l'entropie

L'égalité (II.21) permet aussi de déterminer la *distribution d'équilibre* pour un système isolé.

En effet, toute distribution indépendante du temps doit donner

$$\frac{dS}{dt} = 0 \; , \tag{II.23}$$

c'est-à-dire rendre l'entropie *stationnaire.* Or ceci ne peut se produire que si *chacun des termes* de la double somme, dans (II.21), *est séparément nul,* puisqu'ils sont tous non négatifs. On prend bien sûr nulles les probabilités des états qui ne sont pas accessibles dans les conditions fixées au système, de sorte que les termes correspondants sont constamment nuls. Pour des états (ℓ) et (m) accessibles, nous supposons comme plus haut (§ A.2.c) que les coefficients $a^*_{\ell m}$ sont tous non nuls[26].

---

24. La formule (II.20.a) découle immédiatement de (II.18) : il suffit d'y remplacer $a^*_{\ell m}$ par $a^*_{m\ell}$.
Si l'on échange les rôles des deux indices muets (ℓ) et (m), on écrit cette même formule sous la forme

$$\frac{dS}{dt} = k \sum_{(m),(\ell)} a^*_{\ell m}(P_m - P_\ell) \ln P_m \; .$$

Il suffit d'utiliser à nouveau la propriété de symétrie (II.19) pour aboutir à (II.20.b).

25. Il existe une démonstration du théorème *H* un peu différente, qui n'utilise pas la symétrie des coefficients mais seulement l'hypothèse que l'équation maîtresse admet une solution stationnaire (réf. 1, chapitre V, § 5).

26. En fait, il n'est pas indispensable que les coefficients $a^*_{\ell m}$ vérifient une condition aussi stricte. Il suffit comme plus haut que l'ensemble des états accessibles ne soit pas séparé en deux ou plusieurs sous-ensembles sans communication entre eux.

L'annulation de dS/dt ne peut alors avoir lieu que pour

$$P_\ell = P_m \quad \text{quels que soient } (\ell) \text{ et } (m) \text{ accessibles,} \tag{II.24}$$

c'est-à-dire pour la *distribution microcanonique* $\{P^*_\ell\}$.

Ce point stationnaire est un *maximum* : nous savons en effet que l'entropie statistique (II.16) est maximale lorsque les probabilités P_ℓ sont toutes égales (I, § II.B.3). Il est de plus unique; comme nous l'avons annoncé au paragraphe A.2.c, le théorème H assure l'*unicité de la solution d'équilibre* de l'équation maîtresse.

Les résultats des paragraphes a et b peuvent être résumés de la façon très simple suivante : *lorsqu'un système isolé est abandonné à lui-même dans une situation hors d'équilibre, son évolution spontanée s'accompagne d'une croissance monotone de son entropie; cette évolution s'arrête lorsque l'entropie est maximale, le système ayant alors atteint son état d'équilibre, caractérisé par la distribution microcanonique.* Si l'on rapproche ces résultats et les arguments du chapitre II, on constate en premier lieu que *le postulat microcanonique est en réalité une conséquence du postulat d'évolution,* puisque la distribution microcanonique est la solution d'équilibre unique de l'équation maîtresse. En outre, le *«principe du désordre maximum»* (II, § I.B.3.b) acquiert une *interprétation dynamique* intéressante : l'entropie microcanonique S^*, valeur maximale permise à l'entropie statistique S pour un système isolé, est atteinte à la fin du processus d'évolution temporelle, comme aboutissement de la croissance continuelle de S.

3. SYSTÈME EN CONTACT AVEC UN THERMOSTAT

a. Version modifiée du théorème *H*

Lorsqu'on étudie un système $\mathcal{S}$ en contact avec un thermostat $\mathcal{T}$, le théorème H est toujours valable, mais pour le système global isolé $\mathcal{S} \cup \mathcal{T}$. Si l'on restreint son attention au *système $\mathcal{S}$ lui-même, son entropie peut éventuellement diminuer* : il suffit que l'entropie du thermostat augmente davantage durant le même temps. La validité générale du théorème H n'est pas remise en cause, mais on ne peut pas l'appliquer tel quel au système $\mathcal{S}$, qui n'est pas isolé.

Considérons, non plus l'entropie du système, mais son *énergie libre*[27] F, que l'on définit hors d'équilibre par une simple extension de la formule (II.27) du chapitre III :

$$F \mathrel{\widehat{=}} \overline{E} - TS\,, \tag{II.25}$$

où T est la température du thermostat, S l'entropie statistique (II.16) et $\overline{E}$ la valeur moyenne de l'énergie dans la distribution $\{P_\ell(t)\}$:

$$\overline{E}(t) = \sum_{(\ell)} P_\ell(t)\, E_\ell\,. \tag{II.26}$$

On démontre alors que, *pour un système maintenu en contact avec un thermostat, l'énergie libre est une fonction non croissante du temps :*

$$\frac{dF}{dt} \leqslant 0\,. \tag{II.27}$$

27. Au chapitre II, l'entropie (microcanonique) est apparue comme la grandeur centrale dans l'étude d'un système isolé (à l'équilibre). Pour un système en contact avec un thermostat, elle est remplacée dans ce rôle par l'énergie libre (canonique) F (voir chapitre III). C'est la même substitution qui est opérée ici pour les états hors d'équilibre.

On peut en effet écrire[28] la dérivée de F par rapport au temps sous une forme analogue à (II.21) :

$$\frac{dF}{dt} = -\frac{kT}{2} \sum_{(\ell),(m)} \tilde{a}^T_{m\ell}(\widetilde{P}_\ell - \widetilde{P}_m)(\ln \widetilde{P}_\ell - \ln \widetilde{P}_m) , \tag{II.28}$$

où les coefficients $\tilde{a}^T_{m\ell}$ sont définis par

$$\tilde{a}^T_{m\ell} \widehat{=} a^T_{m\ell}\, e^{-E_\ell/kT} , \tag{II.29.a}$$

et les $\widetilde{P}_\ell$ par

$$\widetilde{P}_\ell \widehat{=} P_\ell\, e^{+E_\ell/kT} \tag{II.29.b}$$

Dans la double somme de la formule (II.28), tous les termes sont comme plus haut non négatifs, ce qui conduit au résultat (II.27).

b. Corollaire : l'équilibre comme minimum de l'énergie libre

Ici aussi, la formule (II.28) permet de déterminer la *distribution d'équilibre pour*

---

28. On déduit facilement des formules (II.25), (II.16) et (II.26) que

$$F = \sum_{(\ell)} P_\ell[E_\ell + kT \ln P_\ell] . \tag{1}$$

Dérivons cette expression par rapport au temps :

$$\frac{dF}{dt} = \sum_{(\ell)} \frac{dP_\ell}{dt}[E_\ell + kT \ln P_\ell] + kT \sum_{(\ell)} \frac{dP_\ell}{dt} \tag{2}$$

$$= \sum_{(\ell)} \frac{dP_\ell}{dt}[E_\ell + kT \ln P_\ell] \tag{3}$$

par suite de la condition de normalisation des probabilités (note 23, formule (2)).

Utilisons l'équation maîtresse, dont les coefficients $a^T_{\ell m}$ vérifient ici la relation

$$a^T_{m\ell}\, e^{-E_\ell/kT} = a^T_{\ell m}\, e^{-E_m/kT} . \tag{4}$$

Pour pouvoir répéter la démonstration du théorème H, posons

$$\tilde{a}^T_{m\ell} \widehat{=} a^T_{m\ell}\, e^{-E_\ell/kT} , \tag{5}$$

de sorte que l'équation maîtresse s'écrit

$$\frac{dP_\ell}{dt} = \sum_{(m)} [\tilde{a}^T_{\ell m}\, e^{+E_m/kT} P_m - \tilde{a}^T_{m\ell}\, e^{+E_\ell/kT} P_\ell] , \tag{6}$$

les nouveaux coefficients $\tilde{a}^T_{\ell m}$ étant maintenant symétriques :

$$\tilde{a}^T_{m\ell} = \tilde{a}^T_{\ell m} . \tag{7}$$

On peut ensuite remarquer que, dans la formule (3), l'expression entre crochets n'est rien d'autre que

$$E_\ell + kT \ln P_\ell = kT \ln[P_\ell\, e^{+E_\ell/kT}] , \tag{8}$$

où l'on retrouve le produit qui intervient dans (6). Si donc on pose

$$\widetilde{P}_\ell \widehat{=} P_\ell\, e^{+E_\ell/kT} , \tag{9}$$

la dérivée de F prend la forme

$$\frac{dF}{dt} = kT \sum_{(\ell),(m)} [\tilde{a}^T_{\ell m}\widetilde{P}_m - \tilde{a}^T_{m\ell}\widetilde{P}_\ell] \ln \widetilde{P}_\ell . \tag{10}$$

Un raisonnement calqué sur celui du paragraphe 2.a donne alors

$$\frac{dF}{dt} = -\frac{kT}{2} \sum_{(\ell),(m)} \tilde{a}^T_{m\ell}(\widetilde{P}_\ell - \widetilde{P}_m)(\ln \widetilde{P}_\ell - \ln \widetilde{P}_m) . \tag{11}$$

.../...

un système en contact avec un thermostat de température T. En effet, l'énergie libre est indépendante du temps :

$$\frac{dF}{dt} = 0 \tag{II.30}$$

seulement si chaque terme de la somme s'annule séparément. En supposant que les probabilités de transition $a^T_{\ell m}$ sont non nulles entre tous les états (ℓ) et (m) qui nous intéressent[29], nous en déduisons

$$\widetilde{P}_\ell = \widetilde{P}_m \quad \text{quels que soient } (\ell) \text{ et } (m), \tag{II.31}$$

c'est-à-dire d'après (II.29.b)

$$\frac{P_\ell}{P_m} = \frac{e^{-E_\ell/kT}}{e^{-E_m/kT}}. \tag{II.32}$$

Autrement dit, seule la *distribution canonique,* qui est ici la distribution d'équilibre, rend l'énergie libre stationnaire.

Les résultats des paragraphes a et b peuvent donc encore être résumés de façon simple et intuitivement « parlante » : *lorsqu'un système maintenu en contact avec un thermostat est abandonné dans une situation hors d'équilibre, son évolution spontanée s'accompagne d'une décroissance monotone de son énergie libre; cette évolution s'arrête lorsque l'énergie libre est minimale, le système ayant alors atteint son état d'équilibre* (distribution canonique). Il apparaît ainsi que l'équilibre d'un système en contact avec un thermostat correspond au *minimum d'énergie libre* compatible avec les autres conditions extérieures, et que ce minimum est atteint à la fin du processus d'évolution temporelle du système.

4. IRRÉVERSIBILITÉ DE L'ÉVOLUTION SPONTANÉE

a. Équation maîtresse et irréversibilité

Prenons pour fixer les idées un *système isolé* (des considérations tout à fait analogues vaudraient pour un système en contact avec un thermostat). Le théorème H et les conséquences que nous en avons tirées permettent de saisir la signification de la non invariance de l'équation maîtresse par rapport au renversement du temps (§ I.B.3.b).

Supposons en effet que, par un procédé qu'il n'est pas nécessaire de préciser, le système ait été amené à l'instant t_0 dans un état macroscopique caractérisé par la distribution $\{P_\ell(t_0)\}$, et qu'il soit alors abandonné à lui-même. Connaissant $\{P_\ell(t_0)\}$, on peut déduire de l'équation maîtresse les dérivées dP_ℓ/dt à cet instant initial t_0. Si $\{P_\ell(t_0)\}$ se trouve être la distribution d'équilibre (ici microcanonique), toutes ces dérivées sont nulles et le système reste indéfiniment dans son état d'équilibre macroscopique. Dans tous les autres cas, les dP_ℓ/dt sont différentes de zéro (§ A.2.c) et le système évolue. Le théorème H montre que cette évolution fait constamment croître l'entropie statistique S (§ 2.a); par conséquent, *le système ne repassera jamais (spontanément)*[30] *par l'état de départ* $\{P_\ell(t_0)\}$: l'évolution spontanée d'un système macroscopique est *irréversible.*

Cette irréversibilité a son origine dans l'équation maîtresse. Le théorème H explicite seulement, de manière particulièrement frappante, une des propriétés de cette

29. C'est-à-dire tous les états microscopiques compatibles avec les valeurs fixées pour les paramètres extérieurs (nombre de particules, volume, ...) autres que la température T. Ici aussi, la condition est en fait moins stricte (cf. note 26).

30. Bien entendu, on peut ramener le système à l'état initial en agissant sur lui. Nous étudions ici l'évolution *spontanée* d'un système *isolé.*

équation. Plus précisément, c'est le fait que l'équation maîtresse est du premier ordre par rapport au temps qui permet de calculer de façon univoque la dérivée dS/dt.

b. Analyse physique d'un exemple simple

Considérons tout d'abord une particule unique enfermée dans un récipient de volume V macroscopique et supposons pour simplifier que son mouvement puisse être correctement décrit par la mécanique classique. A l'instant t_0, cette particule se trouve au point $\vec{r}_0$ et sa vitesse est $\vec{v}_0$ (fig. 1). Son mouvement l'amène en $\vec{r}_1$ à l'instant $t_1 = t_0 + \tau$, et sa vitesse est alors $\vec{v}_1$. L'équation de la dynamique newtonienne, nous l'avons déjà souligné (§ I.B.3.b), est invariante par renversement du temps. Ceci implique que, si la même particule est placée à l'instant t_0 au point $\vec{r}_1$ et qu'on lui communique la vitesse $\vec{v}\,'_1 = -\vec{v}_1$, elle se trouvera en $\vec{r}_0$ à l'instant $t'_1 = t_0 + \tau$, avec la vitesse $\vec{v}\,'_0 = -\vec{v}_0$. Pour saisir ce fait de façon intuitive, imaginons que l'on ait pu filmer le mouvement « direct » de la particule de $\vec{r}_0$ à $\vec{r}_1$, entre les instants t_0 et t_1; si l'on passe ensuite le film à l'envers, c'est le mouvement inverse, de $\vec{r}_1$ à $\vec{r}_0$ avec des vitesses opposées, que l'on verra sur l'écran de projection. Aucun spectateur ne trouvera ce mouvement inverse anormal, car il est parfaitement compatible avec les lois physiques.

Examinons maintenant le comportement d'un gaz constitué d'un grand nombre de particules, en supposant toujours qu'il est décrit par la mécanique classique. A l'instant t_0, le récipient qui contient le gaz est séparé en deux compartiments par une cloison (fig. 2.a); toutes les particules se trouvent dans la moitié gauche du récipient, la moitié droite étant vide. On retire alors la cloison. Nous savons que le gaz va se répandre dans le compartiment initialement vide, de façon à occuper finalement de façon homogène la totalité du volume disponible (fig. 2.b). Ceci sera réalisé à un

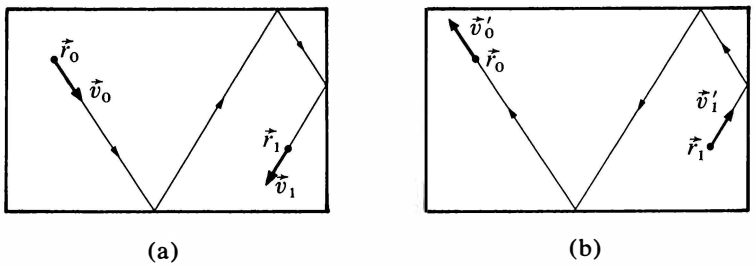


FIGURE 1

Les équations du mouvement microscopique sont invariantes par renversement du temps : si le mouvement de $(\vec{r}_0, \vec{v}_0)$ à $(\vec{r}_1, \vec{v}_1)$ est possible (a), alors le mouvement inverse de $(\vec{r}_1, \vec{v}\,'_1 = -\vec{v}_1)$ à $(\vec{r}_0, \vec{v}\,'_0 = -\vec{v}_0)$ l'est aussi (b).

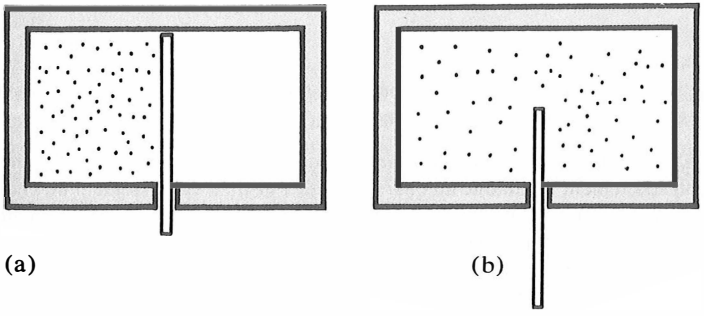


FIGURE 2

Expansion irréversible d'un gaz constitué d'un grand nombre de molécules.

instant $t_1 = t_0 + \tau$, si l'on choisit τ grand devant le temps de relaxation τ_r du gaz[31]. L'irréversibilité que nous avons définie de façon générale au paragraphe a prend ici un aspect très simple et concret : jamais le gaz ne reviendra spontanément se confiner tout entier dans la moitié gauche du récipient[32]. Imaginons comme plus haut que l'on ait pu filmer les molécules du gaz au cours de son évolution spontanée entre t_0 et t_1, et que l'on passe ensuite le film à l'envers. Personne ne sera dupe : l'évolution inverse qui se déroulera sur l'écran paraîtra aussi saugrenue que le spectacle d'un plongeur qui bondirait hors de l'eau à reculons vers le plongeoir, ou de tout autre phénomène évidemment contraire aux lois naturelles.

c. D'où vient l'irréversibilité?

Comment l'irréversibilité de l'évolution macroscopique peut-elle découler de lois microscopiques invariantes par renversement du temps? Dans l'exemple ci-dessus, pourquoi n'est-il pas possible, en changeant de signe les vitesses de toutes les molécules du gaz, de faire parcourir à celui-ci le « chemin » inverse de son évolution spontanée? La réponse à cette deuxième question est relativement simple : le nombre des molécules est tellement énorme qu'il est impensable de chercher à contrôler la vitesse de chacune d'elles (I, § I.B.2). Notons que, dans l'état initial où le gaz n'occupe que le compartiment de gauche, on ne possède pas une information détaillée sur la position précise de chaque molécule, mais seulement une information globale, à l'échelle macroscopique : elles se trouvent toutes dans la moitié gauche. L'évolution du gaz a pour effet de « diluer » cette information au niveau microscopique, où elle devient inaccessible : chaque molécule garde en principe, dans sa position et sa vitesse, le « souvenir » d'une parcelle de l'information initiale, mais celle-ci ne peut pas être reconstituée car ces parcelles sont trop ténues et trop nombreuses. C'est cette *perte d'information* que mesure l'accroissement d'entropie prévu par le théorème *H*. C'est elle aussi qui empêche le retour en arrière.

Mais ceci ne répond que partiellement à la question primordiale concernant le passage d'équations fondamentales symétriques dans l'inversion du temps à une évolution irréversible du système. Cette question a fait l'objet de nombreux travaux, et elle n'est pas encore parfaitement élucidée à l'heure actuelle (voir complément IV.A). Nous nous contenterons ici d'indiquer l'idée physique générale qui sert de base à la compréhension du phénomène.

Le problème de l'irréversibilité est intimement lié à celui du déterminisme : *les équations fondamentales sont déterministes,* en ce sens qu'elles fournissent de façon univoque l'état microscopique du système à l'instant t si l'on se donne cet état à un instant antérieur t_0; au contraire, *l'évolution macroscopique* du système *est stochastique,* c'est-à-dire que sa description est fondée sur les probabilités. Dans l'exemple du gaz analysé au paragraphe b, l'impossibilité de « retourner en arrière » est due au fait que l'état final du système n'est pas exactement connu au niveau microscopique, ce qui empêche d'envisager de changer le signe de toutes les vitesses; cette méconnaissance de l'état microscopique final provient à son tour de la nécessité

31. Ce temps de relaxation dépend de la façon dont on retire la cloison et de la taille du récipient, mais il excède rarement quelques secondes.

32. Le nombre de particules qui se trouvent dans le compartiment de droite initialement vide ne croît pas de façon strictement monotone au cours du temps jusqu'à établissement de l'équilibre : les aléas du mouvement microscopique peuvent amener ce nombre à diminuer légèrement durant certains intervalles de temps restreints. De tels effets ne sont pas en contradiction avec le théorème *H* ou l'équation maîtresse, qui s'expriment en termes de probabilités : des fluctuations autour de la valeur attendue sont elles-mêmes attendues. Cependant, pour un système macroscopique tel qu'un gaz, ces fluctuations sont très petites en valeur relative, et leur probabilité est une fonction très rapidement décroissante de leur amplitude.

de traiter de façon probabiliste l'évolution du système. Donc les antinomies réversible/irréversible et déterministe/stochastique constituent deux aspects d'une seule et même question fondamentale.

La clef du problème est fournie par *l'extrême sensibilité des équations microscopiques par rapport à de petits changements* dans les conditions initiales ou de faibles perturbations exercées sur le système au cours de son évolution. Pour préciser ce point fondamental, nous allons utiliser à nouveau les exemples simples du paragraphe b.

Reprenons d'abord le cas d'une particule unique (fig. 1). La situation est en réalité moins simple que la description que nous en avons donnée. En effet, la position et la vitesse initiales de la particule ne peuvent pas être connues avec une précision parfaite : on sait seulement que l'état initial appartient à un petit domaine de l'espace des phases (espace à 6 dimensions dont les points ont pour coordonnées $\vec{r}$ et $\vec{p}$, position et impulsion de la particule). Examinons alors le comportement du système lorsqu'on choisit successivement comme état initial les divers points de ce domaine. A partir de chacun de ces points, l'équation du mouvement (déterministe) donne de façon univoque l'état de la particule aux instants ultérieurs. Chacun de ces mouvements est représenté par une trajectoire dans l'espace des phases. On pourrait penser que deux trajectoires issues de points initiaux voisins vont rester continuellement voisines. Eh bien, non : *ces deux trajectoires s'écartent irrémédiablement l'une de l'autre.* On peut le comprendre facilement, pour une particule microscopique, si l'on tient compte du fait que les parois du récipient sont loin d'être lisses et planes à l'échelle microscopique : même si l'on admet pour simplifier que la particule rebondit spéculairement[33], les irrégularités de la paroi vont accentuer la différence entre deux trajectoires voisines (fig. 3). Ainsi, les chocs successifs de la particule sur les parois du récipient ont pour effet de « diluer » l'information initiale en élargissant de façon « explosive » le domaine d'incertitude dans l'espace des phases[34]. N'oublions pas qu'en outre le système étudié, même s'il s'agit

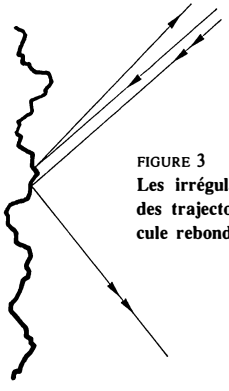


FIGURE 3
Les irrégularités microscopiques des parois ont pour effet d'écarter l'une de l'autre des trajectoires initialement voisines (on a supposé ici, pour simplifier, que la particule rebondit comme une balle sur la paroi).

33. Le choc d'une particule microscopique sur la surface d'un solide est en réalité un phénomène beaucoup plus complexe, dont les détails dépendent des interactions entre la particule et les atomes constituant la paroi. Assez souvent, la particule reste un certain temps adsorbée sur la paroi avant de se libérer à nouveau et de repartir; dans ce cas, le choc désoriente davantage encore sa vitesse et l'effet décrit ici en est renforcé.

34. Cette évolution respecte le théorème de Liouville (appendice III, § II.3.b) : le volume total du domaine de l'espace des phases dans lequel doit se trouver le point représentatif du système reste constant au cours du temps. Mais ce domaine s'éparpille dans l'ensemble de l'espace des phases accessible comme le fait une goutte d'encre dans un verre d'eau (le volume d'encre ne varie pas, mais il se répartit de façon uniforme dans l'eau).

d'une particule unique, est soumis à de très faibles mais multiples influences extérieures (§ I.A.1 et complément IV.B), dont l'importance est considérablement accrue par la sensibilité intrinsèque des équations du mouvement. Ces deux effets conjugués font qu'on devient incapable, au bout d'un certain temps qui peut être assez court (complément IV.B, § I.3), de préciser la position de la particule dans le volume du récipient. On doit dès lors, comme en mécanique statistique, avoir recours aux probabilités pour analyser son mouvement ultérieur : si par exemple on divise le récipient en deux parties d'égal volume, la particule a la probabilité 1/2 de se trouver dans l'un de ces deux compartiments, et la probabilité 1/2 de se trouver dans l'autre. *L'évolution déterministe a fait place à une évolution stochastique.* Il est bien entendu impossible également de revenir en arrière, c'est-à-dire de retrouver le domaine initial limité de l'espace des phases à partir de la description probabiliste finale : *l'évolution est devenue irréversible.*

Dans un système de N particules, s'ajoute aux précédents l'effet des *collisions entre particules,* qui agissent dans le même sens[35]. L'irréversibilité se manifeste alors de façon plus évidente : dans l'exemple de la figure 2, la probabilité pour que, après diffusion, les N molécules du gaz se retrouvent toutes dans la moitié gauche du récipient vaut $(1/2)^N$; pour N de l'ordre du nombre d'Avogadro, cette probabilité est si fantastiquement petite que l'événement correspondant est véritablement impossible. Mais il est intéressant de souligner que, contrairement à ce que pourraient suggérer les arguments du paragraphe b, les systèmes comportant un grand nombre de particules ne sont pas seuls à évoluer de façon irréversible : l'instabilité des équations du mouvement par rapport à de légères modifications des conditions initiales ou extérieures, qui substitue une évolution stochastique irréversible au déterminisme réversible que l'on attendrait, semble être une propriété quasi-universelle des systèmes dynamiques[36].

Références

1. N.G. van Kampen, *Stochastic Processes in Physics and Chemistry,* North-Holland, Amsterdam (1981), chapitre V.
2. H.J. Kreuzer, *Non-equilibrium Thermodynamics and its Statistical Foundations,* Clarendon Press, Oxford (1981), chapitres 10 et 11.
3. A.D. Dolgov et Ya. B. Zeldovich, « Cosmology and Elementary Particles », *Reviews of Modern Physics 53,* 1 (1981).

35. Nous avons mis jusqu'ici en avant les influences extérieures qui s'exercent sur le système : choc des particules sur les parois, forces créées par des objets étrangers au système, ... Mais il est probable que, même pour un système rigoureusement isolé, deux trajectoires dans l'espace des phases initialement issues de deux points voisins de cet espace s'écartent l'une de l'autre de façon exponentielle au cours du temps. Cette propriété a effectivement été démontrée pour le gaz de « sphères dures » (modèle dans lequel les molécules du gaz sont assimilées à des sphères sans interaction à distance mais rigoureusement impénétrables les unes aux autres, se comportant lors des chocs comme des boules de billard).

36. Cette propriété joue également un rôle primordial en ce qui concerne le *problème ergodique* (I, § IV.B.3 et II, § I.B.3.a). En effet, même si l'on arrivait à préparer un système isolé dans un état macroscopique initial restreint à un groupe très étroit d'états microscopiques, l'évolution de ces divers états les amènerait rapidement à couvrir uniformément tout le domaine des états microscopiques accessibles (cf. note 34). C'est ce qui explique que le comportement d'un système particulier suive en fait les lois probabilistes définies en termes d'ensemble statistique. Voir complément IV.A.

COMPLÉMENTS DU CHAPITRE IV

IV.A. Notions succinctes de théorie ergodique	Discussion théorique du «principe ergodique», déjà évoqué aux chapitres précédents (I, § IV.B.3 et II, § I.B.3). Assez facile, car reste à un niveau élémentaire. Le modèle du § II.3.b est amusant.
IV.B. Perturbations aléatoires d'un système «isolé» : discussion physique et ordres de grandeur	Analyse les influences extérieures susceptibles de changer l'état microscopique d'un système macroscopique «isolé». Assez facile. Les ordres de grandeur sont surprenants.
IV.C. Propriétés de l'équation maîtresse	Quelques démonstrations (assez faciles car semblables à celles qu'on utilise couramment en mécanique quantique) précisant et prolongeant l'étude de l'équation maîtresse.
IV.D. «Postulat d'évolution» et équation de Schrödinger	Aborde une question fondamentale mais difficile : peut-on démontrer le «postulat d'évolution»? Le § I peut être lu pour lui-même : il introduit de façon générale la notion de processus markovien.
IV.E. Évolution d'un gaz classique dilué : l'équation de Boltzmann	Introduit la célèbre équation fondamentale proposée en 1872 par Boltzmann pour décrire l'évolution d'un gaz vers l'équilibre macroscopique, puis en donne quelques propriétés essentielles (notamment la version originale du «théorème *H*»). Important mais difficile.
IV.F. Équations générales d'évolution d'un fluide classique	Déduit de l'équation de Liouville un système d'équations d'évolution exactes pour un fluide, et en présente deux applications simples. Pas très difficile.

COMPLÉMENT IV.A

Notions succinctes de théorie ergodique

I. Le problème ergodique

II. La « hiérarchie ergodique » des systèmes classiques
1. Systèmes ergodiques
2. La propriété de « mélange »
3. « K-systèmes » et systèmes de Bernoulli
a. Définition physique
b. La « transformation du boulanger »

III. Le problème ergodique en mécanique quantique

I. LE PROBLÈME ERGODIQUE

Dès le chapitre I (§ IV.B.3), nous avons énoncé le *« principe ergodique »,* dont nous avons constamment, depuis, admis la validité : pour les systèmes que nous étudions en mécanique statistique, nous supposons que, *à l'équilibre, les valeurs moyennes des diverses grandeurs, évaluées à l'aide d'un ensemble statistique de systèmes identiques, coïncident avec les valeurs moyennes de ces mêmes grandeurs calculées sur les fluctuations d'un système unique au cours du temps.*

Dans le chapitre IV, nous avons étudié *l'évolution des systèmes macroscopiques vers leur état d'équilibre,* et avons constaté que l'*irréversibilité* de cette évolution est elle aussi liée au « principe ergodique ».

Les recherches sur ces questions, que l'on regroupe sous le vocable de *« problème ergodique »,* ont montré que la réalité est plus complexe et plus riche qu'on ne le pensait au début (et que nous ne l'avons indiqué jusqu'ici). Nous allons décrire succinctement, sans entrer dans les détails techniques, les principales notions qui ont été dégagées; le sujet est d'ailleurs loin d'être clos car des problèmes importants demeurent posés.

Nous nous plaçons délibérément dans le cadre de la *mécanique classique,* où les recherches sont plus avancées et les résultats plus faciles à exposer qu'en mécanique quantique. Nous dirons quelques mots, dans la dernière partie, sur les difficultés particulières que présente la théorie ergodique quantique.

II. LA « HIÉRARCHIE ERGODIQUE » DES SYSTÈMES CLASSIQUES

Les systèmes physiques se classent, du point de vue du problème ergodique, en quatre catégories qui s'« emboîtent » l'une dans l'autre : tout « système de Bernoulli » (§ 3) est un « K-système », tout « K-système » est « mélangeant » (§ 2), et tout système « mélangeant » est « ergodique » (§ 1).

1. SYSTÈMES ERGODIQUES

La première catégorie comprend les systèmes ergodiques au sens restreint du terme, pour lesquels moyennes d'ensemble et moyennes temporelles coïncident.

L'exemple le plus simple d'un tel système est l'*oscillateur harmonique à une dimension.* Ceci peut se comprendre simplement. La position $x(t)$ d'un oscillateur harmonique de pulsation ω est donnée en fonction du temps t par une expression de la forme

$$x(t) = x_0 \cos(\omega t + \varphi) \; ; \tag{A.1}$$

x_0 est l'amplitude du mouvement, φ la phase à l'origine.

Choisissons un oscillateur particulier. La *valeur moyenne dans le temps* d'une grandeur $f(t)$ peut être calculée par la formule

$$\widehat{f} = \frac{1}{\tau} \int_0^\tau f(t)\,\mathrm{d}t \,, \tag{A.2}$$

où $\tau = 2\pi/\omega$ est la période de l'oscillateur. Si nous cherchons par exemple cette valeur moyenne temporelle pour la position x et son carré x^2, nous obtenons facilement

$$\widehat{x} = \frac{1}{\tau} \int_0^\tau x_0 \cos(\omega t + \varphi)\,\mathrm{d}t = 0 \,, \tag{A.3.a}$$

$$\widehat{x^2} = \frac{1}{\tau} \int_0^\tau x_0^2 \cos^2(\omega t + \varphi)\,\mathrm{d}t = \frac{1}{2} x_0^2 \,. \tag{A.3.b}$$

Prenons maintenant un *ensemble statistique* d'oscillateurs harmoniques de même pulsation ω et de *même énergie totale.* Ceci implique qu'ils aient tous même amplitude x_0, mais ils peuvent différer les uns des autres par la phase à l'origine φ. Si nous supposons (comme dans le postulat microcanonique) que toutes les valeurs de la phase φ (comprise entre 0 et 2π) sont *également probables*[1], la moyenne d'ensemble de la grandeur f s'écrit

$$\bar{f} = \frac{1}{2\pi} \int_0^{2\pi} f(\varphi)\,\mathrm{d}\varphi \,. \tag{A.4}$$

---

1. Pour un système caractérisé par une seule variable de position x, l'espace des phases (I, § III.C.1) est à deux dimensions, x et p (*attention* à la terminologie : ne pas confondre l'« espace des phases » et la « phase » d'un oscillateur harmonique). La distribution microcanonique (complément II.F, § I) donne même probabilité à tous les éléments de surface $\mathrm{d}x.\mathrm{d}p$ du domaine de l'espace des phases correspondant à une énergie comprise entre E et $E + \delta E$.

Pour l'oscillateur harmonique de masse m et de pulsation ω, les variables canoniques x et p prennent la forme

$$x = x_0 \cos(\omega t + \varphi) \,, \tag{1.a}$$

$$p = -m\omega x_0 \sin(\omega t + \varphi) \,. \tag{1.b}$$

Effectuons alors le changement de variables défini par

$$x = x_0 \cos \Phi \,, \tag{2.a}$$

$$p = -m\omega x_0 \sin \Phi \,, \tag{2.b}$$

avec $x_0 \geqslant 0$,

$$0 \leqslant \Phi < 2\pi \,. \tag{3}$$

L'énergie de l'oscillateur dépend seulement de x_0 :

$$\frac{p^2}{2m} + \frac{1}{2} m\omega^2 x^2 = \frac{1}{2} m\omega^2 x_0^2 \,, \tag{4}$$

et la surface élémentaire de l'espace des phases s'écrit simplement

$$\mathrm{d}x\,\mathrm{d}p = m\omega x_0\,\mathrm{d}x_0\,\mathrm{d}\Phi \,. \tag{5}$$

En effet, le jacobien de la transformation (2) vaut

$$\begin{vmatrix} \dfrac{\partial x}{\partial x_0} & \dfrac{\partial x}{\partial \Phi} \\ \dfrac{\partial p}{\partial x_0} & \dfrac{\partial p}{\partial \Phi} \end{vmatrix} = \begin{vmatrix} \cos \Phi & -x_0 \sin \Phi \\ -m\omega \sin \Phi & -m\omega x_0 \cos \Phi \end{vmatrix} = -m\omega x_0 \,. \tag{6}$$

Fixer l'énergie à la valeur E (δE près) revient donc à fixer x_0 (δx_0 près). Une distribution de probabilité uniforme dans le domaine ainsi délimité de l'espace des phases se traduit par une distribution uniforme de la variable Φ sur l'intervalle $[0, 2\pi]$.

Si comme ci-dessus nous calculons ainsi les valeurs moyennes de x et x^2, nous obtenons les mêmes résultats qu'en (A.3) :

$$\bar{x} = \frac{1}{2\pi} \int_0^{2\pi} x_0 \cos(\omega t + \varphi)\, d\varphi = 0 , \tag{A.5.a}$$

$$\overline{x^2} = \frac{1}{2\pi} \int_0^{2\pi} x_0^2 \cos^2(\omega t + \varphi)\, d\varphi = \frac{1}{2} x_0^2 . \tag{A.5.b}$$

Au contraire, *un système de plusieurs oscillateurs harmoniques indépendants n'est pas ergodique*[2]. En effet, si l'on fixe l'énergie *totale* du système, il existe un grand nombre de façons de la répartir entre les divers oscillateurs. Dans un ensemble statistique, toutes ces répartitions seront réalisées, et de manière équiprobable. Si l'on choisit par contre un système particulier, l'énergie de chaque oscillateur est fixée par les conditions initiales et se conserve au cours du temps; ce système ne passe donc, au cours de son évolution temporelle, que par une sous-classe très restreinte des états représentés dans l'ensemble statistique. Il ne peut donc y avoir égalité des moyennes temporelle et statistique[3].

On a cru pendant longtemps que la non ergodicité des systèmes d'oscillateurs harmoniques était liée à l'indépendance des divers oscillateurs; une faible anharmonicité (cf. note 2) rétablirait aussitôt l'ergodicité, pensait-on, en permettant le transfert d'énergie d'un oscillateur à l'autre. Un théorème, dû à A. Kolmogorov, V. Arnold et J. Moser (le *«théorème KAM»*) affirme qu'il n'en est rien : si elle reste petite devant l'énergie totale du système d'oscillateurs, une perturbation anharmonique ne le rend pas ergodique. Le comportement des systèmes d'oscillateurs franchement anharmoniques s'avère compliqué; il est difficile de dire avec certitude s'il est ou non ergodique[4].

---

2. Un système constitué d'un nombre fini d'oscillateurs harmoniques *couplés* est équivalent à un système d'oscillateurs harmoniques *indépendants,* qui sont les *modes propres* du système initial. Pour obtenir un couplage entre les modes propres, il faut introduire dans l'hamiltonien de départ des termes anharmoniques, c'est-à-dire de degré supérieur à 2 dans les positions des oscillateurs.

3. Plus concrètement, pour un système de deux oscillateurs harmoniques, l'ensemble statistique fait varier de façon aléatoire leurs phases φ_1 et φ_2, mais aussi leurs amplitudes x_1 et x_2, en respectant cependant la condition

$$\frac{1}{2} m_1 \omega_1^2 x_1^2 + \frac{1}{2} m_2 \omega_2^2 x_2^2 = E \text{ fixée}$$

(m_1 et m_2 sont les masses des deux oscillateurs, ω_1 et ω_2 leurs pulsations). Au contraire, lorsqu'on choisit un système particulier, les amplitudes x_1 et x_2 restent constantes au cours du temps, et les états où par exemple la position du premier oscillateur est supérieure à l'amplitude x_1 ainsi fixée ne sont jamais atteints.

4. Des calculs sur ordinateur ont toutefois permis d'explorer ce comportement. Pour une énergie fixée, certains états initiaux donnent lieu dans l'espace des phases à des trajectoires qui se cantonnent à un domaine restreint de l'hypersurface d'énergie constante, d'autres à des trajectoires « erratiques » qui s'étalent sur l'ensemble de cette hypersurface. L'existence du premier type de trajectoires permet de conclure que le système n'est pas ergodique. Toutefois, les domaines couverts par ces trajectoires non ergodiques se rétrécissent lorsque l'énergie croît, pour disparaître semble-t-il au-dessus d'une certaine énergie critique; s'il en est bien ainsi, le système devient alors ergodique.

Le problème se présente différemment pour un gaz. Un gaz rigoureusement parfait (aucune interaction entre les constituants) ne serait pas ergodique, pour la même raison que ci-dessus : l'énergie de chaque molécule serait une constante du mouvement dans chaque système particulier, et varierait d'un système à l'autre dans un ensemble statistique. Mais le théorème KAM ne s'applique pas ici, essentiellement à cause de la *forte interaction répulsive entre deux molécules aux courtes distances* (cf. complément I.A, figure 1). Puisque cette caractéristique de l'interaction entre molécules paraît cruciale, on a étudié le modèle du *« gaz de sphères dures »* : dans ce modèle, les constituants du gaz interagissent comme des boules de billard, c'est-à-dire seulement lorsqu'ils arrivent au contact; ils sont alors rigoureusement impénétrables l'un à l'autre. On a pu montrer (Ya. Sinai - 1970) que le gaz de sphères dures est un système ergodique, et qu'il possède même les propriétés plus fortes des systèmes « mélangeants » (§ 2), et jusqu'à celles des « K-systèmes » (§ 3).

2. LA PROPRIÉTÉ DE « MÉLANGE »

Nous ne nous sommes préoccupés jusqu'ici que des propriétés d'équilibre : si un système est ergodique, on peut décrire son état d'équilibre macroscopique dans le cadre d'un ensemble statistique, c'est-à-dire à l'aide d'une distribution de probabilité dans l'espace des phases (lorsque le système est isolé, cette distribution est uniforme dans le sous-espace d'énergie E fixée, à δE près). Mais nous voulons également savoir quels types de systèmes sont susceptibles de vérifier le « postulat d'évolution » que nous avons posé au chapitre IV. Nous allons examiner ici la question plus générale suivante : à quelle condition un système préparé dans un état quelconque va-t-il évoluer vers son état d'équilibre macroscopique?

L'exemple de l'oscillateur harmonique à une dimension montre à l'évidence que la propriété d'ergodicité n'est pas suffisante pour assurer l'évolution irréversible vers l'équilibre. Considérons en effet un ensemble statistique d'oscillateurs harmoniques à une dimension, de pulsation ω et d'amplitude x_0, comme au paragraphe 1. Supposons qu'initialement ($t = 0$) la phase φ des divers oscillateurs de l'ensemble ne soit pas distribuée de façon équiprobable. L'ensemble statistique décrit alors un état hors d'équilibre macroscopique. Mais la différence de phase entre deux oscillateurs quelconques de l'ensemble ne varie pas au cours du temps; l'évolution dans le temps de l'ensemble statistique ne manifeste donc aucune tendance à le rapprocher de la situation d'équilibre, dans laquelle la phase serait distribuée de façon équiprobable.

Pour qu'on puisse affirmer d'un système qu'il évoluera vers l'équilibre si on l'abandonne à lui-même dans une situation hors d'équilibre, il faut qu'il possède une propriété supplémentaire, que l'on appelle la *« propriété de mélange »*.

D'une façon générale, un état macroscopique est caractérisé par une densité de probabilité dans l'espace des phases associé au système; cette densité de probabilité w est non nulle seulement dans le *sous-espace $\mathcal{E}_E$ délimité par les valeurs des paramètres extérieurs* imposés au système, notamment son *énergie* E (fixée à δE près) s'il est isolé. Dans un état hors d'équilibre, w n'est pas uniforme sur $\mathcal{E}_E$. Si le système possède la propriété de mélange, son évolution va progressivement étaler la densité de probabilité sur la totalité du sous-espace $\mathcal{E}_E$, pour aboutir finalement à la distribution microcanonique, uniforme sur $\mathcal{E}_E$.

Ce comportement est *analogue à celui de deux liquides que l'on mélange* en les agitant. Imaginons par exemple que l'on verse quelques gouttes d'encre dans un verre d'eau. Au début, toute l'encre est concentrée près de l'endroit où on l'a déposée, et c'est de l'eau pure qui remplit le reste du verre (distribution nettement non uniforme). Si l'on agite, on homogénéise le mélange : tout le liquide apparaît bientôt uniformément coloré. Soulignons que cet état final d'équilibre est obtenu quelle que soit la façon dont on procède, c'est-à-dire quel que soit le mouvement que l'on imprime au système des deux liquides; on aboutirait d'ailleurs au même état sans agiter, en attendant

suffisamment longtemps pour que l'encre ait diffusé dans tout le verre. C'est à cause de cette analogie que la propriété décrite ici est appelée « mélange »[5].

On peut montrer qu'un système « mélangeant » est *a fortiori* ergodique. Le prototype d'un tel système est le *gaz de sphères dures* évoqué plus haut. L'origine physique du comportement « mélangeant » réside dans l'*extrême sensibilité,* et même l'*instabilité, des équations du mouvement* par rapport à de petits changements dans les conditions initiales ou les forces : les trajectoires, dans l'espace des phases, de deux systèmes identiques partant de conditions initiales très voisines ne restent pas voisines, mais s'écartent au contraire l'une de l'autre de façon « explosive », c'est-à-dire exponentiellement dans le temps (cf. IV, § II.B.4.c)[6]. Dans le cas du gaz de sphères dures, on peut assez facilement comprendre d'où vient cette instabilité : lorsqu'une bille frappe une surface *convexe* (ici, celle d'une autre bille), une légère modification de sa trajectoire, c'est-à-dire du point d'impact ou de la direction d'incidence, est aussitôt amplifiée par le choc (voir le calcul du complément IV.B, § I.2).

3. « *K*-SYSTÈMES » ET SYSTÈMES DE BERNOULLI

Il existe des systèmes dont le comportement est encore plus aléatoire et imprévisible, pourrait-on dire, que ceux des deux catégories précédentes.

---

5. L'analogie avec le mélange de deux liquides permet de comprendre une caractéristique importante de l'évolution des systèmes « mélangeants ». La quantité d'encre, c'est-à-dire le nombre de molécules colorées, se conserve dans l'opération, comme le fait la densité dans l'espace des phases d'après le théorème de Liouville (appendice III, paragraphe II.3.b). Supposons pour fixer les idées qu'il y ait une molécule colorée pour 1 000 molécules d'eau. Pour pouvoir constater que l'encre est répartie de façon uniforme, il faut examiner des volumes de liquide comprenant un nombre de molécules très grand devant 1 000 : tous ces volumes comporteront, aux fluctuations près, la même proportion de molécules colorées; si au contraire on choisissait des volumes beaucoup plus petits, on y trouverait parfois une molécule d'encre, mais le plus souvent aucune. Ces considérations se transposent à la densité de probabilité dans l'espace des phases d'un système « mélangeant » : la densité d'équilibre obtenue en fin d'évolution n'est uniforme que si on la moyenne sur des domaines élémentaires de l'espace des phases qui ne soient pas trop petits. Il faut donc introduire dans l'espace des phases une *« granulation »* (« coarse graining » en anglais), et ne considérer les variations de la densité de probabilité qu'à une échelle supérieure à la taille des « grains ».

6. Un théorème célèbre de Poincaré établit qu'un système mécanique (classique) finit toujours, si l'on attend suffisamment longtemps, par revenir à un état (microscopique) aussi voisin que l'on veut de son état initial. Ce théorème semble exclure toute évolution vers l'équilibre.

Soulignons tout d'abord que, dans le cas d'un système macroscopique, les temps nécessaires pour que le théorème s'applique sont démesurés : l'estimation qu'en a tentée Boltzmann l'a conduit à un ordre de grandeur très supérieur à 10 élevé à la puissance 10^{10}, en années, pour 100 cm^3 de gaz ordinaire.

Il faut en outre comprendre que nous étudions ici l'évolution dans le temps d'un *ensemble statistique,* constitué par un très grand nombre de systèmes préparés de façon identique à l'échelle macroscopique. C'est la distribution de ces systèmes dans l'ensemble qui tend, s'ils possèdent la propriété de mélange, vers la distribution microcanonique. Chacun des systèmes de l'ensemble obéit au théorème de Poincaré, mais leur retour à leur état initial se produit à des moments différents, de sorte que l'ensemble ne revient pas, lui, à la distribution initiale. En d'autres termes, l'état microscopique initial d'un système particulier fait partie des états accessibles; en tant que tel, il doit être à chaque instant représenté dans l'ensemble statistique (mais à égalité avec tous les autres lorsque l'équilibre microcanonique est atteint); que ce soit le même système de l'ensemble statistique ou un autre qui se trouve à tel instant dans cet état microscopique est sans importance du point de vue des probabilités.

a. Définition physique

Pour le comprendre, imaginons la situation suivante, qui schématise les caractéristiques d'une situation expérimentale réelle : le sous-espace accessible $\mathcal{E}_E$ de l'espace des phases est divisé en un certain nombre (pas nécessairement grand) de domaines $\mathcal{E}_0, \mathcal{E}_1, \ldots, \mathcal{E}_n$ qui en constituent une partition; les appareils de mesure dont on dispose permettent seulement de déterminer lequel de ces domaines contient le point représentatif de l'état (microscopique) du système, sans plus de précision. On peut en donner un exemple simple : n particules sont enfermées dans une boîte, et on peut seulement savoir, à un instant donné, combien d'entre elles se trouvent dans l'une des moitiés de la boîte (leurs vitesses n'étant pas observées). Ainsi, chaque mesure donne pour résultat un nombre entier, le numéro k du domaine $\mathcal{E}_k$ dans lequel se trouve le point représentatif du système (ou, dans l'exemple des n particules, le nombre k de celles qui sont situées dans la moitié de la boîte que l'on examine).

Supposons que l'on effectue la mesure à intervalles réguliers, c'est-à-dire aux instants $0, \tau, 2\tau, \ldots$ On obtient ainsi une suite d'entiers, tous compris entre 0 et n. On s'attend à ce que les divers nombres de cette suite soient corrélés entre eux. En effet, les équations du mouvement microscopique déterminent l'état du système à un instant donné s'il est connu à un instant antérieur quelconque. Certes, les mesures effectuées ici sont grossières, et ne permettent donc pas des prédictions précises sur les états ultérieurs. Mais elles sont répétées périodiquement, de sorte que les informations qu'elles fournissent s'affinent en s'accumulant. A la limite, pour certains systèmes au moins, on doit pouvoir choisir la partition en domaines $\{\mathcal{E}_0, \mathcal{E}_1, \ldots, \mathcal{E}_n\}$ et l'intervalle de temps τ entre deux mesures successives de façon que la connaissance des premiers nombres de la suite détermine parfaitement les suivants[7]. Autrement dit, si $\delta k(q)$ est l'incertitude sur le résultat de la q-ième mesure lorsqu'on tient compte des $(q-1)$ résultats antérieurs, on s'attend à ce que $\delta k(q)$ tende vers zéro quand q tend vers l'infini, au moins pour certains choix de la partition $\{\mathcal{E}_k\}$ et de la périodicité τ des mesures.

Pour un « K-système », au contraire, $\delta k(q)$ reste toujours bornée inférieurement par un nombre au moins égal à 1, pour tout choix de la partition en domaines et tout intervalle τ. Le comportement des « K-systèmes » est donc intrinsèquement et fondamentalement aléatoire.

Celui d'un *« système de Bernoulli »* l'est davantage encore. Pour un tel système, il existe au moins une partition $\{\mathcal{E}_0, \mathcal{E}_1, \ldots, \mathcal{E}_n\}$ possédant la propriété suivante : la suite de résultats obtenue dans des mesures successives sur un système donné est constituée de *nombres totalement incorrélés,* comme celle que fournirait un dé ou une roulette de casino. Notons bien que seules certaines partitions bien particulières possèdent cette propriété; quant aux autres partitions possibles, elles vérifient toutes la condition qui définit un K-système. On peut en effet démontrer qu'un système de Bernoulli est nécessairement un K-système, la réciproque n'étant pas vraie.

On a pu prouver qu'une particule ponctuelle qui se meut en rebondissant sur des obstacles convexes fixes (cf. complément IV.B, § I.2), constitue un système de Bernoulli. Mais l'exemple constamment cité est celui de la « transformation du boulanger ».

b. La « transformation du boulanger »

Cet exemple doit sa célébrité à sa simplicité : on y voit clairement et concrètement comment des équations du mouvement parfaitement déterministes et invariantes par renversement du temps (quoique peu réalistes du point de vue physique) produisent une évolution irréversible vers un état totalement aléatoire.

---

7. Pour un oscillateur harmonique par exemple (système ergodique), on peut choisir pour τ la période de son mouvement. Dans ce cas, la mesure donnera à chaque fois le même résultat, puisque l'oscillateur revient dans le même état aux instants $\tau, 2\tau, \ldots$

Considérons un système dont l'état est caractérisé par une seule coordonnée généralisée q et son moment conjugué p, que nous prendrons sans dimensions; l'espace des phases est limité au carré

$$0 \leqslant q < 1 , \tag{A.6.a}$$

$$0 \leqslant p < 1 . \tag{A.6.b}$$

La variable de temps t est discrétisée : les instants successifs sont repérés par la suite des entiers, négatifs, nul, puis positifs. Le mouvement est régi par les équations suivantes :

$$\text{si} \quad 0 \leqslant q(t) < \frac{1}{2} , \qquad \begin{cases} q(t+1) = 2q(t) & \text{(A.7.a)} \\ p(t+1) = \dfrac{1}{2}\, p(t) , & \text{(A.7.b)} \end{cases}$$

$$\text{si} \quad \frac{1}{2} \leqslant q(t) < 1 , \qquad \begin{cases} q(t+1) = 2q(t) - 1 & \text{(A.7.c)} \\ p(t+1) = \dfrac{1}{2}\, p(t) + \dfrac{1}{2} . & \text{(A.7.d)} \end{cases}$$

Ces équations portent le nom de *«transformation du boulanger»* parce qu'elles évoquent la façon dont on pétrit un morceau de pâte (fig. 1).

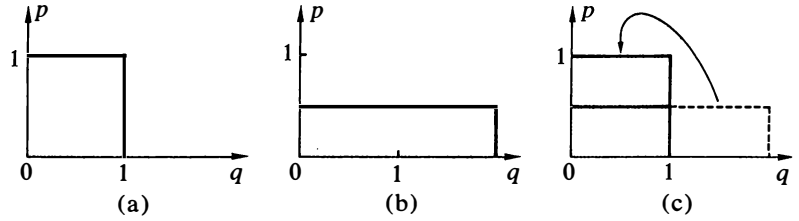


FIGURE 1

La «transformation du boulanger» (formules (A.7)) peut être décomposée en deux étapes : le carré initial (a) est d'abord étiré de façon à multiplier par 2 sa dimension horizontale et diviser par 2 sa dimension verticale :

$$(q, p) \longrightarrow \left(2q, \frac{p}{2}\right) ,$$

ce qui donne le rectangle (b); celui-ci est ensuite scindé en deux parties égales par une coupure verticale, et on amène la moitié droite sur la moitié gauche pour reconstituer finalement le carré (c) :

$$\left(2q, \frac{1}{2}\, p\right) \longrightarrow \begin{cases} \left(2q, \dfrac{1}{2}\, p\right) & \text{si} \quad 2q \leqslant 1 , \\ \left(2q-1, \dfrac{1}{2}\, p + \dfrac{1}{2}\right) & \text{si} \quad 2q > 1 . \end{cases}$$

Les équations du mouvement (A.7) sont *parfaitement déterministes :* si l'on se donne l'état $\{q(t), p(t)\}$ à un instant t, l'état $\{q(t+1), p(t+1)\}$ à l'instant suivant est déterminé de façon unique, et donc de proche en proche l'état à tous les instants ultérieurs[8]. Ces équations sont en outre *invariantes par renversement du temps.* En effet, si l'on schématise la transformation (A.7) par

$$\{q(t), p(t)\} \longmapsto \{q(t+1), p(t+1)\} , \tag{A.8}$$

8. Notons que la transformation (A.7) possède également une propriété importante des équations du mouvement en mécanique classique (appendice III, paragraphe II.3.b) : elle conserve le «volume» dans l'espace des phases, c'est-à-dire ici l'aire $dq\,.\,dp$ puisque cet espace est à deux dimensions.

il est facile de voir qu'inversement

$$\{p(t+1),\ q(t+1)\} \longmapsto \{p(t),\ q(t)\}\ , \tag{A.9}$$

où la flèche symbolise les *mêmes formules* dans les deux cas[9]. Donc, si un mouvement du système l'amène de $(q_0,\ p_0)$ à l'instant $t=0$ à $(q_n,\ p_n)$ à l'instant $t=n$, *le mouvement inverse est régi par les mêmes équations du mouvement* : il suffit de prendre $(p_n,\ q_n)$ comme état initial et d'itérer à nouveau n fois la transformation du boulanger pour aboutir à l'état $(p_0,\ q_0)$[10].

Pour montrer qu'un système régi par la transformation du boulanger est un système de Bernoulli, commençons par écrire q et p en *notation binaire* :

$$q = 0,\ \alpha_1\alpha_2\alpha_3\ \ldots\ , \tag{A.10.a}$$

$$p = 0,\ \beta_1\beta_2\beta_3\ \ldots\ , \tag{A.10.b}$$

où les α_i et les β_i sont des *chiffres,* égaux à 0 ou 1. On se convainc aisément[11] que, dans cette notation, la transformation (A.7) consiste simplement à enlever à $q(t)$ le premier chiffre après la virgule et à le reporter à la même place dans $p(t)$: si $q(t)$ et $p(t)$ sont écrits sous la forme (A.10), alors

$$q(t+1) = 0,\ \alpha_2\alpha_3\ \ldots\ , \tag{A.11.a}$$

$$p(t+1) = 0,\ \alpha_1\beta_1\beta_2\beta_3\ \ldots\ . \tag{A.11.b}$$

---

9. Les formules (A.7) s'inversent facilement pour donner $q(t)$ et $p(t)$ à partir de $q(t+1)$ et $p(t+1)$:

$$\begin{cases} q(t) = \dfrac{1}{2}\, q(t+1) \\ p(t) = 2p(t+1) \end{cases} \quad \text{si} \quad 0 \leqslant p(t+1) < \frac{1}{2}\ , \tag{1.a}$$

et

$$\begin{cases} q(t) = \dfrac{1}{2}\, q(t+1) + \dfrac{1}{2} \\ p(t) = 2p(t+1) - 1 \end{cases} \quad \text{si} \quad \frac{1}{2} \leqslant p(t+1) < 1\ . \tag{1.b}$$

Écrivons ces formules inverses sous la forme

$$\text{si} \quad 0 \leqslant p(t+1) < \frac{1}{2}\ , \quad \begin{cases} p(t) = 2p(t+1) \\ q(t) = \dfrac{1}{2}\, q(t+1)\ , \end{cases} \tag{2.a}$$

$$\text{si} \quad \frac{1}{2} \leqslant p(t+1) < 1\ , \quad \begin{cases} p(t) = 2p(t+1) - 1 \\ q(t) = \dfrac{1}{2}\, q(t+1) + \dfrac{1}{2}\ . \end{cases} \tag{2.b}$$

On reconnaît alors la *même transformation* qu'en (A.7), avec simplement les changements

$$\begin{aligned} q(t) &\longleftrightarrow p(t+1)\ , \\ p(t) &\longleftrightarrow q(t+1)\ . \end{aligned} \tag{3}$$

10. L'interversion de q_n et p_n est l'analogue du changement de sens de la vitesse qu'il faut effectuer pour faire rebrousser chemin à une particule sur sa trajectoire (voir IV, § II.B.4.b).

11. En notation binaire, le nombre deux s'écrit 10. Multiplier un nombre par deux revient donc simplement à déplacer la virgule d'un cran vers la droite sans modifier les chiffres; diviser par deux revient à la déplacer d'un cran vers la gauche.

Si $q(t)$ est compris entre zéro et un demi, le chiffre α_1 de son écriture binaire (A.10.a) est 0; la formule (A.7.a) conduit bien au résultat (A.11.a), et la formule (A.7.b) au résultat (A.11.b), où α_1 est 0. Si $q(t)$ est compris entre un demi et un, le chiffre α_1 est 1; mais comme on applique maintenant les formules (A.7.c) et (A.7.d), les expressions (A.11) sont encore valables.

Divisons maintenant l'espace des phases du système en *deux domaines* $\mathcal{E}_o$ et $\mathcal{E}_1$, définis par

$$\begin{aligned} &(q, p) \in \mathcal{E}_0 \quad \text{si} \quad 0 \leqslant q < \frac{1}{2}, \\ &(q, p) \in \mathcal{E}_1 \quad \text{si} \quad \frac{1}{2} \leqslant q < 1. \end{aligned} \tag{A.12}$$

Autrement dit, $\mathcal{E}_0$ est la moitié gauche du carré unité constituant l'espace des phases accessible, $\mathcal{E}_1$ la moitié droite. En notation binaire, il suffit de regarder le *premier chiffre de q après la virgule* : si ce chiffre est 0, le point (q, p) est dans $\mathcal{E}_0$; si c'est 1, le point est dans $\mathcal{E}_1$. Supposons que nous sachions par exemple que, à l'instant t, le système se trouve dans $\mathcal{E}_1$: α_1 vaut alors 1. Mais ceci n'autorise aucune prédiction concernant la position du système à l'instant $t+1$, puisque celle-ci est déterminée d'après (A.11) par le chiffre α_2, sur lequel nous n'avons pas d'information. Ainsi, si l'on « mesure » aux instants successifs dans lequel des deux domaines se trouve le point représentatif du système, les résultats obtenus ne présentent entre eux aucune corrélation : à chaque nouvelle mesure, les deux résultats possibles sont également probables, comme dans le jeu de pile ou face. On a donc bien affaire à un système de Bernoulli.

On peut également voir dans cet exemple simple comment un état initial (q_0, p_0), s'il n'est pas connu avec une précision infinie, évolue vers l'état d'équilibre macroscopique. Supposons que la coordonnée q_0 soit donnée, en notation binaire, avec m chiffres significatifs. L'évolution sera alors déterministe et réversible[12] entre $t=0$ et $t=m-1$, car la transformation (A.11) pourra être itérée $(m-1)$ fois sans ambiguïté (quoique la précision sur la coordonnée q diminue à chaque fois). A partir de $t=m$, la valeur de q deviendra totalement indéterminée. Celle de p, qui curieusement gagne en précision entre $t=0$ et $t=m$[13], devient ensuite elle aussi aléatoire puisque les premiers chiffres après la virgule ne sont plus connus. L'état obtenu ne peut plus être décrit qu'en termes probabilistes, et la distribution de probabilité devient bientôt uniforme sur le carré accessible de l'espace des phases. Il n'est plus alors possible de retrouver l'état initial par renversement du temps.

III. LE PROBLÈME ERGODIQUE EN MÉCANIQUE QUANTIQUE

Le problème ergodique se présente de façon *radicalement différente* en mécanique quantique. Les propriétés que nous avons décrites au paragraphe II et la « hiérarchie ergodique » qui en résulte pour les systèmes classiques ne paraissent pas avoir d'équivalent quantique. Ceci est surprenant, mais surtout très préoccupant pour

12. Inverser le sens du temps revient à échanger les rôles de q et p (voir notes 9 et 10) : pour « remonter le temps » d'une unité, il suffit d'ôter le premier chiffre de p après la virgule et de le reporter à cette même place dans q. Tant que les chiffres significatifs de q_0 ne sont pas épuisés, on peut ainsi revenir en arrière (voir cependant la note 13).

13. Ce gain de précision sur p est en réalité illusoire. En effet, les incertitudes δq et δp sur les valeurs de q et p sont nécessairement non nulles, si l'on veut que cet exemple académique ait quelques traits réalistes (autrement dit, l'espace des phases est « granulé », cf. note 5). Dans les itérations successives du procédé (A.11), les chiffres significatifs de p_0, puis ceux de q_0, sont progressivement perdus quand ils entrent dans la zone affectée par l'incertitude δp.

la théorie ergodique : c'est la mécanique quantique, nous l'avons souligné à maintes reprises, qui régit de façon générale l'évolution des systèmes physiques à l'échelle microscopique; le comportement aléatoire et l'évolution irréversible des systèmes macroscopiques ne seront donc vraiment compris que si l'on parvient à les décrire et les expliquer dans le cadre quantique.

La première difficulté que l'on rencontre lorsqu'on tente de le faire provient de ce que *l'évolution d'un système quantique fini est «quasi-périodique», ou «récurrente».* Les niveaux d'énergie d'un tel système sont en effet *discrets,* même s'ils sont très resserrés pour un système de taille macroscopique. La valeur moyenne quantique d'une grandeur physique quelconque s'écrit alors comme une somme de termes périodiques, oscillant dans le temps aux diverses *«fréquences de Bohr»* $|E_i - E_j|/h$ associées aux différences entre deux valeurs permises E_i et E_j de l'énergie. On peut démontrer de façon générale (réf. 4, appendice D) un théorème selon lequel le système évolue nécessairement de manière récurrente, revenant au bout d'un temps fini à son état initial. Ce résultat est l'analogue quantique du théorème de Poincaré que nous avons évoqué pour les systèmes classiques (note 6). Il est cependant beaucoup plus contraignant ici : alors qu'en mécanique classique il concerne seulement l'état microscopique d'un système particulier, il est valable en mécanique quantique pour les états purs mais aussi pour les mélanges statistiques[14]. C'est la distribution statistique des états microscopiques (c'est-à-dire l'ensemble statistique) qui doit se reproduire quasiment à l'identique après un certain laps de temps; elle ne peut donc pas évoluer vers la distribution d'équilibre. Certes, pour les systèmes de taille macroscopique, les temps dont il est question sont démesurés, supérieurs à l'âge actuel de l'Univers. Il n'en reste pas moins que l'existence de ce théorème de récurrence complique considérablement la compréhension du comportement des systèmes quantiques en mécanique statistique.

Toutefois, les résultats que nous venons d'évoquer supposent que le système est *totalement isolé,* c'est-à-dire que son hamiltonien est parfaitement bien défini et invariable. Or tout système macroscopique est soumis à d'innombrables influences extérieures très faibles et impossibles à répertorier de façon exhaustive, mais qui n'en sont pas moins capables de modifier profondément son comportement (voir complément IV.B). Il semble que ce sont ces perturbations incontrôlables et aléatoires qui conduisent à l'évolution irréversible vers l'équilibre et à l'ergodicité. C'est ce qu'indiquent certains modèles simples, bien qu'aucune démonstration générale n'existe actuellement de cette conjecture.

Références

1. J. L. Lebowitz et O. Penrose, *Physics Today,* février 1973, p. 23.
2. J. L. Lebowitz, «Ergodic Theory and Statistical Mechanics», in *Transport Phenomena* (Sitges International School of Statistical Mechanics, juin 1974), p. 202.
3. Ya. G. Sinai, «Ergodic Theory», in *The Boltzmann Equation : Theory and Applications* (E. G. D. Cohen et W. Thirrings eds), Springer-Verlag (1973), p. 575.
4. A. Hobson, *Concepts in Statistical Mechanics,* Gordon and Breach, Londres (1971).
5. G. Casati, «La Théorie Ergodique et le Comportement Stochastique dans les Systèmes non linéaires», in *La Pensée physique contemporaine,* Fresnel, Paris (1982).

14. Cette différence essentielle provient apparemment de ce que les distributions de probabilité sont continues en mécanique classique (densités de probabilité dans l'espace des phases) et discrètes en mécanique quantique puisque la base des états propres de l'hamiltonien dans l'espace des états est dénombrable.

COMPLÉMENT IV.B

Perturbations aléatoires d'un système « isolé » : discussion physique et ordres de grandeur

I. Systèmes classiques
1. Quelques données numériques sur un gaz ordinaire
2. Évaluation de l'influence d'une perturbation extérieure
3. Conclusion

II. Systèmes quantiques
1. La relation fondamentale à utiliser
2. Exemple : le gaz d'électrons de conduction dans un métal
3. Évaluation de l'influence d'une perturbation extérieure
4. Conclusion

Nous invoquons à plusieurs reprises dans ce livre (notamment au IV, § I.A.1 et à la fin du complément IV.A) l'influence de perturbations extérieures très faibles sur l'évolution d'un système macroscopique « isolé », dont l'énergie est en principe fixée mais n'est en réalité définie qu'avec une incertitude δE non nulle. Nous nous proposons de préciser ici les ordres de grandeur caractéristiques de ces effets.

I. SYSTÈMES CLASSIQUES

Intéressons-nous tout d'abord à des systèmes dont le comportement peut être compris dans le cadre de la mécanique classique : ceci nous permettra de raisonner plus intuitivement.

1. QUELQUES DONNÉES NUMÉRIQUES SUR UN GAZ ORDINAIRE

Considérons un échantillon de gaz, par exemple un litre d'azote dans les conditions normales : pression d'une atmosphère et température de 0 °C. La masse molaire de l'azote est 28 g.

On sait qu'une mole de gaz, dans les conditions normales, occupe un volume de 22,4 l. Le système que nous étudions comporte donc environ

$$N = \frac{N_A}{22{,}4} \sim 3 . 10^{22} \tag{B.1}$$

molécules, de masse

$$m = \frac{28}{N_A} \text{ g} \sim 5 . 10^{-23} \text{ g} . \tag{B.2}$$

En moyenne, deux molécules voisines sont distantes l'une de l'autre de

$$d = \left(\frac{V}{N}\right)^{1/3} \sim 30 \text{ Å} . \tag{B.3}$$

L'ordre de grandeur des *vitesses d'agitation thermique* peut être évalué à partir du théorème d'équipartition de l'énergie (III, § V.C) :

$$\frac{1}{2} m\overline{v^2} = \frac{3}{2} kT ; \tag{B.4}$$

pour $T \sim 300$ K, on trouve

$$v \sim \sqrt{\overline{v^2}} \sim 500 \text{ m/s} . \tag{B.5}$$

Soit σ la section efficace de collision entre deux molécules[1]. Un raisonnement simple[2] montre que le *nombre de chocs par unité de temps,* pour une molécule donnée, est de l'ordre de

$$n = \frac{N\sigma v}{V}, \tag{B.6}$$

et le *libre parcours moyen* ℓ d'une molécule, c'est-à-dire la valeur moyenne de la distance qu'elle parcourt en ligne droite entre deux chocs consécutifs, est donné par

$$\ell = \frac{V}{N\sigma}. \tag{B.7}$$

Pour déterminer l'ordre de grandeur de la section efficace σ, imaginons des molécules sphériques de rayon ρ_0 et supposons qu'elles se comportent comme des boules de billard; deux telles molécules entrent en collision si et seulement si leur mouvement relatif tend à amener leurs centres à une distance inférieure à $2\rho_0$ (fig. 1). On voit que, dans ce cas,

$$\sigma \sim \pi(2\rho_0)^2. \tag{B.8}$$

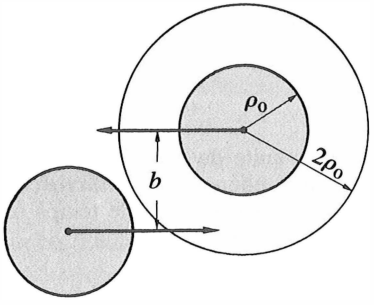


FIGURE 1

Évaluation de la section efficace de collision pour deux molécules sphériques qui se comporteraient comme des boules de billard : dans le repère du centre de masse, il y a collision seulement si le « paramètre d'impact » b est inférieur à $2\rho_0$.

1. La section efficace de collision est définie à partir de la situation standard suivante : un faisceau de particules incidentes, de flux f (nombre de particules franchissant par unité de temps une surface unité perpendiculaire à la direction du faisceau), entre en collision avec une cible fixe, comprenant N particules. On suppose que chaque collision affecte seulement une particule du faisceau et une de la cible, et que les diverses collisions sont indépendantes les unes des autres. Dans ces conditions, le nombre n de collisions par unité de temps s'écrit, par définition de la section efficace σ :

$$n \mathrel{\widehat{=}} Nf\sigma.$$

La section efficace a donc les dimensions d'une surface.

2. Fixons notre attention sur une molécule particulière. Pendant le temps δt, elle parcourt une distance de l'ordre de $v\delta t$, et l'on peut considérer qu'elle balaie un domaine de l'espace de volume $\sigma v\delta t$, de telle sorte qu'elle entre en collision avec les autres molécules dont le centre est situé dans ce domaine. Ces dernières sont au nombre de

$$\sigma v\delta t \frac{N}{V}. \tag{1}$$

On en déduit le nombre de chocs par unité de temps : $n = \sigma v \frac{N}{V}$, (2)

et le libre parcours moyen (qui correspond a une seule collision) : $\ell = \frac{v\delta t}{n\delta t} = \frac{v}{n} = \frac{V}{N\sigma}$. (3)

Les molécules d'azote ne sont pas sphériques, et leurs interactions sont plus complexes que celles de boules de billard; on obtient cependant un ordre de grandeur réaliste en prenant

$$\rho_0 \sim 1{,}5 \text{ Å} . \tag{B.9}$$

Il vient alors

$$\ell \sim 1\,400 \text{ Å} \tag{B.10}$$

et

$$n \sim 4.10^9 \text{ collisions par seconde.} \tag{B.11}$$

Bien qu'environ 50 fois supérieur à la distance moyenne entre molécules voisines, le libre parcours moyen reste petit à l'échelle macroscopique; comme les vitesses d'agitation thermique sont assez considérables, il en résulte qu'une molécule donnée subit chaque seconde un *nombre énorme de chocs.*

Évaluons également l'*énergie* de notre échantillon de gaz. L'énergie potentielle d'interaction entre les molécules est faible devant leur énergie cinétique (complément I.A, § III.3). Pour des molécules diatomiques comme celles de l'azote, une énergie cinétique de rotation s'ajoute à l'énergie cinétique de translation (voir complément III.B. Les vibrations des molécules d'azote ne sont pratiquement pas excitées à la température ambiante). Ceci donne pour valeur approximative de l'énergie totale

$$E \simeq N \times \frac{5}{2} kT , \tag{B.12}$$

c'est-à-dire, numériquement,

$$E \sim 250 \text{ J} . \tag{B.13}$$

Enfin, il sera utile pour la suite de connaître l'ordre de grandeur des *intervalles de temps* τ_M caractéristiques des phénomènes *macroscopiques* pouvant se produire dans un gaz ordinaire. On prendra par exemple le temps mis par le son pour traverser le récipient contenant l'échantillon d'azote :

$$\tau_M \sim \frac{0{,}1 \text{ m}}{340 \text{ m/s}} \sim 3.10^{-4} \text{ s} . \tag{B.14}$$

Les durées caractéristiques d'expériences ou d'observations courantes que l'on peut effectuer sur un tel échantillon sont, quant à elles, plutôt de l'ordre de la seconde, ou supérieures.

2. ÉVALUATION DE L'INFLUENCE D'UNE PERTURBATION EXTÉRIEURE

Tentons maintenant de réaliser un *système isolé* à partir de l'échantillon d'azote considéré. Nous l'enfermons pour cela dans un récipient à parois rigides et adiabatiques. Toutefois, ces parois ne sont jamais parfaites, c'est-à-dire que le travail des forces de pression (dû à des variations de volume du système) et les échanges de chaleur avec l'extérieur ne sont jamais rigoureusement nuls. Le système peut également être plongé dans un champ électromagnétique dont les sources lui sont extérieures, et qui est susceptible de fournir de l'énergie aux molécules[3]. S'il s'agit d'un champ électromagné-

3. Les mécanismes d'échange d'énergie entre un ensemble de molécules et un champ électromagnétique sont multiples. Une onde électromagnétique de fréquence appropriée peut par exemple faire passer une molécule de son niveau fondamental à l'un de ses niveaux excités (excitation des vibrations de la molécule ou excitation de son cortège électronique). Un champ électrique, même statique, oriente les molécules polaires (c'est-à-dire possèdant un moment dipolaire permanent, ce qui n'est pas le cas des molécules d'azote) et polarise les molécules non polaires. Un champ magnétique interagit de façon analogue avec les moments magnétiques, soit que ceux-ci préexistent dans les atomes ou molécules considérés (corps paramagnétiques), soit que le champ magnétique les induise lui-même (diamagnétisme).

tique variable au cours du temps (onde électromagnétique), on peut en protéger le système en l'enfermant dans une boîte à parois conductrices, qui réfléchissent et absorbent les ondes électromagnétiques[4]. Le même dispositif annule à l'intérieur tout champ électrique statique produit par des charges extérieures (théorème des écrans électriques). Il peut subsister un champ magnétique statique, mais il est en principe possible de le compenser en produisant un champ exactement opposé à l'aide de circuits électriques judicieusement choisis et disposés à l'extérieur du système. Sauf situations vraiment exceptionnelles, les systèmes que l'on considère couramment comme « isolés » ne sont pas entourés de tant de soins et de précautions, de sorte que leurs échanges d'énergie avec l'extérieur ne sont pas vraiment nuls, bien qu'ils soient le plus souvent négligeables.

Mais il existe un type d'interactions qu'il est impossible d'écranter ou de compenser : les *interactions gravitationnelles.* Elles sont cependant très faibles dans les systèmes considérés habituellement en mécanique statistique, et ne prennent de l'importance que dans les conditions astrophysiques (étoiles, galaxies, Univers, ...)[5]. Pour donner une idée de l'importance que peuvent avoir des perturbations extérieures faibles et incontrôlées sur un système ordinaire, évaluons par exemple l'influence, sur l'échantillon de gaz du paragraphe 1, d'un observateur de masse $M = 75$ kg et situé à $D = 1$ m environ du récipient.

Calculons en premier lieu l'énergie potentielle d'interaction W entre l'observateur et l'échantillon d'azote :

$$W \simeq -\frac{GM \,.\, Nm}{D}\,, \tag{B.15}$$

où $$G = 6{,}67 \,.\, 10^{-11}\ \mathrm{m^3 \,.\, kg^{-1} \,.\, s^{-2}} \tag{B.16}$$

est la constante gravitationnelle. Le résultat,

$$W \sim -6 \times 10^{-12}\ \mathrm{J}\,, \tag{B.17}$$

est totalement négligeable devant l'énergie interne du gaz (formule (B.13)).

Analysons maintenant la perturbation qu'apporte la présence de l'observateur dans le mouvement des molécules du gaz. Considérons pour cela deux molécules, (1) et (2), entrant en collision. Elles sont initialement distantes d'environ ℓ, et se dirigent l'une vers l'autre avec une vitesse de l'ordre de v. Les forces gravitationnelles $\vec{F}_1$ et $\vec{F}_2$ exercées par l'observateur sur ces deux molécules sont légèrement différentes : en ordre de grandeur,

$$F_1 - F_2 \sim \frac{GMm}{D^2}\,\frac{\ell}{D}\,, \tag{B.18}$$

de sorte que le système des deux particules est perturbé par la différence d'accélération

$$\delta\gamma = \gamma_1 - \gamma_2 \sim \frac{GM\ell}{D^3}\,. \tag{B.19}$$

---

4. La boîte conductrice doit entourer l'enceinte rigide et adiabatique, de façon que la chaleur produite par absorption de l'énergie électromagnétique ne se transmette pas au système.

5. Le cas du champ de pesanteur est intermédiaire : il est la plupart du temps négligeable, mais son influence sur certains systèmes particuliers peut être mise en évidence (cf. complément V.B, paragraphe II.1.b).

Cette perturbation décale la position relative des deux molécules au moment du choc (fig. 2) d'une distance

$$\delta\lambda_0 \sim \delta\gamma\, t^2 \sim \frac{GM\ell}{D^3}\left(\frac{\ell}{v}\right)^2 \tag{B.20}$$

(car $t \sim \ell/v$ est le temps pendant lequel agit l'accélération $\delta\gamma$ avant le choc), ce qui correspond à un décalage angulaire

$$\delta\theta_0 \sim \frac{\delta\lambda_0}{\ell} \sim \frac{GM}{D^3}\left(\frac{\ell}{v}\right)^2 . \tag{B.21}$$

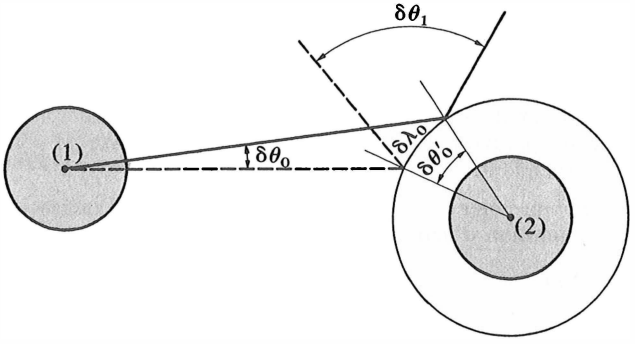


FIGURE 2

Analyse semi-quantitative de la perturbation apportée au choc de deux molécules par une force extérieure. Dans le repère où la molécule (2) est au repos, le centre de la molécule (1) suivrait, en l'absence de force extérieure, la trajectoire représentée en traits tiretés; cette trajectoire est modifiée suivant le trait plein par la perturbation.

Intéressons-nous pour simplifier à la seule molécule (1), et raisonnons comme si les autres étaient au repos. Après le choc, la trajectoire de la molécule (1) est déviée, par rapport à la direction qu'elle aurait eue en l'absence de la perturbation considérée, de[6]

$$\delta\theta_1 \sim \frac{\delta\lambda_0}{\rho_0} \sim \delta\theta_0 \frac{\ell}{\rho_0} , \tag{B.22}$$

où ρ_0 est le rayon des molécules (comme au paragraphe 1, nous assimilons les molécules à des sphères). Pour le choc suivant de la molécule (1), il faudrait combiner la déviation $\delta\theta_0$ due directement à l'action de la force gravitationnelle pendant l'intervalle entre deux chocs et le décalage $\delta\theta_1$ qui vient d'être calculé. Mais, ℓ/ρ_0 étant grand devant 1 (formules (B.9) et (B.10)), l'ordre de grandeur de la perturbation angulaire de la trajectoire après le deuxième choc est donné par

$$\delta\theta_2 \sim \delta\theta_1 \frac{\ell}{\rho_0} \sim \delta\theta_0 \left(\frac{\ell}{\rho_0}\right)^2 ; \tag{B.23}$$

6. Comme ℓ est grand devant ρ_0 (formules (B.9) et (B.10)), l'angle $\delta\theta_0 \sim \delta\lambda_0/\ell$ est petit devant $\delta\theta'_0 \sim \delta\lambda_0/\rho_0$ (voir figure 2). On peut donc raisonner comme si les deux trajectoires incidentes étaient parallèles. L'angle entre les trajectoires réfléchies est alors le double de la différence entre les angles d'incidence, qui vaut $\delta\theta'_0$ (dans ces calculs d'ordre de grandeur, nous « oublions » les facteurs 2).

après le q-ième choc, ce sera donc[7]

$$\delta\theta_q \sim \delta\theta_0 \left(\frac{\ell}{\rho_0}\right)^q . \tag{B.24}$$

On peut considérer que la perturbation étudiée aura complètement « désorienté » la molécule choisie après un nombre q de chocs tel que

$$\delta\theta_q \sim 1 \text{ radian} , \tag{B.25}$$

c'est-à-dire

$$q \sim \frac{\ln(1/\delta\theta_0)}{\ln(\ell/\rho_0)} . \tag{B.26}$$

Numériquement, le résultat est surprenant : bien que $\delta\theta_0$ soit infime :

$$\delta\theta_0 \sim 4.10^{-28} \text{ radian} , \tag{B.27}$$

on trouve[8]

$$q \sim 10 . \tag{B.28}$$

Si l'on remplace, dans ce calcul, l'observateur de 75 kg distant de 1 m par un électron ($M \sim 10^{-30}$ kg) situé aux confins de l'univers ($D \sim 10^{26}$ m), la valeur de $\delta\theta_0$ est évidemment considérablement réduite :

$$\delta\theta_0 \sim 5.10^{-138} \text{ radian} , \tag{B.29}$$

mais celle de q est à peine modifiée :

$$q \sim 50 . \tag{B.30}$$

3. CONCLUSION

Bien que très grossier, le calcul précédent montre de façon concrète la fantastique importance de *l'instabilité des équations dynamiques :* c'est le fait que le nombre q de collisions apparaît en exposant, dans la formule (B.24), qui explique la faiblesse des valeurs trouvées en (B.28) et (B.30). Celles-ci sont à comparer avec le nombre n de chocs que subit une molécule par seconde (formule (B.11)); on voit ainsi

---

7. Le raisonnement très simplifié que nous présentons ici indique que la trajectoire perturbée s'écarte de la trajectoire non perturbée de façon exponentielle en fonction du nombre q de chocs. C'est effectivement la conclusion à laquelle aboutit une étude plus rigoureuse du gaz de sphères dures (réf. 3 du complément IV.A). Notre raisonnement n'est pas une démonstration, mais il met qualitativement en évidence l'influence de l'instabilité des équations dynamiques par rapport aux conditions initiales.

8. L'écart initial (B.21) vaut approximativement

$$\delta\theta_0 \sim 6{,}7.10^{-11}.75 \left(\frac{1{,}4.10^{-7}}{500}\right)^2 \sim 3{,}9.10^{-28} .$$

Comme (cf. (B.9) et (B.10))

$$\frac{\ell}{\rho_0} \sim 950 ,$$

il vient facilement

$$q \sim \frac{\ln 10^{28}}{\ln 10^3} \sim 10 .$$

qu'elles correspondent à des intervalles de temps δt très inférieurs au temps τ_M caractéristique des phénomènes macroscopiques (formule (B.14)) :

$$\delta t \leqslant 10^{-8} \text{ à } 10^{-9} \text{ s} \ll \tau_M \,. \tag{B.31}$$

Avec de tels ordres de grandeur, il n'est pas question de prendre au sérieux l'influence de l'observateur ou celle d'électrons de lointaines galaxies : bien d'autres effets sont susceptibles de se manifester de façon analogue. Nous retiendrons que *tout système classique, si bien «isolé» soit-il, est soumis à de multiples perturbations extérieures, totalement négligeables du point de vue énergétique* (comparer les valeurs (B.17) et (B.13)), *mais qui bouleversent constamment son état microscopique* : même si l'on connaissait, à un instant donné, la position et la vitesse des molécules du gaz, on serait incapable de prédire correctement leur trajectoire après quelques chocs, c'est-à-dire au-delà d'un temps très court, de l'ordre de

$$\tau_m \sim 10^{-9} \text{ s} \ll \tau_M \,. \tag{B.32}$$

Étant donné leurs origines très diverses et incontrôlables, ces perturbations varient elles-mêmes dans le temps, de façon *aléatoire*.

On comprend que de tels effets jouent dans la pratique un rôle primordial pour assurer la validité du «principe ergodique». Imaginons en effet un système macroscopique qui ne serait pas ergodique s'il était parfaitement isolé : préparé dans un état initial bien déterminé, son évolution dans le temps ne l'amènerait à passer que par une classe restreinte d'états microscopiques accessibles. Les multiples perturbations extérieures aléatoires que nous venons d'évoquer modifient sans arrêt cette évolution, sans affecter de façon appréciable la valeur de l'énergie : elles «dévient» le système vers des états microscopiques accessibles qu'il n'aurait pas atteints s'il était resté isolé. Pour les mêmes raisons, ces effets sont également d'une grande importance dans l'évolution d'un système «isolé» vers l'équilibre macroscopique. Bien entendu, ces perturbations ne sont vraiment efficaces que si les équations dynamiques amplifient les déviations infimes qu'elles créent.

II. SYSTÈMES QUANTIQUES

1. LA RELATION FONDAMENTALE A UTILISER

La description quantique des phénomènes à l'échelle microscopique, souvent indispensable, se prête moins bien que la description classique à une analyse semi-quantitative détaillée du comportement des systèmes sous l'effet de perturbations extérieures faibles. En revanche, la *«relation d'incertitude temps-énergie»*

$$\Delta E \,.\, \Delta t \geqslant \hbar \tag{B.33}$$

permet d'associer un intervalle de temps caractéristique

$$\delta t \sim \frac{\hbar}{\delta E} \tag{B.34}$$

à une imprécision δE sur la valeur de l'énergie. Pour le problème qui nous intéresse ici, nous pourrons remplacer δE par l'énergie associée à une perturbation extérieure; δt sera alors le temps au bout duquel cette perturbation aura changé l'état microscopique du système[9].

9. Nous nous contentons ici d'une discussion simple sur les ordres de grandeur. Pour une étude plus détaillée, voir le complément IV.D, paragraphe II.

2. EXEMPLE : LE GAZ D'ÉLECTRONS DE CONDUCTION DANS UN MÉTAL

Comme exemple de système quantique, on peut choisir les électrons de conduction d'un métal tel que le cuivre, dans l'approximation où on les considère comme *indépendants* et *libres* à l'intérieur du métal. Leur nombre par unité de volume est d'environ[10]

$$\frac{N}{V} \simeq 8{,}5 \times 10^{28} \text{ par m}^3 , \tag{B.35}$$

de sorte que, si $V = 1$ cm^3,

$$N \simeq 8{,}5 \times 10^{22} . \tag{B.36}$$

Pour un échantillon cubique d'arête $L = 1$ cm, les énergies possibles pour chacun des électrons sont de la forme (appendice II, § II.4.c)

$$\varepsilon_{\ell_x, \ell_y, \ell_z} = \frac{2\pi^2\hbar^2}{mL^2} (\ell_x^2 + \ell_y^2 + \ell_z^2) - u_0 , \tag{B.37}$$

où ℓ_x, ℓ_y, ℓ_z sont des entiers relatifs, et la constante u_0 (positive) représente l'énergie potentielle qui empêche l'électron de sortir du métal. Nous nous intéresserons seulement au premier terme de l'expression (B.37), c'est-à-dire à l'énergie cinétique de l'électron (quantifiée, puisque celui-ci est confiné dans une région de l'espace bornée).

Pour obtenir les états quantiques du gaz d'électrons, et les énergies correspondantes, il faut tenir compte du *principe de Pauli.* Ainsi, l'état fondamental du système, d'énergie la plus faible possible, est obtenu lorsque les N électrons occupent les N états individuels d'énergies les plus basses : tous les états individuels d'énergie cinétique inférieure au *« niveau de Fermi »* μ_0 sont occupés par les N électrons, tous ceux d'énergie cinétique supérieure à μ_0 sont vides. On montre (VI, § I.B.1) que μ_0 est donné par

$$\mu_0 = \frac{\hbar^2}{2m} \left(3\pi^2 \frac{N}{V}\right)^{2/3} , \tag{B.38}$$

et que l'énergie de l'état fondamental du système vaut

$$E_0 = \frac{3}{5} N\mu_0 - Nu_0 . \tag{B.39}$$

On montre également que, à température ordinaire, l'énergie du système est très proche de cette valeur minimale. On obtient sans difficulté

$$\mu_0 \sim 10^{-18} \text{ J} , \tag{B.40}$$

et on en déduit, à l'aide de (B.36), l'énergie cinétique du gaz d'électrons de conduction dans son état fondamental :

$$\frac{3}{5} N\mu_0 \sim 6 . 10^4 \text{ J} . \tag{B.41}$$

Il nous faut également déterminer un temps τ_M caractéristique des phénomènes macroscopiques auxquels participent les électrons de conduction d'un métal. On peut par exemple transposer ici l'évaluation du paragraphe I.1, c'est-à-dire prendre le temps

---

10. La masse volumique du cuivre vaut environ 9 g/cm^3, et sa masse atomique est 63,5 g. Chaque atome de cuivre fournit un électron de conduction, de sorte que 63,5 g de cuivre en contiennent N_A. D'où

$$\frac{N}{V} = \frac{N_A}{63{,}5} \times 9 \sim 8{,}5 . 10^{22} \text{ par cm}^3 .$$

mis par une onde électromagnétique pour traverser l'échantillon de métal; on trouve ainsi

$$\tau_M \sim \frac{10^{-2}\ \text{m}}{3\times 10^{8}\ \text{m/s}} \sim 3\times 10^{-11}\ \text{s}\ . \tag{B.42}$$

On peut également se référer à la constante de temps RC caractéristique de la décharge d'un condensateur de capacité C dans une résistance R; pour $C\sim 1$ picofarad et $R\sim 1$ ohm, on obtient ainsi 10^{-12} s. Nous prendrons donc

$$\tau_M \sim 10^{-11} \text{ à } 10^{-12}\ \text{s}\ . \tag{B.43}$$

3. ÉVALUATION DE L'INFLUENCE D'UNE PERTURBATION EXTÉRIEURE

Calculons ici aussi l'énergie W d'interaction gravitationnelle entre le système considéré et un observateur de masse $M=75$ kg distant de $D=1$ m :

$$W = -\frac{GMNm}{D}\ , \tag{B.44}$$

soit
$$|W| \sim 4.10^{-16}\ \text{J}\ . \tag{B.45}$$

Cette interaction est susceptible de changer l'état microscopique du système au bout d'un temps de l'ordre de

$$\delta t \sim \frac{\hbar}{|W|} \sim 3.10^{-19}\ \text{s}\ . \tag{B.46}$$

La conclusion est analogue à celle du paragraphe I.3 : bien que l'énergie $|W|$ soit totalement négligeable devant celle du système (cf. (B.41)), la perturbation correspondante peut bouleverser l'état microscopique du gaz d'électrons en un temps très court devant le temps τ_M caractéristique des phénomènes à l'échelle macroscopique (formule (B.43)).

Cependant, contrairement à ce qui se passait dans le cas classique (§ I.2), cette conclusion n'est plus valable pour la perturbation gravitationnelle créée par un électron aux confins de l'Univers : cette fois W est tellement faible :

$$|W| \sim 5.10^{-74}\ \text{J} \tag{B.47}$$

que le δt correspondant devient énorme[11] :

$$\delta t \sim 2.10^{39}\ \text{s}\ ! \tag{B.48}$$

En réalité, apparaissent dans un système quantique macroscopique deux intervalles d'énergie caractéristiques, auxquels correspondent par la relation (B.34) deux échelles de temps.

Il est clair en premier lieu qu'une perturbation W ne peut passer inaperçue par rapport à la valeur de l'énergie du système que si elle est négligeable devant l'incertitude $\delta_0 E$ avec laquelle est connue cette valeur :

$$|W| \ll \delta_0 E\ . \tag{B.49}$$

---

11. Ce n'est plus l'instabilité des équations dynamiques, mais la petitesse de la constante de Planck que nous exploitons dans cette partie II. Ce n'est donc plus le logarithme de la perturbation, mais la perturbation elle-même qui figure dans les formules, ce qui explique la variation incomparablement plus rapide de δt.

A cette borne est en fait associé un temps fantastiquement court : pour des énergies de l'ordre de (B.41), l'incertitude peut difficilement être très inférieure à un millijoule :

$$\delta_0 E \geqslant 10^{-3}\ \text{J}\ , \tag{B.50}$$

ce qui donne

$$\delta_0 t \leqslant 10^{-31}\ \text{s}\ . \tag{B.51}$$

A l'autre extrême se situe l'écart entre deux niveaux consécutifs du système : même s'ils sont très serrés pour un système macroscopique, les niveaux d'énergie sont toujours *discrets*. Il est raisonnable de considérer que cet écart fixe une borne inférieure à l'énergie des perturbations que l'on peut envisager :

$$|W| \geqslant |E_{\ell+1} - E_\ell|\ . \tag{B.52}$$

Dans l'exemple du gaz d'électrons de conduction, l'écart entre le niveau fondamental du système et son premier niveau excité est égal à l'énergie d'excitation minimale d'un électron à partir du niveau de Fermi μ_0, c'est-à-dire d'après (B.37)

$$\frac{2\pi^2\hbar^2}{mL^2} \sim 2.10^{-33}\ \text{J}\ . \tag{B.53}$$

L'intervalle de temps correspondant à cette valeur est de quelques centièmes de seconde, c'est-à-dire beaucoup plus grand que le temps caractéristique τ_M des phénomènes macroscopiques (relation (B.43)).

Ainsi, entre la borne (B.50)-(B.51) et l'évaluation (B.43) du temps caractéristique τ_M des phénomènes macroscopiques s'étend un très large intervalle dans lequel se situent des perturbations modifiant l'état microscopique du système en un temps très court devant τ_M sans affecter de façon décelable la valeur de son énergie.

4. CONCLUSION

En résumé, tout système macroscopique « isolé » est soumis à de multiples perturbations extérieures, incontrôlables et aléatoires parce qu'elles donnent lieu à des échanges d'énergie infimes comparés à l'énergie du système, mais provoquant des changements extrêmement rapides dans l'état microscopique du système. Dans le cas typique du gaz d'électrons de conduction d'un métal, nous avons constaté qu'il existe effectivement un large domaine d'énergie dans lequel se situent de telles perturbations, et nous en avons explicité un exemple (l'influence gravitationnelle de l'observateur sur le gaz d'électrons)[12].

12. Cet exemple montre clairement que les conclusions que nous tirons ici valent seulement pour des systèmes *macroscopiques :* si l'on étudiait un électron unique, par exemple celui d'un atome d'hydrogène, son énergie d'interaction gravitationnelle avec l'observateur serait N fois plus faible que celle qui a été calculée en (B.45); elle serait cette fois infime par rapport aux écarts entre niveaux consécutifs de l'électron dans l'atome :

$$|W| \sim 4.10^{-39}\ \text{J} \simeq 3 \times 10^{-20}\ \text{eV} \ll 1\ \text{eV}\ .$$

Une telle perturbation est donc incapable d'affecter l'état quantique de l'électron.

COMPLÉMENT IV.C

Propriétés de l'équation maîtresse

I. Notations et hypothèses

II. Évolution d'une distribution de probabilités
1. Conservation de la normalisation
2. Conservation de la positivité des composantes

III. Comportement des solutions pour les temps longs
1. Méthode d'analyse
2. Résultats

Dans le chapitre IV, nous avons démontré deux propriétés importantes de l'équation maîtresse : la « relation de bilan détaillé » que vérifie la solution d'équilibre et le « théorème H » qui permet de montrer que le système évolue vers l'équilibre. Nous allons compléter ici cette étude : d'une part, nous allons prouver, ce que nous avons admis jusqu'ici, que la solution de l'équation maîtresse correspondant à une distribution initiale de probabilités reste constamment une distribution de probabilités; d'autre part, nous allons analyser de plus près la façon dont une telle solution tend vers la solution d'équilibre.

I. NOTATIONS ET HYPOTHÈSES

Nous nous limiterons au cas d'un *système isolé,* qui sert de point de départ pour traiter tous les autres. L'énergie est donc fixée (à δE près), et nous n'avons à considérer que les états microscopiques (ℓ) ayant cette énergie et vérifiant les autres contraintes extérieures, c'est-à-dire les *états microscopiques accessibles.* Leur nombre Ω est très grand, mais *fini.*

L'équation maîtresse introduite au chapitre IV s'écrit

$$\frac{dP_\ell}{dt} = \sum_{(m)\neq(\ell)} \left[a_{\ell m} P_m(t) - a_{m\ell} P_\ell(t) \right] . \tag{C.1}$$

Si l'on pose (cf. chapitre IV, note 3)

$$a_{\ell\ell} \mathrel{\widehat{=}} - \sum_{(m)\neq(\ell)} a_{m\ell} , \tag{C.2}$$

on la transforme en

$$\frac{dP_\ell}{dt} = \sum_{(m)} a_{\ell m} P_m(t) . \tag{C.3}$$

Il sera commode dans la suite d'utiliser la *notation matricielle* : si $[P]$ désigne le vecteur-colonne de composantes P_ℓ *et* $[a]$ la matrice d'éléments $a_{\ell m}$, l'équation maîtresse prend la forme[1]

$$\frac{d}{dt}[P] = [a][P] . \tag{C.4}$$

1. On remarquera un petit changement dans les notations : pour que l'on puisse mettre l'équation maîtresse sous forme matricielle, il faut qu'à chaque état microscopique soit associée une valeur unique de l'indice ℓ (ou m), que nous écrivons sans parenthèses; en revanche la notation (ℓ) que nous utilisons habituellement représente un ensemble de plusieurs indices. Un ensemble d'un nombre fini d'indices dénombrables étant lui-même dénombrable, ce changement est sans importance pour les raisonnements qui suivent.

Dans le cas qui nous intéresse (système isolé), le vecteur-colonne $[P]$ est de dimension finie Ω, et $[a]$ est une matrice $\Omega \times \Omega$ dont les éléments vérifient les relations[2]

$$a_{\ell m} \geqslant 0 \quad \text{pour } m \neq \ell \text{ quelconques,} \tag{C.5.a}$$

$$a_{\ell m} = a_{m \ell} \quad \text{pour tout couple } \ell, m, \tag{C.5.b}$$

$$\sum_{\ell} a_{\ell m} = 0 \quad \text{pour tout } m. \tag{C.5.c}$$

Nous supposerons en outre que *tous les états accessibles communiquent entre eux* par la matrice $[a]$: il est *impossible,* en changeant simplement l'ordre des indices, de mettre cette matrice sous forme diagonale par blocs, telle que

$$\begin{bmatrix} [a_1] & [0] \\ [0] & [a_2] \end{bmatrix} \tag{C.6}$$

($[0]$ désigne une matrice, en général rectangulaire, dont tous les éléments sont nuls); nous dirons que la matrice $[a]$ est *irréductible*[3]. Si $[a]$ était réductible à la forme (C.6), cela signifierait que l'on peut diviser l'ensemble des états accessibles en deux sous-ensembles totalement séparés : les probabilités de tous les états appartenant à l'un des sous-ensembles pourraient alors rester constamment nulles (si elles l'étaient à l'instant initial), car ces états ne seraient jamais peuplés à partir de ceux de l'autre sous-ensemble. S'il en était ainsi, le comportement du système ne satisferait pas au *principe ergodique* (I, § IV.B et complément IV.A), qui implique que tous les états microscopiques accessibles sont effectivement atteints par le système. Cela signifierait en réalité que, en énumérant les contraintes extérieures délimitant l'ensemble des états accessibles, on aurait oublié une *loi de conservation* interdisant les transitions entre les deux groupes d'états; il faudrait alors rétablir cette loi dans la caractérisation des états accessibles, ce qui aurait pour effet de choisir l'un des deux sous-ensembles et d'exclure l'autre, c'est-à-dire de revenir à une matrice irréductible, $[a_1]$ ou $[a_2]$.

II. ÉVOLUTION D'UNE DISTRIBUTION DE PROBABILITÉS

L'équation maîtresse (C.4) étant une équation différentielle du premier ordre, sa solution est déterminée de façon univoque par la donnée des composantes de $[P]$ à l'instant initial. Nous supposons donc connue la distribution de probabilités à $t = 0$, c'est-à-dire un ensemble de $P_\ell(0)$ telles que

$$P_\ell(0) \geqslant 0 \quad \text{pour tout } \ell, \tag{C.7.a}$$

$$\sum_{\ell} P_\ell(0) = 1 \, . \tag{C.7.b}$$

---

2. Pour $m \neq \ell$, les coefficients $a_{\ell m}$ sont des probabilités de transition par unité de temps (cf. IV, § I.A.2) et sont donc positifs. La propriété de symétrie (C.5.b), valable pour un système isolé, a été postulée au chapitre IV; elle est démontrée dans le complément IV.D (§ II.2). La relation (C.5.c) découle de la définition (C.2) des éléments diagonaux de la matrice; elle exprime en fait la condition de normalisation des probabilités de transition.

3. *Attention :* nous n'envisageons ici que de changer l'ordre des états, pas d'effectuer un changement de base qui combine linéairement ces états. Lorsqu'on permet de tels changements de base, on peut mettre $[a]$ sous forme bloc-diagonale, et on peut même la diagonaliser complètement (cf. § III.1).

Nous allons montrer que la solution correspondante de l'équation maîtresse reste constamment une distribution de probabilités, c'est-à-dire

$$P_\ell(t) \geqslant 0 \qquad \text{pour tout } \ell \text{ et tout } t > 0, \tag{C.8.a}$$

$$\sum_\ell P_\ell(t) = 1 \qquad \text{pour tout } t > 0. \tag{C.8.b}$$

1. CONSERVATION DE LA NORMALISATION

Montrons en premier lieu que la somme des composantes du vecteur $[P]$ reste constante au cours du temps. Calculons pour cela sa dérivée

$$\frac{\mathrm{d}}{\mathrm{d}t}\left\{\sum_\ell P_\ell(t)\right\} = \sum_\ell \frac{\mathrm{d}}{\mathrm{d}t} P_\ell(t) \,. \tag{C.9}$$

Grâce à l'équation maîtresse (C.3), elle s'écrit

$$\frac{\mathrm{d}}{\mathrm{d}t}\left\{\sum_\ell P_\ell(t)\right\} = \sum_\ell \sum_m a_{\ell m} P_m(t) = \sum_m P_m(t)\left[\sum_\ell a_{\ell m}\right] . \tag{C.10}$$

La somme entre crochets étant nulle pour tout m d'après (C.5.c), la dérivée cherchée l'est également. La condition initiale (C.7.b) entraîne donc (C.8.b).

2. CONSERVATION DE LA POSITIVITÉ DES COMPOSANTES

Supposons que, à l'instant t, certaines des composantes de la solution $[P(t)]$ de l'équation maîtresse soient devenues négatives; nous les noterons $P_n(t)$, et $P_p(t)$ celles qui sont restées positives ou nulles. Formons la somme

$$N(t) \widehat{=} \sum_n P_n(t) \tag{C.11}$$

des composantes négatives. Elle est négative par définition, ou nulle s'il n'existe aucune composante négative :

$$N(t) \leqslant 0 \,. \tag{C.12}$$

Cherchons sa dérivée par rapport au temps :

$$\frac{\mathrm{d}}{\mathrm{d}t} N(t) = \sum_n \frac{\mathrm{d}}{\mathrm{d}t} P_n(t) = \sum_n \sum_\ell a_{n\ell} P_\ell(t) \,. \tag{C.13}$$

Explicitons la somme sur ℓ, c'est-à-dire sur toutes les composantes, en séparant les composantes négatives :

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t} N(t) &= \sum_n \left\{\sum_{n'} a_{nn'} P_{n'} + \sum_{p'} a_{np'} P_{p'}\right\} \\ &= \sum_{n'} P_{n'} \sum_n a_{nn'} + \sum_{p'} P_{p'} \sum_n a_{np'} \,. \end{aligned} \tag{C.14}$$

Utilisons enfin la relation (C.5.c), qui s'écrit dans les notations présentes

$$\sum_n a_{nn'} + \sum_p a_{pn'} = 0 \,. \tag{C.15}$$

Il vient finalement

$$\frac{\mathrm{d}}{\mathrm{d}t} N(t) = -\sum_{n'} P_{n'} \sum_p a_{pn'} + \sum_{p'} P_{p'} \sum_n a_{np'} \,. \tag{C.16}$$

Cette dérivée est donc *positive ou nulle*[4] :

$$\frac{dN}{dt} \geqslant 0 . \tag{C.17}$$

A l'instant initial, toutes les composantes $P_\ell(0)$ sont positives ou nulles, de sorte que

$$N(0) = 0 . \tag{C.18}$$

C'est la valeur maximale que puisse atteindre $N(t)$ (cf. (C.12)). Comme cette fonction ne peut pas décroître d'après (C.17), elle reste nécessairement égale à sa valeur maximale :

$$N(t) \equiv 0 . \tag{C.19}$$

L'hypothèse selon laquelle certaines composantes de $[P(t)]$ sont négatives est donc exclue, ce qui démontre (C.8.a)[5], [6].

III. COMPORTEMENT DES SOLUTIONS POUR LES TEMPS LONGS

1. MÉTHODE D'ANALYSE

Si la matrice $[a]$ était diagonale, la dérivée de P_ℓ serait proportionnelle à P_ℓ elle-même, ce qui donnerait une solution où chaque composante serait une fonction exponentielle du temps. La matrice $[a]$ n'est pas diagonale, mais elle est *diagonalisable par un changement de base,* puisque symétrique et de dimension finie. Nous allons utiliser cette propriété pour résoudre formellement l'équation maîtresse et déterminer le comportement de ses solutions pour t tendant vers l'infini.

---

4. Chacun des deux termes du second membre est positif : les coefficients $a_{p'n'}$ et $a_{n'p'}$ sont forcément non diagonaux et donc positifs ou nuls, et $P_{p'}$ et $P_{n'}$ ont par définition un signe déterminé. La dérivée est nulle si et seulement si aucune composante n'est négative.

5. Il est facile de voir que, si l'on part de conditions initiales où les $P_\ell(0)$ ne sont pas toutes positives ou nulles, on peut obtenir des solutions de l'équation maîtresse dont certaines composantes sont positives ou nulles et d'autres négatives, la répartition entre ces deux catégories variant au cours du temps. Il suffit de prendre deux solutions $[P_1(t)]$ et $[P_2(t)]$ à composantes constamment positives ou nulles, et de considérer $[P_1(t)]-[P_2(t)]$, qui est aussi une solution de l'équation maîtresse, puisque celle-ci est linéaire homogène.

6. Supposons que toutes les composantes de la solution considérée soient positives, et imaginons que l'une d'entre elles, P_{ℓ_0}, s'annule à l'instant t :

$$P_{\ell_0}(t) = 0 . \tag{1}$$

La démonstration qui vient d'être développée implique que P_{ℓ_0} ne peut pas devenir négative. On peut s'en assurer explicitement en examinant la dérivée de P_{ℓ_0} par rapport au temps, donnée par l'équation maîtresse :

$$\frac{dP_{\ell_0}}{dt} = \sum_{\ell} a_{\ell_0 \ell} P_\ell(t) = \sum_{\ell \neq \ell_0} a_{\ell_0 \ell} P_\ell(t) \geqslant 0 . \tag{2}$$

Le seul terme négatif de la somme (celui qui correspond à $\ell = \ell_0$) s'annule d'après (1) à l'instant t, de sorte que la composante P_{ℓ_0} ne peut pas devenir négative.

Nous introduisons donc les *valeurs propres* de la matrice $[a]$; elles sont *réelles*, puisque $[a]$ est hermitique (réelle et symétrique); nous les noterons $-\alpha_r$ car elles vont s'avérer négatives ou nulles, et $[\varphi^r]$ les *vecteurs propres* qui leur sont associés[7] :

$$[a][\varphi^r] = -\alpha_r [\varphi^r] . \tag{C.20}$$

Comme $[a]$ est réelle, symétrique et de dimension finie Ω, elle admet Ω vecteurs propres linéairement indépendants, qui constituent donc une *base* dans l'espace vectoriel de dimension Ω :

$$\{[\varphi^r];\ r = 1, 2, \ldots, \Omega\} . \tag{C.21}$$

Cet espace vectoriel est facilement muni d'un produit scalaire : si $[\psi]$ et $[\chi]$ sont deux vecteurs-colonnes, de composantes ψ_e et χ_m, leur produit scalaire est défini de la façon habituelle[8] :

$$\langle [\psi], [\chi] \rangle \widehat{=} \sum_e \psi_e \chi_e . \tag{C.22}$$

Avec cette définition, on montre facilement (cf. C.D.L., p. 137) que deux vecteurs propres de $[a]$ associés à des valeurs propres différentes sont nécessairement orthogonaux, et que l'on peut choisir l'ensemble des $[\varphi^r]$ orthonormé :

$$\langle [\varphi^r], [\varphi^{r'}] \rangle = \delta_{rr'} . \tag{C.23}$$

Revenons maintenant à l'équation maîtresse. Nous pouvons exprimer sa solution $[P(t)]$ sous forme d'une combinaison linéaire des vecteurs $[\varphi^r]$:

$$[P(t)] = \sum_r c_r(t) [\varphi^r] , \tag{C.24}$$

formule qui s'inverse facilement grâce à (C.23) :

$$c_r(t) = \langle [\varphi^r], [P(t)] \rangle . \tag{C.25}$$

Si l'on reporte l'expression (C.24) dans l'équation maîtresse (C.4) et que l'on tient compte de (C.20) puis de (C.23), on trouve des équations différentielles séparées[9] pour les divers coefficients c_r :

$$\frac{\mathrm{d}}{\mathrm{d}t} c_r = -\alpha_r c_r , \tag{C.26}$$

dont la solution est immédiate :

$$c_r(t) = c_r(0) \mathrm{e}^{-\alpha_r t} . \tag{C.27}$$

---

7. Il peut se faire que plusieurs de ces vecteurs propres correspondent à la même valeur propre (qui est dans ce cas dégénérée). Pour simplifier les notations, nous gardons toujours un seul indice r variant de 1 à Ω, et dirons, s'il y a dégénérescence, que plusieurs valeurs propres $-\alpha_r$ coïncident.

8. Le formalisme que nous utilisons est analogue à celui que la mécanique quantique a rendu familier. La seule différence est que les quantités manipulées ici sont toutes *réelles*.

9. En effet, on a d'une part

$$\frac{\mathrm{d}}{\mathrm{d}t} [P(t)] = \sum_r \frac{\mathrm{d}c_r}{\mathrm{d}t} [\varphi^r] , \tag{1}$$

et d'autre part

$$[a][P(t)] = \sum_r c_r [a][\varphi^r] = -\sum_r c_r \alpha_r [\varphi^r] . \tag{2}$$

Comme les $[\varphi^r]$ forment une base dans l'espace vectoriel utilisé, l'égalité de (1) et (2) implique les Ω égalités (C.26).

Formellement, donc, la solution de l'équation maîtresse correspondant à une distribution initiale $[P(0)]$ donnée s'obtient très facilement si l'on connaît les vecteurs propres et valeurs propres de la matrice $[a]$: on calcule les Ω coefficients

$$c_r(0) = \langle [\varphi^r], [P(0)] \rangle , \tag{C.28}$$

et l'on écrit directement

$$[P(t)] = \sum_r c_r(0)\, e^{-\alpha_r t} [\varphi^r] . \tag{C.29}$$

2. RÉSULTATS

Un premier vecteur propre de $[a]$ et la valeur propre associée découlent immédiatement de la relation (C.5.c) : un vecteur-colonne dont les composantes sont *toutes égales* entre elles est vecteur propre de $[a]$ avec la *valeur propre zéro*[10]. Ce vecteur propre correspond évidemment à la *distribution microcanonique :* on peut le normaliser de telle sorte que ses composantes s'écrivent

$$P^e_\ell = \frac{1}{\Omega} \quad \text{pour tout } \ell. \tag{C.30}$$

Il décrit la *distribution d'équilibre,* puisqu'il est une solution indépendante du temps de l'équation maîtresse.

Nous attribuerons à ce vecteur propre l'indice $r = 1$. Le coefficient correspondant, dans le développement (C.24) de la solution générale $[P(t)]$, est indépendant du temps puisque α_1 est nul (formule (C.27)); il est facile de voir qu'il est toujours égal à 1, quelle que soit la distribution initiale de probabilités[11]; la formule (C.29) s'écrit donc plus précisément

$$[P(t)] = [P^e] + \sum_{r=2}^{\Omega} c_r(0)\, e^{-\alpha_r t} [\varphi^r] . \tag{C.31}$$

---

10. Considérons en effet le vecteur $[\varphi^1]$, de composantes

$$\varphi^1_\ell = C \tag{1}$$

où C est indépendant de ℓ. Appliquons-lui la matrice $[a]$: le vecteur obtenu a pour composantes

$$([a][\varphi^1])_\ell = \sum_m a_{\ell m} \varphi^1_m = C \sum_m a_{\ell m} = 0 \tag{2}$$

d'après la relation (C.5.c). Donc

$$[a][\varphi^1] = 0 \ : \tag{3}$$

$[\varphi^1]$ est vecteur propre de $[a]$ avec la valeur propre zéro.

11. Sous sa forme (C.28), la formule donnant les coefficients $c_r(0)$ suppose que les vecteurs propres $[\varphi^r]$ soient normalisés selon (C.23), et non comme des probabilités; dans le cas qui nous occupe, $[\varphi^1]$ est ainsi normalisé si

$$\varphi^1_\ell = \frac{1}{\sqrt{\Omega}} \quad \text{pour tout } \ell. \tag{1}$$

On a alors

$$c_1(0) = \sum_\ell \varphi^1_\ell P_\ell(0) = \frac{1}{\sqrt{\Omega}} \sum_\ell P_\ell(0) = \frac{1}{\sqrt{\Omega}} \tag{2}$$

(puisque les probabilités initiales sont, elles, normalisées de la façon habituelle). Le premier terme de la somme (C.29) a donc pour composantes

$$c_1(0)\, \varphi^1_\ell = \frac{1}{\Omega} \quad \text{pour tout } \ell \quad = P^e_\ell . \tag{3}$$

Nous allons montrer maintenant que *toutes les autres valeurs propres de la matrice* $[a]$ *sont négatives.* Pour cela, prenons un vecteur-colonne $[\varphi]$ quelconque, de composantes φ_ℓ, et calculons

$$\langle [\varphi], [a][\varphi] \rangle = \sum_{\ell, m} \varphi_\ell a_{\ell m} \varphi_m \, . \tag{C.32}$$

Les propriétés de la matrice $[a]$ permettent de mettre ce produit scalaire sous la forme[12]

$$\langle [\varphi], [a][\varphi] \rangle = -\frac{1}{2} \sum_{\ell, m} a_{\ell m} (\varphi_\ell - \varphi_m)^2 \, , \tag{C.33}$$

où l'on voit aussitôt qu'il est toujours négatif ou nul. En prenant pour $[\varphi]$ l'un quelconque des vecteurs propres de $[a]$ (normalisés selon (C.23)), on en déduit

$$\langle [\varphi^r], [a][\varphi^r] \rangle = -\alpha_r \leqslant 0 \, . \tag{C.34}$$

Mais l'égalité (C.33) implique également que le produit scalaire du premier membre ne peut s'annuler que si chacun des termes, dans la somme du second membre, est séparément nul; comme la matrice $[a]$ est irréductible, ceci ne peut se produire que si toutes les composantes du vecteur $[\varphi]$ sont égales. Autrement dit, dans la formule (C.34), l'égalité à zéro a lieu seulement pour le premier vecteur propre : *seule la première valeur propre, celle qui correspond à la distribution d'équilibre* (C.30), *est nulle; toutes les autres sont négatives.*

Ce résultat remarquable, qui découle des seules propriétés générales (C.5) de la matrice $[a]$, permet d'exploiter la formule (C.31) : les termes dépendant du temps tendent tous vers zéro lorsque t tend vers l'infini, de sorte que

$$[P(t)] \underset{t \to +\infty}{\longrightarrow} [P^e] \, . \tag{C.35}$$

Donc, *quelle que soit la distribution de probabilités initiale, la solution de l'équation maîtresse atteint comme limite, après un temps suffisamment long, la distribution d'équilibre microcanonique.* Ainsi se trouve démontré à nouveau, de façon plus directe que par le théorème H (IV, § II.B.2)[13], le fait qu'un système macroscopique isolé tend vers son état d'équilibre.

---

12. Écrivons d'abord

$$\sum_{\ell, m} \varphi_\ell a_{\ell m} \varphi_m = \sum_{\ell} \varphi_\ell \sum_{m} a_{\ell m} \varphi_m \, . \tag{1}$$

Dans la somme sur m, séparons le terme correspondant à $m = \ell$ et utilisons (C.5.c) pour éliminer $a_{\ell\ell}$:

$$\sum_{\ell, m} \varphi_\ell a_{\ell m} \varphi_m = \sum_{\ell} \varphi_\ell \sum_{m \neq \ell} (a_{\ell m} \varphi_m - a_{m\ell} \varphi_\ell) \, . \tag{2}$$

Dans cette expression, la restriction $m \neq \ell$ n'est pas nécessaire, puisque la quantité à sommer s'annule pour $m = \ell$. En outre, grâce à la symétrie des $a_{\ell m}$, on peut la transformer en

$$\sum_{\ell, m} \varphi_\ell a_{\ell m} \varphi_m = \sum_{\ell, m} \varphi_\ell a_{\ell m} (\varphi_m - \varphi_\ell) \, . \tag{3}$$

On obtient une formule équivalente en intervertissant les rôles des indices muets ℓ et m, puis en utilisant à nouveau la symétrie des coefficients :

$$\sum_{\ell, m} \varphi_\ell a_{\ell m} \varphi_m = \sum_{\ell, m} \varphi_m a_{m\ell} (\varphi_\ell - \varphi_m) = \sum_{\ell, m} \varphi_m a_{\ell m} (\varphi_\ell - \varphi_m) \, . \tag{4}$$

Il suffit alors d'écrire que la quantité étudiée est égale à la demi-somme des deux expressions (3) et (4) pour aboutir à l'égalité (C.33).

13. Remarquer cependant que la démonstration de la note 12, dont le résultat permet d'aboutir à la conclusion (C.35), est très analogue à celle du théorème H et s'appuie comme elle sur la symétrie des coefficients $a_{\ell m}$.

On constate en outre ici que *l'approche vers l'équilibre est exponentielle.* De manière plus précise, supposons que les valeurs propres $-\alpha_r$ de $[a]$ soient ordonnées par valeurs absolues croissantes :

$$\alpha_1 = 0 < \alpha_2 \leqslant \alpha_3 \leqslant \ldots \leqslant \alpha_\Omega \,. \tag{C.36}$$

Dans la formule (C.31), les termes successifs de la somme tendent vers zéro de plus en plus vite, puisque leur décroissance est caractérisée par une suite de temps τ_r tels que

$$\tau_r \widehat{=} \frac{1}{\alpha_r} \,. \tag{C.37}$$

Après une période transitoire au cours de laquelle tous les termes sont présents simultanément puis deviennent tour à tour négligeables, l'approche vers l'équilibre est régie par le terme $r = 2$, qui est le dernier à disparaître[14] :

$$[P(t)] \simeq [P^e] + c_2(0)\, e^{-t/\tau_2} [\varphi^2] \quad \text{pour } t \gg \tau_3. \tag{C.38}$$

Le *temps de relaxation* du système (IV, § II.B.1.b) s'identifie donc avec τ_2; il est déterminé par *la valeur propre de* $[a]$ *la plus petite en valeur absolue,* après $\alpha_1 = 0$.

La valeur propre nulle existe dans tous les cas : c'est une caractéristique générale de l'équation maîtresse, qui découle des propriétés (C.5) définissant la classe des matrices $[a]$ acceptables. Tout système macroscopique tend donc, si l'on attend suffisamment longtemps, vers son état d'équilibre. Cependant, la rapidité de cette évolution dépend des caractéristiques particulières du système considéré : pour évaluer α_2, il faudrait connaître les diverses probabilités de transition par unité de temps $a_{\acute{e}m}$ entre les états du système, et pas seulement leurs propriétés génériques (C.5). L'approche vers l'équilibre peut donc, suivant le système étudié et les perturbations extérieures auxquelles il est soumis, être très rapide ou au contraire extrêmement lente, tous les cas intermédiaires étant possibles (cf. IV, § II.B.1.b.β).

14. Nous supposons que les valeurs propres de $[a]$ forment un spectre partout discret (éventuellement dégénéré). On peut imaginer des cas plus compliqués, où l'approche vers l'équilibre ne serait plus exponentielle.

COMPLÉMENT IV.D

« Postulat d'évolution » et équation de Schrödinger

I. Signification statistique générale de l'équation maîtresse
   1. Notion de processus stochastique
   2. Processus markoviens stationnaires
   3. L'équation maîtresse
   4. Conclusion

II. « Démonstration » de l'équation maîtresse pour un système physique macroscopique
   1. Conditions de validité de l'équation maîtresse
      a. Le problème fondamental : probabilités et amplitudes de probabilité
      b. L'hypothèse du « chaos moléculaire » et la « granulation » du temps
   2. Évaluation approchée des probabilités de transition entre états microscopiques
      a. Notations. Position du problème
      b. Système différentiel équivalent à l'équation de Schrödinger
      c. Solution approchée du système différentiel
      d. Résultat final. Discussion

Nous avons présenté, au début du chapitre IV, l'équation maîtresse et les propriétés de ses coefficients (pour un système isolé) comme un postulat physique. Nous avons cependant indiqué (IV, § I.B.3) que, moyennant certaines hypothèses, ce « postulat d'évolution » peut se déduire de l'équation de Schrödinger qui régit le comportement microscopique du système. Nous allons préciser ici la façon dont peut s'obtenir une telle « démonstration du postulat d'évolution ».

Nous commencerons par expliciter des conditions générales qui, lorsqu'elles sont satisfaites par un système de nature quelconque, permettent d'affirmer que l'évolution de ce système obéit à une équation maîtresse. Nous discuterons ensuite dans quelle mesure les systèmes physiques que nous étudions en mécanique statistique vérifient ces conditions, et nous indiquerons comment les propriétés des coefficients de l'équation maîtresse découlent dans ce cas de l'équation de Schrödinger.

I. SIGNIFICATION STATISTIQUE GÉNÉRALE DE L'ÉQUATION MAITRESSE

1. NOTION DE PROCESSUS STOCHASTIQUE

Considérons un système qui peut se trouver dans l'un des états (ℓ) d'un ensemble discret d'états possibles (les résultats s'étendent sans difficulté aux cas où l'ensemble d'états $\{\alpha\}$ est continu). Nous supposons que la description de ce système est *probabiliste,* et qu'*il évolue au cours du temps.*

Nous noterons ici $P_1(\ell, t)$ la probabilité pour que le système soit dans l'état (ℓ) à l'instant t. Plus généralement, étant donné n instants successifs *ordonnés :*

$$t_1 \leqslant t_2 \leqslant \ldots \leqslant t_n \,, \tag{D.1}$$

$P_n(\ell_1, t_1; \ell_2, t_2; \ldots; \ell_n, t_n)$ sera la *probabilité pour que le système se trouve dans l'état* (ℓ_1) *à l'instant* t_1, *dans l'état* (ℓ_2) *à l'instant* t_2, ..., *et dans l'état* (ℓ_n) *à l'instant* t_n. Sauf cas très particulier, la probabilité composée P_n n'est pas égale au produit de n probabilités P_1 :

$$P_n(\ell_1, t_1; \ell_2, t_2; \ldots; \ell_n, t_n) \neq P_1(\ell_1, t_1).P_1(\ell_2, t_2)\ldots P_1(\ell_n, t_n) \,. \tag{D.2}$$

Si deux instants consécutifs t_i et t_{i+1} coïncident, la probabilité P_n n'est différente de zéro que si les états associés à ces deux instants coïncident aussi : par exemple,

$$P_n(\ell_1, t_2; \ell_2, t_2; \ldots; \ell_n, t_n) = \delta_{\ell_1 \ell_2} P_{n-1}(\ell_2, t_2; \ldots; \ell_n, t_n) . \tag{D.3}$$

L'évolution d'un tel système est appelée un *« processus stochastique »*. De façon générale, un processus stochastique est entièrement connu lorsque l'*ensemble des fonctions* P_n est connu. La probabilité $P_1(\ell, t)$ permet de calculer, à chaque instant, la *valeur moyenne* de n'importe quelle grandeur Y attachée au système :

$$\overline{Y}(t) \widehat{=} \sum_{\ell} P_1(\ell, t)\, Y_\ell \tag{D.4}$$

(Y_ℓ étant la valeur prise par Y dans l'état (ℓ)), ainsi que l'écart quadratique moyen $\Delta Y(t)$ et tous les moments de la distribution statistique de la grandeur Y à l'instant t. En outre, grâce à P_2, P_3, ..., P_n, ..., on a accès aux *corrélations temporelles* entre les diverses grandeurs : par exemple, si Y et Z sont deux grandeurs, qui prennent les valeurs Y_ℓ et Z_ℓ dans l'état (ℓ), leur *« fonction de corrélation »* entre l'instant t_1 et l'instant ultérieur t_2 s'écrit

$$\overline{Y(t_1) Z(t_2)} \widehat{=} \sum_{\ell_1} \sum_{\ell_2} P_2(\ell_1, t_1; \ell_2, t_2)\, Y_{\ell_1} Z_{\ell_2} \tag{D.5}$$

(si Z est identique à Y, on obtient la *« fonction d'autocorrélation »* de la grandeur Y). On calcule de façon analogue, à partir de P_n, des fonctions de corrélation relatives à n instants successifs.

Les fonctions P_n ne sont pas quelconques : comme ce sont des *probabilités,* elles doivent vérifier les *conditions* suivantes :

$$0 \leqslant P_n \leqslant 1 , \quad \forall (\ell_1, t_1; \ell_2, t_2; \ldots; \ell_n, t_n) ; \tag{D.6.a}$$

$$\sum_{\ell} P_1(\ell, t) = 1 , \quad \forall t ; \tag{D.6.b}$$

$$\sum_{\ell_i} P_n(\ell_1, t_1; \ell_2, t_2; \ldots; \ell_i, t_i; \ldots; \ell_n, t_n) = P_{n-1}(\ell_1, t_1; \ldots; \ell_{i-1}, t_{i-1}; \ell_{i+1}, t_{i+1}; \ldots; \ell_n, t_n) , \quad \forall t_i : t_{i-1} \leqslant t_i \leqslant t_{i+1} . \tag{D.6.c}$$

Ces conditions impliquent que les fonctions P_n ne sont *pas indépendantes.*

Il est commode d'introduire également des *probabilités conditionnelles,* que l'on notera $P_{n/q}$: par exemple, $P_{1/q}(\ell, t/\ \ell_1, t_1; \ell_2, t_2; \ldots; \ell_q, t_q)$ est la probabilité pour que le système se trouve dans l'état (ℓ) à l'instant t *sachant qu'il se trouvait dans l'état* (ℓ_1) *à l'instant* t_1, *dans l'état* (ℓ_2) *à l'instant* t_2, ..., *et dans l'état* (ℓ_q) *à l'instant* t_q, avec

$$t_1 \leqslant t_2 \leqslant \ldots \leqslant t_q \leqslant t . \tag{D.7}$$

Les probabilités conditionnelles se déduisent immédiatement des probabilités P_n : par exemple,

$$P_{1/q}(\ell, t/\ \ell_1, t_1; \ell_2, t_2; \ldots; \ell_q, t_q) = \frac{P_{q+1}(\ell_1, t_1; \ell_2, t_2; \ldots; \ell_q, t_q; \ell, t)}{P_q(\ell_1, t_1; \ell_2, t_2; \ldots; \ell_q, t_q)} . \tag{D.8}$$

Si $t = t_q$, on a évidemment

$$P_{1/q}(\ell, t_q/\ \ell_1, t_1; \ldots; \ell_q, t_q) = \delta_{\ell \ell_q} ; \tag{D.9}$$

de plus,

$$\sum_{\ell} P_{1/q}(\ell, t/\ \ell_1, t_1; \ldots; \ell_q, t_q) = 1 . \tag{D.10}$$

2. PROCESSUS MARKOVIENS STATIONNAIRES

Un processus stochastique est dit *markovien* si, pour *toute* suite de $(q+1)$ instants successifs vérifiant les conditions (D.7), on a

$$P_{1/q}(\ell,\ t/\ \ell_1,\ t_1;\ \ell_2,\ t_2;\ \ldots;\ \ell_q,\ t_q) = P_{1/1}(\ell,\ t/\ \ell_q,\ t_q)\ . \tag{D.11}$$

Autrement dit, pour une suite fixée d'instants successifs telle que (D.7), la probabilité conditionnelle $P_{1/q}$ est ici indépendante des états (ℓ_1), (ℓ_2), ..., (ℓ_{q-1}) où se trouvait le système aux instants t_1, t_2, ..., t_{q-1}; elle est déterminée uniquement par l'état (ℓ_q) qu'il avait à l'instant t_q précédant immédiatement l'instant t où l'on cherche à connaître son état (ℓ). *Le fait de fixer l'état* (ℓ_q) *à l'instant* t_q *fait perdre au système toute «mémoire»* des états par lesquels il est passé aux instants antérieurs à t_q.

Un processus markovien est *entièrement caractérisé par deux fonctions*[1] : $P_1(\ell,\ t)$, probabilité pour que le système soit dans l'état (ℓ) à l'instant t, et $P_{1/1}(\ell,\ t/\ \ell',\ t')$, que l'on appelle *«probabilité de transition»* de l'état (ℓ') à l'état (ℓ) entre les instants t' et t.

Les fonctions $P_1(\ell,\ t)$ et $P_{1/1}(\ell,\ t/\ \ell',\ t')$ ne peuvent bien sûr pas être quelconques. Elles doivent d'abord être *constamment comprises entre* 0 *et* 1, et l'on doit avoir, pour $t = t'$,

$$P_{1/1}(\ell,\ t/\ \ell',\ t) = \delta_{\ell\ell'}\ , \quad \forall t\ . \tag{D.12}$$

En outre, la probabilité de transition $P_{1/1}$ doit vérifier une *condition de cohérence,* dite *«équation de Chapman-Kolmogorov» :*

$$P_{1/1}(\ell_3,\ t_3/\ \ell_1,\ t_1) = \sum_{\ell_2} P_{1/1}(\ell_3,\ t_3/\ \ell_2,\ t_2).P_{1/1}(\ell_2,\ t_2/\ \ell_1,\ t_1)\ , \tag{D.13}$$

pour tout t_2 tel que

$$t_1 \leqslant t_2 \leqslant t_3\ ; \tag{D.14}$$

cette condition exprime que, pour aller de l'état (ℓ_1) à l'état (ℓ_3) entre les instants t_1 et t_3, le système doit passer, à un instant intermédiaire t_2, par l'un quelconque des états possibles[2]. Enfin, les deux fonctions doivent être liées par la relation évidente :

$$P_1(\ell_2,\ t_2) = \sum_{\ell_1} P_{1/1}(\ell_2,\ t_2/\ \ell_1,\ t_1).P_1(\ell_1,\ t_1)\ , \quad \forall t_1 \leqslant t_2\ . \tag{D.15}$$

---

1. On peut en effet reconstruire toutes les probabilités P_n à partir des deux fonctions $P_1(\ell,\ t)$ et $P_{1/1}(\ell,\ t/\ \ell',\ t')$. En combinant les égalités (D.8) et (D.11), on obtient une relation de récurrence entre les P_n :

$$P_n(\ell_1,\ t_1;\ \ell_2,\ t_2;\ \ldots;\ \ell_n,\ t_n) = P_{1/1}(\ell_n,\ t_n/\ \ell_{n-1},\ t_{n-1}).P_{n-1}(\ell_1,\ t_1;\ \ldots;\ \ell_{n-1},\ t_{n-1})\ , \tag{1}$$

qui se résout immédiatement :

$$P_2(\ell_1,\ t_1;\ \ell_2,\ t_2) = P_{1/1}(\ell_2,\ t_2/\ \ell_1,\ t_1).P_1(\ell_1,\ t_1)\ , \tag{2.a}$$

$$P_3(\ell_1,\ t_1;\ \ell_2,\ t_2;\ \ell_3,\ t_3) = P_{1/1}(\ell_3,\ t_3/\ \ell_2,\ t_2).P_{1/1}(\ell_2,\ t_2/\ \ell_1,\ t_1).P_1(\ell_1,\ t_1)\ , \tag{2.b}$$

et ainsi de suite.

2. L'équation de Chapman-Kolmogorov (D.13) se déduit aisément des formules (2) de la note 1. Prenons la deuxième d'entre elles (2.b), et sommons ses deux membres sur ℓ_2. D'après (D.6.c), on obtient ainsi la relation

$$P_2(\ell_1,\ t_1;\ \ell_3,\ t_3) = P_1(\ell_1,\ t_1).\sum_{\ell_2} P_{1/1}(\ell_3,\ t_3/\ \ell_2,\ t_2).P_{1/1}(\ell_2,\ t_2/\ \ell_1,\ t_1)\ . \tag{1}$$

Utilisons ensuite l'autre formule, (2.a), pour transformer le premier membre :

$$P_2(\ell_1,\ t_1;\ \ell_3,\ t_3) = P_{1/1}(\ell_3,\ t_3/\ \ell_1,\ t_1).P_1(\ell_1,\ t_1)\ . \tag{2}$$

Il suffit alors de diviser par $P_1(\ell_1,\ t_1)$ pour aboutir à (D.13).

Un processus markovien est dit *stationnaire* si la probabilité de transition $P_{1/1}(\ell, t/\ \ell', t')$ ne dépend que de la différence

$$\tau \widehat{=} t - t' , \tag{D.16}$$

et pas des instants t' et t séparément.

Un *processus markovien stationnaire* est donc caractérisé par la donnée des deux fonctions $P_1(\ell, t)$ et $P_{1/1}(\ell, \tau/\ \ell', 0)$. Dans ce cas, nous reprendrons les notations habituelles et poserons

$$P_\ell(t) \widehat{=} P_1(\ell, t) , \tag{D.17.a}$$

$$\mathcal{T}_{\ell m}(\tau) \widehat{=} P_{1/1}(\ell, t+\tau/\ m, t) \equiv P_{1/1}(\ell, \tau/\ m, 0), \quad \tau \geqslant 0 . \tag{D.17.b}$$

$P_\ell(t)$ est la probabilité pour que le système se trouve dans l'état (ℓ) à l'instant t; on a donc

$$0 \leqslant P_\ell(t) \leqslant 1 , \quad \forall(\ell) \text{ et } t , \tag{D.18.a}$$

$$\sum_\ell P_\ell(t) = 1 , \quad \forall t . \tag{D.18.b}$$

$\mathcal{T}_{\ell m}(\tau)$ est la probabilité de transition de l'état (m) à l'état (ℓ) durant l'intervalle de temps τ; ces probabilités de transition doivent être telles que[3]

$$0 \leqslant \mathcal{T}_{\ell m}(\tau) \leqslant 1 , \quad \forall(\ell), (m), \tau \geqslant 0 ; \tag{D.19.a}$$

$$\mathcal{T}_{\ell m}(0) = \delta_{\ell m} ; \tag{D.19.b}$$

$$\sum_\ell \mathcal{T}_{\ell m}(\tau) = 1 , \quad \forall(m), \tau \geqslant 0 . \tag{D.19.c}$$

En outre, elles doivent satisfaire à l'*équation de Chapman-Kolmogorov* (D.13) qui s'écrit, pour un processus stationnaire et dans les nouvelles notations,

$$\mathcal{T}_{\ell k}(\tau + \tau') = \sum_m \mathcal{T}_{\ell m}(\tau') \mathcal{T}_{mk}(\tau) , \quad \forall(\ell), (k), \tau \geqslant 0, \tau' \geqslant 0 . \tag{D.20}$$

Enfin, les $P_\ell(t)$ et $\mathcal{T}_{\ell m}(\tau)$ doivent vérifier la relation (D.15) :

$$P_\ell(t+\tau) = \sum_m \mathcal{T}_{\ell m}(\tau) P_m(t) , \quad \forall(\ell), t, \tau \geqslant 0 . \tag{D.21}$$

3. L'ÉQUATION MAITRESSE

Nous allons montrer maintenant que les probabilités $P_\ell(t)$ de *tout processus markovien stationnaire*[4] vérifient une équation maîtresse de la forme IV (I.6).

La *première étape* de la démonstration consiste à *remplacer l'équation de Chapman-Kolmogorov* (D.20), valable pour τ et τ' quelconques (non négatifs), *par une équation différentielle,* obtenue en choisissant τ' infiniment petit. Développons pour cela $\mathcal{T}_{\ell m}(\tau')$ près de $\tau' = 0$, en supposant que sa dérivée à l'origine est définie, et non nulle au moins pour certains couples (ℓ, m) : compte tenu de (D.19.b), nous obtenons

$$\mathcal{T}_{\ell m}(\tau') = \delta_{\ell m} + \tau' a_{\ell m} , \quad \text{au premier ordre en } \tau' > 0 . \tag{D.22}$$

Si $\ell \neq m$, le premier terme est nul; $a_{\ell m}$ est alors la *probabilité de transition par unité de temps* de l'état (m) à l'état (ℓ), nécessairement positive :

$$a_{\ell m} \geqslant 0 \quad \text{pour} \quad (\ell) \neq (m) . \tag{D.23.a}$$

---

3. Ces relations s'interprètent très simplement. La dernière, par exemple, exprime que le système, dont on sait qu'il part de l'état (m), aboutit certainement, au bout du temps τ, à l'un quelconque de ses états (ℓ) possibles. *Attention :* la somme sur m de $\mathcal{T}_{\ell m}(\tau)$ n'est pas égale à 1.

4. A condition toutefois que les fonctions $P_\ell(t)$ et $\mathcal{T}_{\ell m}(\tau)$ soient dérivables par rapport à t et τ respectivement.

Pour $m = \ell$, la probabilité $\mathcal{T}_{\ell m}(\tau')$ est proche de 1 (pour τ' petit), mais forcément inférieure, de sorte que

$$a_{\ell\ell} \leqslant 0 \ . \tag{D.23.b}$$

En outre, la condition (D.19.c) implique

$$\sum_{\ell} a_{\ell m} = 0 \ , \quad \forall(m) \ . \tag{D.23.c}$$

Cette dernière relation s'écrit sous une forme plus commode si l'on sépare le terme correspondant à $\ell = m$ dans la somme :

$$\sum_{\ell \neq m} a_{\ell m} = -a_{mm} \ ; \tag{D.24}$$

on peut ainsi n'utiliser que les probabilités de transition par unité de temps, reliant des *états distincts*.

Pour τ' infiniment petit (nous le noterons désormais $\mathrm{d}\tau$), l'équation (D.20) s'écrit

$$\mathcal{T}_{\ell k}(\tau + \mathrm{d}\tau) = \sum_{m} \mathcal{T}_{\ell m}(\mathrm{d}\tau) . \mathcal{T}_{mk}(\tau) \ . \tag{D.25}$$

En utilisant la définition (D.22), on la transforme facilement en l'équation différentielle[5]

$$\frac{\mathrm{d}}{\mathrm{d}\tau} \mathcal{T}_{\ell k}(\tau) = \sum_{m} a_{\ell m} \mathcal{T}_{mk}(\tau) \ . \tag{D.26}$$

Dans une *deuxième étape*, nous allons *éliminer les probabilités de transition* $\mathcal{T}_{\ell k}$ *et* $\mathcal{T}_{mk}$ *et les remplacer par les probabilités* P_ℓ *et* P_m. Il suffit pour cela de multiplier les deux membres de la relation (D.26) par $P_k(t)$ et de sommer sur k. On obtient ainsi[6]

$$\frac{\mathrm{d}}{\mathrm{d}t} P_\ell(t) = \sum_{m} a_{\ell m} P_m(t) \ . \tag{D.27}$$

---

5. On obtient en effet, lorsqu'on reporte (D.22) dans (D.25),

$$\mathcal{T}_{\ell k}(\tau + \mathrm{d}\tau) = \sum_{m} [\delta_{\ell m} + \mathrm{d}\tau . a_{\ell m}] \mathcal{T}_{mk}(\tau) = \mathcal{T}_{\ell k}(\tau) + \mathrm{d}\tau \sum_{m} a_{\ell m} \mathcal{T}_{mk}(\tau) \ . \tag{1}$$

L'équation différentielle (D.26) s'en déduit aussitôt :

$$\frac{1}{\mathrm{d}\tau} [\mathcal{T}_{\ell k}(\tau + \mathrm{d}\tau) - \mathcal{T}_{\ell k}(\tau)] = \sum_{m} a_{\ell m} \mathcal{T}_{mk}(\tau) \ . \tag{2}$$

6. La manipulation indiquée donne en effet

$$\sum_{k} \frac{\mathrm{d}}{\mathrm{d}\tau} \mathcal{T}_{\ell k}(\tau) P_k(t) = \sum_{m} a_{\ell m} \sum_{k} \mathcal{T}_{mk}(\tau) P_k(t) \ . \tag{1}$$

On reconnaît au second membre (cf. (D.21))

$$P_m(t + \tau) = \sum_{k} \mathcal{T}_{mk}(\tau) P_k(t) \ . \tag{2}$$

Pour transformer le premier membre, dérivons par rapport à τ la relation (D.21) :

$$\frac{\partial}{\partial \tau} P_\ell(t + \tau) = \sum_{m} \frac{\mathrm{d}}{\mathrm{d}\tau} \mathcal{T}_{\ell m}(\tau) P_m(t) \ . \tag{3}$$

L'égalité (1) s'écrit donc

$$\frac{\partial}{\partial \tau} P_\ell(t + \tau) = \sum_{m} a_{\ell m} P_m(t + \tau) \ . \tag{4}$$

Il suffit de changer de variable : $t + \tau \rightsquigarrow t$, pour aboutir à l'équation (D.27). (5)

Cette équation différentielle n'est autre que l'*équation maîtresse :* en séparant dans la somme le terme $m = \ell$ et en utilisant la relation (D.24), on la met sous la forme plus « parlante » que nous avons introduite dans le chapitre IV[7],[8] :

$$\frac{\mathrm{d}}{\mathrm{d}t} P_\ell(t) = \sum_m \left[a_{\ell m} P_m(t) - a_{m\ell} P_\ell(t) \right] . \qquad \text{(D.28)}$$

4. CONCLUSION

Nous ne reviendrons pas ici sur l'interprétation très simple de l'équation maîtresse comme *bilan probabiliste* du « peuplement » et du « dépeuplement » des états du système, puisque nous l'avons développée dans le chapitre IV (§ I.B.1). Mais nous insisterons sur les *conditions de validité* de la démonstration précédente : elle suppose que l'évolution dans le temps du système étudié puisse être assimilée à un *processus markovien stationnaire.* Lorsque c'est le cas, l'équation maîtresse donne, si l'on arrive à l'intégrer, *le comportement du système sur des intervalles de temps quelconques,* éventuellement très longs, *à partir des probabilités de transition par unité de temps* entre les divers états, c'est-à-dire des probabilités de transition sur un *intervalle de temps très court* (les ordres de grandeur dépendent évidemment du phénomène que l'on étudie).

II. « DÉMONSTRATION » DE L'ÉQUATION MAITRESSE POUR UN SYSTÈME PHYSIQUE MACROSCOPIQUE

Nous cherchons, pour les systèmes physiques de taille macroscopique, à déduire le « postulat d'évolution », que nous avons posé au chapitre IV (§ I.B), de leur description microscopique. Si nous parvenons à montrer que leur évolution à l'échelle macroscopique est un processus markovien stationnaire, nous pourrons en conclure (grâce aux résultats du paragraphe I) que cette évolution est régie par une équation maîtresse. Cette démonstration est délicate; elle n'est pas à l'heure actuelle parfaitement satisfaisante, même si on peut la trouver suffisamment convaincante. Il faudra ensuite étudier les probabilités de transition par unité de temps entre états microscopiques pour établir les deux propriétés que nous leur avons supposées.

1. CONDITIONS DE VALIDITÉ DE L'ÉQUATION MAITRESSE

Nous nous posons donc en premier lieu la question suivante : *les conditions de validité de l'équation maîtresse,* dans la démonstration que nous en avons donnée au paragraphe I, *sont-elles satisfaites* par l'évolution d'un système physique macroscopique?

---

7. Comme le terme $m = \ell$ de la somme (D.28) est nul, on peut indifféremment considérer que celle-ci porte sur tous les états (m) possibles, ou seulement sur $(m) \neq (\ell)$.

8. Si les états du système forment un continuum repéré par la (ou les) variable(s) α, on remplace $P_\ell(t)$ par $w(\alpha, t)$, *densité de probabilité* pour que l'état du système soit (α) à l'instant t, et $a_{\ell m}$ par $a(\alpha, \alpha')$. La somme sur tous les états possibles devient une intégrale, de sorte que l'équation maîtresse prend la forme

$$\frac{\partial}{\partial t} w(\alpha, t) = \int \mathrm{d}\alpha' \left[a(\alpha, \alpha') w(\alpha', t) - a(\alpha', \alpha) w(\alpha, t) \right]$$

(les diverses étapes de la démonstration se transposent sans difficulté majeure du cas discret au cas continu).

a. Le problème fondamental : probabilités et amplitudes de probabilité

La réponse à cette question est au premier abord *négative : au niveau microscopique, l'évolution dans le temps d'un système décrit par la mécanique quantique n'est pas un processus stochastique;* il ne peut donc *a fortiori* s'agir d'un processus markovien stationnaire.

En effet, l'équation de Schrödinger

$$i\hbar \frac{\mathrm{d}}{\mathrm{d}t} |\psi(t)\rangle = H|\psi(t)\rangle \tag{D.29}$$

qui régit cette évolution est parfaitement *déterministe* : si l'on se donne l'état $|\psi(t_0)\rangle$ du système à l'instant initial t_0, l'équation différentielle du premier ordre (D.29) détermine sans ambiguïté l'état $|\psi(t)\rangle$ à tout instant ultérieur, si l'hamiltonien H est connu. Cette équation est également *invariante par renversement du temps* (nous l'avons montré, dans un cas particulier simple, au chapitre IV, note 7), alors que l'équation maîtresse que nous recherchons ne possède pas cette propriété (cf. IV, § I.B.3.b)[9].

On peut se demander pourquoi la description quantique d'un système, qui fait pourtant intervenir des probabilités, n'est pas de type stochastique. C'est que l'équation de Schrödinger ne fournit pas directement des probabilités, mais des *amplitudes de probabilité.* Plus précisément, l'état $|\psi(t)\rangle$ du système peut à chaque instant être développé sur la base constituée par les états propres $|\ell\rangle$ de l'hamiltonien H :

$$|\psi(t)\rangle = \sum_{\ell} c_\ell(t)|\ell\rangle \,. \tag{D.30}$$

On dit alors parfois que la probabilité pour que le système se trouve dans l'état (ℓ) à l'instant t est proportionnelle à $|c_\ell(t)|^2$. Mais cette probabilité ne se manifeste vraiment en tant que telle que *si l'on effectue une mesure* : lorsque le système est dans l'état $|\psi(t)\rangle$ donné par (D.30), $|c_\ell(t)|^2$ caractérise la probabilité pour qu'une mesure de l'énergie donne la valeur E_ℓ associée à (ℓ) (si cette valeur est non dégénérée); cependant, tant qu'une telle mesure n'est pas effectuée, c'est sur les amplitudes de probabilité $c_\ell(t)$ qu'il convient de raisonner. Ces amplitudes sont des nombres complexes et leurs *différences de phase,* qui varient au cours du temps, jouent un rôle crucial dans l'évolution des propriétés du système. C'est donc l'intervention des amplitudes de probabilité qui distingue fondamentalement la mécanique quantique d'une théorie purement probabiliste.

b. L'hypothèse du « chaos moléculaire » et la « granulation » du temps

Mais souvenons-nous que nous considérons ici des systèmes *macroscopiques;* leur étude ne peut pas et ne doit pas prendre en compte toutes les caractéristiques de leurs propriétés et de leur comportement à l'échelle microscopique (cf. I, § I.B.2). L'équation de Schrödinger et le formalisme quantique, qui s'appliquent au niveau microscopique, donnent beaucoup trop de détails sur l'évolution d'un système macroscopique; il nous faut éliminer presque toute cette *information pléthorique,* pour n'en garder que la faible fraction qui nous intéresse vraiment.

Nous savons également qu'un système macroscopique, si bien isolé soit-il, est constamment soumis à des influences extérieures incontrôlables (complément IV.B).

---

9. Ces conclusions valent aussi, évidemment, pour les systèmes dont l'évolution microscopique peut être décrite par la mécanique classique. Mais le problème se présente alors de façon assez différente, et somme toute un peu plus simple (cf. chapitre IV, paragraphe II.B.4 et complément IV.A). C'est pourquoi nous analysons ici la question dans le cadre de la mécanique quantique.

Dans l'exemple typique d'un gaz ordinaire, dont l'évolution microscopique peut être décrite par la *mécanique classique,* nous avons constaté que ces perturbations, bien que mettant en jeu des énergies extrêmement faibles, bouleversent néanmoins l'état microscopique du gaz en un temps τ à peine supérieur à celui qui sépare en moyenne deux chocs successifs d'une molécule sur ses voisines : les chocs entre molécules, extrêmement nombreux pendant tout intervalle de temps appréciable à l'échelle macroscopique, amplifient exponentiellement l'effet des perturbations extérieures. Ainsi, au bout d'un temps τ très court devant les temps caractéristiques d'une évolution macroscopique du gaz, celui-ci « perd la mémoire » de ses états microscopiques antérieurs.

Dans le cadre de la *description quantique* d'un système macroscopique, on peut également définir un intervalle de temps

$$\tau \sim \frac{\hbar}{|W|}, \tag{D.31}$$

où $|W|$ est une énergie typique associée aux perturbations extérieures incontrôlables qui agissent sur le système. Nous savons (complément IV.B, § II) qu'ici aussi cette durée τ est très courte à l'échelle des phénomènes macroscopiques. On peut alors supposer que les relations de phase que maintiendrait l'équation de Schrödinger entre les états microscopiques successifs du système (§ a) sont détruites, par les multiples perturbations extérieures inconnues, au bout d'un temps de l'ordre de τ : c'est ce que l'on appelle l'*« hypothèse des phases aléatoires »*. Ainsi, non seulement les influences extérieures induisent des transitions entre les états $|\ell\rangle$ (qui seraient des états stationnaires si l'hamiltonien H du système était parfaitement connu), mais en outre ces transitions, si on les analyse sur des temps longs devant τ, peuvent être décrites *directement en termes de probabilités* (puisque les différences de phase entre amplitudes de probabilité sont « brouillées » au bout d'un temps de l'ordre de τ)[10].

On traduit les hypothèses ou résultats précédents en disant qu'un *« chaos moléculaire »* règne dans les systèmes macroscopiques, tant quantiques que classiques. L'idée remonte à Boltzmann lui-même, qui l'a introduite sous le vocable de « Stosszahlansatz ». Nous l'exprimerons de la façon suivante : *un système macroscopique « perd la mémoire » de ses états microscopiques antérieurs au bout d'un temps τ très court par rapport à l'échelle de temps caractéristique de son évolution macroscopique.*

Donc, pour étudier l'évolution du système au niveau macroscopique, on considère seulement des intervalles de temps grands devant τ et l'on s'interdira d'analyser les phénomènes à une échelle de temps inférieure ou de l'ordre de τ : *le temps est en quelque sorte « granulé »* (« coarse-grained » en anglais), les « grains » ayant pour taille τ.

En résumé, l'hypothèse du « chaos moléculaire » implique que *l'évolution du système à l'échelle macroscopique est un processus markovien stationnaire, à condition que tous les intervalles de temps $|t_i - t_j|$ considérés soient grands devant τ* (« granulation » du temps). En effet, puisque les relations de phase entre états quantiques sont constamment détruites (à l'échelle de temps τ), l'évolution macroscopique doit être décrite en termes de probabilités : c'est un processus stochastique. En outre, puisque le système « oublie » régulièrement (au bout d'un temps de l'ordre de τ) les états microscopiques par lesquels il est passé, le processus est markovien. Il est enfin stationnaire, parce que l'équation de Schrödinger (D.29) est invariante par translation dans le temps : la relation entre $|\psi(t)\rangle$ et $|\psi(t_0)\rangle$ fait intervenir seulement l'intervalle de temps $t - t_0$, et non pas t et t_0 séparément.

10. L'état du système n'est plus alors décrit par un vecteur d'état $|\psi\rangle$, mais par un mélange statistique d'états purs (voir chapitre I, paragraphe III.A et complément I.H).

2. ÉVALUATION APPROCHÉE DES PROBABILITÉS DE TRANSITION ENTRE ÉTATS MICROSCOPIQUES

Considérons un système macroscopique isolé (au sens habituel du terme). L'analyse du paragraphe 1 nous a conduits à la conclusion que son évolution à l'échelle macroscopique est régie par une équation maîtresse. Nous voulons maintenant étudier les coefficients de cette équation et montrer qu'ils possèdent les propriétés que nous avons posées au chapitre IV (§ I.B.2). Nous allons pour cela suivre une méthode de perturbation, proposée à l'origine par Pauli (1928).

a. Notations. Position du problème

S'il était parfaitement isolé, le système aurait pour hamiltonien H, d'états propres $|\ell\rangle$ et de valeurs propres E_ℓ :

$$H|\ell\rangle = E_\ell|\ell\rangle . \tag{D.32}$$

Nous supposons l'ensemble des kets $|\ell\rangle$ orthonormé :

$$\langle\ell|\ell'\rangle = \delta_{\ell\ell'} . \tag{D.33}$$

Il constitue une base dans l'espace des états du système.

En fait, le véritable hamiltonien H_e qui régit l'évolution du système à l'échelle microscopique est de la forme

$$H_e = H + W , \tag{D.34}$$

où W décrit les *perturbations* extérieures, que nous ne connaissons pas en détail. Le terme supplémentaire W est petit devant H. Nous noterons $W_{\ell\ell'}$ ses éléments de matrice :

$$W_{\ell\ell'} \widehat{=} \langle\ell|W|\ell'\rangle . \tag{D.35}$$

Nous ne tiendrons compte explicitement de la « granulation » du temps qu'à la fin du calcul (§ d). Pour le moment, nous imaginons que le système est à l'instant $t=0$ dans l'un des états propres de H, que nous notons $|\ell_0\rangle$, et nous cherchons la probabilité pour qu'une mesure effectuée à un instant t ultérieur le trouve dans un état $|\ell\rangle$ différent.

b. Système différentiel équivalent à l'équation de Schrödinger

L'état du système à l'instant t s'obtient par intégration de l'équation de Schrödinger

$$i\hbar \frac{\mathrm{d}}{\mathrm{d}t}|\psi(t)\rangle = (H+W)|\psi(t)\rangle , \tag{D.36}$$

assortie de la condition initiale

$$|\psi(0)\rangle = |\ell_0\rangle . \tag{D.37}$$

Nous allons chercher $|\psi(t)\rangle$ sous la forme d'une combinaison linéaire des états $|\ell\rangle$:

$$|\psi(t)\rangle = \sum_\ell c_\ell(t)|\ell\rangle . \tag{D.38}$$

Il est commode de poser

$$c_\ell(t) \widehat{=} \gamma_\ell(t)\, \mathrm{e}^{-iE_\ell t/\hbar} , \tag{D.39}$$

et de prendre comme inconnues les coefficients γ_ℓ : ceux-ci seraient en effet indépendants du temps si W était nul[11], de sorte qu'ils sont bien adaptés au traitement approché que nous voulons mettre en œuvre. En reportant (D.38) et (D.39) dans l'équation de Schrödinger (D.36), on obtient pour les $\gamma_\ell(t)$ le système suivant d'équations différentielles[12] :

$$i\hbar \frac{\mathrm{d}}{\mathrm{d}t} \gamma_\ell(t) = \sum_{\ell'} W_{\ell\ell'} \, \mathrm{e}^{i(E_\ell - E_{\ell'})t/\hbar} \, \gamma_{\ell'}(t) \, . \tag{D.40}$$

Jusqu'ici, pas d'approximation : le système (D.40) est parfaitement équivalent à l'équation de Schrödinger (D.36). La condition initiale (D.37) prend la forme

$$\gamma_\ell(t=0) = \delta_{\ell\ell_0} \, . \tag{D.41}$$

c. Solution approchée du système différentiel

Comme la dépendance en temps des $\gamma_\ell(t)$ est d'autant plus faible que W est plus petit, nous allons appliquer une méthode de *solution par approximations successives;* nous discuterons ensuite (§ d) la validité du résultat obtenu.

A l'ordre zéro dans la perturbation (c'est-à-dire si W était nul), les équations (D.40) se réduiraient à

$$\frac{\mathrm{d}}{\mathrm{d}t} \gamma_\ell(t) = 0 \, , \tag{D.42}$$

dont la solution serait, compte tenu de la condition initiale (D.41),

$$\gamma_\ell^{(0)} = \delta_{\ell\ell_0} \, . \tag{D.43}$$

Pour trouver la solution à l'ordre suivant, on reporte (D.43) dans le second membre des équations différentielles (D.40), qui s'intègrent alors facilement pour donner la solution $\gamma_\ell^{(1)}$ à l'ordre un en W. Et ainsi de suite : la solution à l'ordre n s'obtient en remplaçant $\gamma_{\ell'}$, dans les équations (D.40), par la solution d'ordre $n-1$.

---

11. Il est effectivement facile de vérifier que, pour $W=0$, la solution générale de l'équation de Schrödinger s'écrit

$$|\psi(t)\rangle = \sum_\ell c_\ell(0) \, \mathrm{e}^{-iE_\ell t/\hbar} |\ell\rangle \, .$$

12. En effet, on a d'une part

$$i\hbar \frac{\mathrm{d}}{\mathrm{d}t} |\psi(t)\rangle = \sum_{\ell'} \left[i\hbar \frac{\mathrm{d}}{\mathrm{d}t} \gamma_{\ell'} + E_{\ell'} \gamma_{\ell'} \right] \mathrm{e}^{-iE_{\ell'} t/\hbar} |\ell'\rangle \, , \tag{1}$$

et d'autre part

$$(H+W)|\psi(t)\rangle = \sum_{\ell''} \gamma_{\ell''} \, \mathrm{e}^{-iE_{\ell''} t/\hbar} (E_{\ell''} + W) |\ell''\rangle \, . \tag{2}$$

Multiplions ces deux expressions à gauche par le bra $\langle \ell |$. Il vient, compte tenu de la relation d'orthonormalisation (D.33),

$$\left(i\hbar \frac{\mathrm{d}}{\mathrm{d}t} \gamma_\ell + E_\ell \gamma_\ell \right) \mathrm{e}^{-iE_\ell t/\hbar} = \gamma_\ell \, \mathrm{e}^{-iE_\ell t/\hbar} E_\ell + \sum_{\ell''} \gamma_{\ell''} \, \mathrm{e}^{-iE_{\ell''} t/\hbar} \langle \ell | W | \ell'' \rangle \, , \tag{3}$$

c'est-à-dire finalement

$$i\hbar \frac{\mathrm{d}}{\mathrm{d}t} \gamma_\ell = \mathrm{e}^{iE_\ell t/\hbar} \sum_{\ell'} \gamma_{\ell'} \, \mathrm{e}^{-iE_{\ell'} t/\hbar} W_{\ell\ell'} \tag{4}$$

(noter que la simplification est due au changement de fonctions (D.39)).

Nous nous contenterons ici de l'*approximation d'ordre un,* qui consiste donc à remplacer le système d'équations exactes (D.40) par le système approché

$$i\hbar \frac{d}{dt} \gamma_\ell(t) \simeq \sum_{\ell'} W_{\ell\ell'} \, e^{i(E_\ell - E_{\ell'})t/\hbar} \, \gamma_{\ell'}^{(0)} . \tag{D.44}$$

Ce système s'écrit d'après (D.43)

$$i\hbar \frac{d}{dt} \gamma_\ell(t) \simeq W_{\ell\ell_0} \, e^{i(E_\ell - E_{\ell_0})t/\hbar} . \tag{D.45}$$

Il s'intègre aussitôt, si l'on admet que W est indépendant du temps[13] : pour vérifier la condition initiale (D.41), il faut prendre[14]

$$\gamma_\ell(t) \simeq \delta_{\ell\ell_0} - W_{\ell\ell_0} \frac{1}{E_\ell - E_{\ell_0}} [e^{i(E_\ell - E_{\ell_0})t/\hbar} - 1] . \tag{D.46}$$

La probabilité $\mathcal{P}_{\ell\ell_0}(t)$ pour que, à l'instant t, une mesure trouve le système dans un état $|\ell\rangle$ *différent* de $|\ell_0\rangle$ s'obtient en prenant le carré du module de $c_\ell(t)$, c'est-à-dire de $\gamma_\ell(t)$ (voir formules (D.38) et (D.39)) :

$$\mathcal{P}_{\ell\ell_0}(t) \simeq |W_{\ell\ell_0}|^2 \frac{4 \sin^2[(E_\ell - E_{\ell_0}) t/2\hbar]}{(E_\ell - E_{\ell_0})^2} \quad \text{pour } (\ell) \neq (\ell_0). \tag{D.47}$$

Notons que cette formule approchée n'assure pas que la probabilité trouvée soit inférieure à 1 : la valeur maximum de l'expression (D.47), obtenue lorsque $E_\ell = E_{\ell_0}$, est proportionnelle à t^2 et peut donc croître indéfiniment avec t. Nous reviendrons plus loin sur ce problème.

d. Résultat final. Discussion

Grâce à l'hypothèse des phases aléatoires (§ 1.b), nous allons interpréter $\mathcal{P}_{\ell\ell_0}(t)$ comme la *probabilité de transition* de $|\ell_0\rangle$ à $|\ell\rangle$ pendant l'intervalle de temps t. L'expression (D.47) s'annule comme il se doit lorsque t tend vers zéro, mais sa dérivée par rapport à t s'annule également, ce qui est plus inattendu et pourrait être rédhibitoire car la probabilité de transition par unité de temps correspondante paraît nulle. C'est là qu'intervient la *«granulation» du temps* : la probabilité de transition par unité de temps $a_{\ell m}$ que nous cherchons est en réalité définie par

$$a_{\ell m} \widehat{=} \frac{1}{\tau} \mathcal{P}_{\ell m}(\tau) . \tag{D.48}$$

La formule (D.47) donne alors

$$a_{\ell m} = \frac{2\pi}{\hbar} |W_{\ell m}|^2 . f_{2\hbar/\tau}(E_\ell - E_m) , \tag{D.49}$$

où nous avons posé

$$f_\alpha(\varepsilon) \widehat{=} \frac{\alpha}{\pi} \frac{\sin^2 \varepsilon/\alpha}{\varepsilon^2} . \tag{D.50}$$

---

13. Les perturbations extérieures que décrit W n'ont pas de raison d'être vraiment constantes au cours du temps. Nous supposons pour simplifier que c'est le cas sur un intervalle de temps de l'ordre du temps τ caractéristique du «chaos moléculaire», puisque c'est pour de telles valeurs de t que nous allons utiliser le résultat de ce calcul.

14. La formule (D.46) est valable même si $E_\ell = E_{\ell_0}$: il suffit de prendre la limite de l'expression pour $E_\ell - E_{\ell_0}$ tendant vers zéro.

La fonction $f_\alpha(\varepsilon)$ est représentée sur la figure 1; elle est pratiquement nulle en dehors d'un intervalle de largeur $2\pi\alpha$ autour de $\varepsilon = 0$; on peut montrer que son intégrale vaut 1 quel que soit α positif :

$$\int_{-\infty}^{+\infty} f_\alpha(\varepsilon)\,d\varepsilon = 1 \,. \tag{D.51}$$

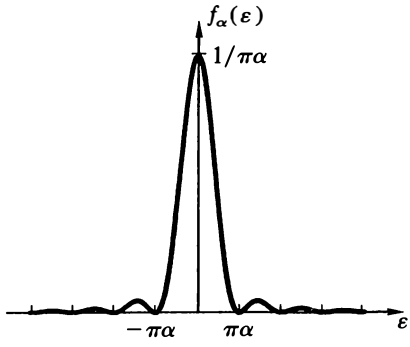


FIGURE 1

Allure de la fonction $f_\alpha(\varepsilon)$ définie par la relation (D.50). Cette fonction, symétrique par rapport à $\varepsilon = 0$, est pratiquement nulle en dehors de l'intervalle $(-\pi\alpha, +\pi\alpha)$: le premier maximum secondaire vaut 4,8 % environ du maximum principal, le suivant 1,6 %.

Au chapitre IV, nous avons écrit une équation maîtresse régissant l'évolution dans le temps t des probabilités $P_\ell(t)$ des états propres $|\ell\rangle$ de l'hamiltonien H; les arguments développés ici s'appliquent, eux, à des moyennes résultant de la « granulation » du temps. L'équation maîtresse doit donc être utilisée *seulement pour décrire l'évolution du système à l'échelle macroscopique;* on doit s'interdire d'y chercher la réponse à des questions plus « fines » concernant le comportement du système à une échelle de temps inférieure à celle qui caractérise la « granulation ».

Ceci étant admis, les deux propriétés fondamentales des coefficients de l'équation maîtresse, que nous avions postulées au chapitre IV (§ I.B.2), apparaissent clairement sur le résultat (D.49)-(D.50). Tout d'abord, *les probabilités de transition par unité de temps sont symétriques* dans l'échange des états initial et final :

$$a_{m\ell} = a_{\ell m} \,. \tag{D.52}$$

En effet, la fonction $f_\alpha(\varepsilon)$ est paire en ε et la perturbation W doit être hermitique (c'est, d'après (D.34), la différence entre deux hamiltoniens), ce qui s'exprime par l'égalité

$$W_{m\ell} = W^*_{\ell m} \tag{D.53}$$

(l'étoile notant ici le complexe conjugué), entraînant

$$|W_{m\ell}|^2 = |W_{\ell m}|^2 \,. \tag{D.54}$$

En second lieu, *les transitions* entre états microscopiques *conservent l'énergie* à $\hbar/\tau$ près, puisque la fonction $f_\alpha(\varepsilon)$ est pratiquement nulle si $|\varepsilon| > \pi\alpha$ (voir figure 1). D'après la définition (D.31) du temps τ, la précision $\hbar/\tau$ de la conservation de l'énergie est de l'ordre de $|W|$. Ce résultat est parfaitement naturel : comme W représente une perturbation incontrôlée du système « isolé » que nous étudions, l'énergie de ce système ne peut pas être déterminée avec une incertitude inférieure à $|W|$. Inversement, si

l'énergie du système est connue à δE près, les perturbations W peuvent atteindre cet ordre de grandeur et donc induire des transitions entre états microscopiques dont les énergies diffèrent de δE environ.

Mais le calcul qui conduit au résultat (D.49) est seulement approché (cf. § c). Pour que cette approximation à l'ordre le plus bas soit valable, il faudrait au moins que la probabilité de transition entre deux états pendant le temps τ (nombre sans dimension) reste toujours très petite devant 1. Or cette probabilité s'écrit

$$a_{\ell m} \cdot \tau \simeq \frac{2\pi\tau}{\hbar} |W_{\ell m}|^2 f_{2\hbar/\tau}(E_\ell - E_m) , \tag{D.55}$$

avec (fig. 1)

$$f_{2\hbar/\tau}(E_\ell - E_m) \leqslant \frac{\tau}{2\pi\hbar} . \tag{D.56}$$

Le produit $a_{\ell m} \cdot \tau$ est donc de l'ordre de 1 selon la relation (D.31), et non pas petit devant 1. On est forcé de conclure que *l'évaluation précédente des probabilités de transition par unité de temps n'est pas vraiment valable;* elle permet seulement, à partir d'un calcul simple, de comprendre qualitativement d'où peuvent provenir les propriétés génériques des coefficients $a_{\ell m}$ de l'équation maîtresse. Pour l'améliorer, on ne peut pas envisager de poursuivre aux ordres supérieurs la solution itérative amorcée au paragraphe c : les termes successifs sont tous du même ordre de grandeur[15], de sorte qu'il est impossible de tronquer le processus après un nombre fini d'itérations. Il faut faire appel à des techniques plus élaborées (voir par exemple réf. 2). On constate alors que l'hermiticité de W ne suffit pas à assurer la symétrie des coefficients $a_{\ell m}$[16], et qu'il faut invoquer l'invariance par renversement du temps (réf. 3).

Références

1. N.G. van Kampen, *Stochastic Processes in Physics and Chemistry,* North Holland, Amsterdam (1981).
2. H.J. Kreuzer, *Nonequilibrium Thermodynamics and its Statistical Foundations,* Clarendon Press, Oxford (1981), chapitre 10.
3. T.Y. Wu et T. Ohmura, *Quantum Theory of Scattering,* Prentice Hall, Englewood Cliffs (1962), section W-3.

---

15. L'approximation $\gamma_\ell^{(n)}(t)$ résultant des n premières itérations est donnée à partir des $\gamma_\ell^{(n-1)}$ par les équations (D.40) :

$$i\hbar \frac{\mathrm{d}}{\mathrm{d}t} \gamma_\ell^{(n)}(t) = \sum_{\ell'} W_{\ell\ell'} \, \mathrm{e}^{i(E_\ell - E_{\ell'})t/\hbar} \gamma_{\ell'}^{(n-1)}(t) . \tag{1}$$

Le premier membre a pour ordre de grandeur

$$\left| i\hbar \frac{\mathrm{d}}{\mathrm{d}t} \gamma_\ell^{(n)} \right| \sim \frac{\hbar}{\tau} |\gamma_\ell^{(n)}| , \tag{2}$$

de sorte que le rapport

$$\frac{|\gamma_\ell^{(n)}|}{|\gamma_\ell^{(n-1)}|} \sim \frac{|W|\tau}{\hbar} \tag{3}$$

est de l'ordre de 1 d'après la relation (D.31).

16. Ce n'est qu'au premier ordre dans la pertrubation W que son hermiticité entraîne la symétrie des probabilités de transition. On s'en apercevrait déjà si l'on poussait au second ordre la méthode itérative décrite au paragraphe c.

COMPLÉMENT IV.E

Évolution d'un gaz classique dilué : l'équation de Boltzmann

I. Démonstration de l'équation de Boltzmann
  1. Position du problème et ordres de grandeur
  2. Étude préliminaire des collisions entre les molécules
    a. Hypothèses et approximations
    b. Cinématique d'une collision élastique
    c. Section efficace différentielle de collision
  3. Bilan d'évolution dans l'espace des phases à une particule
  4. Évaluation de l'intégrale de collision

II. Propriétés générales de l'équation de Boltzmann
  1. Équations de conservation locales déduites de l'équation de Boltzmann
  2. Évolution de l'entropie : le « théorème H »
    a. Équation d'évolution pour la densité locale d'entropie
    b. Contribution des collisions à l'évolution de l'entropie
    c. Conclusion
  3. Distributions d'équilibre

Dans l'équation maîtresse, l'évolution du système est envisagée de façon globale : les états (ℓ) sont les états microscopiques du système pris dans son ensemble. Il serait intéressant de connaître plus en détail le comportement de ses constituants microscopiques au cours de l'évolution macroscopique, c'est-à-dire de développer ce que l'on nomme traditionnellement la *« théorie cinétique »* de cette évolution (cf. complément III.C). Mais ce problème est certainement plus difficile; une solution éventuelle ne peut venir que d'une étude spécifique de chaque type de système : on ne peut espérer comprendre le mouvement des particules constituant le système que si l'on sait analyser, au moins de façon statistique, l'effet de leurs interactions. Ce sont évidemment les gaz qui se prêtent le plus facilement à une telle étude : les interactions y sont peu intenses, à courte portée, et l'on peut réduire leur influence en augmentant la dilution du gaz.

C'est encore L. Boltzmann qui a ouvert la voie en proposant, en 1872, une équation cinétique décrivant l'évolution des gaz dilués. Bien qu'elle ait été abondamment critiquée (surtout au début), analysée, discutée et utilisée, l'équation de Boltzmann conserve après plus d'un siècle sa forme initiale (à peine modernisée) et sa puissance explicative.

I. DÉMONSTRATION DE L'ÉQUATION DE BOLTZMANN

Il existe plusieurs façons de démontrer l'équation de Boltzmann. La méthode que nous choisissons ici présente le double avantage d'être relativement simple et de bien dégager les hypothèses physiques sous-jacentes.

1. POSITION DU PROBLÈME ET ORDRES DE GRANDEUR

Nous considérons un *gaz ordinaire,* pur, et suffisamment *dilué* pour être *presque parfait :* la distance moyenne entre molécules voisines est supposée grande devant la portée des forces qu'elles exercent les unes sur les autres, de telle sorte qu'elles sont la plupart du temps libres et indépendantes. Nous allons chercher à analyser *l'influence des collisions,* qui sont, dans ces conditions, suffisamment rares pour

apporter une correction relativement faible au modèle des gaz parfaits. Mais, même rares, *les collisions jouent un rôle essentiel dans l'évolution du gaz vers l'équilibre macroscopique.* On le comprend aisément : dans un ensemble de molécules rigoureusement indépendantes, il n'y aurait entre elles aucun échange, et donc en particulier aucun mécanisme de redistribution de l'énergie totale susceptible d'uniformiser la température entre les diverses parties macroscopiques du système.

Le gaz considéré, enfermé dans un récipient de volume V, est *isolé.* Nous admettons néanmoins[1] que les molécules puissent être soumises à une *force extérieure* $\vec{F}(\vec{r})$ (par exemple la force de pesanteur), dépendant éventuellement de leur position $\vec{r}$.

Nous décrivons le mouvement des molécules par la *mécanique classique,* et nous les supposons *monoatomiques.* Cette hypothèse a deux conséquences importantes : d'une part, l'état microscopique du gaz à un instant donné peut être caractérisé par les seules positions $\vec{r}_j$ et impulsions $\vec{p}_j$ des N molécules; d'autre part, *les collisions sont élastiques,* c'est-à-dire qu'elles conservent l'énergie cinétique totale des molécules qu'elles mettent en jeu, car tout transfert d'énergie à des degrés de liberté internes est exclu[2].

Si nous prenons l'exemple de l'hélium à la température ambiante et à la pression atmosphérique, les ordres de grandeur sont analogues à ceux du complément IV.B (§ I.1) : la portée des forces intermoléculaires vaut environ

$$\rho_0 \sim 3 \text{ Å} , \tag{E.1}$$

la distance moyenne entre molécules voisines

$$d \sim 30 \text{ Å} , \tag{E.2}$$

la vitesse des molécules

$$v \sim 1\,000 \text{ m/s} , \tag{E.3}$$

leur libre parcours moyen

$$\ell \sim 1\,500 \text{ Å} \tag{E.4}$$

et l'intervalle de temps moyen entre deux collisions successives d'une même molécule

$$\tau_m \sim 10^{-10} \text{ s} . \tag{E.5}$$

2. ÉTUDE PRÉLIMINAIRE DES COLLISIONS ENTRE LES MOLÉCULES

a. Hypothèses et approximations

Nous allons considérer seulement les *collisions binaires* et *négliger celles qui font intervenir plus de deux molécules à la fois.* Avec les ordres de grandeur rappelés au paragraphe 1, ceci est parfaitement légitime. En effet, une molécule donnée passe un temps de l'ordre de ℓ/v entre deux collisions successives, qui durent chacune environ ρ_0/v. Le rapport de ces deux temps,

$$\frac{\rho_0/v}{\ell/v} \sim 2 \times 10^{-3} , \tag{E.6}$$

donne l'ordre de grandeur de la probabilité pour qu'une molécule soit en cours de collision, c'est-à-dire la proportion moyenne de molécules qui sont en train de participer

1. Le gaz reste *thermiquement isolé* (pas d'échanges de chaleur avec l'extérieur) et enfermé dans un récipient à *parois rigides* (pas d'échange de travail à travers les forces de pression).

2. Les atomes constituant le gaz se trouvent tous, pour ce qui concerne leur structure interne, dans leur état fondamental : aux températures envisagées (de l'ordre de la température ordinaire), les degrés de liberté électroniques sont « gelés » (complément III.B, § II.1), et l'énergie des collisions est beaucoup trop faible pour pouvoir les exciter.

à une collision à un instant donné. Cette proportion est déjà faible. Grossièrement, on peut estimer que la fraction de ces collisions qui mettent en jeu trois molécules à la fois est à son tour de l'ordre de grandeur (E.6). L'influence des collisions ternaires est donc tout à fait négligeable. Il en est *a fortiori* de même pour les collisions à plus de trois particules.

Pour que deux molécules entrent en collision, il faut qu'elles s'approchent l'une de l'autre à une distance de l'ordre de ρ_0. De telles distances sont inappréciables à l'échelle macroscopique où nous allons raisonner. Nous considérerons donc que *deux molécules n'entrent en collision que si elles se trouvent au même point $\vec{r}$ au même instant.* Notons que les impulsions de ces deux molécules peuvent au contraire être très différentes, et qu'elles seront en règle générale considérablement modifiées par le choc (voir § b).

Avant la collision, les deux molécules qui s'approchent l'une de l'autre proviennent de régions de l'espace relativement éloignées, puisque chacune d'elles a parcouru depuis son dernier choc une distance de l'ordre du libre parcours moyen ℓ, c'est-à-dire grande devant ρ_0 et même devant la distance moyenne d entre molécules (voir paragraphe 1); en tout cas, les dernières collisions qu'elles ont subies avant d'aborder celle que nous examinons ont fait intervenir d'autres molécules, qui se trouvent maintenant ailleurs et ont elles-mêmes subi (ou vont subir) d'autres collisions. Il en résulte qu'*on peut négliger les corrélations entre les deux molécules qui entrent en collision.* C'est la fameuse *hypothèse du « chaos moléculaire »* (« Stosszahlansatz »), formulée par Boltzmann en 1872. Nous avons indiqué au complément IV.B (§ I.3) comment cette hypothèse devient plausible lorsqu'on analyse l'influence sur le système considéré de perturbations extérieures très faibles et incontrôlables. Elle implique que les variations dans le temps des grandeurs macroscopiques ne seront analysées que sur des intervalles de temps $t' - t$ tels que

$$|t' - t| \geqslant \tau_m \,. \tag{E.7}$$

b. Cinématique d'une collision élastique

Les collisions que nous avons à étudier sont élastiques et pratiquement ponctuelles. Analysons comment une telle collision affecte les impulsions des molécules qui y prennent part.

Soient $\vec{p}_0$ et $\vec{p}_1$ les impulsions des deux molécules avant la collision, $\vec{p}\,'_0$ et $\vec{p}\,'_1$ leurs impulsions après la collision. La *conservation de l'impulsion totale*[3] s'écrit

$$\vec{p}\,'_0 + \vec{p}\,'_1 = \vec{p}_0 + \vec{p}_1 \,. \tag{E.8}$$

---

3. L'impulsion totale se conserve si le système constitué par les deux particules est *isolé*. Ici, le gaz est suffisamment dilué pour que les interactions des deux molécules avec les $(N-2)$ autres durant la collision soient négligeables. Cependant, dans le cas où il existe une force extérieure $\vec{F}$, il faut que le produit $F.\rho_0/v$ de son intensité par la durée d'une collision soit très petit devant l'impulsion échangée dans la collision. Cette condition est très largement satisfaite par la force gravitationnelle, car

$$mg.\frac{\rho_0}{v} \ll mv \tag{1}$$

est équivalent à

$$\frac{v^2}{\rho_0} \gg g \,, \tag{2}$$

et les ordres de grandeur du paragraphe 1 donnent

$$\frac{v^2}{\rho_0} \sim 3 \times 10^{15}\ \mathrm{ms}^{-2} \,. \tag{3}$$

Une collision *élastique* conserve en outre *l'énergie cinétique totale;* si l'on multiplie ses deux membres par $2m$, cette seconde équation de conservation prend la forme

$$\vec{p}\,'^2_0 + \vec{p}\,'^2_1 = \vec{p}^{\,2}_0 + \vec{p}^{\,2}_1 \,. \tag{E.9}$$

Il est commode d'introduire l'impulsion $\vec{p}_G$ du centre de masse des deux molécules, qui reste inchangée dans le choc :

$$\vec{p}_G \widehat{=} \vec{p}_0 + \vec{p}_1 = \vec{p}\,'_0 + \vec{p}\,'_1 \,, \tag{E.10}$$

et les impulsions relatives

$$\vec{p}_r \widehat{=} (\vec{p}_0 - \vec{p}_1)/2 \,, \quad \vec{p}\,'_r \widehat{=} (\vec{p}\,'_0 - \vec{p}\,'_1)/2 \,. \tag{E.11}$$

Ces formules s'inversent facilement pour donner

$$\vec{p}_0 = \vec{p}_G/2 + \vec{p}_r \,, \quad \vec{p}_1 = \vec{p}_G/2 - \vec{p}_r \,, \tag{E.12.a}$$

$$\vec{p}\,'_0 = \vec{p}_G/2 + \vec{p}\,'_r \,, \quad \vec{p}\,'_1 = \vec{p}_G/2 - \vec{p}\,'_r \,. \tag{E.12.b}$$

La conservation de l'énergie cinétique (formule (E.9)) impose alors

$$\vec{p}\,'^2_r = \vec{p}^{\,2}_r \,. \tag{E.13}$$

En résumé, lorsque deux molécules d'impulsions $\vec{p}_0$ et $\vec{p}_1$ entrent en collision, *leur centre de masse est animé d'un mouvement de translation uniforme qui n'est pas affecté par le choc; celui-ci ne change pas non plus le module de l'impulsion relative des deux particules, mais seulement sa direction.*

Il est intéressant de noter que les lois de conservation (E.8)-(E.9) sont symétriques dans l'échange des couples d'impulsions $(\vec{p}_0, \vec{p}_1)$ et $(\vec{p}\,'_0, \vec{p}\,'_1)$. Ceci implique que, si $\vec{p}\,'_0$ et $\vec{p}\,'_1$ sont les impulsions initiales dans une collision, celle-ci peut donner $\vec{p}_0$ et $\vec{p}_1$ comme impulsions finales (bien entendu, les impulsions après collision ne sont pas déterminées de façon unique par la donnée des impulsions initiales).

c. Section efficace différentielle de collision

Nous connaissons maintenant les modifications qu'une collision est susceptible d'apporter aux impulsions des molécules, mais il nous faut encore évaluer les probabilités des divers changements possibles.

D'après les résultats du paragraphe b, seul le mouvement relatif des deux molécules est affecté par la collision. Plaçons-nous donc dans le *référentiel du centre de masse* pour le système des deux molécules : c'est le référentiel animé par rapport au laboratoire du même mouvement de translation uniforme, de vitesse $\vec{v}_G = \vec{p}_G/2m$, que le centre de masse. Dans ce référentiel, la dynamique du système est équivalente à celle d'une particule a de masse $\mu = m/2$ qui interagirait avec une particule A constamment immobile à l'origine des coordonnées, la force d'interaction étant celle qui existe réellement entre les deux molécules[4]. Dans cette dynamique équivalente,

---

4. Les deux molécules interagissent par des forces dépendant uniquement de leur position relative $\vec{r}_1 - \vec{r}_2$ et obéissant au principe d'égalité de l'action et de la réaction :

$$\vec{F}_{12} = -\vec{F}_{21} = \vec{F}\,(\vec{r}_1 - \vec{r}_2) \,. \tag{1}$$

On en déduit que le centre de masse a un mouvement rectiligne uniforme. D'autre part, la variable relative

$$\vec{r} \widehat{=} \vec{r}_1 - \vec{r}_2 \tag{2}$$

obéit à l'équation

$$\frac{m}{2}\frac{\mathrm{d}^2\vec{r}}{\mathrm{d}t^2} = \vec{F}\,(\vec{r}) \,; \tag{3}$$

c'est celle qui régit le mouvement d'une particule de masse $\mu = m/2$ soumise à la force $\vec{F}\,(\vec{r})$.

la collision est vue de la façon suivante : la particule a se dirige, avec l'impulsion initiale $\vec{p}_r$, vers le centre de force A immobile; le choc la dévie (A restant toujours immobile), de sorte que son impulsion finale est $\vec{p}\,'_r$; la collision étant élastique, l'énergie cinétique de la particule a est conservée, c'est-à-dire que $\vec{p}\,'_r$ a même module que $\vec{p}_r$.

Pour décrire cette collision, on introduit la section efficace différentielle, définie de la façon habituelle : on imagine qu'un faisceau de particules a, ayant toutes la même impulsion $\vec{p}_r$, de flux Φ (nombre de particules franchissant, pendant l'unité de temps, une surface unité normale à $\vec{p}_r$), est dirigé vers une cible constituée de ν particules A; le nombre $\mathrm{d}n_c/\mathrm{d}t$ de collisions, par unité de temps, dans lesquelles une particule a est déviée dans l'angle solide infinitésimal $\mathrm{d}\Omega$ autour d'une direction donnée s'écrit

$$\frac{\mathrm{d}n_c}{\mathrm{d}t} = \Phi . \nu . \sigma . \mathrm{d}\Omega . \tag{E.14}$$

La *section efficace différentielle* σ dans la direction considérée ne dépend pas, pour des molécules monoatomiques, de l'orientation de l'impulsion initiale $\vec{p}_r$ (isotropie de l'espace), mais seulement de son module p_r (c'est-à-dire de l'énergie cinétique); de même, elle ne dépend que de l'angle de déflection

$$\theta \widehat{=} (\widehat{\vec{p}_r, \vec{p}\,'_r}) , \quad 0 \leqslant \theta \leqslant \pi \tag{E.15}$$

et pas de l'angle azimutal φ qui repère avec θ la direction de l'impulsion finale $\vec{p}\,'_r$.

3. BILAN D'ÉVOLUTION DANS L'ESPACE DES PHASES A UNE PARTICULE

Nous allons caractériser ici l'état macroscopique du gaz par la «*fonction de répartition*»[5] $f(\vec{r}, \vec{p}\,; t)$, définie de telle sorte que $f(\vec{r}, \vec{p}\,; t)\, \mathrm{d}^3r\, \mathrm{d}^3p$ est le nombre moyen[6] de particules qui, à l'instant t, se trouvent dans le volume d^3r autour du point $\vec{r}$ et possèdent l'impulsion $\vec{p}$ à d^3p près. La fonction de répartition vérifie donc la condition de normalisation

$$\int \mathrm{d}^3r . \mathrm{d}^3p\, f(\vec{r}, \vec{p}\,; t) = N . \tag{E.16}$$

Elle caractérise la répartition des particules du système dans l'*espace des phases individuel à 6 dimensions.*

Considérons dans cet espace un domaine δD_0 *fixe,* de volume $\delta^3 r_0 . \delta^3 p_0$ infinitésimal (à l'échelle macroscopique; voir note 6), et dont l'un des sommets se situe au point $(\vec{r}_0, \vec{p}_0)$ (fig. 1). A l'instant t, le domaine δD_0 contient les points représentatifs de

$$\delta N(\vec{r}_0, \vec{p}_0; t) = f(\vec{r}_0, \vec{p}_0; t)\, \delta^3 r_0 \delta^3 p_0 \tag{E.17}$$

---

5. Cette description n'est pas complète : $f(\vec{r}, \vec{p}\,; t)$ est en fait la «fonction de répartition à une particule», et on peut définir des «fonctions de répartition à n particules» pour $1 \leqslant n \leqslant N$ (complément IV.F, § II.1).

6. Comme toujours en mécanique statistique, la moyenne dont il s'agit ici est prise sur un ensemble statistique d'un très grand nombre $\mathcal{N}$ de systèmes identiques (I, § IV.B.2). Toutefois, si le nombre moyen $f . \mathrm{d}^3r . \mathrm{d}^3p$ est suffisamment grand (ce qui demande en particulier que le volume $\mathrm{d}^3r . \mathrm{d}^3p$ de l'espace des phases à une particule, infinitésimal à l'échelle macroscopique, soit grand à l'échelle microscopique), les fluctuations pourront être négligées. C'est ce que nous supposerons dans la suite.

particules. Entre les instants t et $t+\mathrm{d}t$, ce nombre varie de

$$\mathrm{d}[\delta N]=\frac{\partial f}{\partial t}(\vec{r}_0, \vec{p}_0; t)\,\delta^3 r_0 \delta^3 p_0 . \mathrm{d}t . \tag{E.18}$$

Pour calculer $\mathrm{d}[\delta N]$, on peut compter les particules qui entrent dans le domaine considéré et celles qui le quittent pendant l'intervalle de temps $(t, t+\mathrm{d}t)$. Nous nous limiterons au premier ordre dans chacun des infiniments petits δx_0, δy_0, δz_0, δp_{0x}, δp_{0y}, δp_{0z} et $\mathrm{d}t$.

Supposons dans un premier temps que *les interactions entre les particules soient négligeables.* Le mouvement de l'une d'elles obéit alors aux équations

$$\vec{r}(t+\mathrm{d}t)=\vec{r}(t)+\frac{1}{m}\vec{p}(t)\mathrm{d}t , \tag{E.19.a}$$

$$\vec{p}(t+\mathrm{d}t)=\vec{p}(t)+\vec{F}(\vec{r})\mathrm{d}t . \tag{E.19.b}$$

On en déduit facilement (fig. 1) que les $\delta N(\vec{r}_0, \vec{p}_0; t)$ particules dont le point représentatif appartient à l'instant t au domaine δD_0 aboutissent à l'instant $t+\mathrm{d}t$ dans un domaine de même volume repéré par $(\vec{r}_0+\vec{p}_0\,\mathrm{d}t/m, \vec{p}_0+\vec{F}(\vec{r}_0)\mathrm{d}t)$, si l'on néglige des termes d'ordre supérieur dans les infiniment petits choisis plus haut. De même, les particules se trouvant à $t+\mathrm{d}t$ dans le domaine δD_0 sont celles que contenait à l'instant t un domaine de même volume situé en $(\vec{r}_0-\vec{p}_0\mathrm{d}t/m, \vec{p}_0-\vec{F}(\vec{r}_0)\mathrm{d}t)$; leur nombre est donc donné par

$$\delta N(\vec{r}_0, \vec{p}_0; t+\mathrm{d}t)=\delta N\left(\vec{r}_0-\frac{\vec{p}_0}{m}\mathrm{d}t, \vec{p}_0-\vec{F}(\vec{r}_0)\mathrm{d}t; t\right) . \tag{E.20}$$

En comparant ce résultat avec le nombre initial (E.17), on trouve aussitôt (toujours au premier ordre dans les infiniment petits)

$$\begin{aligned}\mathrm{d}[\delta N] &\widehat{=} \delta N(\vec{r}_0, \vec{p}_0; t+\mathrm{d}t)-\delta N(\vec{r}_0, \vec{p}_0; t)\\ &=\delta^3 r_0 . \delta^3 p_0 . \mathrm{d}t\left[-\frac{\vec{p}_0}{m}\cdot\frac{\partial}{\partial \vec{r}} f(\vec{r}_0, \vec{p}_0; t)-\vec{F}(\vec{r}_0)\cdot\frac{\partial}{\partial \vec{p}} f(\vec{r}_0, \vec{p}_0; t)\right]\end{aligned} \tag{E.21}$$

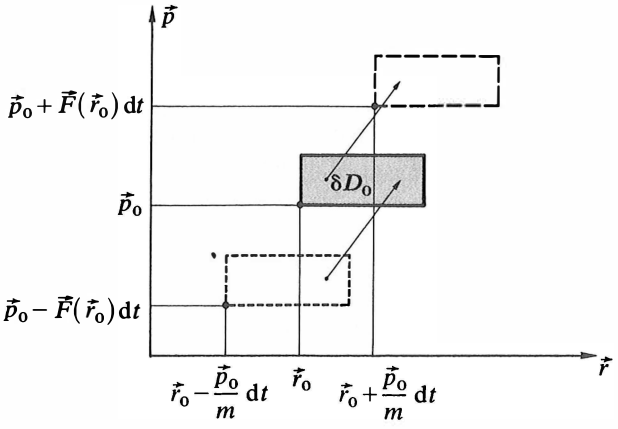


FIGURE 1

Représentation schématique de l'espace des phases à une particule. Le domaine infinitésimal fixe δD_0 sur lequel on raisonne est dessiné en trait plein. Les particules qu'il contient à l'instant t se retrouvent à $t+\mathrm{d}t$ dans le domaine tireté; elles sont remplacées par celles qui se trouvaient à l'instant t dans le domaine en pointillés (ces trois domaines peuvent se recouvrir partiellement, pour des valeurs de $\mathrm{d}t$ plus petites que celle qui a été choisie ici).

(nous notons $\partial/\partial\vec{r}$ l'opérateur gradient par rapport à $\vec{r}$, $\partial/\partial\vec{p}$ l'opérateur gradient par rapport à $\vec{p}$). Il suffit ensuite d'égaler les deux expressions (E.18) et (E.21) pour trouver une *équation d'évolution pour la fonction de répartition f.*

Mais nous avons jusqu'ici négligé totalement les *interactions entre les molécules,* qui se manifestent par une *contribution supplémentaire* au bilan que nous voulons établir. Nous allons évaluer cette contribution au paragraphe suivant, et écrivons provisoirement l'équation d'évolution

$$\left[\frac{\partial}{\partial t}+\frac{\vec{p}}{m}\cdot\frac{\partial}{\partial\vec{r}}+\vec{F}(\vec{r})\cdot\frac{\partial}{\partial\vec{p}}\right]f(\vec{r},\vec{p}\,;t)=I_c\,, \tag{E.22}$$

où I_c est l'*« intégrale de collision »*.

4. ÉVALUATION DE L'INTÉGRALE DE COLLISION

Le raisonnement du paragraphe précédent, qui nous a fourni le *premier membre de l'équation d'évolution* (E.22), n'a fait appel à *aucune approximation* (cf. complément IV.F, § II.2). Nous allons maintenant *calculer de façon approchée le second membre* de cette équation, en nous appuyant sur les hypothèses et ordres de grandeur présentés aux paragraphes 1 et 2.a.

Les molécules du domaine δD_0 qui subissent entre t et $t+\mathrm{d}t$ une collision avec n'importe quelle autre molécule quittent δD_0, puisque leur impulsion est changée. Pour déterminer leur nombre, raisonnons dans l'espace ordinaire, car deux molécules n'entrent en collision que si elles occupent pratiquement la même position au même instant. Nous considérons donc que les $f(\vec{r}_0,\vec{p}_0;t)\,\delta^3r_0.\delta^3p_0$ molécules, animées de l'impulsion $\vec{p}_0$, viennent de pénétrer à l'instant t dans le volume δ^3r_0 de l'espace ordinaire, où elles sont rejointes par $f(\vec{r}_0,\vec{p}_1;t)\,\delta^3r_0.\delta^3p_1$ molécules animées de l'impulsion $\vec{p}_1$. Ces deux ensembles de molécules sont alors suffisamment proches pour entrer en collision. Nous les traitons comme indépendants, selon l'hypothèse du « chaos moléculaire » (§ 2.a). Pour nous ramener à la formule standard (E.14), nous prenons l'un des deux nombres précédents comme nombre ν de particules-cibles; l'autre donnera le flux Φ de particules incidentes : il faut le diviser par δ^3r_0 pour obtenir le nombre de particules par unité de volume, puis multiplier ce résultat par le module $|\vec{v}_0-\vec{v}_1|=|\vec{p}_0-\vec{p}_1|/m$ de la vitesse relative[7]. La formule (E.14) s'écrit alors

$$\mathrm{d}n_c=\frac{1}{m}\,|\vec{p}_0-\vec{p}_1|\,f(\vec{r}_0,\vec{p}_0;t)f(\vec{r}_0,\vec{p}_1;t)\,\delta^3r_0.\delta^3p_0.\delta^3p_1$$
$$\times\sigma(|\vec{p}_0-\vec{p}_1|/2,\theta)\,\mathrm{d}\Omega.\mathrm{d}t\,. \tag{E.23}$$

---

7. Si une région de l'espace ordinaire contient des particules, en nombre n par unité de volume, toutes animées de la même vitesse $\vec{v}$, cet ensemble de particules constitue un faisceau. Le nombre de particules franchissant pendant le temps $\mathrm{d}t$ une surface d'aire unité normale à $\vec{v}$ est égal à

$$nv\,\mathrm{d}t;$$

d'où le flux

$$\Phi=nv$$

du faisceau.

Dans la situation envisagée ici, c'est uniquement le *mouvement relatif* des deux ensembles de particules qu'il faut prendre en compte pour analyser les collisions; v doit donc être remplacée par la vitesse relative $|\vec{v}_0-\vec{v}_1|$. En outre, on peut indifféremment choisir l'un ou l'autre des deux nombres comme représentant le nombre ν de particules cibles, et calculer le flux Φ à partir de l'autre, puisque seul le produit $\Phi\nu$ intervient dans la formule (E.14).

Elle donne le nombre dn_c de collisions qui, entre t et $t+dt$, changent les impulsions $\vec{p}_0$ et $\vec{p}_1$ de deux des molécules considérées pour les transformer en $\vec{p}\,'_0$ et $\vec{p}\,'_1$ vérifiant les relations (cf. § 2.b et formule (E.15)) :

$$\vec{p}\,'_0=\frac{1}{2}(\vec{p}_0+\vec{p}_1)+\vec{p}\,'_r\,,\quad \vec{p}\,'_1=\frac{1}{2}(\vec{p}_0+\vec{p}_1)-\vec{p}\,'_r\,, \tag{E.24}$$

avec $$|\vec{p}\,'_r|=\frac{1}{2}|\vec{p}_0-\vec{p}_1|\,, \tag{E.25.a}$$

$$\widehat{(\vec{p}_0-\vec{p}_1,\vec{p}\,'_r)}=\theta,\quad 0\leqslant\theta\leqslant\pi\,. \tag{E.25.b}$$

Le nombre total de molécules qui disparaissent entre t et $t+dt$ du domaine δD_0 s'obtient à partir de (E.23) en sommant sur toutes les impulsions $\vec{p}_1$ (puisque l'impulsion initiale des « partenaires » des molécules qui nous intéressent peut être quelconque) et sur toutes les directions possibles des impulsions finales (puisque seul importe le fait que $\vec{p}_0$ ait été modifiée[8]). On obtient ainsi une première contribution (négative) $I_c^{(1)}$ à l'« intégrale de collision » :

$$I_c^{(1)}(\vec{r}_0,\vec{p}_0;t)\delta^3r_0.\delta^3p_0.dt=-\delta^3r_0.\delta^3p_0.dt\int d^3p_1\int d\Omega\frac{1}{m}|\vec{p}_0-\vec{p}_1|$$
$$f(\vec{r}_0,\vec{p}_0;t)f(\vec{r}_0,\vec{p}_1;t)\sigma(|\vec{p}_0-\vec{p}_1|/2,\theta)\,. \tag{E.26}$$

Il existe évidemment une deuxième contribution, de signe opposé : le domaine δD_0 est peuplé par des molécules qui avaient initialement une impulsion différente de $\vec{p}_0$ mais à qui une collision donne $\vec{p}_0$ comme impulsion finale à $t+dt$. Pour obtenir le nombre de ces molécules, choisissons d'abord $\vec{p}_1$ comme impulsion *finale* de la deuxième molécule participant à la collision. Comme nous l'avons souligné à la fin du paragraphe 2.b, ceci est possible si les impulsions *initiales* des deux molécules sont $\vec{p}\,'_0$ et $\vec{p}\,'_1$, données par les relations (E.24)-(E.25). Nous pouvons donc encore utiliser la formule (E.23), à condition d'y remplacer $\vec{p}_0$ par $\vec{p}\,'_0$ et $\vec{p}_1$ par $\vec{p}\,'_1$:

$$dn'_c=\frac{1}{m}|\vec{p}\,'_0-\vec{p}\,'_1|f(\vec{r}_0,\vec{p}\,'_0;t)f(\vec{r}_0,\vec{p}\,'_1;t)\delta^3r_0.\delta^3p'_0.\delta^3p'_1$$
$$\times\sigma(|\vec{p}\,'_0-\vec{p}\,'_1|/2,\theta)d\Omega.dt\,. \tag{E.27}$$

Or nous savons (§ 2.b) que

$$|\vec{p}\,'_0-\vec{p}\,'_1|=|\vec{p}_0-\vec{p}_1|\,; \tag{E.28}$$

en outre, l'angle θ, compris par définition entre 0 et π, reste lui aussi inchangé quand on intervertit $\vec{p}_r$ et $\vec{p}\,'_r$[9].

---

8. La probabilité n'est pas nulle pour que les impulsions des particules restent inchangées dans la collision (la section efficace ne s'annule pas « vers l'avant », c'est-à-dire pour $\theta=0$). Mais, comme dans l'équation maîtresse, la contribution correspondante va aussi figurer dans le terme de « source » (E.31), avec le signe opposé.

9. Les impulsions $\vec{p}_0$ et $\vec{p}_1$ peuvent se calculer, à partir de $\vec{p}\,'_0$ et $\vec{p}\,'_1$, par des formules semblables à (E.24)-(E.25) :

$$\vec{p}_0=\frac{1}{2}(\vec{p}\,'_0+\vec{p}\,'_1)+\vec{p}_r\,,\quad \vec{p}_1=\frac{1}{2}(\vec{p}\,'_0+\vec{p}\,'_1)-\vec{p}_r\,, \tag{1}$$

avec $$|\vec{p}_r|=\frac{1}{2}|\vec{p}\,'_0-\vec{p}\,'_1|\,, \tag{2.a}$$

$$\widehat{(\vec{p}\,'_0-\vec{p}\,'_1,\vec{p}_r)}=\theta\,,\quad 0\leqslant\theta\leqslant\pi\,. \tag{2.b}$$

Enfin, on montre facilement[10] que

$$d^3p'_0 . d^3p'_1 = d^3p_0 . d^3p_1 \,. \tag{E.29}$$

On peut donc écrire

$$dn'_c = \frac{1}{m} |\vec{p}_0 - \vec{p}_1| f(\vec{r}_0, \vec{p}\,'_0; t) f(\vec{r}_0, \vec{p}\,'_1; t) \delta^3 r_0 . \delta^3 p_0 . \delta^3 p_1 \times \sigma(|\vec{p}_0 - \vec{p}_1|/2, \theta) d\Omega . dt \,. \tag{E.30}$$

Finalement, une intégration sur toutes les impulsions $\vec{p}_1$ et tous les angles de déflection possibles (équivalente à une intégration sur toutes les impulsions initiales $\vec{p}\,'_0$ et $\vec{p}\,'_1$ susceptibles de donner $\vec{p}_0$ comme impulsion finale) fournit la deuxième contribution $I_c^{(2)}$ à l'« intégrale de collision » :

$$I_c^{(2)}(\vec{r}_0, \vec{p}_0; t) \delta^3 r_0 . \delta^3 p_0 . dt = \delta^3 r_0 . \delta^3 p_0 . dt \times \int d^3p_1 \int d\Omega \frac{1}{m} |\vec{p}_0 - \vec{p}_1| . \sigma(|\vec{p}_0 - \vec{p}_1|/2, \theta) f(\vec{r}_0, \vec{p}\,'_0; t) f(\vec{r}_0, \vec{p}\,'_1; t) \,. \tag{E.31}$$

Dans le bilan du paragraphe 3, la somme des expressions (E.26) et (E.31), dues aux collisions entre particules, donne le second membre de l'équation d'évolution pour f. Pour alléger les notations, on pose traditionnellement

$$f \widehat{=} f(\vec{r}, \vec{p}\,; t) \,, \quad f_1 \widehat{=} f(\vec{r}, \vec{p}_1; t) \,, \quad f' \widehat{=} f(\vec{r}, \vec{p}\,'; t) \,, \quad f'_1 \widehat{=} f(\vec{r}, \vec{p}\,'_1; t) \,. \tag{E.32}$$

L'équation de Boltzmann s'écrit alors :

$$\boxed{\frac{\partial f}{\partial t} + \frac{\vec{p}}{m} \cdot \frac{\partial f}{\partial \vec{r}} + \vec{F} \cdot \frac{\partial f}{\partial \vec{p}} = \int d^3p_1 \int d\Omega \frac{1}{m} |\vec{p} - \vec{p}_1| . \sigma(|\vec{p} - \vec{p}_1|/2, \theta) \times (f' f'_1 - f f_1)} \tag{E.33}$$

Il faut se souvenir que, dans f' et f'_1, figurent des impulsions $\vec{p}\,'$ et $\vec{p}\,'_1$ qui sont des fonctions de $\vec{p}$, $\vec{p}_1$ et θ :

$$\vec{p}\,' = \frac{1}{2}(\vec{p} + \vec{p}_1) + \vec{p}\,'_r \,, \quad \vec{p}\,'_1 = \frac{1}{2}(\vec{p} + \vec{p}_1) - \vec{p}\,'_r \,, \tag{E.34}$$

avec $$|\vec{p}\,'_r| = \frac{1}{2} |\vec{p} - \vec{p}_1| \,, \tag{E.35.a}$$

$$\widehat{(\vec{p} - \vec{p}_1, \vec{p}\,'_r)} = \theta \,, \quad 0 \leqslant \theta \leqslant \pi \,. \tag{E.35.b}$$

---

10. Il suffit d'utiliser les formules (E.12). Le jacobien de la transformation (E.12.a) vaut -1, de sorte que

$$d^3p_0 . d^3p_1 = d^3p_G . d^3p_r \,. \tag{1}$$

De même,

$$d^3p'_0 . d^3p'_1 = d^3p_G . d^3p'_r \tag{2}$$

découle de (E.12.b). Or $\vec{p}\,'_r$ s'obtient à partir de $\vec{p}_r$ par une rotation d'angle θ, ce qui implique

$$d^3p'_r = d^3p_r \,. \tag{3}$$

L'égalité (E.29) s'en déduit aussitôt.

II. PROPRIÉTÉS GÉNÉRALES DE L'ÉQUATION DE BOLTZMANN

L'équation de Boltzmann joue un rôle de premier plan dans la compréhension théorique des *«phénomènes de transport»*, c'est-à-dire de la réponse d'un système maintenu dans des conditions extérieures de déséquilibre (différence de potentiel, de température, de concentration, ou de vitesse, imposée entre deux points du système). Nous n'aborderons pas ici ces questions. Nous nous contenterons de dégager quelques propriétés fondamentales de l'équation de Boltzmann, et de montrer qu'elles sont analogues à celles de l'équation maîtresse, étudiées au chapitre IV.

1. ÉQUATIONS DE CONSERVATION LOCALES DÉDUITES DE L'ÉQUATION DE BOLTZMANN

Les collisions conservent évidemment le *nombre de particules,* l'*impulsion* totale et l'*énergie cinétique* totale. Chacune de ces grandeurs vérifie donc, dans un gaz évoluant selon l'équation de Boltzmann, une *équation de conservation locale* : en chaque point $\vec{r}$, la variation dans le temps de la *densité volumique* de la grandeur considérée est donnée par la divergence du vecteur *densité de courant* associé. Pour comprendre comment ces équations de conservation locales découlent de l'équation de Boltzmann, prenons *l'exemple simple du nombre de particules*[11].

Considérons un élément de volume d^3r de l'espace ordinaire, repéré par le vecteur $\vec{r}$. Il est infinitésimal à l'échelle macroscopique, mais contient pourtant à chaque instant un grand nombre de particules, donné par

$$dN(\vec{r}\,;t)=d^3r\int d^3p\, f(\vec{r},\vec{p}\,;t)\,. \tag{E.36}$$

La densité volumique $n(\vec{r},t)$ correspondante s'écrit donc

$$n(\vec{r},t)=\int d^3p\, f(\vec{r},\vec{p}\,;t)\,. \tag{E.37}$$

Pour déduire de l'équation de Boltzmann une équation d'évolution pour $n(\vec{r},t)$, il suffit par conséquent d'intégrer les deux membres sur les impulsions $\vec{p}$.

Or *l'intégrale sur $\vec{p}$ du second membre de l'équation de Boltzmann est nulle.* Examinons en effet l'intégrale du terme $f'f'_1$:

$$I_{f'f'_1}(\vec{r},t)\widehat{=}\int d^3p\int d^3p_1\int d\Omega\,\frac{1}{m}\,|\vec{p}-\vec{p}_1|\sigma(|\vec{p}-\vec{p}_1|/2,\theta).f'f'_1\,. \tag{E.38}$$

Remplaçons les éléments différentiels $d^3p.d^3p_1$ par $d^3p'.d^3p'_1$ (relation (E.29)) et $|\vec{p}-\vec{p}_1|$ par $|\vec{p}\,'-\vec{p}\,'_1|$ (relation (E.28)); il vient

$$I_{f'f'_1}=\int d^3p'\int d^3p'_1\int d\Omega\,\frac{1}{m}\,|\vec{p}\,'-\vec{p}\,'_1|\sigma(|\vec{p}\,'-\vec{p}\,'_1|/2,\theta).f'f'_1\,. \tag{E.39}$$

Si maintenant nous changeons les variables muettes $\vec{p}\,'$ et $\vec{p}\,'_1$ en $\vec{p}$ et $\vec{p}_1$, $f'f'_1$ se change d'après les définitions (E.32) en ff_1, de sorte que

$$I_{f'f'_1}(\vec{r},t)=I_{ff_1}(\vec{r},t)\,. \tag{E.40}$$

Donc le second membre de l'équation de Boltzmann ne contribue pas à l'équation d'évolution de la densité de particules $n(\vec{r},t)$.

11. Pour les autres grandeurs, voir référence 3, chapitre 7 et référence 4, chapitre IX.

Dans le premier membre, le terme proportionnel à $\vec{F}(\vec{r})$ donne aussi une contribution nulle, et l'on obtient finalement[12] *l'équation de conservation locale du nombre de particules*

$$\frac{\partial}{\partial t} n(\vec{r}, t) + \operatorname{div} \vec{j}_N(\vec{r}, t) = 0 , \tag{E.41}$$

où la « densité de courant de particules » $\vec{j}_N$ est proportionnelle à la densité volumique d'impulsion :

$$\vec{j}_N(\vec{r}, t) \widehat{=} \frac{1}{m} \int d^3p \, \vec{p} f(\vec{r}, \vec{p} ; t) . \tag{E.42}$$

En intégrant cette équation locale dans un volume fixe V_0, on montre facilement que la dérivée par rapport au temps du nombre de particules contenu dans V_0 est égale, au signe près, au flux du vecteur $\vec{j}_N$ sortant à travers la surface s_0 qui limite ce volume :

$$\frac{d}{dt} \int_{V_0} d^3r \, n(\vec{r}, t) + \int_{s_0} \vec{ds} \cdot \vec{j}_N(\vec{r}, t) = 0 . \tag{E.43}$$

Lorsqu'on prend pour V_0 le volume total V du gaz, le second terme est nul car les molécules ne peuvent pas quitter le récipient; on retrouve ainsi que le nombre total N de particules est conservé, et donc que la fonction de répartition reste constamment normalisée selon (E.16) :

$$\frac{d}{dt} \int_V d^3r \, n(\vec{r}, t) = \frac{d}{dt} \int d^3r \int d^3p \, f(\vec{r}, \vec{p} ; t) = 0 . \tag{E.44}$$

2. ÉVOLUTION DE L'ENTROPIE : LE « THÉORÈME *H* »

Mais *l'équation de Boltzmann,* si elle conserve le nombre de particules, l'impulsion et l'énergie, *ne conserve pas l'entropie :* c'est en se fondant sur l'équation qu'il venait de proposer que Boltzmann établit le fameux « théorème *H* » concernant l'évolution de l'entropie d'un système hors d'équilibre. Nous allons reprendre ici cette démonstration.

---

12. Comme $\vec{F}$ ne dépend pas de $\vec{p}$, la contribution du troisième terme s'écrit

$$\vec{F} \cdot \int d^3p \, \frac{\partial f}{\partial \vec{p}} , \tag{1}$$

et les trois composantes de l'intégrale sont nulles. En effet, calculons par exemple

$$\int d^3p \, \frac{\partial}{\partial p_x} f(\vec{r}, \vec{p} ; t) = \int_{-\infty}^{+\infty} dp_y \int_{-\infty}^{+\infty} dp_z \int_{-\infty}^{+\infty} dp_x \, \frac{\partial}{\partial p_x} f(\vec{r}, \vec{p} ; t)$$

$$= \int_{-\infty}^{+\infty} dp_y \int_{-\infty}^{+\infty} dp_z \left[f(\vec{r}, p_x = +\infty, p_y, p_z ; t) - f(\vec{r}, p_x = -\infty, p_y, p_z ; t) \right] . \tag{2}$$

La fonction qui reste à intégrer est identiquement nulle car, pour être normalisable, f doit tendre vers zéro lorsque l'une quelconque des composantes de $\vec{r}$ ou $\vec{p}$ tend vers l'infini.

Les deux premiers termes de l'équation de Boltzmann (E.33) donnent ensuite facilement l'équation (E.41), car $\partial/\partial t$ et $\partial/\partial \vec{r}$ sortent de l'intégrale sur $\vec{p}$.

a. Équation d'évolution pour la densité locale d'entropie

L'entropie du gaz que nous considérons, dont l'état macroscopique est décrit par la fonction $f(\vec{r}, \vec{p}\,; t)$, est donnée (à une constante près) par la formule[13]

$$S(t) = -k \int \mathrm{d}^3r.\mathrm{d}^3p\, f \ln f\,. \tag{E.45}$$

Il est instructif d'étudier plus précisément le comportement de la *densité locale d'entropie,* c'est-à-dire de la quantité

$$H(\vec{r}, t) \widehat{=} -k \int \mathrm{d}^3p\, f(\vec{r}, \vec{p}\,; t) \ln f(\vec{r}, \vec{p}\,; t)\,. \tag{E.46}$$

Sa dérivée par rapport au temps s'écrit

$$\frac{\partial H}{\partial t} = -k \int \mathrm{d}^3p\, (1 + \ln f) \frac{\partial f}{\partial t}\,, \tag{E.47}$$

où $\partial f/\partial t$ est donnée par l'équation de Boltzmann :

$$\frac{\partial f}{\partial t} = -\frac{\vec{p}}{m} \cdot \frac{\partial f}{\partial \vec{r}} - \vec{F} \cdot \frac{\partial f}{\partial \vec{p}} + I_c\,. \tag{E.48}$$

Examinons d'abord la contribution à (E.47) des deux premiers termes du second membre de (E.48). Le plus simple est le deuxième. Comme la force $\vec{F}$ est indépendante de $\vec{p}$, sa contribution prend la forme

$$k\vec{F} \cdot \int \mathrm{d}^3p\, (1 + \ln f) \frac{\partial f}{\partial \vec{p}} = k\vec{F} \cdot \int \mathrm{d}^3p\, \frac{\partial}{\partial \vec{p}} (f \ln f)\,. \tag{E.49}$$

---

13. Nous savons (complément I.G, § III.2) que la définition de l'entropie statistique pose problème dans le cas de probabilités continues. Mais comme nous nous intéressons ici seulement à la *variation dans le temps* de l'entropie $S(t)$ du gaz, il nous suffit de la connaître à une constante près.

Pour bien comprendre l'origine de la formule (E.45), plaçons-nous dans le cadre plus général décrit au complément IV.F (§ II.1). L'entropie statistique associée à une densité $w(\vec{r}^{(N)}, \vec{p}^{(N)}\,; t)$ dans l'espace des phases à $6N$ dimensions du gaz s'écrit

$$S(t) = -k \frac{1}{N!} \int \mathrm{d}^{3N}r.\mathrm{d}^{3N}p\, w(\vec{r}^{(N)}, \vec{p}^{(N)}\,; t) \ln w(\vec{r}^{(N)}, \vec{p}^{(N)}\,; t) + C\,, \tag{1}$$

où C est une constante indéterminée. Si l'état macroscopique du gaz est caractérisé par la seule fonction $f(\vec{r}, \vec{p}\,; t)$, les N particules sont traitées comme indépendantes, c'est-à-dire que la densité w dans l'espace des phases du système s'écrit simplement

$$w(\vec{r}^{(N)}, \vec{p}^{(N)}\,; t) = \frac{N!}{N^N} \prod_{i=1}^{N} f(\vec{r}_i, \vec{p}_i; t) \tag{2}$$

(le facteur $N!/N^N$ provient de l'indiscernabilité des particules; voir formules (F.1) et (F.12) du complément IV.F). On calcule alors facilement

$$\begin{aligned} S(t) = C - k \frac{1}{N^N} \int \mathrm{d}^{3N}r.\mathrm{d}^{3N}p\, & f(\vec{r}_1, \vec{p}_1; t) \times \ldots \times f(\vec{r}_N, \vec{p}_N; t) \\ & \times [\ln f(\vec{r}_1, \vec{p}_1; t) + \ldots + \ln f(\vec{r}_N, \vec{p}_N; t) + \ln N! - N \ln N] \\ = C - k \left[\frac{1}{N} \int \mathrm{d}^3r.\mathrm{d}^3p\, f(\vec{r}, \vec{p}\,; t) \ln f(\vec{r}, \vec{p}\,; t)\right] & \times N - k(\ln N! - N \ln N) \end{aligned} \tag{3}$$

(on a utilisé la relation (E.16) de normalisation de f). On aboutit bien, à une constante additive près, à la formule (E.45).

Or, pour être normalisable, f doit tendre vers zéro pour les grandes valeurs de l'impulsion; il en résulte que l'intégrale précédente est nulle (cf. note 12).

Étudions ensuite le premier terme de l'expression (E.48). L'intégrale correspondante se transforme aisément :

$$\frac{k}{m}\int \mathrm{d}^3p\,(1+\ln f)\,\vec{p}\cdot\frac{\partial f}{\partial \vec{r}}=\frac{k}{m}\int \mathrm{d}^3p\,\vec{p}\cdot\frac{\partial}{\partial \vec{r}}\,(f\ln f)$$

$$=\frac{k}{m}\frac{\partial}{\partial \vec{r}}\cdot\int \mathrm{d}^3p\,\vec{p}f\ln f\,. \tag{E.50}$$

Elle est donc égale à la divergence d'un champ de vecteurs : si l'on pose

$$\vec{j}_H(\vec{r},t)\widehat{=}-\frac{k}{m}\int \mathrm{d}^3p\,\vec{p}f\ln f\,, \tag{E.51}$$

les égalités (E.47) et (E.48) se combinent pour donner

$$\frac{\partial H}{\partial t}+\operatorname{div}\vec{j}_H=-k\int \mathrm{d}^3p\,(1+\ln f)\,I_c\,. \tag{E.52}$$

Le premier membre de cette équation évoque une équation de conservation locale comme celle de la masse (§ 1) : si le second membre était nul (collisions totalement négligeables), l'entropie totale du gaz resterait constante au cours du temps comme la normalisation de f (formules (E.43)-(E.44)). Analysons donc ce second membre.

b. Contribution des collisions à l'évolution de l'entropie

Compte tenu de l'expression de l'intégrale de collision I_c (formule (E.33)), le second membre de l'équation d'évolution (E.52) de la densité locale d'entropie s'écrit

$$\left(\frac{\partial H}{\partial t}\right)_c=-k\int \mathrm{d}^3p\int \mathrm{d}^3p_1\int \mathrm{d}\Omega\,\frac{1}{m}\,|\vec{p}-\vec{p}_1|\,\sigma(|\vec{p}-\vec{p}_1|/2,\theta)$$

$$\times(1+\ln f)(f'f'_1-ff_1)\,. \tag{E.53}$$

Une méthode analogue à celle du paragraphe II.B.2.a du chapitre IV permet de mettre cette formule sous une forme plus symétrique, où son signe apparaît clairement.

Intervertissons d'abord les rôles des variables d'intégration $\vec{p}$ et $\vec{p}_1$ (ce qui intervertit aussi $\vec{p}\,'$ et $\vec{p}\,'_1$). Nous transformons ainsi (E.53) en

$$\left(\frac{\partial H}{\partial t}\right)_c=-k\int \mathrm{d}^3p_1\int \mathrm{d}^3p\int \mathrm{d}\Omega\,\frac{1}{m}\,|\vec{p}_1-\vec{p}\,|\,\sigma(|\vec{p}_1-\vec{p}\,|/2,\theta)$$

$$\times(1+\ln f_1)(f'_1f'-f_1f)\,, \tag{E.54}$$

et nous pouvons effectuer une première symétrisation en prenant la demi-somme des égalités (E.53) et (E.54) :

$$\left(\frac{\partial H}{\partial t}\right)_c=-\frac{k}{2}\int \mathrm{d}^3p\int \mathrm{d}^3p_1\int \mathrm{d}\Omega\,\frac{1}{m}\,|\vec{p}-\vec{p}_1|\,\sigma(|\vec{p}-\vec{p}_1|/2,\theta)$$

$$\times(2+\ln ff_1)(f'f'_1-ff_1)\,. \tag{E.55}$$

Effectuons ensuite le changement de variables

$$(\vec{p},\vec{p}_1)\rightsquigarrow(\vec{p}\,',\vec{p}\,'_1)\,. \tag{E.56}$$

D'après (E.29), les éléments d'intégration sont inchangés, de sorte que

$$\left(\frac{\partial H}{\partial t}\right)_c=-\frac{k}{2}\int \mathrm{d}^3p'\int \mathrm{d}^3p'_1\int \mathrm{d}\Omega\,\frac{1}{m}\,|\vec{p}-\vec{p}_1|\,\sigma(|\vec{p}-\vec{p}_1|)/2,\theta)$$

$$\times(2+\ln ff_1)(f'f'_1-ff_1)\,. \tag{E.57}$$

Dans la fonction à intégrer, c'est maintenant $\vec{p}$ et $\vec{p}_1$ qui sont des fonctions des nouvelles variables d'intégration $\vec{p}\,'$, $\vec{p}\,'_1$ et θ (cf. note 9). Mais ces variables sont muettes, ce qui permet de changer leur nom; si nous appelons $\vec{p}$ et $\vec{p}_1$ les impulsions sur lesquelles on intègre dans (E.57), les fonctions $\vec{p}$ et $\vec{p}_1$ qui figurent dans l'intégrale deviennent précisément $\vec{p}\,'$ et $\vec{p}\,'_1$, définies par les formules (E.34)-(E.35). Ceci ne change ni $|\vec{p}-\vec{p}_1|$ ni la section efficace; la formule (E.57) s'écrit donc sous la forme équivalente

$$\left(\frac{\partial H}{\partial t}\right)_c = -\frac{k}{2}\int d^3p \int d^3p_1 \int d\Omega \frac{1}{m}|\vec{p}-\vec{p}_1|\,\sigma(|\vec{p}-\vec{p}_1|/2,\theta) \times (2+\ln f'f'_1)(ff_1 - f'f'_1)\,. \qquad \text{(E.58)}$$

L'expression finale s'obtient en prenant à nouveau la demi-somme des égalités (E.55) et (E.58) (prendre garde aux signes!) :

$$\left(\frac{\partial H}{\partial t}\right)_c = \frac{k}{4}\int d^3p \int d^3p_1 \int d\Omega \frac{1}{m}|\vec{p}-\vec{p}_1|\,\sigma(|\vec{p}-\vec{p}_1|/2,\theta) \times (\ln f'f'_1 - \ln ff_1)(f'f'_1 - ff_1)\,. \qquad \text{(E.59)}$$

On constate sur cette expression que

$$\left(\frac{\partial H}{\partial t}\right)_c \geqslant 0\,, \qquad \text{(E.60)}$$

car la fonction à intégrer est partout positive ou nulle[14].

c. Conclusion

La densité locale d'entropie $H(\vec{r}, t)$ définie en (E.46) obéit donc à l'équation d'évolution

$$\frac{\partial}{\partial t}H(\vec{r}, t) + \operatorname{div}\vec{j}_H(\vec{r}, t) = \left(\frac{\partial H}{\partial t}\right)_c \geqslant 0\,. \qquad \text{(E.61)}$$

Si l'on intègre cette équation dans un volume V_0 fixe, on obtient la relation

$$\frac{d}{dt}\int_{V_0} H(\vec{r}, t)\,d^3r + \int_{s_0}\vec{j}_H(\vec{r}, t).\overrightarrow{ds} = \int_{V_0}\left(\frac{\partial H}{\partial t}\right)_c d^3r \geqslant 0\,, \qquad \text{(E.62)}$$

où s_0 est la surface limitant le volume V_0 et $\overrightarrow{ds}$ le vecteur élément de surface courant sur s_0. Cette relation s'interprète de la façon suivante : la variation dans le temps de l'entropie contenue dans le volume V_0 (premier terme) a pour origine, d'une part un *flux d'entropie* qui traverse la surface s_0 (deuxième terme, qui provoque une diminution d'entropie dans V_0 si le flux sortant est positif), et d'autre part une *création d'entropie* à l'intérieur de V_0 (deuxième membre de la relation, toujours non négatif), dont sont responsables les collisions.

Si l'on prend pour V_0 le volume total V du récipient contenant le gaz (que nous avons supposé isolé thermiquement), le flux d'entropie à travers les parois du récipient est nul, et l'on retrouve la conclusion générale du « théorème H » :

$$\frac{dS}{dt} \geqslant 0\,: \qquad \text{(E.63)}$$

l'entropie totale S du gaz ne peut pas décroître.

14. La fonction logarithme étant constamment croissante, $(\ln f'f'_1 - \ln ff_1)$ a partout même signe que $(f'f'_1 - ff_1)$.

Mais l'équation d'évolution (E.61) est plus riche : en premier lieu, elle fournit le *bilan local* du comportement de l'entropie dans le gaz hors d'équilibre, et introduit les notions de *flux d'entropie* et de *sources locales d'entropie;* en outre, elle précise *l'origine physique de la création d'entropie,* qui réside dans les *collisions* entre molécules. Cette richesse provient évidemment de l'équation de Boltzmann, qui la tire elle-même du *modèle physique relativement détaillé* dont on l'a déduite (§ I.1 et § I.2.a). En particulier, le rôle des collisions comme source d'entropie a pour origine directe le phénomène du *« chaos moléculaire », qui détruit constamment de l'information* : après un choc, les positions et les impulsions des deux molécules qui y ont pris part sont évidemment corrélées; mais chacune de ces deux molécules va entrer ensuite en collision avec d'autres molécules, provenant d'autres régions de l'espace et sortant elles-mêmes d'autres collisions; les perturbations extérieures vont rapidement, après quelques chocs seulement, « désorienter » les deux molécules considérées et donc effacer leurs corrélations. On comprend comment ce processus introduit une *irréversibilité dans le temps* : il y a dissymétrie entre « avant la collision », où les deux molécules ne sont pas corrélées, et « après la collision », où elles le sont, quoique de façon éphémère[15].

Les conclusions précédentes ne sont valables, à strictement parler, que pour le système particulier qui a été étudié (gaz classique) et dans le cadre des approximations restrictives que cette étude a nécessitées. Il est cependant vraisemblable que les notions introduites à cette occasion ont une portée plus générale : l'équation de Boltzmann est le prototype d'équations d'évolution plus détaillées que l'équation maîtresse globale du chapitre IV, et elle sert de guide pour la compréhension de systèmes plus difficiles à analyser qu'un gaz classique dilué.

3. DISTRIBUTIONS D'ÉQUILIBRE

Comme dans le cas du « théorème H » général (IV, § II.B.2.b), l'expression (E.59) du terme de création d'entropie permet également de trouver les caractéristiques de l'équilibre : ce terme est positif ou nul parce que la fonction figurant sous l'intégrale est partout positive ou nulle; l'entropie aura donc atteint son maximum lorsque cette fonction sera partout nulle, c'est-à-dire lorsque

$$f'f'_1 - ff_1 \equiv 0 \ . \tag{E.64}$$

On peut écrire cette condition sous la forme

$$\ln f' + \ln f'_1 \equiv \ln f + \ln f_1 \ . \tag{E.65}$$

Souvenons-nous que f' et f'_1 contiennent les impulsions finales $\vec{p}\,'$ et $\vec{p}\,'_1$ d'une collision dans laquelle $\vec{p}$ et $\vec{p}_1$, qui figurent dans f et f_1, sont les impulsions initiales. La condition (E.65) exprime donc que *le logarithme de la distribution d'équilibre f_e est une grandeur additive conservée dans les collisions élastiques.* Nous connaissons cinq fonctions de l'impulsion $\vec{p}$ qui sont des grandeurs additives conservées : la fonction constante, les trois composantes de $\vec{p}$ (conservation de l'impulsion), et la fonction $\vec{p}^{\,2}$ (conservation de l'énergie cinétique). On peut montrer[16] que toute autre fonction de

15. La comparaison avec l'équation de Vlasov (complément IV.F, § III.1) permet d'apprécier l'importance physique de l'équation de Boltzmann : il s'agit *a priori* dans les deux cas d'évaluer de façon approchée l'influence des interactions entre particules sur l'évolution du système; mais l'équation de Vlasov reste invariante par renversement du temps, alors que l'« approximation » tirée de l'hypothèse du « chaos moléculaire » conduit à une évolution irréversible.

16. Voir par exemple la référence 2 (p. 307), qui reproduit une démonstration de H. Grad (*Comm. Pure Appl. Maths. 2,* 311, 1949).

$\vec{p}$ ainsi conservée est nécessairement une *combinaison linéaire des cinq fonctions précédentes :*

$$\ln f_e = a + \vec{b} \cdot \vec{p} + c\vec{p}^{\,2} , \tag{E.66}$$

où les cinq coefficients a, $\vec{b}$ et c sont *a priori* des fonctions de $\vec{r}$ et t. Un changement simple de notations permet de mettre la formule précédente sous la forme

$$\ln f_e = \alpha - \gamma (\vec{v} - \vec{u})^2 , \tag{E.67}$$

où nous avons introduit la *vitesse* $\vec{v} = \vec{p}/m$, les cinq fonctions de $\vec{r}$ et t étant cette fois α, γ et $\vec{u}$.

La distribution d'équilibre est donc la distribution de Maxwell (complément III.C, § I)[17] :

$$f_e(\vec{r}, \vec{v} ; t) = e^{\alpha}\, e^{-\gamma(\vec{v}-\vec{u})^2} , \tag{E.68}$$

que l'on peut réécrire, en changeant à nouveau de notations,

$$f_e(\vec{r}, \vec{v} ; t) = n \left(\frac{m\beta}{2\pi}\right)^{3/2} e^{-\frac{1}{2}\beta m (\vec{v}-\vec{u})^2} . \tag{E.69}$$

Ici, n est le nombre de molécules par unité de volume[18], $\beta = 1/kT$ est le facteur habituel de température, et $\vec{u}$ est la vitesse d'ensemble des molécules; mais ces paramètres sont *a priori* fonctions de $\vec{r}$ et t, de sorte que cette distribution caractérise un *équilibre local et dépendant du temps.*

La distribution décrivant le véritable *équilibre macroscopique* du gaz doit être de la forme (E.69) et vérifier en outre l'équation de Boltzmann, que nous écrivons ici

$$\left(\frac{\partial}{\partial t} + \vec{v} \cdot \frac{\partial}{\partial \vec{r}} + \frac{1}{m} \vec{F} \cdot \frac{\partial}{\partial \vec{v}}\right) f_e(\vec{r}, \vec{v} ; t) = 0 , \tag{E.70}$$

puisque f_e annule le second membre. On peut montrer[19] que, pour un gaz isolé thermiquement et enfermé dans un récipient rigide immobile dans le référentiel (galiléen) choisi, cette équation implique que la vitesse d'ensemble $\vec{u}$ des molécules doit être nulle en tous points et à tout instant :

$$u_e(\vec{r}, t) \equiv 0 . \tag{E.71}$$

---

17. On comprend sans doute mieux, physiquement, comment les collisions parviennent à uniformiser la distribution des vitesses si l'on sait que la différence entre les énergies cinétiques de deux particules est en moyenne diminuée par une collision (exercice IV.11).

18. Dans la formule (E.69), les facteurs ont été choisis de telle sorte que

$$\int d^3v f_e(\vec{r}, \vec{v} ; t) = n(\vec{r}, t) , \tag{1}$$

la condition (E.16) de normalisation de f s'écrivant toujours

$$\int_V d^3r\, n(\vec{r}, t) = N . \tag{2}$$

Noter que la nécessité que f_e soit normalisable impose au coefficient c de la formule (E.66) d'être négatif, et à γ et β d'être positifs.

19. Le calcul, un peu long mais sans difficulté majeure, est décomposé en ses principales étapes dans l'exercice IV.12.

Si l'on admet ce résultat, on déduit facilement de l'équation (E.70) les conditions suivantes[20] :

$$\overrightarrow{\text{grad}}\ \gamma = \vec{0}\ , \tag{E.72.a}$$

$$\frac{\partial \gamma}{\partial t} = 0\ , \tag{E.72.b}$$

$$\overrightarrow{\text{grad}}\ \alpha - 2\gamma \frac{\vec{F}}{m} = \vec{0}\ , \tag{E.72.c}$$

$$\frac{\partial \alpha}{\partial t} = 0\ . \tag{E.72.d}$$

L'équilibre n'est donc possible que si *la force extérieure* $\vec{F}$ *est indépendante du temps.* Il est caractérisé par une *température uniforme et indépendante du temps* (γ constante). La répartition spatiale des molécules est déterminée par la relation (E.72.c); si la force $\vec{F}$ dérive d'une énergie potentielle $\Phi(\vec{r}\,)$:

$$\vec{F}\,(\vec{r}\,) = -\overrightarrow{\text{grad}}\ \Phi(\vec{r}\,)\ , \tag{E.73}$$

cette relation implique

$$\alpha(\vec{r}\,) = -\frac{2\gamma}{m}\,\Phi(\vec{r}\,) + \alpha_0\ , \tag{E.74}$$

où α_0 est une constante. La formule (E.68) s'écrit donc en définitive

$$f_e(\vec{r}, \vec{v}\,) = n_0 \left(\frac{m\beta}{2\pi}\right)^{3/2} e^{-\beta\left[\frac{1}{2}m\vec{v}^2 + \Phi(\vec{r}\,)\right]}\ , \tag{E.75}$$

où cette fois n_0 et β sont constantes[21].

La notion d'équilibre local suggérée par la formule (E.69) est à la base de la compréhension des *phénomènes de transport :* lorsqu'un système est maintenu dans une situation de déséquilibre forcé, par des conditions extérieures inhomogènes (qui imposent par exemple un gradient permanent de température ou de densité de particules), son évolution le conduit à une situation d'équilibre local où les paramètres n, β et $\vec{u}$ dépendent du point $\vec{r}$, mais elle ne peut pas atteindre son terme qui serait la situation d'équilibre global : le jeu des conditions aux limites inhomogènes imposées de l'extérieur au système et de sa tendance propre à évoluer vers l'équilibre macroscopique aboutit à un état stationnaire où les propriétés du système sont

---

20. Il suffit d'écrire que l'équation

$$\left(\frac{\partial}{\partial t} + \vec{v}\cdot\frac{\partial}{\partial \vec{r}} + \frac{1}{m}\,\vec{F}\cdot\frac{\partial}{\partial \vec{v}}\right) \ln f_e = 0\ , \tag{1}$$

c'est-à-dire, d'après (E.67) et (E.71),

$$\left(\frac{\partial}{\partial t} + \vec{v}\cdot\frac{\partial}{\partial \vec{r}} + \frac{1}{m}\,\vec{F}\cdot\frac{\partial}{\partial \vec{v}}\right)(\alpha - \gamma\vec{v}^{\,2}) = 0\ , \tag{2}$$

est vérifiée identiquement en $\vec{v}$.

21. On peut s'étonner de trouver une distribution de type canonique dans un système isolé thermiquement. C'est que les molécules situées dans le volume infinitésimal d^3r sont en contact thermique avec le reste du gaz (le couplage étant assuré par les collisions), qui constitue pour elles un thermostat. La température T est donc la température microcanonique de l'ensemble du gaz, qu'il impose à chaque molécule prise individuellement, comme dans la situation canonique décrite au chapitre III.

indépendantes du temps mais où il est parcouru par des *flux permanents* (de chaleur ou de particules, par exemple). Nous n'aborderons pas ici l'étude de ces phénomènes (voir références 1 à 4).

Références

1. F. Reif, *Fundamentals of Statistical and Thermal Physics,* McGraw-Hill, New York (1965), chapitres 12 à 14.
2. A. Sommerfeld, *Thermodynamics and Statistical Mechanics,* Academic Press, New York (1964), chapitre V.
3. H. J. Kreuzer, *Non-equilibrium Thermodynamics and its Statistical Foundations,* Clarendon Press, Oxford (1981), chapitre 7 et chapitres 2, 3 et 9.
4. S. R. de Groot et P. Mazur, *Non-equilibrium Thermodynamics,* Dover, New York (1984), chapitre IX et chapitres XI à XIII.
5. *The Boltzmann Equation : Theory and Applications* (E.G.D. Cohen et W. Thirring eds.), Springer-Verlag, New York (1973).

COMPLÉMENT IV.F

Équations générales d'évolution d'un fluide classique

I. L'équation de Liouville

II. La « hiérarchie *BBGKY* »
1. Définition des fonctions de répartition à n particules
2. Équations d'évolution des fonctions de répartition

III. Applications simples
1. L'équation de Vlasov
2. Équations locales de conservation

L'équation de Boltzmann que nous avons étudiée au complément IV.E régit l'évolution dans le temps de la « fonction de répartition à une particule » dans le cas simple d'un gaz dilué. Nous allons établir ici des équations d'évolution *exactes* pour les *« fonctions de répartition à plusieurs particules »* d'un fluide classique *a priori* quelconque. Outre leur intérêt théorique, ces équations peuvent se prêter à des approximations différentes, en principe plus systématiques et mieux contrôlées que celles qui conduisent à l'équation de Boltzmann.

I. L'ÉQUATION DE LIOUVILLE

Nous considérons un fluide constitué de N particules identiques dont le mouvement, à l'intérieur d'un récipient de volume V, peut être décrit par la *mécanique classique;* nous supposons que ces particules n'ont *pas de structure interne.* L'état macroscopique de ce système est alors caractérisé par une *densité de probabilité* $w(\vec{r}_1, \vec{p}_1; \vec{r}_2, \vec{p}_2; \ldots; \vec{r}_N, \vec{p}_N; t)$ dans son *espace des phases* à $6N$ dimensions (appendice III, paragraphe II.3.a). Pour alléger l'écriture des formules, nous noterons souvent $\vec{r}^{\,(N)}$ l'ensemble des N variables de position $\vec{r}_i$ ($i = 1, 2, \ldots, N$) et $\vec{p}^{\,(N)}$ l'ensemble des N impulsions $\vec{p}_i$, et de même $\mathrm{d}^{3N}r \,.\, \mathrm{d}^{3N}p$ le volume infinitésimal de l'espace des phases. Les particules étant identiques, la fonction $w(\vec{r}^{\,(N)}, \vec{p}^{\,(N)}; t)$ est *symétrique* par rapport à l'échange de deux couples $(\vec{r}_i, \vec{p}_i)$ et $(\vec{r}_j, \vec{p}_j)$ quelconques; nous la normalisons par la condition[1]

$$\frac{1}{N!} \int w(\vec{r}^{\,(N)}, \vec{p}^{\,(N)}; t)\, \mathrm{d}^{3N}r \,.\, \mathrm{d}^{3N}p = 1 \,, \tag{F.1}$$

1. Les particules sont *indiscernables :* nous divisons donc par $N!$ le volume de l'espace des phases (cf. III, § V.A.2.b), de sorte que

$$\frac{1}{N!}\, w(\vec{r}_1, \vec{p}_1; \vec{r}_2, \vec{p}_2; \ldots; \vec{r}_N, \vec{p}_N)\, \mathrm{d}^3r_1 \,.\, \mathrm{d}^3p_1 \,.\, \mathrm{d}^3r_2 \,.\, \mathrm{d}^3p_2 \ldots \mathrm{d}^3r_N \,.\, \mathrm{d}^3p_N$$

est la probabilité pour que l'une quelconque des N particules se trouve au point $\vec{r}_1$ à d^3r_1 près avec l'impulsion $\vec{p}_1$ à d^3p_1 près, une autre au point $\vec{r}_2$ avec l'impulsion $\vec{p}_2$, etc.

où l'intégrale porte sur tout l'espace des phases (chaque variable de position est en fait limitée à l'intérieur du volume V du récipient, mais il sera commode de considérer simplement que w devient nulle lorsque l'une des variables de position sort du volume V).

La densité de probabilité dans l'espace des phases vérifie *l'équation de Liouville* (appendice III, paragraphe II.3.b) :

$$\frac{\mathrm{d}w}{\mathrm{d}t} \widehat{=} \frac{\partial w}{\partial t} + \sum_{i=1}^{N} \left(\frac{\partial w}{\partial \vec{r}_i} \cdot \frac{\mathrm{d}\vec{r}_i}{\mathrm{d}t} + \frac{\partial w}{\partial \vec{p}_i} \cdot \frac{\mathrm{d}\vec{p}_i}{\mathrm{d}t} \right) = 0 , \tag{F.2}$$

où $\partial/\partial \vec{r}_i$ note l'opérateur gradient par rapport à $\vec{r}_i$, de composantes $\partial/\partial x_i$, $\partial/\partial y_i$ et $\partial/\partial z_i$, et $\partial/\partial \vec{p}_i$ note de même le gradient par rapport à $\vec{p}_i$. Si $\mathcal{H}(\vec{r}^{(N)}, \vec{p}^{(N)})$ est la fonction de Hamilton du système[2], associée à son énergie totale, les équations du mouvement s'écrivent de façon générale (appendice III, formules (14))

$$\frac{\mathrm{d}\vec{r}_i}{\mathrm{d}t} = \frac{\partial \mathcal{H}}{\partial \vec{p}_i} ,$$

$$\frac{\mathrm{d}\vec{p}_i}{\mathrm{d}t} = -\frac{\partial \mathcal{H}}{\partial \vec{r}_i} , \quad i = 1, 2, \ldots, N . \tag{F.3}$$

On définit traditionnellement le *«crochet de Poisson»* $\{A, B\}$ de deux fonctions $A(\vec{r}^{(N)}, \vec{p}^{(N)}; t)$ et $B(\vec{r}^{(N)}, \vec{p}^{(N)}; t)$ comme

$$\{A, B\} \widehat{=} \sum_{i=1}^{N} \left(\frac{\partial A}{\partial \vec{r}_i} \cdot \frac{\partial B}{\partial \vec{p}_i} - \frac{\partial A}{\partial \vec{p}_i} \cdot \frac{\partial B}{\partial \vec{r}_i} \right) . \tag{F.4}$$

L'équation de Liouville (F.2) prend alors, compte tenu des équations du mouvement (F.3), la forme compacte

$$\frac{\partial w}{\partial t} + \{w, \mathcal{H}\} = 0 . \tag{F.5}$$

Nous supposons désormais que les N particules interagissent seulement par des *forces à deux corps* qui dérivent d'une *énergie potentielle* $u(\vec{r}_i - \vec{r}_j)$ ne dépendant que de la position relative des deux particules; nous noterons pour simplifier

$$u_{ij} \widehat{=} u(\vec{r}_i - \vec{r}_j) . \tag{F.6}$$

Les particules peuvent en outre être soumises à des *forces extérieures* dérivant d'une énergie potentielle $\Phi(\vec{r})$:

$$\vec{F}(\vec{r}) = -\frac{\partial}{\partial \vec{r}} \Phi(\vec{r}) . \tag{F.7}$$

La fonction de Hamilton du système considéré s'écrit alors

$$\mathcal{H}(\vec{r}^{(N)}, \vec{p}^{(N)}) = \sum_{i=1}^{N} \left[\frac{\vec{p}_i^{\,2}}{2m} + \Phi(\vec{r}_i) \right] + \sum_{\substack{i,j=1 \\ i<j}}^{N} u_{ij} . \tag{F.8}$$

---

2. Nous supposons que la fonction de Hamilton ne dépend pas explicitement du temps :

$$\frac{\partial \mathcal{H}}{\partial t} = 0 ,$$

ce qui implique que l'énergie totale du système se conserve au cours du temps.

II. LA « HIÉRARCHIE *BBGKY* »

1. DÉFINITION DES FONCTIONS DE RÉPARTITION A *n* PARTICULES

A partir de la densité de probabilité globale $w(\vec{r}^{(N)}, \vec{p}^{(N)}; t)$, on peut définir des *« fonctions de répartition réduites »* f_n à n particules ($n \leq N$) en posant

$$f_n(\vec{r}^{(n)}, \vec{p}^{(n)}; t) \equiv f_n(\vec{r}_1, \vec{p}_1; \vec{r}_2, \vec{p}_2; \ldots; \vec{r}_n, \vec{p}_n; t)$$

$$\widehat{=} \frac{1}{(N-n)!} \int w(\vec{r}^{(N)}, \vec{p}^{(N)}; t)\, \mathrm{d}^3 r_{n+1} . \mathrm{d}^3 p_{n+1} \ldots . \mathrm{d}^3 r_N . \mathrm{d}^3 p_N . \qquad \text{(F.9)}$$

Ainsi, f_n est proportionnelle à la densité de probabilité de trouver à l'instant t une particule en $\vec{r}_1$ avec l'impulsion $\vec{p}_1$, une autre en $\vec{r}_2$ avec $\vec{p}_2$, ... et une n-ième en $\vec{r}_n$ avec $\vec{p}_n$, quelles que soient par ailleurs les positions et impulsions des $(N-n)$ autres particules; d'après (F.1), la normalisation de cette fonction est telle que

$$\frac{1}{n!} \int f_n(\vec{r}^{(n)}, \vec{p}^{(n)}; t)\, \mathrm{d}^{3n} r . \mathrm{d}^{3n} p = \frac{N!}{n!(N-n)!} \qquad \text{(F.10)}$$

(le second membre est égal au nombre de façons distinctes de choisir n particules parmi N). On vérifie en outre aisément que

$$\int f_{n+1} \mathrm{d}^3 r_{n+1} \mathrm{d}^3 p_{n+1} = (N-n) f_n . \qquad \text{(F.11)}$$

La *« fonction de répartition à une particule »* $f_1(\vec{r}, \vec{p}\,; t)$, qui est d'après (F.10) normalisée de telle sorte que

$$\int f_1(\vec{r}, \vec{p}\,; t)\, \mathrm{d}^3 r\, \mathrm{d}^3 p = N , \qquad \text{(F.12)}$$

s'interprète comme la densité moyenne[3] de particules au point $(\vec{r}, \vec{p}\,)$ de l'espace des phases individuel, à l'instant t. La fonction de répartition à $n = N$ particules coïncide évidemment avec $w(\vec{r}^{(N)}, \vec{p}^{(N)}; t)$.

2. ÉQUATIONS D'ÉVOLUTION DES FONCTIONS DE RÉPARTITION

L'équation régissant l'évolution de la fonction de répartition à n particules f_n s'obtient évidemment en intégrant l'équation de Liouville (F.5) sur les positions et impulsions des $(N-n)$ autres particules, comme l'indique la définition (F.9).

Commençons par établir cette équation pour la fonction de répartition à une particule $f_1(\vec{r}_1, \vec{p}_1; t)$. Nous multiplions pour cela l'équation de Liouville par

$$\frac{1}{(N-1)!} \mathrm{d}^3 r_2 . \mathrm{d}^3 p_2 . \ \ldots \ . \mathrm{d}^3 r_N . \mathrm{d}^3 p_N$$

et intégrons sur les $(N-1)$ positions $\vec{r}_i$ $(i = 2, 3, \ldots, N)$ et les $(N-1)$ impulsions $\vec{p}_i$ correspondantes. Nous obtenons ainsi

$$\frac{\partial f_1}{\partial t} = \frac{1}{(N-1)!} \int \left[\prod_{i=2}^{N} \mathrm{d}^3 r_i . \mathrm{d}^3 p_i \right] \{\mathcal{H}, w\} . \qquad \text{(F.13)}$$

---

3. La moyenne se réfère comme toujours à un ensemble statistique de systèmes identiques (I, § IV.B.2). Cependant, nous supposerons que le nombre considéré est suffisamment grand pour que ses fluctuations soient négligeables (il faut pour cela que l'élément de volume $\mathrm{d}^3 r . \mathrm{d}^3 p$ de l'espace des phases soit très grand à l'échelle microscopique); alors, $f_1(\vec{r}, \vec{p}\,; t)\mathrm{d}^3 r . \mathrm{d}^3 p$ est simplement *le nombre de particules* situées à l'instant t au point $\vec{r}$ à $\mathrm{d}^3 r$ près et animées de l'impulsion $\vec{p}$ à $\mathrm{d}^3 p$ près.

L'expression (F.8) de la fonction de Hamilton permet d'expliciter cette équation sous la forme[4]

$$\frac{\partial f_1}{\partial t} = I_1 + I_2 \,, \tag{F.14}$$

avec $$I_1 \widehat{=} \frac{1}{(N-1)!} \int \left[\prod_{i=2}^{N} \mathrm{d}^3 r_i \,.\, \mathrm{d}^3 p_i \right] \left\{ \frac{\vec{p}_1^{\,2}}{2m} + \Phi(\vec{r}_1),\ w \right\} , \tag{F.15.a}$$

$$I_2 \widehat{=} \frac{1}{(N-1)!} \int \left[\prod_{i=2}^{N} \mathrm{d}^3 r_i \,.\, \mathrm{d}^3 p_i \right] \left\{ \sum_{j=2}^{N} u_{1j},\ w \right\} . \tag{F.15.b}$$

---

4. La forme (F.14)-(F.15) de l'équation (F.13) découle des deux relations suivantes :

$$\int \mathrm{d}^3 r_i \, \mathrm{d}^3 p_i \left\{ \frac{p_i^2}{2m} + \Phi(\vec{r}_i),\ w \right\} = 0 \,, \tag{1}$$

$$\int \mathrm{d}^3 p_i \, \mathrm{d}^3 p_j \{ u_{ij},\ w \} = 0 \,. \tag{2}$$

Pour démontrer la première, remarquons que l'une des fonctions figurant dans le crochet de Poisson dépend seulement d'un couple de variables $(\vec{r}_i, \vec{p}_i)$, de sorte que la somme de la définition (F.4) se réduit à un seul terme :

$$\left\{ \frac{\vec{p}_i^{\,2}}{2m} + \Phi(\vec{r}_i),\ w \right\} = \frac{\partial \Phi}{\partial \vec{r}_i} \cdot \frac{\partial w}{\partial \vec{p}_i} - \frac{\vec{p}_i}{m} \cdot \frac{\partial w}{\partial \vec{r}_i} \,. \tag{3}$$

En outre, $\partial \Phi / \partial \vec{r}_i$ est indépendant de $\vec{p}_i$ et $\vec{p}_i / m$ évidemment indépendant de $\vec{r}_i$. On peut donc écrire

$$\int \mathrm{d}^3 r_i \, \mathrm{d}^3 p_i \left\{ \frac{\vec{p}_i^{\,2}}{2m} + \Phi(\vec{r}_i),\ w \right\} = \int \mathrm{d}^3 r_i \frac{\partial \Phi}{\partial \vec{r}_i} \cdot \int \mathrm{d}^3 p_i \frac{\partial w}{\partial \vec{p}_i} - \int \mathrm{d}^3 p_i \frac{\vec{p}_i}{m} \cdot \int \mathrm{d}^3 r_i \frac{\partial w}{\partial \vec{r}_i} \,. \tag{4}$$

Cette expression est une somme de termes nuls. En effet, prenons par exemple

$$\begin{aligned} \int \mathrm{d}^3 p_i \frac{\partial w}{\partial p_{ix}} &= \int_{-\infty}^{+\infty} \mathrm{d}p_{iy} \int_{-\infty}^{+\infty} \mathrm{d}p_{iz} \int_{-\infty}^{+\infty} \mathrm{d}p_{ix} \frac{\partial w}{\partial p_{ix}} \\ &= \int_{-\infty}^{+\infty} \mathrm{d}p_{iy} \int_{-\infty}^{+\infty} \mathrm{d}p_{iz} \left[w(\ldots, p_{ix} = +\infty, p_{iy}, p_{iz}, \ldots) - w(\ldots, p_{ix} = -\infty, p_{iy}, p_{iz}, \ldots) \right] . \end{aligned} \tag{5}$$

La fonction qui reste à intégrer est nulle car, pour être normalisable, w doit tendre vers zéro (suffisamment vite) lorsque l'une quelconque de ses variables (t excepté) tend vers l'infini. On montrerait de même que

$$\int \mathrm{d}^3 r_i \frac{\partial w}{\partial \vec{r}_i} = \vec{0} \tag{6}$$

(ici, w est strictement nulle dès que $\vec{r}_i$ sort du volume V).

Dans la relation (2), l'énergie d'interaction $u_{ij} \equiv u(\vec{r}_i - \vec{r}_j)$ dépend de deux variables de position différentes, mais elle est indépendante des impulsions. La définition (F.4) du crochet de Poisson donne donc ici

$$\{ u_{ij},\ w \} = \frac{\partial u_{ij}}{\partial \vec{r}_i} \cdot \frac{\partial w}{\partial \vec{p}_i} + \frac{\partial u_{ij}}{\partial \vec{r}_j} \cdot \frac{\partial w}{\partial \vec{p}_j} \,. \tag{7}$$

On montre facilement, comme ci-dessus, que l'intégrale sur $\mathrm{d}^3 p_i$ de $\partial w / \partial \vec{p}_i$ et l'intégrale sur $\mathrm{d}^3 p_j$ de $\partial w / \partial \vec{p}_j$ sont nulles, ce qui établit la formule (2).

Le premier terme I_1 se calcule sans difficulté : puisque l'intégration porte sur des variables que le crochet de Poisson n'affecte pas, il y apparaît à nouveau la fonction de répartition f_1[5] :

$$I_1 = \frac{\partial \Phi}{\partial \vec{r}_1} \cdot \frac{\partial}{\partial \vec{p}_1} f_1(\vec{r}_1, \vec{p}_1; t) - \frac{\vec{p}_1}{m} \cdot \frac{\partial}{\partial \vec{r}_1} f_1(\vec{r}_1, \vec{p}_1; t) . \tag{F.16}$$

Le deuxième terme I_2 est une somme de $(N-1)$ intégrales portant sur $\{u_{1j}, w\}$ pour $j = 2, 3, \ldots, N$; comme la fonction w est symétrique par rapport à l'échange des deux couples de variables $(\vec{r}_j, \vec{p}_j)$ et $(\vec{r}_2, \vec{p}_2)$, ces $(N-1)$ intégrales sont égales entre elles et l'on a

$$I_2 = \frac{1}{(N-2)!} \int \left[\prod_{i=2}^{N} \mathrm{d}^3 r_i \, \mathrm{d}^3 p_i \right] \{u_{12}, w\} . \tag{F.17}$$

Pour calculer cette expression, il faut évidemment traiter à part l'intégration sur $\vec{r}_2$ et $\vec{p}_2$, ce qui donne[6]

$$I_2 = \int \mathrm{d}^3 r_2 \, \mathrm{d}^3 p_2 \, \frac{\partial u_{12}}{\partial \vec{r}_1} \cdot \frac{\partial}{\partial \vec{p}_1} f_2(\vec{r}_1, \vec{p}_1; \vec{r}_2, p_2; t) . \tag{F.18}$$

Ainsi, *l'équation d'évolution pour* $f_1(\vec{r}_1, \vec{p}_1; t)$ *fait intervenir la fonction de répartition à 2 particules* $f_2(\vec{r}_1, \vec{p}_1; \vec{r}_2, \vec{p}_2; t)$. Cette équation s'écrit en définitive, si l'on regroupe (F.16) et (F.18) et qu'on fait apparaître la force extérieure (F.7),

$$\left[\frac{\partial}{\partial t} + \frac{\vec{p}_1}{m} \cdot \frac{\partial}{\partial \vec{r}_1} + \vec{F}(\vec{r}_1) \cdot \frac{\partial}{\partial \vec{p}_1} \right] f_1(\vec{r}_1, \vec{p}_1; t) = \int \mathrm{d}^3 r_2 \, \mathrm{d}^3 p_2 \, \frac{\partial u(\vec{r}_1 - \vec{r}_2)}{\partial \vec{r}_1} \cdot \frac{\partial}{\partial \vec{p}_1} f_2(\vec{r}_1, \vec{p}_1; \vec{r}_2, \vec{p}_2; t) . \tag{F.19}$$

---

5. Comme dans la note 4 (formule (3)), on a

$$\left\{ \frac{\vec{p}_1^{\,2}}{2m} + \Phi(\vec{r}_1), w \right\} = \frac{\partial \Phi}{\partial \vec{r}_1} \cdot \frac{\partial w}{\partial \vec{p}_1} - \frac{\vec{p}_1}{m} \cdot \frac{\partial w}{\partial \vec{r}_1} . \tag{1}$$

Lorsqu'on intègre sur $\vec{r}_2$, $\vec{p}_2$, $\vec{r}_3$, $\vec{p}_3$, ..., $\vec{r}_N$, $\vec{p}_N$, les facteurs $\partial \Phi / \partial \vec{r}_1$ et $\vec{p}_1/m$ sortent de l'intégrale, ainsi que les dérivées $\partial / \partial \vec{p}_1$ et $\partial / \partial \vec{r}_1$ qui s'appliquent à w : par exemple,

$$\frac{1}{(N-1)!} \int \left[\prod_{i=2}^{N} \mathrm{d}^3 r_i \, \mathrm{d}^3 p_i \right] \frac{\partial \Phi}{\partial \vec{r}_1} \cdot \frac{\partial w}{\partial \vec{p}_1} = \frac{\partial \Phi}{\partial \vec{r}_1} \cdot \frac{\partial}{\partial \vec{p}_1} \frac{1}{(N-1)!} \int \left[\prod_{i=2}^{N} \mathrm{d}^3 r_i \, \mathrm{d}^3 p_i \right] w . \tag{2}$$

L'intégrale restante n'est autre que f_1.

6. Le crochet de Poisson intervenant ici se calcule comme dans la note 4, formule (7) :

$$\{u_{12}, w\} = \frac{\partial u_{12}}{\partial \vec{r}_1} \cdot \frac{\partial w}{\partial \vec{p}_1} + \frac{\partial u_{12}}{\partial \vec{r}_2} \cdot \frac{\partial w}{\partial \vec{p}_2} . \tag{1}$$

Le deuxième terme donne une contribution nulle lorsqu'on intègre sur $\vec{p}_2$ (formule (5) de la note 4). Quant au premier, son intégrale s'écrit

$$\frac{1}{(N-2)!} \int \left[\prod_{i=2}^{N} \mathrm{d}^3 r_i \, \mathrm{d}^3 p_i \right] \frac{\partial u_{12}}{\partial \vec{r}_1} \cdot \frac{\partial w}{\partial \vec{p}_1} = \int \mathrm{d}^3 r_2 \, \mathrm{d}^3 p_2 \, \frac{\partial u_{12}}{\partial \vec{r}_1} \cdot \frac{\partial}{\partial \vec{p}_1} \frac{1}{(N-2)!} \int \left[\prod_{i=3}^{N} \mathrm{d}^3 r_i \, \mathrm{d}^3 p_i \right] w , \tag{2}$$

où l'on reconnaît la définition de la fonction de répartition f_2.

Le calcul précédent se généralise sans difficulté majeure pour donner l'équation d'évolution concernant la fonction de répartition f_n sous la forme

$$\frac{\partial f_n}{\partial t} + \{f_n,\ \mathcal{H}_n\} = \int \mathrm{d}^3 r_{n+1}\, \mathrm{d}^3 p_{n+1} \sum_{i=1}^{n} \frac{\partial u_{i\,n+1}}{\partial \vec{r}_i} \cdot \frac{\partial}{\partial \vec{p}_i} f_{n+1}, \tag{F.20}$$

où
$$\mathcal{H}_n \widehat{=} \sum_{i=1}^{n} \left[\frac{\vec{p}_i^{\,2}}{2m} + \Phi(\vec{r}_i)\right] + \sum_{\substack{i,\ j=1 \\ i<j}}^{n} u_{ij} \tag{F.21}$$

est la fonction de Hamilton du système réduit à n particules. Ainsi, aucune de ces équations (mise à part la dernière, pour $n = N$, qui n'est autre que l'équation de Liouville (F.5)) n'est « fermée » : il faut connaître f_{n+1} pour pouvoir étudier l'évolution de f_n. Le système des équations (F.20), pour n variant de 1 à N, constitue ce qu'on appelle la *« hiérarchie BBGKY »,* car y sont attachés les noms de Bogoliubov, Born et Green, Kirkwood, et Yvon.

Ce système d'équations est jusqu'ici *exact,* et parfaitement équivalent à l'équation de Liouville. Son principal intérêt est cependant qu'il se prête mieux que l'équation de Liouville complète aux *approximations,* qu'il permet d'introduire de façon systématique. En effet, pour pouvoir véritablement exploiter la « hiérarchie *BBGKY* », il est nécessaire de la *tronquer à un certain niveau,* c'est-à-dire de la réduire à un système fermé de n équations faisant intervenir seulement les n premières fonctions de répartition. Pour cela, on peut dans certains cas, en se fondant sur des données empiriques ou sur l'intuition physique, exprimer de façon approchée la fonction de répartition f_{n+1} à partir des fonctions de répartition à un nombre de particules inférieur ou égal à n. Dans d'autres cas, on postule tout simplement que le second membre de l'équation (F.20) est négligeable, quitte à le vérifier *a posteriori* en se fondant sur le résultat obtenu. Le système d'équations tronqué permet ensuite de calculer f_n, puis f_{n-1}, ...; il est évidemment très difficile en pratique de dépasser $n = 2$. Une illustration très simple de cette méthode d'approximation est donnée au paragraphe III.1 ci-dessous; l'équation de Boltzmann, que nous avons démontrée par des arguments plus concrets au complément IV.E (§ I), peut aussi être obtenue de cette façon (réf. 1, p. 179).

III. APPLICATIONS SIMPLES

1. L'ÉQUATION DE VLASOV

Dans l'équation (F.19) qui régit l'évolution de la fonction de répartition à une particule, supposons que la fonction f_2 puisse être remplacée au second membre par l'expression approchée[7]

$$f_2(\vec{r}_1, \vec{p}_1; \vec{r}_2, \vec{p}_2; t) \simeq f_1(\vec{r}_1, \vec{p}_1; t).f_1(\vec{r}_2, \vec{p}_2; t)\ . \tag{F.22}$$

Cette hypothèse considère les particules comme *indépendantes* les unes des autres : la probabilité pour trouver, à un instant t quelconque, deux particules l'une en $\vec{r}_1$ et l'autre en $\vec{r}_2$ avec les impulsions respectives $\vec{p}_1$ et $\vec{p}_2$, est alors le produit de la probabilité pour trouver à cet instant une particule en $\vec{r}_1$ avec l'impulsion $\vec{p}_1$ par la probabilité pour trouver au même instant une particule en $\vec{r}_2$ avec $\vec{p}_2$; *on néglige donc les corrélations* entre ces deux événements simultanés.

7. Pour respecter la condition de normalisation (F.10), il faudrait introduire au second membre le facteur $(N-1)/N$; mais ce facteur est pratiquement égal à 1 pour N grand.

L'équation (F.19) ne fait plus alors intervenir que la fonction de répartition à une particule $f_1(\vec{r}, \vec{p}\,; t)$; elle s'écrit[8]

$$\left[\frac{\partial}{\partial t}+\frac{\vec{p}}{m}\cdot\frac{\partial}{\partial \vec{r}}+\left[\vec{F}(\vec{r})-\frac{\partial}{\partial \vec{r}}\,\bar{u}(\vec{r}, t)\right]\cdot\frac{\partial}{\partial \vec{p}}\right] f_1(\vec{r}, \vec{p}\,; t)=0\,, \tag{F.23}$$

avec $$\bar{u}(\vec{r}, t)\widehat{=}\int \mathrm{d}^3r'\,\mathrm{d}^3p'\,u(\vec{r}-\vec{r}\,')f_1(\vec{r}\,', \vec{p}\,'; t)\,. \tag{F.24}$$

C'est l'équation proposée par Vlasov en 1938. L'hypothèse d'indépendance des particules a pour effet d'approximer leurs interactions par un *potentiel moyen* $\bar{u}(\vec{r}, t)$: lorsqu'une particule se trouve au point $\vec{r}$, elle y ressent globalement les effets des $(N-1)$ autres particules à travers ce potentiel $\bar{u}(\vec{r}, t)$, qui s'ajoute alors simplement au potentiel extérieur $\Phi(\vec{r})$ (cependant, comme l'équation est destinée à décrire une répartition des particules évoluant au cours du temps, le potentiel $\bar{u}$ dépend en général du temps t). Ce type d'approximation, dit de « champ moyen », est très souvent utilisé en mécanique statistique (cf. compléments III.G, III.H, III.J et III.K). Notons un point important : le potentiel moyen $\bar{u}$ dépend lui-même de la fonction de répartition f_1, que l'on cherche à déterminer[9]. On est donc confronté à un *« problème d'auto-cohérence »* : la solution f_1 de l'équation (F.23), où figure le potentiel moyen $\bar{u}$, doit redonner ce même potentiel lorsqu'on la reporte dans la formule (F.24).

L'équation de Vlasov est utilisée pour décrire l'évolution des plasmas dilués hors d'équilibre : le potentiel d'interaction $u(\vec{r}-\vec{r}\,')$ a dans ce cas la forme bien connue déduite de la loi de Coulomb, et le gradient du potentiel moyen $\bar{u}$ est à un facteur près le champ électrique (dépendant du point $\vec{r}$ et du temps t) produit par la distribution non uniforme des charges dans le plasma; ce champ, en s'ajoutant à celui que l'on applique éventuellement au plasma de l'extérieur, influe à son tour sur l'évolution des charges et donc de leur distribution.

Soulignons que, contrairement à l'équation de Boltzmann (complément IV.E, § II.2.c), l'équation de Vlasov est *invariante par renversement du temps* comme

---

8. Le second membre de l'équation (F.19) devient ici, compte tenu de l'hypothèse (F.22),

$$\int \mathrm{d}^3r_2\,\mathrm{d}^3p_2\,\frac{\partial u(\vec{r}_1-\vec{r}_2)}{\partial \vec{r}_1}\cdot\frac{\partial}{\partial \vec{p}_1}\,f_2(\vec{r}_1, \vec{p}_1; \vec{r}_2, \vec{p}_2; t)$$
$$=\frac{\partial}{\partial \vec{p}_1}\,f_1(\vec{r}_1, \vec{p}_1; t)\cdot\frac{\partial}{\partial \vec{r}_1}\int \mathrm{d}^3r_2\,\mathrm{d}^3p_2\,u(\vec{r}_1-\vec{r}_2)\,f_1(\vec{r}_2, \vec{p}_2; t)\,. \tag{1}$$

Avec la définition (F.24) de $\bar{u}(\vec{r}_1, t)$, cette expression prend la forme

$$\frac{\partial}{\partial \vec{p}_1}\,f_1(\vec{r}_1, \vec{p}_1; t)\cdot\frac{\partial}{\partial \vec{r}_1}\,\bar{u}(\vec{r}_1, t)\,, \tag{2}$$

qui n'est autre que le dernier terme de l'équation (F.23).

9. Il dépend seulement, en fait, de la répartition des particules dans l'espace ordinaire, et pas de la distribution des impulsions. En effet, si l'on définit la densité locale de particules

$$n(\vec{r}, t)\widehat{=}\int \mathrm{d}^3p\,f_1(\vec{r}, \vec{p}\,; t)\,, \tag{1}$$

on constate que le potentiel moyen $\bar{u}$ s'écrit

$$\bar{u}(\vec{r}, t)=\int \mathrm{d}^3r'\,u(\vec{r}-\vec{r}\,')\,n(\vec{r}\,', t)\,. \tag{2}$$

l'équation de Liouville[10] : si l'on change t en $-t$ dans l'équation de Vlasov (F.23), il faut changer simultanément $\vec{p}$ en $-\vec{p}$ (le renversement du temps transforme une vitesse en son opposée) et l'on constate que $f_1(\vec{r}, -\vec{p}\,; -t)$ vérifie la même équation que $f_1(\vec{r}, \vec{p}\,; t)$. Donc, si l'équation de Vlasov permet de rendre compte correctement de l'évolution d'un plasma dilué hors d'équilibre, elle ne peut pas décrire le processus irréversible qui amène ce plasma à l'équilibre macroscopique, lorsqu'il est abandonné à lui-même sans actions extérieures.

2. ÉQUATIONS LOCALES DE CONSERVATION

Lorsque, en hydrodynamique classique, on considère le fluide comme un milieu continu, les lois de conservation de la masse, de l'énergie, de l'impulsion et du moment cinétique se traduisent chacune par une *«équation de continuité»* reliant les *densités locales* de ces grandeurs et les *densités de courant* correspondantes.

Prenons l'exemple simple de la *conservation de la masse.* En hydrodynamique macroscopique, on définit la *densité volumique de masse* $\rho(\vec{r}, t)$ en posant que la masse de fluide contenue à l'instant t dans le volume infinitésimal d^3r situé autour du point $\vec{r}$ est égale à $\rho(\vec{r}, t)\mathrm{d}^3r$. Cette densité locale de masse va se déduire ici de la fonction de répartition à une particule f_1. En effet, choisissons des éléments de volume d^3r petits à l'échelle macroscopique mais de taille suffisante pour contenir à chaque instant un nombre de particules grand devant 1, et supposons que *la fonction de répartition* $f_1(\vec{r}, \vec{p}\,; t)$ *ne varie de façon appréciable que sur des distances d'ordre macroscopique.* Alors, le nombre de particules situées à l'instant t dans l'élément de volume d^3r, quelle que soit par ailleurs leur impulsion, s'écrit (cf. § II.1)

$$\mathrm{d}N(\vec{r}, t) = \mathrm{d}^3r \int \mathrm{d}^3p\, f_1(\vec{r}, \vec{p}\,; t)\,; \tag{F.25}$$

la masse $\rho(\vec{r}, t)\mathrm{d}^3r$ est simplement égale à $m\mathrm{d}N$, où m est la masse d'une particule, d'où

$$\rho(\vec{r}, t) = m \int \mathrm{d}^3p\, f_1(\vec{r}, \vec{p}\,; t)\,. \tag{F.26}$$

Cette égalité suggère que l'on peut obtenir une équation d'évolution pour ρ à partir de la première équation (F.19) de la hiérarchie *BBGKY*, en intégrant les deux membres par rapport à $\vec{p}_1$ après les avoir multipliés par m. *Le second membre donne une contribution nulle, de même que le terme proportionnel à* $\vec{F}(\vec{r}_1)$[11].

---

10. L'équation de Liouville incorpore en principe complètement les équations du mouvement microscopiques. Le manque d'information associé à la distribution de probabilité w reste donc constant au cours du temps (on pourrait le vérifier explicitement par une méthode analogue à celle du complément IV.E, § II.2.a). Effectivement, le changement de t en $-t$ et de $\vec{p}_i$ en $-\vec{p}_i$ laisse l'équation de Liouville (F.5) inchangée car le crochet de Poisson est du premier ordre par rapport aux dérivées $\partial/\partial\vec{p}_i$.

11. On a en effet, pour la même raison que dans la formule (5) de la note 4,

$$\int \mathrm{d}^3p_1 \frac{\partial}{\partial \vec{p}_1} f_2(\vec{r}_1, \vec{p}_1; \vec{r}_2, \vec{p}_2; t) = \vec{0} \tag{1}$$

et

$$\int \mathrm{d}^3p_1 \frac{\partial}{\partial \vec{p}_1} f_1(\vec{r}_1, \vec{p}_1; t) = \vec{0}\,. \tag{2}$$

Il reste donc[12]

$$\frac{\partial}{\partial t}\rho(\vec{r},t)+\operatorname{div}\vec{j}(\vec{r},t)=0\,, \tag{F.27}$$

où la *densité de courant de masse* $\vec{j}$, définie par

$$\vec{j}(\vec{r},t)\widehat{=}\int \mathrm{d}^3p\,\vec{p}\,f_1(\vec{r},\vec{p}\,;t)\,, \tag{F.28}$$

s'identifie à la *densité volumique d'impulsion.* On reconnaît dans (F.27) la forme habituelle d'une *équation de conservation locale,* ici celle de la masse, que l'on écrit plutôt en hydrodynamique

$$\frac{\partial}{\partial t}\rho(\vec{r},t)+\operatorname{div}[\rho(\vec{r},t)\,\vec{v}\,(\vec{r},t)]=0\,, \tag{F.29}$$

où $\vec{v}\,(\vec{r},t)$ est le *champ de vitesses* dans le fluide. Ce champ de vitesses se déduit ici de la densité d'impulsion : la vitesse d'ensemble d'un élément de fluide d^3r situé au point $\vec{r}$, dont la masse est $\rho(\vec{r},t)\mathrm{d}^3r$ à l'instant t et l'impulsion $\vec{j}(\vec{r},t)\mathrm{d}^3r$, est donnée par

$$\vec{v}\,(\vec{r},t)=\frac{1}{\rho(\vec{r},t)}\,\vec{j}(\vec{r},t)\,. \tag{F.30}$$

On pourrait démontrer de façon analogue des équations de conservation locale pour les *autres grandeurs conservées :* énergie, impulsion et moment cinétique (réf. 1, p. 158). Les densités que l'on définit alors font intervenir les deux premières fonctions de répartition f_1 et f_2 (celle-ci étant nécessaire pour la densité d'énergie potentielle). On établit ainsi un lien direct entre l'hydrodynamique macroscopique, qui traite le fluide comme un milieu continu, et la mécanique statistique des particules qui le constituent à l'échelle microscopique.

Références

1. H. J. Kreuzer, *Non-equilibrium Thermodynamics and its Statistical Foundations,* Clarendon Press, Oxford (1981), chapitre 7.

2. S. R. de Groot et P. Mazur, *Non-equilibrium Thermodynamics,* Dover, New York (1984), chapitre II.

12. Dans le premier terme, l'opérateur $\partial/\partial t$ sort de l'intégrale, qui donne ensuite directement $\rho(\vec{r},t)$. On traite de même l'opérateur vectoriel $\partial/\partial\vec{r}_1$ du deuxième terme, qui s'écrit alors

$$\frac{\partial}{\partial\vec{r}_1}\cdot\int \mathrm{d}^3p_1\,\vec{p}_1 f_1(\vec{r}_1,\vec{p}_1;t)=\operatorname{div}\vec{j}(\vec{r}_1,t)\,.$$

Exercices IV

ÉQUATION MAITRESSE

IV.1. Soit un système isolé. On suppose que les probabilités de transition par unité de temps entre ses Ω états accessibles sont toutes égales :

$$a^*_{\ell m} = a \qquad \forall(\ell) \text{ et } (m) \text{ accessibles.}$$

Écrire et résoudre l'équation maîtresse donnant la probabilité $P_\ell(t)$ de chacun des états accessibles.

IV.2. On considère un système microscopique ayant trois états possibles (1), (2), (3), d'énergies ε_1, ε_2, ε_3 (ce peut être par exemple une particule de spin 1 plongée dans un champ magnétique), en contact avec un thermostat de température T. On note P_ℓ la probabilité que le système se trouve dans l'état (ℓ), et $a_{\ell m}$ la probabilité de transition par unité de temps de l'état (m) à l'état (ℓ); on pose

$$a \widehat{=} a_{23}\, e^{-\varepsilon_3/kT}, \qquad b \widehat{=} a_{31}\, e^{-\varepsilon_1/kT}, \qquad c \widehat{=} a_{12}\, e^{-\varepsilon_2/kT}.$$

1) Écrire sous forme matricielle l'équation maîtresse donnant l'évolution dans le temps des probabilités P_ℓ.

2) On suppose d'abord $\varepsilon_1 = \varepsilon_2 = \varepsilon_3$.

a) Calculer, en fonction de a, b, c, les valeurs propres de la matrice 3×3 des coefficients de l'équation maîtresse. Montrer que l'une de ces valeurs propres est nulle et les deux autres négatives.

b) Quelles sont les valeurs d'équilibre des probabilités P_ℓ? Avec quel temps de relaxation sont-elles atteintes à partir d'un état hors d'équilibre?

3) On reprend le problème dans le cas général où ε_1, ε_2 et ε_3 sont différentes.

a) Montrer que la matrice des coefficients admet toujours une valeur propre nulle. Quel est le vecteur propre correspondant?

b) Écrire l'équation donnant les deux autres valeurs propres. Celles-ci sont-elles toujours réelles?(*) Si elles sont réelles, sont-elles toujours négatives?

c) Quelles sont les valeurs des probabilités P_ℓ à l'équilibre? Quelle est l'expression du temps de relaxation?

(*) Pour s'en assurer on pourra, après avoir posé

$$e_i \widehat{=} e^{\epsilon_i/kT}, \qquad i = 1, 2, 3,$$

effectuer le changement de variables

$$\frac{b}{a} \widehat{=} \beta, \quad \frac{c}{a} \widehat{=} \gamma, \quad \frac{e_2}{e_1} \widehat{=} \beta + x, \quad \frac{e_3}{e_1} \widehat{=} \gamma + y,$$

qui permet de ramener l'étude du discriminant de l'équation considérée à celle d'un trinôme du second degré en $X \widehat{=} x/y$.

APPLICATIONS PHYSIQUES

IV.3. Résonance magnétique. Un système de N particules de spin 1/2 *indépendantes,* dont chacune porte un moment magnétique $\vec{\mu}$, est plongé dans un champ magnétique uniforme $\vec{B}$ et placé en contact avec un thermostat de température T. On ne s'intéresse pas au mouvement des particules, seulement à leurs 2 états de spin possibles, (+) et (−). On notera P_+ et P_- les probabilités de ces deux états, a_{+-} et a_{-+} les probabilités de transition entre eux, par unité de temps, dues à l'interaction avec le thermostat.

1) Écrire les équations d'évolution de $P_+(t)$ et $P_-(t)$. En déduire celles de $P_+ + P_-$ et de $P \widehat{=} P_+ - P_-$.

2) Quelle relation doivent vérifier a_{+-} et a_{-+} lorsque le champ magnétique vaut B?

3) *a*) On choisit d'abord $B = 0$. Quelles sont les valeurs de P_+ et P_- à l'équilibre? Avec quel temps de relaxation τ_0 sont-elles atteintes?

b) Mêmes questions pour B non nul; on notera $\tau_1(T, B)$ le temps de relaxation.

4) On applique maintenant au système, plongé dans un champ B et toujours en contact avec le thermostat, un champ électromagnétique résonnant, c'est-à-dire de pulsation $\omega_0 = 2|\mu B|/\hbar$ (radiofréquence); celui-ci induit entre les états (+) et (−) des transitions dont la probabilité par unité de temps est la même dans les deux sens (cf. complément VI.E, § II) :

$$b_{+-} = b_{-+} = b \ .$$

a) Écrire la nouvelle équation d'évolution de $P(t)$.

b) Quelle est la valeur de P en régime stationnaire, et quel est le temps caractéristique τ d'établissement de ce régime? Comparer ces valeurs avec celles qui ont été obtenues à la question 3.*b*; discuter en fonction de $b\tau_1$. Montrer en particulier que, si a_{+-} et a_{-+} sont négligeables devant b, le champ de radiofréquence égalise les populations des deux états (cf. complément II.B, § III.2).

c) Quelle puissance le système des N spins reçoit-il, en moyenne, de la part du champ de radiofréquence? de la part du thermostat? Analyser le bilan d'énergie en régime stationnaire, et examiner la limite où $b\tau_1 \gg 1$.

IV.4. Polarisation dynamique par effet Overhauser. On considère un cristal contenant des atomes dont le noyau a un spin $I = 1/2$ et qui possèdent en outre un électron célibataire, de spin $S = 1/2$. Le moment magnétique μ_e de l'électron est de l'ordre du magnéton de Bohr $\mu_B \sim 10^{-23}$ J/T; celui μ_n du noyau est environ 1 000 fois plus petit, car de l'ordre du « magnéton nucléaire ».

On notera $P_+^{(e)}$ ou $P_-^{(e)}$ la probabilité pour que l'électron se trouve dans l'état de spin (+) ou (−), et $P_+^{(n)}$ et $P_-^{(n)}$ les probabilités analogues pour le noyau. On définit les « taux de polarisation » électronique $\eta^{(e)}$ et nucléaire $\eta^{(n)}$ par

$$\eta^{(e,n)} \widehat{=} \left| P_+^{(e,n)} - P_-^{(e,n)} \right| \ . \tag{1}$$

Pour réaliser certaines expériences de physique des particules, on cherche à *« polariser » les noyaux,* c'est-à-dire à obtenir un taux $\eta^{(n)}$ aussi grand que possible.

1) Pour cela, on tente d'abord d'appliquer un champ magnétique B aussi intense que possible et d'amener le système à une température T aussi basse que possible. Montrer, en prenant $B \sim 1$ tesla et $T \sim 1$ K, que l'effet obtenu à l'équilibre thermique est appréciable pour les électrons mais très faible pour les noyaux.

2) On utilise en réalité une *méthode de polarisation « dynamique »,* fondée sur l'« effet Overhauser ».

La principale (sinon la seule) interaction d'un spin nucléaire avec l'extérieur provient de l'hamiltonien de « structure hyperfine » de l'atome (C.D.L. pp. 1209 et

1220), qui couple le spin du noyau et celui de l'électron. Cet hamiltonien est essentiellement proportionnel à $\vec{I} \cdot \vec{S}$, de sorte que la somme $I_z + S_z$ des projections des deux spins le long du champ magnétique est conservée (C.D.L. p. 994); ainsi, le spin nucléaire ne peut « se retourner », par exemple de (+) à (−), que si celui de l'électron « se retourne » simultanément en sens inverse, de (−) à (+) (c'est ensuite le couplage des électrons avec le réseau, c'est-à-dire avec le thermostat, qui permet l'établissement de l'équilibre thermique, tant pour les noyaux que pour les électrons). Donc, pour ce qui concerne un spin nucléaire, les seules probabilités de transition (par unité de temps) appréciables sont $a_{+-;-+}$ et $a_{-+;+-}$ (on a indiqué en premier l'état du noyau, en second celui de l'électron).

a) L'atome (noyau + électron) étant en contact avec un thermostat de température T, montrer que

$$\frac{a_{-+;+-}}{a_{+-;-+}} = e^{-2(\mu_n - \mu_e)B/kT} . \tag{2}$$

b) L'interaction de structure hyperfine amène les spins nucléaires et les spins électroniques à l'équilibre entre eux, de sorte que la relation de bilan détaillé

$$a_{+-;-+} P_-^{(n)} P_+^{(e)} = a_{-+;+-} P_+^{(n)} P_-^{(e)} \tag{3}$$

est satisfaite. Vérifier que c'est bien le cas lorsque l'équilibre total est atteint avec le thermostat.

c) A l'aide d'un champ de radiofréquence résonnant (de pulsation $\omega_0 = 2|\mu_e B|/\hbar$), on « sature » la polarisation des électrons, c'est-à-dire qu'on réalise l'égalité

$$P_+^{(e)} = P_-^{(e)} \tag{4}$$

(cf. exercice IV.3, question 4.*b*). En admettant que la relation (3) reste valable, montrer que la polarisation des noyaux est alors celle qu'ils auraient, à l'équilibre thermique, si leur moment magnétique était μ_e et non μ_n. C'est ce résultat remarquable qui constitue l'effet Overhauser.

ÉQUATIONS D'ÉVOLUTION D'UN CORPUSCULE BROWNIEN

IV.5. L'équation de Langevin. Pour analyser le mouvement brownien (cf. complément I.D, § III), Langevin introduisit l'équation

$$m \frac{d\vec{v}}{dt} = \vec{F}_{ex}(t) - \alpha \vec{v} + \vec{F}_f(t) , \tag{1}$$

qui exprime la relation fondamentale de la dynamique appliquée au corpuscule (macroscopique, mais de très petite taille) dont on veut étudier le mouvement à l'intérieur d'un fluide : m est la masse du corpuscule, $\vec{v}$ sa vitesse; $\vec{F}_{ex}$ est une force d'origine extérieure (pesanteur, force électrique si le corpuscule est chargé et soumis à un champ électrique); l'action des molécules du fluide sur le corpuscule engendre une force fluctuante dont la valeur moyenne est la force de type visqueux $-\alpha\vec{v}$ et la partie aléatoire est représentée par le terme $\vec{F}_f(t)$; par définition, la valeur moyenne (prise sur un ensemble statistique) de ce dernier terme est nulle à chaque instant :

$$\overline{\vec{F}_f(t)} = \vec{0} . \tag{2}$$

Nous allons pour simplifier raisonner à une dimension, et tirer de l'équation de Langevin l'expression de l'écart quadratique moyen sur la position du corpuscule à l'instant t, c'est-à-dire du coefficient de diffusion D (complément I.D, note 12).

1) On se place dans le cas où F_{ex} est identiquement nulle.

a) Montrer que

$$m\frac{d}{dt}(xv) = mv^2 - \alpha xv + xF_f(t) . \tag{3}$$

b) On veut prendre la valeur moyenne (sur un ensemble statistique) des deux membres de cette égalité. Montrer que l'opération de valeur moyenne commute avec la dérivation par rapport au temps :

$$\frac{\overline{df}}{dt} = \frac{d}{dt}\bar{f} \quad \forall f(t) . \tag{4}$$

c) On suppose que la position x du corpuscule et la force aléatoire F_f ne sont pas corrélées, c'est-à-dire que

$$\overline{xF_f} = \bar{x}.\bar{F}_f \quad \text{pour tout } t \tag{5}$$

(cf. exercice IV.7). En utilisant le théorème d'équipartition de l'énergie, déduire de l'égalité (3) l'équation différentielle suivante pour la fonction $\overline{xv}(t)$:

$$m\frac{d}{dt}(\overline{xv}) = kT - \alpha\overline{xv} . \tag{6}$$

2) On suppose que les corpuscules de l'ensemble statistique partent tous de $x = 0$ à l'instant $t = 0$, et que leurs vitesses initiales sont réparties suivant la distribution d'équilibre à la température T.

a) Montrer que ces hypothèses impliquent, par raison de symétrie, que la valeur moyenne $\bar{x}(t)$ reste constamment nulle. Résoudre l'équation différentielle (6) avec ces conditions initiales.

b) En intégrant une nouvelle fois, trouver $\overline{x^2}$ en fonction du temps t.

3) *a*) Montrer que, pour $t \ll m/\alpha$, l'écart quadratique moyen Δx croît proportionnellement au temps t. Évaluer numériquement le temps caractéristique m/α (on utilisera les ordres de grandeur donnés au complément I.D (note 13); pour un corpuscule sphérique de rayon a, $\alpha = 6\pi\eta a$, où η est le coefficient de viscosité du fluide : $\eta \simeq 0{,}028$ g cm^{-1}s^{-1}). Comparer le résultat au temps τ_m caractéristique des fluctuations de $F_f(t)$ (cf. complément IV.B, § I.1).

b) Pour $t \gg m/\alpha$, Δx croît comme $\sqrt{t}$. Trouver l'expression du coefficient de diffusion D, donné par

$$(\Delta x)^2 = 2Dt , \quad t \gg m/\alpha . \tag{7}$$

4) On suppose maintenant que les corpuscules étudiés portent une charge électrique q, et qu'ils sont soumis à un champ électrique $\vec{\mathcal{E}}$ uniforme et indépendant du temps.

a) Montrer que la vitesse moyenne $\bar{v}$ des corpuscules le long de $\vec{\mathcal{E}}$ atteint une valeur limite $\bar{v}_{lim}$ proportionnelle à $\mathcal{E}$. En déduire la « mobilité » μ des corpuscules, définie par

$$\mu = \frac{\bar{v}_{lim}}{\mathcal{E}} . \tag{8}$$

b) Montrer que le coefficient de diffusion D et la mobilité μ vérifient la « relation d'Einstein »

$$\frac{\mu}{D} = \frac{q}{kT} . \tag{9}$$

IV.6. Équation de Langevin et théorème de fluctuation-dissipation. On étudie le mouvement d'un gros corpuscule, de masse m, dans un fluide en équilibre thermodynamique à la température T. A l'instant initial, la vitesse du corpuscule est fixée à la valeur $\vec{v}_0$. La vitesse $\vec{v}(t)$ à l'instant t est donnée par l'équation de Langevin (cf. exercice IV.5), qui s'écrit en l'absence de force extérieure

$$m\frac{\mathrm{d}\vec{v}}{\mathrm{d}t} = -\alpha\vec{v} + \vec{F}_f(t)\,. \tag{1}$$

On définit la « fonction d'autocorrélation » de la force fluctuante entre les instants t et t' par

$$C(t, t') \widehat{=} \overline{\vec{F}_f(t).\vec{F}_f(t')}\,, \tag{2}$$

la moyenne étant prise sur un ensemble statistique. $\vec{F}_f(t)$ est une fonction très rapidement variable, de moyenne statistique nulle; il est raisonnable de supposer que $C(t, t')$ ne dépend que de $|t-t'|$ et n'est différente de zéro que pour des intervalles de temps inférieurs à un « temps de corrélation » τ, lui-même très inférieur au temps m/α caractéristique du mouvement du corpuscule : $\tau \ll m/\alpha$.

1) Montrer que

$$\vec{v}(t) = \mathrm{e}^{-\alpha t/m}\left[\vec{v}_0 + \frac{1}{m}\int_0^t \mathrm{d}t'\,\mathrm{e}^{\alpha t'/m}\vec{F}_f(t')\right]. \tag{3}$$

En déduire la vitesse moyenne $\overline{\vec{v}(t)}$ (prise sur un ensemble statistique pour $\vec{v}_0$ fixée) et montrer que, pour $t \to \infty$, elle tend vers la valeur correspondant à l'équilibre statistique.

2) Montrer que le carré de la vitesse quadratique moyenne (toujours pour $\vec{v}_0$ fixée) vaut à l'instant t

$$\overline{\vec{v}^{\,2}(t)} = \mathrm{e}^{-2\alpha t/m}\left[\vec{v}_0^{\,2} + \frac{1}{m^2}\int_0^t \mathrm{d}t' \int_0^t \mathrm{d}t''\,\mathrm{e}^{\alpha(t'+t'')/m} C(t'-t'')\right]. \tag{4}$$

3) *a*) Montrer, grâce au changement de variables

$$x = t' + t'', \quad y = t' - t'',$$

que $\overline{\vec{v}^{\,2}(t)}$ peut encore s'écrire

$$\overline{\vec{v}^{\,2}(t)} = \mathrm{e}^{-2\alpha t/m}\left[\vec{v}_0^{\,2} + \frac{1}{2m^2}\int_0^t \mathrm{d}x\,\mathrm{e}^{\alpha x/m}\int_{-x}^{+x}\mathrm{d}y\,C(y) + \frac{1}{2m^2}\int_t^{2t}\mathrm{d}x\,\mathrm{e}^{\alpha x/m}\int_{x-2t}^{2t-x}\mathrm{d}y\,C(y)\right]. \tag{5}$$

b) Montrer que, pour $t \gg \tau$, on peut remplacer dans les deux intégrales sur y les bornes d'intégration par $-\infty$ et $+\infty$. Calculer $\overline{\vec{v}^{\,2}(t)}$ dans ce domaine.

c) Lorsque $t \to \infty$, le corpuscule se trouve en équilibre thermique avec le fluide. Utiliser alors l'équipartition de l'énergie pour montrer que le « coefficient de friction » α qui figure dans l'équation de Langevin est relié à la fonction d'autocorrélation $C(t)$ de la force fluctuante par

$$\alpha = \frac{1}{6kT}\int_{-\infty}^{+\infty} C(t)\,\mathrm{d}t\,. \tag{6}$$

Ce résultat est l'expression, dans cet exemple simple, du fameux « théorème de fluctuation-dissipation ».

IV.7. Coefficient de diffusion d'une particule brownienne. On veut utiliser les résultats de l'exercice précédent pour déterminer le coefficient de diffusion d'une particule brownienne de façon plus rigoureuse que dans l'exercice IV.5.

1) Montrer que la position du corpuscule à l'instant t se calcule aisément à partir de l'expression de sa vitesse $\vec{v}(t)$ et peut s'écrire (après une intégration par parties)

$$\vec{r}(t)=\frac{m}{\alpha}\vec{v}_0(1-e^{-\alpha t/m})+\frac{1}{\alpha}\int_0^t dt'(1-e^{-\alpha(t-t')/m})\vec{F}_f(t')\,,$$

si la position initiale est choisie comme origine.

2) Exprimer son carré statistique moyen $\overline{\vec{r}^2(t)}$ à l'aide de la fonction d'autocorrélation $C(t)$.

3) Pour calculer $\overline{\vec{r}^2(t)}$ dans la limite $t \gg \tau$, on peut utiliser la même méthode que pour $\overline{\vec{v}^2(t)}$ dans l'exercice IV.6. On peut aussi remarquer que, dans cette limite, tout se passe comme si $C(t)$ était proportionnelle à la fonction de Dirac $\delta(t)$, le coefficient de proportionnalité étant donné par la relation (6) de l'exercice IV.6 :

$$C(t)\simeq 6kT\alpha\,\delta(t)\,.$$

Montrer que, pour $t \gg \tau$,

$$\overline{\vec{r}^2(t)}=\frac{m^2}{\alpha^2}\vec{v}_0^{\,2}(1-e^{-\alpha t/m})^2+\frac{3kTm}{\alpha^2}\left[\frac{2\alpha t}{m}-3+4\,e^{-\alpha t/m}-e^{-2\alpha t/m}\right].$$

4) Montrer que, pour $t \gg m/\alpha$ (et pour $\vec{v}_0^{\,2}$ de l'ordre de kT/m), $\overline{\vec{r}^2(t)}$ est proportionnel à t; déterminer le coefficient de diffusion défini par

$$\overline{x^2(t)}\simeq\overline{y^2(t)}\simeq\overline{z^2(t)}\simeq\frac{1}{3}\overline{\vec{r}^2(t)}\widehat{=}2Dt\,,\qquad t\gg\frac{m}{\alpha}\,.$$

5) Contrairement aux apparences, le résultat obtenu pour $\overline{\vec{r}^2(t)}$ n'est pas en désaccord avec le calcul de $\overline{x^2(t)}$ effectué dans l'exercice IV.5. Supposons en effet le corpuscule en équilibre thermique avec le fluide à l'instant initial et prenons, dans l'expression de $\overline{\vec{r}^2(t)}$, une moyenne statistique sur les valeurs de $\vec{v}_0$. Montrer que le résultat obtenu est alors en accord avec celui de l'exercice IV.5, question 3.*b*. Pour qu'un tel calcul soit justifié, il faut que les deux opérations de moyenne, aux instants t et 0, soient indépendantes, c'est-à-dire que l'intervalle de temps qui les sépare doit être beaucoup plus grand que le temps de corrélation (soit, là encore, $t \gg \tau$).

IV.8. Marche au hasard et équation de la diffusion. Une particule se déplace au hasard sur un axe Ox en effectuant des « pas » de longueur fixe Δ, soit vers la droite, soit vers la gauche, avec une égale probabilité; deux pas successifs sont séparés par un intervalle de temps τ, toujours le même. Les pas sont statistiquement indépendants. On note $P(n, s)$ la probabilité pour que la particule se trouve au point d'abscisse $n\Delta$ au temps $s\tau$, sachant qu'elle se trouve à l'origine des coordonnées au temps $t=0$.

1) Exprimer $P(n, s)$ en fonction des $P(n', s-1)$.

2) On suppose que Δ et τ tendent simultanément vers zéro de telle façon que $\Delta^2/2\tau$ tende vers une constante D non nulle. On pose $n\Delta = x$ et $s\tau = t$. En développant jusqu'au premier ordre significatif la relation obtenue en 1, montrer que la densité de probabilité $w(x, t)$ pour que la particule se trouve au point x (à dx près) à l'instant t vérifie « l'équation de la diffusion »

$$\frac{\partial w}{\partial t}=D\frac{\partial^2 w}{\partial x^2}\,.$$

3) En déduire les équations d'évolution pour la position moyenne $\bar{x}(t)$ de la particule à l'instant t et pour l'écart quadratique moyen $\Delta x(t)$ correspondant. Calculer $\bar{x}(t)$ et $\Delta x(t)$ (on supposera que $w(x, t)$ décroît suffisamment vite quand $|x| \to \infty$ pour que $\bar{x}(t)$ et $\Delta x(t)$ restent définis).

4) Vérifier que la distribution gaussienne $w_G(x, t)$ ayant pour centre et pour écart quadratique moyen les valeurs $\bar{x}(t)$ et $\Delta x(t)$ calculées ci-dessus satisfait l'équation de la diffusion.

IV.9. Marche au hasard d'une particule liée élastiquement. Équation de Smoluchowski. Une particule se déplace au hasard sur un axe Ox, en effectuant des pas de longueur Δ séparés par des intervalles de temps τ, comme dans l'exercice IV.8. Mais ici la probabilité qu'elle se déplace vers la droite ou la gauche dépend de la position de la particule : lorsqu'elle se trouve au point d'abscisse $m\Delta$, elle a la probabilité $p = (1 - m/N)/2$ d'effectuer le pas suivant vers la droite et $q = (1 + m/N)/2$ de l'effectuer vers la gauche, N étant un entier donné. On note $P(n, s; n_0)$ la probabilité de trouver la particule au point d'abscisse $n\Delta$ au temps $s\tau$, sachant qu'elle est partie du point $n_0\Delta$ à l'instant $t = 0$.

1) Exprimer $P(n, s; n_0)$ en fonction des $P(n', s - 1; n_0)$.

2) On suppose que $\Delta \to 0$, $\tau \to 0$, $N \to \infty$ de telle façon que $\Delta^2/2\tau \to D$ et $1/N\tau \to \gamma$, où D et γ sont deux constantes non nulles. Soit alors $w(x, t; x_0)$ la densité de probabilité pour que la particule se trouve au point d'abscisse x $(= n\Delta)$ à l'instant t $(= s\tau)$ si elle est partie du point x_0 $(= n_0\Delta)$ au temps $t = 0$. Montrer, à partir de la relation obtenue en 1, que w vérifie «l'équation de Smoluchowski»

$$\frac{\partial w}{\partial t} = \gamma \frac{\partial (xw)}{\partial x} + D \frac{\partial^2 w}{\partial x^2} .$$

3) En déduire les équations d'évolution pour les valeurs moyennes de $x(t)$ et $x^2(t)$. Calculer $\bar{x}(t)$ et l'écart quadratique moyen $\Delta x(t)$.

4) Montrer que la distribution gaussienne $w_G(x, t; x_0)$ ayant pour centre et pour écart quadratique moyen les valeurs $\bar{x}(t)$ et $\Delta x(t)$ calculées ci-dessus satisfait l'équation de Smoluchowski.

COLLISIONS ET ÉQUATION DE BOLTZMANN

IV.10. Un solide en équilibre à la température T_S est plongé dans un gaz en équilibre à la température T_G. On supposera que les atomes constituant ce système sont indépendants les uns des autres et qu'ils peuvent être décrits par la mécanique classique.

1) Un atome de gaz, de masse m et de vitesse $\vec{v}$, entre en collision élastique avec un atome du solide, de masse M et de vitesse $\vec{V}$. Quelle est l'énergie $\delta\varepsilon$ transférée du solide vers le gaz au cours d'une telle collision?

2) Calculer la moyenne $\overline{\delta\varepsilon}$ de ce transfert d'énergie sur un grand nombre de collisions.

3) Montrer que l'énergie passe du corps chaud au corps froid, et que l'équilibre est atteint lorsque le solide et le gaz ont même température.

IV.11. On étudie une collision élastique entre deux particules identiques; les notations sont celles du complément IV.E, paragraphe I.2.b.

1) Exprimer la différence entre les énergies cinétiques initiales des deux particules en fonction du produit scalaire $\vec{p}_G \cdot \vec{p}_r$, et la différence entre les énergies cinétiques finales en fonction de $\vec{p}_G \cdot \vec{p}_r'$.

2) On choisit l'axe Oz le long de $\vec{p}_r$, et l'axe Ox dans le plan défini par $\vec{p}_G$ et $\vec{p}_r$. On pose

$$\widehat{(\vec{p}_r, \vec{p}_G)} = \theta_G , \quad 0 \leqslant \theta_G \leqslant \pi .$$

Calculer $\vec{p}_G \cdot \vec{p}_r$ et $\vec{p}_G \cdot \vec{p}_r'$ en fonction des modules p_G et p_r, de θ_G, et des angles θ et φ repérant la direction de $\vec{p}_r'$.

3) La probabilité pour que, p_r étant fixé, la direction de $\vec{p}_r'$ se trouve dans l'angle solide infinitésimal $d\Omega = \sin\theta \, d\theta \, d\varphi$ autour de la direction (θ, φ) est égale à $\sigma(p_r, \theta) d\Omega$, où σ est la section efficace de collision. Montrer que la valeur moyenne de $\vec{p}_G \cdot \vec{p}_r'$ s'écrit

$$\overline{\vec{p}_G \cdot \vec{p}_r'} = \vec{p}_G \cdot \vec{p}_r \, \overline{\cos\theta} .$$

En déduire qu'une collision élastique diminue en moyenne la différence entre les énergies cinétiques des deux particules qui y prennent part.

IV.12. Conditions d'équilibre d'un gaz dilué (complément IV.E). On reprend la condition d'équilibre global (formule (E.70) du complément IV.E) dans le cas général où $\vec{u}$ n'est pas nulle *a priori*, et on veut montrer que cette condition implique la nullité de $\vec{u}$ pour un gaz enfermé dans un récipient à parois adiabatiques et rigides, immobile dans le référentiel du laboratoire.

1) *Relations vérifiées par les composantes de* $\vec{u}$

Expliciter les termes du troisième et du deuxième ordre en v de l'équation

$$\left(\frac{\partial}{\partial t} + \vec{v} \cdot \frac{\partial}{\partial \vec{r}} + \frac{1}{m} \vec{F} \cdot \frac{\partial}{\partial \vec{v}} \right) \ln f_e = 0 , \tag{1}$$

où $\ln f_e$ a la forme générale (E.67). En déduire que la condition

$$\overrightarrow{\text{grad}}\, \gamma = \vec{0} \tag{2}$$

reste inchangée, et que les composantes de $\vec{u}$ doivent vérifier

$$-\frac{\partial \gamma}{\partial t} + 2\gamma \frac{\partial u_x}{\partial x} = 0 , \tag{3.a}$$

$$\frac{\partial u_y}{\partial x} + \frac{\partial u_x}{\partial y} = 0 , \tag{3.b}$$

et les quatre autres conditions qui s'en déduisent par permutation circulaire de x, y, z.

2) *Solution générale des conditions précédentes*

a) Montrer que les trois conditions du type (3.a) impliquent que les composantes de $\vec{u}$ sont de la forme

$$u_x = \frac{\gamma'}{2\gamma} x + g_1(y, z, t) , \tag{4.a}$$

$$u_y = \frac{\gamma'}{2\gamma} y + g_2(z, x, t) , \tag{4.b}$$

$$u_z = \frac{\gamma'}{2\gamma} z + g_3(x, y, t) , \tag{4.c}$$

où γ' est la dérivée de la fonction $\gamma(t)$.

b) En appliquant les conditions (3.b) à ces expressions, écrire les trois relations que doivent vérifier les dérivées partielles des fonctions g_1, g_2, g_3.

c) Montrer que ces trois relations imposent aux fonctions g_1, g_2, g_3 d'être de la forme

$$g_1(y, z, t) = A_1(t)\, yz + B_1(t)\, y + C_1(t)\, z + D_1(t) \tag{5}$$

(les expressions de g_2 et g_3 étant obtenues par permutation circulaire simultanée de x, y, z et 1, 2, 3), et qu'elles donnent en outre les égalités

$$A_1 + A_2 = A_2 + A_3 = A_3 + A_1 = 0\ , \tag{6.a}$$

$$B_1 + C_2 = C_3 + B_2 = C_1 + B_3 = 0\ . \tag{6.b}$$

d) En déduire que la solution générale des équations (3) s'écrit finalement

$$\vec{u} = \frac{\gamma'}{2\gamma}\,\vec{r} + \vec{a}\,(t) \wedge \vec{r} + \vec{b}\,(t)\ , \tag{7}$$

où $\vec{a}$ et $\vec{b}$ sont deux vecteurs arbitraires pouvant dépendre du temps.

3) *Solution particulière prenant en compte les conditions aux limites*

Supposons pour simplifier que le récipient rigide et immobile contenant le gaz est parallélépipédique, les axes étant choisis parallèles à ses arêtes. La composante u_x de $\vec{u}$ doit alors s'annuler, pour $x = \pm L_1$, quels que soient y, z et t (les composantes u_y et u_z vérifiant des conditions analogues sur les faces perpendiculaires à Oy et Oz).

Montrer que ces conditions aux limites imposent la nullité de γ', $\vec{a}\,(t)$ et $\vec{b}\,(t)$ pour tout t.

IV.13. L'équation de Boltzmann dans l'approximation du temps de relaxation. On considère un plasma dilué et peu ionisé, et on s'intéresse à un type d'ions, dont chacun porte la charge e. On note $f(\vec{r}, \vec{v}, t)$ la fonction de répartition de ces ions dans l'espace des phases individuel à 6 dimensions (on utilise ici la vitesse $\vec{v}$ à la place de l'impulsion $\vec{p}$).

1) On veut d'abord établir pour f une *équation de Boltzmann simplifiée,* dans laquelle la contribution des collisions est traitée de façon approchée.

Le plasma étant peu ionisé, c'est presque exclusivement avec des molécules neutres que les particules chargées que nous étudions entrent en collision. On pose l'hypothèse simple que chaque ion a la probabilité $\mathrm{d}t/\tau$ (où τ est un temps caractéristique) de subir une collision pendant l'intervalle de temps infinitésimal $\mathrm{d}t$.

a) En reprenant le bilan d'évolution du nombre d'ions situés dans un domaine infinitésimal fixe $\delta^3 r_0 \delta^3 v_0$ de l'espace des phases (complément IV.E, § I.3), montrer que la contribution des collisions au dépeuplement de ce domaine entre t et $t + \mathrm{d}t$ s'écrit

$$\mathrm{d}n_c = -f(\vec{r}_0, \vec{v}_0, t)\, \delta^3 r_0 . \delta^3 v_0 . \frac{\mathrm{d}t}{\tau}\ . \tag{1}$$

b) Ce traitement très grossier des collisions ne permet pas de calculer directement le terme de peuplement qu'elles induisent. On le déduit d'une nouvelle hypothèse : la contribution totale des collisions s'annule (ainsi que le premier membre de l'équation de Boltzmann; cf. question 3) lorsque la fonction de répartition coïncide avec celle d'équilibre, $f_e(\vec{r}, \vec{v})$, que l'on se donne par ailleurs. Montrer que l'équation de Boltzmann s'écrit alors, dans cette *« approximation du temps de relaxation »,*

$$\frac{\partial f}{\partial t} + \vec{v} \cdot \frac{\partial f}{\partial \vec{r}} + \frac{\vec{F}}{m} \cdot \frac{\partial f}{\partial \vec{v}} = -\frac{1}{\tau}\,(f - f_e)\ . \tag{2}$$

En l'absence de force extérieure $\vec{F}$, la fonction de répartition à l'équilibre est indépendante de $\vec{r}$; on sait par ailleurs (complément III.C, § I.1) que c'est la distribution de Maxwell :

$$f_e^0(\vec{v}) = \frac{N}{V}\left(\frac{m}{2\pi kT}\right)^{3/2} e^{-m\vec{v}^2/2kT}, \tag{3}$$

si le système est maintenu à la température T.

2) La densité volumique de charge $\rho(\vec{r}, t)$ et la densité de courant $\vec{j}(\vec{r}, t)$ associées aux ions que nous considérons s'écrivent

$$\rho(\vec{r}, t) = e \int d^3v f(\vec{r}, \vec{v}, t), \tag{4.a}$$

$$\vec{j}(\vec{r}, t) = e \int d^3v \vec{v} f(\vec{r}, \vec{v}, t). \tag{4.b}$$

Vérifier (cf. complément IV.E, § II.1) que l'équation de conservation locale de la charge, qui relie ρ et $\vec{j}$, découle de l'équation de Boltzmann (2).

3) On suppose que le plasma est enfermé dans un récipient. On lui applique un *champ électrique* indépendant du temps et on attend que s'établisse un nouvel *état d'équilibre.* Soit $\vec{\mathcal{E}}(\vec{r})$ le champ électrique total au point $\vec{r}$, et $\varphi(\vec{r})$ le potentiel dont il dérive. Montrer que, dans ce cas,

$$f_e(\vec{r}, \vec{v}) = A \exp -\frac{1}{kT}\left[\frac{1}{2} m\vec{v}^2 + e\varphi(\vec{r})\right], \tag{5}$$

où A est indépendant de $\vec{r}$ et $\vec{v}$, est solution d'équilibre de l'équation de Boltzmann.

4) Les charges ne sont plus maintenant confinées, mais le plasma est inséré dans un *circuit électrique fermé.* Le champ électrique $\vec{\mathcal{E}}$ est *uniforme* et dirigé suivant Ox. On s'intéresse au *régime permanent* qui s'établit dans ces conditions, c'est-à-dire qu'on cherche une fonction de répartition $f(\vec{v})$ qui soit, comme $\vec{\mathcal{E}}$, indépendante de $\vec{r}$ et t. On fait l'hypothèse que les collisions sont pratiquement inchangées par la présence du champ électrique, de sorte que c'est f_e^0 (formule (3)) qui figure au second membre de l'équation de Boltzmann (2).

a) Montrer que l'équation (2) s'écrit alors

$$\frac{e\mathcal{E}}{m}\frac{\partial f}{\partial v_x} = -\frac{1}{\tau}(f - f_e^0). \tag{6}$$

b) On suppose $\mathcal{E}$ assez faible pour que ses effets puissent être traités au premier ordre. On pose

$$f = f_e^0 + f_1, \tag{7}$$

où f_1 est la correction du premier ordre en $\mathcal{E}$. Déduire de l'équation (6) l'expression de f_1.

c) Montrer que f_1 reste effectivement petite devant f_e^0 si

$$\frac{e\mathcal{E}\tau v_x}{kT} \ll 1. \tag{8}$$

Interpréter physiquement cette condition.

5) Calculer, dans ce domaine des champs faibles, la conductivité électrique σ du plasma, définie par

$$\vec{j} \widehat{=} \sigma\vec{\mathcal{E}}. \tag{9}$$

ANALYSE DE L'IRRÉVERSIBILITÉ DANS DES MODÈLES SIMPLES

IV.14. Modèle d'Ehrenfest. Pour étudier l'évolution irréversible d'un système macroscopique vers l'équilibre (et les « paradoxes » qu'elle soulève), Paul et Tatiana Ehrenfest imaginèrent le modèle suivant. $2N$ boules numérotées de 1 à $2N$ sont distribuées dans deux boîtes A et B. Dans l'état initial, la boîte A contient $N+n_0$ boules et la boîte B $N-n_0$ ($-N \leqslant n_0 \leqslant N$). On tire alors au hasard un nombre compris entre 1 et $2N$ et on change de boîte la boule qui porte ce numéro. Le processus est répété régulièrement, et on étudie la probabilité $P(n_0, n; s)$ pour que, après s tirages, $N+n$ boules se trouvent dans A(*).

1) Exprimer $P(n_0, n; s)$ en fonction des probabilités antérieures $P(n_0, n'; s-1)$ (équation maîtresse).

2) Montrer que, si les boules sont réparties au hasard entre les deux récipients, la probabilité pour qu'il y ait $N+n$ boules dans A est

$$P_0(n) = \frac{(2N)!}{(N+n)!(N-n)!}\left(\frac{1}{2}\right)^{2N}.$$

Vérifier que $P_0(n)$ est une solution stationnaire (c'est-à-dire indépendante de s) de l'équation maîtresse et correspond donc à la distribution d'équilibre quelle que soit la distribution initiale.

3) A l'aide de l'équation maîtresse, exprimer la valeur moyenne $\bar{n}(s)$ de n après s tirages en fonction de $\bar{n}(s-1)$. En déduire que

$$\bar{n}(s) = n_0\left(1-\frac{1}{N}\right)^s.$$

4) Soit τ l'intervalle de temps entre deux tirages. Montrer que, si $N \to \infty$ et $\tau \to 0$, de façon que $N\tau$ tende vers une constante α non nulle, $\bar{n}(t = s\tau)$ tend exponentiellement vers sa valeur d'équilibre.

5) Soit $P'(n_0, n; s)$ la probabilité pour que, au temps $t = s\tau$, la boîte A contienne *pour la première fois* $N+n$ boules. En particulier, $P'(n_0, n_0; s)$ représente la probabilité pour que la boîte A revienne à son état initial pour la première fois à l'instant $t = s\tau$; on dit encore que $P'(n_0, n_0; s)$ est la probabilité pour que le « temps de récurrence » (ou « temps de Poincaré ») soit $s\tau$. On peut montrer(**) que

$$\sum_{s=1}^{\infty} P'(n_0, n_0; s) = 1 ;$$

autrement dit, quel que soit l'état de départ, la probabilité pour que le système revienne dans cet état à un moment ou à un autre est égale à 1 (théorème de Poincaré). On peut aussi montrer(**) que le temps moyen de récurrence vaut

$$T_R = \tau \frac{(N+n_0)!(N-n_0)!}{(2N)!}\, 2^{2N}.$$

Évaluer T_R pour $N = n_0 = 10\,000$, puis pour $N = 10\,000$ et $n_0 = 0$ (on pourra prendre $\tau \sim 1$ s). On montre ainsi que, si $N+n_0$ et $N-n_0$ diffèrent beaucoup, T_R est

(*) On peut considérer que les boules sont les molécules d'un gaz contenues dans deux récipients A et B. A l'instant initial, les pressions sont différentes dans les deux récipients. On ménage alors un trou dans la paroi qui les sépare. Le modèle ci-dessus permet une description simplifiée de l'évolution du gaz.

(**) On trouvera les démonstrations dans le très bel article de M. Kac (*American Mathematical Monthly*, vol. *54*, n° 7), repris dans *Selected Papers on Noise and Stochastic Processes* (N. Wax, ed), Dover, New York (1954).

gigantesque; le processus d'évolution apparaîtra alors pratiquement comme irréversible. Si au contraire $N+n_0$ et $N-n_0$ sont presque égaux, T_R reste faible et cela n'a pas de sens de parler d'irréversibilité; dans ce dernier cas il s'agit simplement, pour les grands systèmes, des fluctuations autour de la valeur d'équilibre.

IV.15. L'anneau de Kac. Pour analyser le lien entre des équations microscopiques invariantes par renversement du temps et l'évolution irréversible des systèmes macroscopiques, M. Kac a proposé (1956) un modèle simplifié de système mécanique possédant ces deux types de propriétés.

N particules ($N \gg 1$) sont localisées sur un cercle en des points A_i ($i = 1, 2, \ldots, N$) régulièrement espacés (fig. 1). Chacune de ces particules a 2 états possibles, que nous noterons (+) et (−). A intervalles réguliers dans le temps, c'est-à-dire aux instants $t_n = n\tau_0$ ($n = 1, 2, \ldots$), les N particules avancent simultanément d'un cran dans le sens positif, celle qui se trouvait au point A_i passant en A_{i+1} (avec $A_{N+1} \equiv A_1$). Certains des arcs A_iA_{i+1}, en nombre M au total ($M < N$), possèdent la propriété d'inverser l'état des particules qui les franchissent; ces arcs, que nous qualifierons d'« actifs », sont en traits gras sur la figure. Les $N-M$ autres arcs ne changent pas l'état des particules.

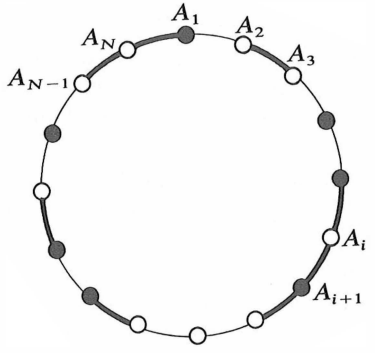


FIGURE 1

L'anneau de Kac : les arcs « actifs » sont en traits gras, et les particules, localisées aux points A_i ($i = 1, 2, \ldots, N$), ont deux états possibles, symbolisés par des cercles blancs ou noirs.

On imagine un *ensemble statistique* de tels systèmes. Tous les systèmes de l'ensemble possèdent le même nombre M d'arcs « actifs », mais la répartition de ces derniers sur le cercle varie de façon totalement aléatoire dans l'ensemble statistique. Il s'agit d'étudier l'évolution dans le temps de cet ensemble statistique.

1) Montrer que la loi d'évolution « microscopique » qui est posée pour chacun des systèmes est invariante par renversement du temps. Montrer en outre que chaque système possède un « cycle de Poincaré », c'est-à-dire qu'il existe un temps T au bout duquel le système revient à son état initial. Montrer que T est le même pour tous les systèmes et donner sa valeur.

2) Soit, pour un système particulier, $N_+(n)$ le nombre de particules qui se trouvent dans l'état (+) à l'instant t_n, $N_-(n)$ le nombre de celles qui se trouvent dans l'état (−) à cet instant. Définissons également $Q_+(n)$ (et $Q_-(n)$) comme le nombre de particules se trouvant dans l'état (+) (ou dans l'état (−)) et situées, à l'instant t_n, à l'origine d'un arc « actif ».

a) Écrire les « lois de conservation » que vérifient N_+ et N_- d'une part, Q_+ et Q_- d'autre part.

b) Exprimer $N_+(n+1)$ et $N_-(n+1)$ à partir de $N_+(n)$, $N_-(n)$, $Q_+(n)$ et $Q_-(n)$.

c) Exprimer de même $N_+(n-1)$ et $N_-(n-1)$ à partir de $N_+(n)$, $N_-(n)$, $R_+(n)$ et $R_-(n)$, où $R_+(n)$ et $R_-(n)$ sont les nombres de particules dans l'état (+) et l'état (−) et situées à l'instant t_n à l'extrémité d'un arc « actif ».

3) On pose comme hypothèse « physique » qu'il n'existe pas de corrélation, dans l'ensemble statistique, entre l'état d'une particule et le fait qu'elle soit située à l'origine d'un arc « actif »; plus précisément, on suppose que les valeurs moyennes d'ensemble des nombres N_+, N_-, Q_+ et Q_- vérifient les relations

$$\frac{\overline{Q}_+(n)}{\overline{N}_+(n)} = \frac{\overline{Q}_-(n)}{\overline{N}_-(n)} = \frac{M}{N} = \alpha \,. \tag{1}$$

a) Montrer que cette hypothèse permet de résoudre, en valeur moyenne, les équations d'évolution de la question 2.*b*, qui donnent ainsi

$$\overline{N}_+(n) - \overline{N}_-(n) = (1-2\alpha)^n \left[\overline{N}_+(0) - \overline{N}_-(0)\right] . \tag{2}$$

b) Analyser la solution (2) : montrer que

$$\left|\overline{N}_+(n) - \overline{N}_-(n)\right| \underset{n\to\infty}{\longrightarrow} 0 \,, \tag{3}$$

quelle que soit la distribution de départ dans l'ensemble; ce comportement était-il attendu à partir des propriétés « microscopiques » de la question 1? La contradiction entre le comportement (3) et l'existence d'un cycle de Poincaré constitue le « paradoxe de Zermélo ».

c) Comment s'écrit la solution des équations d'évolution moyennes si l'on remplace l'hypothèse (1) par l'hypothèse analogue concernant R_+ et R_- au lieu de Q_+ et Q_-? Comment se comporte cette nouvelle solution lorsque n devient grand?

4) Pour préciser l'analyse, on introduit les notations suivantes : sur un anneau particulier, l'état de la particule située au point A_i à l'instant t_n est caractérisé par $\eta_i(n)$, égal à +1 ou −1 suivant que cet état est (+) ou (−); on pose en outre $\xi_i = -1$ si l'arc A_iA_{i+1} est « actif », et $\xi_i = +1$ dans le cas contraire.

a) Montrer que les « équations du mouvement microscopique » prennent alors la forme simple

$$\eta_{i+1}(n+1) = \xi_i \eta_i(n) \,, \tag{4}$$

et que leur solution s'écrit

$$\eta_i(n) = \xi_{i-1} \cdot \xi_{i-2} \cdot \ldots \cdot \xi_{i-n} \eta_{i-n}(0) \,, \tag{5}$$

où les indices sont définis modulo N. Vérifier que cette solution possède bien la période T trouvée à la question 1.

b) Calculer la différence $N_+ - N_-$ à l'instant t_n en fonction des $\eta_i(n)$, puis utiliser la solution (5) pour l'exprimer à partir des $\eta_j(0)$ caractérisant l'état initial pour l'anneau particulier considéré ici.

5) On veut maintenant calculer la moyenne d'ensemble de $N_+ - N_-$ en fonction du temps t_n.

On admet que, dans l'ensemble statistique, les M arcs « actifs » sont répartis de façon équiprobable sur les N arcs que comporte au total un anneau, et que la répartition des $\eta_j(0)$ est ensuite choisie de façon *indépendante* de la précédente.

a) Montrer que la valeur moyenne de $N_+ - N_-$ prend alors la forme

$$\overline{N}_+(n) - \overline{N}_-(n) = \overline{\xi_1 \cdot \xi_2 \cdot \ldots \cdot \xi_n} \times \left[\overline{N}_+(0) - \overline{N}_-(0)\right] . \tag{6}$$

b) On suppose $n < N$. Quelle est, dans l'ensemble statistique, la probabilité pour qu'un ξ_i particulier, pour i quelconque mais fixé, soit égal à -1? En déduire la probabilité $P(n, r)$ pour que, parmi les n premiers ξ_i ($i = 1, 2, \ldots, n$), il s'en trouve r qui soient égaux à -1.

c) Montrer qu'on trouve ainsi, pour $n < N$,

$$\overline{\xi_1 \cdot \xi_2 \cdot \ldots \cdot \xi_n} = \sum_{r=0}^{n} (-1)^r P(n, r) = (1 - 2\alpha)^n , \tag{7}$$

ce qui établit le résultat (2) sans faire appel à une hypothèse supplémentaire.

d) Si n dépasse N, certains ξ_i apparaissent 2 fois dans le produit de n facteurs dont on cherche la valeur moyenne. Montrer que, si $N \leqslant n \leqslant 2N$, il faut remplacer n par $2N - n$ dans le résultat (7), ce qui donne une solution du type de celle trouvée en 3.*c*. Que se passe-t-il si l'on continue à augmenter n à N fixé?

6) L'analyse « physique » des résultats précédents peut être amorcée comme suit.

a) Établir le lien entre la propriété trouvée en 5.*d* et l'existence d'un « cycle de Poincaré » (question 1). Est-il important de ce point de vue que le cycle ait exactement la même durée pour tous les systèmes de l'ensemble? Cet aspect du modèle est-il « réaliste »?

b) Qu'est devenue, dans la question 5, l'invariance par renversement du temps?

c) Est-il « raisonnable » de supposer que n reste inférieur à N (pour s'en tenir à la solution (2))? Si l'on prend N de l'ordre du nombre d'Avogadro et l'intervalle de temps caractéristique $\tau_0 = t_n/n$ de l'ordre d'une fraction de seconde, comment se situe par rapport à N la valeur de n correspondant à l'âge de l'Univers ($\sim 10^{10}$ ans)?

d) Rapprocher l'hypothèse (1), qui fournit très simplement la solution « correcte », de celle du « chaos moléculaire » proposée par Boltzmann (complément IV.E, § I.2.a).

e) Les valeurs moyennes sont prises sur un ensemble statistique. Peut-on imaginer, dans le cadre du modèle de Kac, une façon d'introduire le principe ergodique?

V

I. SITUATION ET DISTRIBUTION GRAND-CANONIQUES

A. Situation grand-canonique : définitions
   1. Réservoir de particules
   2. Température et potentiel chimique grand-canoniques
   3. Comparaison avec la situation microcanonique

B. Distribution grand-canonique des états microscopiques
   1. Probabilité d'un état microscopique
   2. L'ensemble grand-canonique

C. Distribution statistique des variables internes
   1. Distribution statistique du nombre de particules
      a. Formule générale
      b. Cas d'un système macroscopique
   2. Distribution d'une variable interne quelconque

II. GRANDE FONCTION DE PARTITION ET GRAND POTENTIEL

A. Définition
   1. La grande fonction de partition
      a. Définition générale
      b. Relation avec la fonction de partition
   2. Le grand potentiel

B. Rôle fondamental : calcul des propriétés du système
   1. Caractéristiques de la distribution statistique du nombre de particules et de l'énergie
      a. Nombre moyen de particules
      b. Énergie moyenne
      c. Écarts quadratiques moyens et moments d'ordre supérieur
   2. Distribution d'une variable interne
   3. Calcul de l'entropie grand-canonique
   4. Pression grand-canonique

C. Système constitué de sous-systèmes indépendants
   1. Factorisation de la grande fonction de partition et additivité du grand potentiel
   2. Application au cas particulier d'un fluide simple

D. Évolution spontanée du système après relâchement d'une contrainte

III. LE FORMALISME GRAND-CANONIQUE A LA LIMITE THERMODYNAMIQUE

A. Limite thermodynamique de la description grand-canonique
   1. Description grand-canonique à l'approximation gaussienne
   2. Passage à la limite thermodynamique

B. Équivalence des trois descriptions « canoniques » dans la limite thermodynamique

IV. DESCRIPTION GRAND-CANONIQUE D'UN SYSTÈME DE PARTICULES INDÉPENDANTES IDENTIQUES ET INDISCERNABLES

A. Formules générales
   1. Factorisation de la grande fonction de partition
   2. Propriétés du système

B. Les statistiques quantiques
   1. Statistique de Fermi-Dirac
   2. Statistique de Bose-Einstein
      a. Condition sur le potentiel chimique
      b. Distribution de Bose

C. La statistique classique de Maxwell-Boltzmann
   1. Limite classique des statistiques quantiques
   2. Description grand-canonique des gaz parfaits classiques monoatomiques

CHAPITRE V

Système en équilibre avec un réservoir de particules. Description grand-canonique

Au chapitre III, nous avons étudié des systèmes dont l'énergie n'est pas fixée, en nous plaçant dans la situation la plus simple : le système considéré est en contact thermique avec un réservoir d'énergie beaucoup plus gros que lui, le thermostat, dont l'influence est caractérisée par un paramètre unique, la température T. C'est une démarche analogue que nous entreprenons dans le présent chapitre pour des systèmes dont le nombre de particules et l'énergie peuvent varier : nous supposons que le réservoir d'énergie est aussi un *réservoir de particules,* c'est-à-dire qu'il peut fournir ou prendre au système, outre de l'énergie, un nombre *a priori* quelconque de particules. A l'équilibre, nous dirons que le système étudié se trouve dans la *situation grand-canonique.* L'influence du réservoir est alors caractérisée par sa température T et son potentiel chimique μ, qu'il impose, de l'extérieur, au système considéré.

Il s'agit donc ici de généraliser les notions et résultats du chapitre III. Nous allons pour cela reprendre les principales rubriques de ce dernier, mais nous n'entrerons dans les détails que pour souligner les différences et les points nouveaux.

I. SITUATION ET DISTRIBUTION GRAND-CANONIQUES

A. Situation grand-canonique : définitions

Considérons un système $\mathcal{S}$ constitué de particules que nous supposons pour simplifier toutes identiques[1]; il peut échanger librement ces particules avec un autre système $\mathcal{R}$ beaucoup plus gros que lui, et constitué du même type de particules.

---

1. Les raisonnements qui suivent se généralisent sans peine au cas où le système $\mathcal{S}$ comprend des *particules de différentes sortes.* On le met alors en contact avec autant de réservoirs $\mathcal{R}_i$ ($i = 1, 2, \ldots, r$) qu'il a de sortes de particules. Ces réservoirs doivent tous être à la *même température,* pour que l'équilibre thermique soit possible; mais ils imposent *un potentiel chimique* μ_i *pour chaque type de particules.* Les définitions (I.5) se généralisent donc sous la forme

$$T \widehat{=} T^*_{\mathcal{R}_1} = T^*_{\mathcal{R}_2} = \ldots = T^*_{\mathcal{R}_r},$$

$$\mu_i \widehat{=} \mu^*_{\mathcal{R}_i}; \quad i = 1, 2, \ldots, r.$$

Insistons d'emblée sur le fait qu'il s'agit bien sûr de *particules indiscernables,* de sorte que l'état du système[2] $\mathcal{S}$ reste inchangé si l'une de ses particules est simplement remplacée par une particule de $\mathcal{R}$ (à condition que ce remplacement ne s'accompagne d'aucune autre modification). Notons également que ces *échanges de particules* s'accompagnent forcément d'*échanges d'énergie* : les particules qui franchissent la frontière entre $\mathcal{S}$ et $\mathcal{R}$ transportent nécessairement de l'énergie cinétique, et de plus l'énergie potentielle d'une assemblée de particules varie avec leur nombre. Nous supposons néanmoins que le couplage ainsi introduit entre $\mathcal{S}$ et $\mathcal{R}$ reste faible (II, § III.A.1.b. et § III.B.3).

1. RÉSERVOIR DE PARTICULES

Le système global formé par la réunion de $\mathcal{S}$ et $\mathcal{R}$ est isolé et à l'équilibre[3]. L'hypothèse de couplage faible implique que son énergie totale E_{tot}, fixée, est la somme de l'énergie E du système $\mathcal{S}$ que nous voulons étudier et de celle $E_{\mathcal{R}}$ du système $\mathcal{R}$:

$$E + E_{\mathcal{R}} = E_{\text{tot}} \,. \tag{I.1.a}$$

De même, son nombre de particules N_{tot} est la somme[4] du nombre de particules N de $\mathcal{S}$ et de celui $N_{\mathcal{R}}$ de $\mathcal{R}$:

$$N + N_{\mathcal{R}} = N_{\text{tot}} \,. \tag{I.1.b}$$

La démarche est alors tout à fait analogue à celle du chapitre III (§ I.A.1), où nous avons défini la notion de réservoir d'énergie (thermostat).

Le système $\mathcal{R}$ étant beaucoup plus gros que $\mathcal{S}$, l'énergie E et le nombre de particules N sont toujours beaucoup plus petits que $E_{\mathcal{R}}$ et $N_{\mathcal{R}}$, respectivement :

$$E \ll E_{\mathcal{R}} \,, \tag{I.2.a}$$

$$N \ll N_{\mathcal{R}} \,. \tag{I.2.b}$$

Les échanges entre $\mathcal{S}$ et $\mathcal{R}$ peuvent modifier de façon appréciable l'état macroscopique de $\mathcal{S}$ sans que celui de $\mathcal{R}$ en soit pratiquement affecté. De façon plus précise, soit $S^*_{\mathcal{R}}$ l'entropie microcanonique du système $\mathcal{R}$; elle dépend de l'énergie E et du nombre de particules N du système $\mathcal{S}$:

$$S^*_{\mathcal{R}} = S^*_{\mathcal{R}}(E_{\mathcal{R}} = E_{\text{tot}} - E,\ N_{\mathcal{R}} = N_{\text{tot}} - N,\ \ldots) \,. \tag{I.3}$$

---

2. Maintenant qu'il n'est plus constitué par un ensemble de particules délimité une fois pour toutes, la définition du système $\mathcal{S}$ et sa séparation d'avec $\mathcal{R}$ peuvent dans certains cas n'être pas évidentes. Il n'en est que plus important, nous l'avons déjà souligné, de *bien délimiter le système* $\mathcal{S}$ que l'on considère. Assez souvent, on pourra pour cela utiliser le volume, même si ses parois sont perméables aux particules : le système $\mathcal{S}$ sera constitué des particules situées à l'intérieur d'une enceinte fixe (réelle ou fictive). Mais le formalisme grand-canonique s'applique également à des situations moins simples : molécules d'un gaz adsorbées sur les parois du récipient qui le contient, électrons d'un solide « piégés » par des impuretés (voir complément V.A et VI.A), ... Dans de tels cas, il faudra par exemple définir le système $\mathcal{S}$ comme « l'ensemble des molécules adsorbées » (le réservoir $\mathcal{R}$ étant constitué des autres molécules, restées à l'état gazeux), ou « l'ensemble des électrons occupant les états offerts par les impuretés » (les autres électrons du solide servant de réservoir $\mathcal{R}$), ...

3. Comme au chapitre III, le système $\mathcal{S}$ que l'on étudie peut être *microscopique ou macroscopique* : la seule exigence concerne le système global $\mathcal{S} \cup \mathcal{R}$, qui doit être macroscopique pour qu'on puisse lui appliquer le postulat fondamental du chapitre II.

4. Il est question ici de particules matérielles, dont le nombre total se conserve (dans le domaine non relativiste où nous nous situons). Nous verrons au chapitre VI (§ III.A.3) comment traiter le cas de particules, comme les photons, qui peuvent être créées ou annihilées lors de leur contact avec un autre système.

Par *définition*, $\mathcal{R}$ constitue pour $\mathcal{S}$ un *réservoir de particules et d'énergie si sa température* $T^*_{\mathcal{R}}$ *et son potentiel chimique* $\mu^*_{\mathcal{R}}$ *microcanoniques sont pratiquement indépendants de l'énergie* E *et du nombre de particules* N *du système* $\mathcal{S}$:

$$T^*_{\mathcal{R}}(E_{\mathcal{R}} = E_{\text{tot}} - E,\ N_{\mathcal{R}} = N_{\text{tot}} - N,\ \ldots) \simeq T^*_{\mathcal{R}}\ (E_{\mathcal{R}} = E_{\text{tot}},\ N_{\mathcal{R}} = N_{\text{tot}},\ \ldots)\ , \quad \text{(I.4.a)}$$

$$\mu^*_{\mathcal{R}}(E_{\mathcal{R}} = E_{\text{tot}} - E,\ N_{\mathcal{R}} = N_{\text{tot}} - N,\ \ldots) \simeq \mu^*_{\mathcal{R}}\ (E_{\mathcal{R}} = E_{\text{tot}},\ N_{\mathcal{R}} = N_{\text{tot}},\ \ldots)\ . \quad \text{(I.4.b)}$$

$\mathcal{R}$ est alors à la fois un thermostat et un réservoir de particules pour le système $\mathcal{S}$. Comme au chapitre III, cette définition est relative : c'est $\mathcal{S}$ qui sert de référence lorsqu'on dit que $\mathcal{R}$ doit être « gros ». Elle est de plus approchée : c'est la précision de l'expérience qui décide si tel système $\mathcal{R}$ peut être pris comme réservoir.

2. TEMPÉRATURE ET POTENTIEL CHIMIQUE GRAND-CANONIQUES

Lorsque les hypothèses (I.4) sont satisfaites et que le système global a atteint son état d'équilibre macroscopique, on dit que le système $\mathcal{S}$ se trouve dans la *situation grand-canonique*. On définit sa *température grand-canonique*, que nous noterons encore T, et son *potentiel chimique grand-canonique* μ par

$$T \mathrel{\widehat{=}} T^*_{\mathcal{R}}, \quad \text{(I.5.a)}$$

$$\mu \mathrel{\widehat{=}} \mu^*_{\mathcal{R}}. \quad \text{(I.5.b)}$$

Ces deux paramètres, qui sont *imposés au système, de l'extérieur, par le réservoir* $\mathcal{R}$, suffisent à caractériser le rôle de ce dernier.

3. COMPARAISON AVEC LA SITUATION MICROCANONIQUE

Lorsque, comme au chapitre II, on étudie un système $\mathcal{S}$ *isolé* (situation microcanonique), son état macroscopique d'équilibre est défini par un ensemble de paramètres extérieurs comprenant l'énergie E, le nombre de particules N, le volume V, ... :

paramètres extérieurs microcanoniques : E, N, V, ... (I.6)

Dans la situation grand-canonique, *l'énergie* E *et le nombre de particules* N *du système sont devenus des variables internes.* A l'équilibre, l'état macroscopique du système est maintenant déterminé par la donnée de la température T et du potentiel chimique μ grand-canoniques, fixés de l'extérieur par le réservoir, et qui remplacent l'énergie et le nombre de particules :

paramètres extérieurs grand-canoniques : T, μ, V, ... (I.7)

B. Distribution grand-canonique des états microscopiques

1. PROBABILITÉ D'UN ÉTAT MICROSCOPIQUE

Soit un système $\mathcal{S}$ en équilibre avec un réservoir de particules et d'énergie, à la température T et au potentiel chimique μ. Nous notons comme toujours (ℓ) un état microscopique de $\mathcal{S}$, dont l'énergie est E_ℓ et le nombre de particules N_ℓ; comme le nombre de particules de $\mathcal{S}$ n'est plus fixé, il nous faut considérer ici tous les états (ℓ), pour toutes les valeurs *a priori* possibles de N_ℓ comme de E_ℓ. Les états microscopiques du réservoir $\mathcal{R}$ seront repérés par (L), l'énergie correspondante étant E_L et le nombre de particules N_L. Le raisonnement donnant la probabilité P^G_ℓ pour que, dans ces conditions, le système $\mathcal{S}$ se trouve dans l'état (ℓ) généralise simplement celui du chapitre III (§ I.B.1).

Le *système global* $\mathcal{S} \cup \mathcal{R}$ étant *isolé*, nous lui appliquons le postulat fondamental énoncé au chapitre II. Ses états microscopiques peuvent être caractérisés par (ℓ, L), puisque par hypothèse le couplage entre $\mathcal{S}$ et $\mathcal{R}$ est faible; l'énergie $E_{\ell,L}$ et le nombre de particules $N_{\ell,L}$ d'un tel état sont simplement

$$E_{\ell,L} = E_\ell + E_L \,, \tag{I.8.a}$$

$$N_{\ell,L} = N_\ell + N_L \,. \tag{I.8.b}$$

Pour calculer la probabilité P_ℓ^G qu'a le système $\mathcal{S}$ de se trouver dans l'état (ℓ), l'état du réservoir pouvant être quelconque, il suffit de *compter les états accessibles du système global* $\mathcal{S} \cup \mathcal{R}$ *qui réalisent cette situation,* c'est-à-dire les états (ℓ, L) qui vérifient les conditions

$$\begin{cases} \mathcal{S} \text{ dans l'état } (\ell) \text{ choisi} \,, & \text{(I.9.a)} \\ E_{\text{tot}} \leqslant E_{\ell,L} \leqslant E_{\text{tot}} + \delta\text{E} \,, & \text{(I.9.b)} \\ N_{\ell,L} = N_{\text{tot}} \,, & \text{(I.9.c)} \end{cases}$$

où E_{tot} et N_{tot} sont fixés. Compte tenu de (I.8), ces conditions peuvent être écrites

$$\begin{cases} \mathcal{S} \text{ dans l'état } (\ell) \text{ choisi} \,, & \text{(I.10.a)} \\ E_{\text{tot}} - E_\ell \leqslant E_L \leqslant E_{\text{tot}} - E_\ell + \delta E \,, & \text{(I.10.b)} \\ N_L = N_{\text{tot}} - N_\ell \,, & \text{(I.10.c)} \end{cases}$$

puisque le choix de (ℓ) fixe aussitôt E_ℓ et N_ℓ. La probabilité cherchée est donc telle que

$$P_\ell^G \propto \Omega_{\mathcal{R}}(E_{\mathcal{R}} = E_{\text{tot}} - E_\ell,\ N_{\mathcal{R}} = N_{\text{tot}} - N_\ell,\ \ldots) \,, \tag{I.11}$$

où $\Omega_{\mathcal{R}}(E_{\mathcal{R}},\ N_{\mathcal{R}},\ \ldots)$ est le nombre d'états microscopiques du réservoir $\mathcal{R}$ ayant l'énergie $E_{\mathcal{R}}$ (à δE près) et le nombre de particules $N_{\mathcal{R}}$. Si l'on introduit l'entropie microcanonique $S_{\mathcal{R}}^*$ du réservoir, la formule précédente devient

$$P_\ell^G = C \exp \frac{1}{k} S_{\mathcal{R}}^*(E_{\mathcal{R}} = E_{\text{tot}} - E_\ell,\ N_{\mathcal{R}} = N_{\text{tot}} - N_\ell,\ \ldots) \,, \tag{I.12}$$

la constante C pouvant être déterminée ensuite par normalisation.

La probabilité grand-canonique P_ℓ^G d'un état (ℓ) du système $\mathcal{S}$ dépend donc de l'énergie E_ℓ et du nombre de particules N_ℓ de cet état. Pour préciser cette dépendance, il suffit de développer le second membre de (I.12) autour des valeurs $E_{\mathcal{R}} = E_{\text{tot}}$ et $N_{\mathcal{R}} = N_{\text{tot}}$, en se rappelant que la définition du réservoir (§ A.1) implique que ce développement peut être tronqué au premier ordre en E_ℓ et N_ℓ (cf. III, note 6) :

$$\begin{aligned} S_{\mathcal{R}}^*(E_{\text{tot}} - E_\ell,\ N_{\text{tot}} - N_\ell,\ \ldots) = {} & S_{\mathcal{R}}^*(E_{\text{tot}},\ N_{\text{tot}},\ \ldots) \\ & - E_\ell \frac{\partial S_{\mathcal{R}}^*}{\partial E_{\mathcal{R}}}(E_{\text{tot}},\ N_{\text{tot}},\ \ldots) - N_\ell \frac{\partial S_{\mathcal{R}}^*}{\partial N_{\mathcal{R}}}(E_{\text{tot}},\ N_{\text{tot}},\ \ldots) \,. \end{aligned} \tag{I.13}$$

D'après les définitions (I.5) de la température et du potentiel chimique grand-canoniques, cette formule s'écrit simplement

$$S_{\mathcal{R}}^*(E_{\text{tot}} - E_\ell,\ N_{\text{tot}} - N_\ell,\ \ldots) = S_{\mathcal{R}}^*(E_{\text{tot}},\ N_{\text{tot}},\ \ldots) - \frac{E_\ell}{T} + N_\ell \frac{\mu}{T} \,. \tag{I.14}$$

On la reporte alors dans (I.12), ce qui donne finalement

$$\boxed{P_\ell^G = \frac{1}{\Xi} \mathrm{e}^{-(E_\ell - \mu N_\ell)/kT}} \,, \tag{I.15}$$

où la constante $1/\Xi$ est déterminée par normalisation :

$$\boxed{\Xi = \sum_{(\ell)} \mathrm{e}^{-(E_\ell - \mu N_\ell)/kT}} \tag{I.16}$$

L'expression (I.15) donne la *distribution grand-canonique* des états microscopiques (ℓ) du système $\mathcal{S}$ en fonction de la température T et du potentiel chimique μ imposés par le réservoir. Ξ est appelée *«grande fonction de partition»* du système $\mathcal{S}$ (voir paragraphe II ci-dessous).

2. L'ENSEMBLE GRAND-CANONIQUE

L'ensemble statistique (I, § IV.B.2) associé à la distribution grand-canonique (I.15) est constitué d'un très grand nombre $\mathcal{N}$ de systèmes, tous préparés de la même façon que le système $\mathcal{S}$ étudié, et maintenus par un réservoir de particules et d'énergie à la même température T et au même potentiel chimique μ. La proportion $\mathcal{N}_\ell/\mathcal{N}$ de ces systèmes qui se trouvent dans un état microscopique (ℓ) choisi à l'avance tend vers P_ℓ^G lorsque $\mathcal{N}$ tend vers l'infini.

L'ensemble microcanonique (II, § I.B.2) ne comprenait que des systèmes de même énergie E (à δE près) et même nombre de particules N. Ici par contre, les systèmes constituant l'ensemble grand-canonique ont *des énergies et des nombres de particules différents*[5], la conservation de l'énergie totale et du nombre total de particules étant assurée par le couplage au réservoir. Cependant, la formule (I.15) montre clairement que *la probabilité des états* (ℓ) *varie très rapidement avec leur énergie* E_ℓ *et leur nombre de particules* N_ℓ.

Attention : P_ℓ^G est la probabilité de l'*état* (ℓ); pour en déduire la distribution statistique de l'énergie ou du nombre de particules du système, il faudra tenir compte du fait qu'à une énergie ou un nombre de particules donné sont associés plusieurs états microscopiques (voir paragraphe ci-dessous).

C. Distribution statistique des variables internes

1. DISTRIBUTION STATISTIQUE DU NOMBRE DE PARTICULES

Dans la situation grand-canonique, le nombre N de particules du système $\mathcal{S}$ est devenu une variable interne (II, § I.A.1). Sa distribution statistique se déduit facilement de la distribution grand-canonique.

a. Formule générale

Si l'on choisit une valeur de N, il existe le plus souvent un grand nombre d'états microscopiques (ℓ) dont le nombre de particules N_ℓ est égal à N. La probabilité $P^G(N)$ de cette valeur de N s'obtient en sommant les probabilités grand-canoniques de tous ces états :

$$P^G(N) = \sum_{(\ell)\,:\,N_\ell = N} P_\ell^G \,, \tag{I.17}$$

soit, d'après (I.15),

$$P^G(N) = \frac{1}{\Xi}\, e^{\mu N/kT} \sum_{(\ell)\,:\,N_\ell = N} e^{-E_\ell/kT} \,. \tag{I.18}$$

5. Noter que $N_\ell = 0$ n'est pas exclu : la probabilité (I.15) correspondante n'est pas nulle (même si $E_\ell = 0$ quand $N_\ell = 0$). Ce fait, qui peut paraître surprenant à première vue, est lié à la définition du système $\mathcal{S}$ (cf. note 2).

On reconnaît, dans la somme restant à effectuer au second membre, la *fonction de partition* du système étudié, *pour un nombre de particules fixé et égal à N* :

$$Z(T, N, V, \ldots) \mathrel{\widehat{=}} \sum_{(\mathscr{E}) \,:\, N_\mathscr{E} = N} \mathrm{e}^{-E_\mathscr{E}/kT} . \tag{I.19}$$

La formule (I.18) s'écrit donc (la dépendance de Z par rapport aux autres paramètres extérieurs étant sous-entendue)

$$P^G(N) = \frac{1}{\Xi} \mathrm{e}^{\mu N/kT} Z(T, N) , \tag{I.20}$$

ou bien

$$P^G(N) = \frac{1}{\Xi} \exp\left\{ -\frac{1}{kT} \left[F(T, N) - \mu N \right] \right\} , \tag{I.21}$$

où $F(T, N, V, \ldots)$ est l'*énergie libre* canonique (III, § II.A.2).

Les formules (I.20) et (I.21) s'appliquent dans tous les cas, que le système considéré soit de taille microscopique ou macroscopique. Nous vérifierons plus loin (formule (II.5)) que la distribution $P^G(N)$ est convenablement normalisée.

b. Cas d'un système macroscopique

Si l'on a affaire à un système de taille macroscopique, les valeurs de N ayant une probabilité non négligeable sont très grandes devant 1, et la variable N peut être traitée comme *continue*. Dans ce cas, la probabilité $P^G(N)$ trouvée au paragraphe a peut tout aussi bien être considérée comme une *densité de probabilité*.

En outre, comme celle de toute variable interne, la distribution statistique du nombre de particules a pratiquement, à une excellente approximation, une *forme gaussienne*.

En effet, l'expression (I.21) de $P^G(N)$ est le produit de deux facteurs dépendant de N (Ξ est indépendant de N d'après sa définition (I.16)) :

$$\exp + \frac{\mu N}{kT} \tag{I.22.a}$$

est une fonction croissante de N si μ est positif, décroissante si μ est négatif (rappelons que μ est imposé au système de l'extérieur); quant à

$$\exp\left[-\frac{1}{kT} F(T, N) \right] , \tag{I.22.b}$$

son sens de variation dépend du signe du *potentiel chimique canonique*

$$\mu^c \mathrel{\widehat{=}} \frac{\partial F}{\partial N} , \tag{I.23}$$

signe qui est fixé par les propriétés du système lui-même[6]. Chacun des deux facteurs (I.22) varie très rapidement avec N (fonction exponentielle); ils varient en sens inverse si μ et μ^c sont de même signe[7].

---

6. Pour un système macroscopique tel que ceux que nous considérons ici, le potentiel chimique μ^c a même signe que le potentiel chimique microcanonique μ^* : ils doivent coïncider à la limite thermodynamique (III, § III.C). Le signe du potentiel chimique microcanonique μ^* a été discuté dans la note 45 du chapitre II.

7. Examinons ce qui se passe si l'on cherche à imposer au système $\mathcal{S}$ un potentiel chimique grand-canonique μ de signe contraire à son potentiel chimique canonique μ^c. Supposons pour fixer les idées que μ^c soit négatif et μ positif. Le réservoir $\mathcal{R}$ fournira alors des particules au système $\mathcal{S}$. Comme μ^c est une fonction croissante de N (III, § II.D.3.b), cet apport de particules tendra à rapprocher μ^c de μ (qui reste par hypothèse invariable). Si cet effet peut aller jusqu'à inverser le signe de μ^c et aboutir à l'égalité $\mu^c = \mu$ (cf. (I.25)), l'équilibre pourra s'établir. Si au contraire μ^c ne peut pas changer de signe, l'équilibre grand-canonique sera impossible : $\mathcal{R}$ se videra dans le système $\mathcal{S}$ et ne pourra donc jouer son rôle de réservoir.

La *valeur la plus probable* N_m du nombre de particules est alors donnée par

$$\frac{\partial}{\partial N}\left[F(T, N)-\mu N\right](N=N_m)=0\ , \tag{I.24}$$

c'est-à-dire

$$\mu^c(T, N_m, V, \ldots)=\mu\ . \tag{I.25}$$

N_m est donc telle que le potentiel chimique canonique du système, si son nombre de particules était fixé à N_m, serait précisément égal au potentiel chimique μ imposé par le réservoir; cette égalité n'est d'ailleurs qu'un cas particulier de la condition d'équilibre entre deux systèmes pouvant échanger des particules (III, § II.D.3.a).

Nous avons montré (III, formule (II.52)) que

$$\frac{\partial\mu^c}{\partial N}=\frac{\partial^2 F}{\partial N^2}>0\ . \tag{I.26}$$

Il s'ensuit (fig. 1) que, si l'équation (I.25) a une racine N_m (cf. note 7), cette racine est unique. De plus, elle correspond bien à un *maximum*[8] de $P^G(N)$:

$$\frac{\partial^2 P^G}{\partial N^2}(N_m)<0\ . \tag{I.27}$$

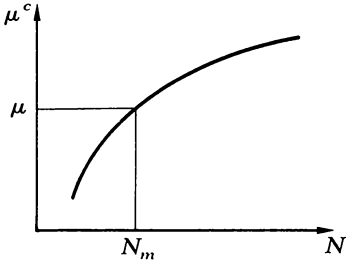


FIGURE 1
Résolution graphique de l'équation (I.25) donnant N_m : μ^c étant une fonction croissante de N, la racine N_m est unique si elle existe (c'est-à-dire si μ est choisi de même signe que μ^c; voir note 7).

Selon la méthode maintenant bien connue (II, § II.A.2 et III, § I.C.1.b), nous déduisons de ce qui précède que $P^G(N)$ est pratiquement de la forme

$$P^G(N)\simeq\frac{1}{\sqrt{2\pi(\Delta N)^2}}\exp\left[-\frac{(N-N_m)^2}{2(\Delta N)^2}\right], \tag{I.28}$$

avec[9]

$$(\Delta N)^2=kT/\frac{\partial^2 F}{\partial N^2}(N_m)\ . \tag{I.29}$$

8. Le signe de $\partial^2 P^G/\partial N^2$ se déduit simplement (cf. formule (I.21)) de celui de

$$\frac{\partial^2}{\partial N^2}\left[F(T, N)-\mu N\right]=\frac{\partial^2 F}{\partial N^2}>0$$

(μ étant bien sûr indépendant de N).

9. Noter que, comme toujours,

$$\Delta N\sim\sqrt{N_m}\ , \quad \text{car} \quad \frac{\partial^2 F}{\partial N^2}\sim\frac{N_m}{N_m^2}=\frac{1}{N_m}\ .$$

2. DISTRIBUTION D'UNE VARIABLE INTERNE QUELCONQUE

Soit y une variable interne, y_ℓ la valeur qu'elle prend dans l'état (ℓ) du système. Sa distribution statistique grand-canonique $P^G(y)$ se déduit de P^G_ℓ par la méthode habituelle :

$$P^G(y)=\frac{1}{\Xi}\sum_{(\ell)\,:\,y_\ell=y} e^{-(E_\ell-\mu N_\ell)/kT}\,. \tag{I.30}$$

Cette formule est *générale* et *exacte.*

Lorsque le système étudié est de *taille macroscopique* et qu'on peut traiter les variables E, N et y comme continues, la distribution de la variable interne y est caractérisée par une *densité de probabilité* $w^G(y)$. Supposons que nous connaissions le *nombre*

$$\rho(E, N, y)\,dE\,dN\,dy \tag{I.31}$$

d'états microscopiques (ℓ) dont l'énergie est E à dE près, le nombre de particules N à dN près, et dans lesquels la grandeur considérée vaut y à dy près. Pour un tel système macroscopique, la formule (I.30) peut être remplacée en pratique par

$$w^G(y)=\frac{1}{\Xi}\int_0^{+\infty} dN \int_{E_0}^{+\infty} dE\,\rho(E, N, y)\,e^{-(E-\mu N)/kT}\,. \tag{I.32}$$

Cette expression est directement normalisée. On a en effet

$$\int dy\,\rho(E, N, y)=\rho(E, N)\,, \tag{I.33}$$

où $\rho(E, N)$ est la *densité d'états* du système, et la grande fonction de partition Ξ, définie en (I.16), peut se calculer dans le cas présent par

$$\Xi=\int_0^{+\infty} dN \int_{E_0}^{+\infty} dE\,\rho(E, N)\,e^{-(E-\mu N)/kT}\,. \tag{I.34}$$

La distribution statistique de l'*énergie* s'obtient de manière analogue :

$$\begin{aligned} P^G(E)&=\frac{1}{\Xi}\sum_{(\ell)\,:\,E_\ell=E} e^{-(E_\ell-\mu N_\ell)/kT} \\ &=\frac{1}{\Xi}\,e^{-E/kT}\sum_{(\ell)\,:\,E_\ell=E} e^{\mu N_\ell/kT}\,. \end{aligned} \tag{I.35}$$

Cette formule rigoureuse peut, pour un système macroscopique, se mettre sous la forme

$$w^G(E)\,dE=\frac{1}{\Xi}\,e^{-E/kT}\,dE\int_0^{+\infty} dN\,\rho(E, N)\,e^{\mu N/kT}\,, \tag{I.36}$$

dont la normalisation est assurée par l'expression (I.34) de Ξ.

II. GRANDE FONCTION DE PARTITION ET GRAND POTENTIEL

A. Définition

Pour un système isolé à l'équilibre, la fonction fondamentale est le nombre $\Omega(E, N, V, \ldots)$ d'états accessibles, ou, ce qui revient au même, l'entropie microcanonique $S^*(E, N, V, \ldots)$; lorsque le système est en équilibre avec un thermostat, c'est la fonction de partition $Z(T, N, V, \ldots)$ ou l'énergie libre $F(T, N, V, \ldots)$. Dans la situation grand-canonique que nous étudions ici, ce rôle est joué par la grande fonction de partition $\Xi(T, \mu, V, \ldots)$ ou par le grand potentiel $J(T, \mu, V, \ldots)$ qui s'en déduit simplement.

1. LA GRANDE FONCTION DE PARTITION

a. Définition générale

La grande fonction de partition Ξ a été définie par (I.16), pour assurer la normalisation de la distribution grand-canonique (I.15). Cette « constante » (indépendante de l'état (ℓ) et de ses caractéristiques E_ℓ et N_ℓ) est en fait une *fonction de la température T et du potentiel chimique μ, ainsi que des autres paramètres extérieurs* tels que le volume V :

$$\Xi(T, \mu, V, \ldots) \widehat{=} \sum_{(\ell)} e^{-(E_\ell - \mu N_\ell)/kT} . \tag{II.1}$$

Ξ est une *grandeur sans dimension.*

Dans le cas d'un *système macroscopique,* la somme sur les états (ℓ) peut pratiquement toujours être remplacée par une intégrale sur l'énergie E et une somme ou une intégrale sur le nombre N de particules, si l'on connaît la *densité d'états* $\rho(E, N, V, \ldots)$ du système :

$$\Xi(T, \mu, V, \ldots) = \int_0^{+\infty} dN \int_{E_0}^{+\infty} dE\, \rho(E, N, V, \ldots)\, e^{-(E-\mu N)/kT} , \tag{II.2}$$

où E_0 est le niveau fondamental du système lorsque celui-ci comporte N particules.

Comme au chapitre III, nous remplacerons parfois la variable température T par la variable

$$\beta \widehat{=} \frac{1}{kT} . \tag{II.3}$$

b. Relation avec la fonction de partition

Il est possible de décomposer, de façon *exacte,* la somme sur l'ensemble des états microscopiques (ℓ) en une somme sur ceux qui ont un nombre de particules N fixé, suivie d'une somme sur toutes les valeurs de N possibles. La grande fonction de partition Ξ prend alors la forme

$$\Xi = \sum_{N=0}^{+\infty} \sum_{(\ell):\, N_\ell = N} e^{-\beta(E_\ell - \mu N_\ell)} = \sum_{N=0}^{+\infty} e^{\beta\mu N} \sum_{(\ell):\, N_\ell = N} e^{-\beta E_\ell} , \tag{II.4}$$

où l'on reconnaît la fonction de partition $Z(\beta, N, V, \ldots)$ associée au système lorsque son nombre de particules est fixé :

$$\Xi(\beta, \mu, V, \ldots) = \sum_{N=0}^{+\infty} e^{\beta\mu N} Z(\beta, N, V, \ldots) = 1 + \sum_{N=1}^{+\infty} e^{\beta\mu N} Z(\beta, N, V, \ldots) . \tag{II.5}$$

Ainsi, la grande fonction de partition apparaît comme une série de puissances de la *fugacité*

$$\varphi \widehat{=} e^{\beta\mu} , \tag{II.6}$$

le coefficient du terme d'ordre N étant la fonction de partition correspondant à un nombre de particules égal à N.

2. LE GRAND POTENTIEL

A partir de la grande fonction de partition Ξ, on définit le grand potentiel J par

$$\boxed{J(T, \mu, V, \ldots) \widehat{=} -kT \ln \Xi(T, \mu, V, \ldots)} \tag{II.7}$$

(cette définition est calquée sur celle de l'énergie libre F à partir de la fonction de partition Z).

Le grand potentiel a les *dimensions d'une énergie.*

B. Rôle fondamental : calcul des propriétés du système

Les principales caractéristiques d'un système en équilibre avec un réservoir de particules et d'énergie se déduisent de façon simple de sa grande fonction de partition ou de son grand potentiel.

1. CARACTÉRISTIQUES DE LA DISTRIBUTION STATISTIQUE DU NOMBRE DE PARTICULES ET DE L'ÉNERGIE

a. Nombre moyen de particules

La valeur moyenne du nombre N de particules du système est par définition

$$\overline{N}^G = \sum_{(\ell)} N_\ell P_\ell^G = \frac{1}{\Xi} \sum_{(\ell)} N_\ell \, e^{-\beta(E_\ell - \mu N_\ell)} . \tag{II.8}$$

Si l'on compare cette expression à celle de Ξ (formule (II.1)), on voit facilement que

$$\boxed{\overline{N}^G = \frac{1}{\beta} \frac{\partial}{\partial \mu} (\ln \Xi) = kT \frac{\partial}{\partial \mu} (\ln \Xi)} , \tag{II.9}$$

ce que l'on peut aussi écrire, compte tenu de la définition (II.7) de J,

$$\overline{N}^G = -\frac{\partial J}{\partial \mu} \tag{II.10}$$

(la dérivation par rapport à μ maintient constants les autres paramètres extérieurs, ici T ou β, V, ...).

b. Énergie moyenne

Un résultat analogue peut être obtenu pour la valeur moyenne de l'énergie. Il vient en effet à partir de (II.1)

$$\frac{\partial}{\partial \beta} (\ln \Xi) = -\frac{1}{\Xi} \sum_{(\ell)} (E_\ell - \mu N_\ell) \, e^{-\beta(E_\ell - \mu N_\ell)} , \tag{II.11}$$

d'où l'on tire aussitôt

$$\boxed{\overline{E}^G - \mu \overline{N}^G = -\frac{\partial}{\partial \beta} (\ln \Xi)} \tag{II.12}$$

Si l'on connaît déjà $\overline{N}^G$, cette formule donne $\overline{E}^G$. On peut d'ailleurs y reporter l'expression (II.9) de $\overline{N}^G$, ce qui donne finalement

$$\overline{E}^G = \frac{1}{\beta} \left[\mu \frac{\partial}{\partial \mu} - \beta \frac{\partial}{\partial \beta} \right] (\ln \Xi) . \tag{II.13}$$

Il est ensuite facile[10] de faire apparaître le grand potentiel J au lieu de $\ln \Xi$:

$$\overline{E}^G = J + \left(\beta \frac{\partial}{\partial \beta} - \mu \frac{\partial}{\partial \mu} \right) J . \tag{II.14}$$

---

10. Il suffit de remplacer $\ln \Xi$ par $-\beta J$ pour obtenir

$$\overline{E}^G = -\frac{1}{\beta} \left[\mu \frac{\partial}{\partial \mu} - \beta \frac{\partial}{\partial \beta} \right] (\beta J) = -\mu \frac{\partial J}{\partial \mu} + J + \beta \frac{\partial J}{\partial \beta} .$$

c. Écarts quadratiques moyens et moments d'ordre supérieur

La forme même de la formule (II.9) suggère d'appliquer une seconde fois l'opérateur différentiel $\frac{1}{\beta}\frac{\partial}{\partial\mu}$. On trouve alors effectivement[11]

$$(\Delta N^G)^2 = \frac{1}{\beta^2}\frac{\partial^2}{\partial\mu^2}(\ln \Xi) . \tag{II.15}$$

Un résultat analogue est valable pour l'énergie, bien que l'expression (II.13) de sa valeur moyenne soit plus compliquée :

$$(\Delta E^G)^2 = \left[\frac{\mu}{\beta}\frac{\partial}{\partial\mu} - \frac{\partial}{\partial\beta}\right]^2 (\ln \Xi) . \tag{II.16}$$

Le procédé se généralise sans peine[11] et permet d'exprimer les moments $\overline{(N-\bar{N})^n}^G$ et $\overline{(E-\bar{E})^n}^G$ d'ordre n (entier positif) quelconque à partir des dérivées n-ièmes (et inférieures) de $\ln \Xi$.

2. DISTRIBUTION D'UNE VARIABLE INTERNE

La formule (I.30), donnant la distribution statistique $P^G(y)$ d'une variable interne y quelconque, peut être interprétée en termes de grandes fonctions de partition, par des arguments tout à fait semblables à ceux du chapitre III (§ II.B.2).

En effet, l'expression

$$\Xi(T, \mu, V, \ldots; y) \widehat{=} \sum_{(\ell) \,:\, y_\ell = y} e^{-(E_\ell - \mu N_\ell)/kT} \tag{II.17}$$

---

11. Les formules (II.9) et (II.13) sont de la forme

$$\bar{A}^G = D(\ln \Xi) , \tag{1}$$

où A désigne N ou E, et D un opérateur différentiel linéaire :

$$D \equiv \frac{1}{\beta}\frac{\partial}{\partial\mu} \quad \text{pour} \quad A \equiv N , \qquad D \equiv \left[\frac{\mu}{\beta}\frac{\partial}{\partial\mu} - \frac{\partial}{\partial\beta}\right] \quad \text{pour} \quad A \equiv E . \tag{2}$$

On en déduit aussitôt

$$(\Delta A^G)^2 = D^2(\ln \Xi) , \tag{3}$$

c'est-à-dire les formules (II.15) et (II.16).

En effet, si nous posons

$$f_\ell \widehat{=} \exp[-\beta(E_\ell - \mu N_\ell)] , \tag{4}$$

Ξ s'écrit

$$\Xi = \sum_{(\ell)} f_\ell , \tag{5}$$

la valeur moyenne de A prend la forme

$$\bar{A}^G = \frac{1}{\Xi}\sum_{(\ell)} A_\ell f_\ell , \tag{6}$$

et la formule (1) découle de

$$D f_\ell = A_\ell f_\ell \tag{7}$$

(que l'on vérifie facilement en appliquant les opérateurs (2) à l'expression (4)). Calculons alors

$$D^2(\ln \Xi) = D[D(\ln \Xi)] . \tag{8}$$

.../...

est celle de *la grande fonction de partition du système dans la situation où la grandeur physique considérée serait fixée à la valeur y.* Ainsi, la formule (I.30) peut s'écrire comme le quotient de cette grande fonction de partition « partielle » par la grande fonction de partition « totale » du système (dans la situation où la variable interne y fluctue librement) :

$$P^G(T, \mu, V, \ldots; y) = \frac{\Xi(T, \mu, V, \ldots; y)}{\Xi(T, \mu, V, \ldots)} . \tag{II.18}$$

Si l'on définit *le grand potentiel du système dans le cas où la grandeur étudiée est fixée à la valeur y* :

$$J(T, \mu, V, \ldots; y) \widehat{=} -kT \ln \Xi(T, \mu, V, \ldots; y) , \tag{II.19}$$

on peut mettre (II.18) sous la forme

$$P^G(T, \mu, V, \ldots; y) = \exp\left\{-\frac{1}{kT}\left[J(T, \mu, V, \ldots; y) - J(T, \mu, V, \ldots)\right]\right\} , \tag{II.20}$$

où $J(T, \mu, V, \ldots)$ est le grand potentiel dans le cas où la grandeur est libre de fluctuer. Autrement dit, *à* $T, \mu, V, \ldots$ *constants,* la dépendance en y de $P^G(y)$ est donnée par

$$P^G(y) \propto \exp\left[-\frac{1}{kT} J(T, \mu, V, \ldots; y)\right] . \tag{II.21}$$

Par la méthode habituelle, on peut déduire de là que, pour un système macroscopique, la distribution de la variable interne y coïncide pratiquement avec une gaussienne.

3. CALCUL DE L'ENTROPIE GRAND-CANONIQUE

En reportant la distribution grand-canonique (I.15) dans la définition générale I (II.2) de l'entropie statistique, on trouve[12] l'expression suivante pour l'entropie grand-canonique S^G :

$$S^G = \frac{1}{T}\left(\overline{E}^G - \mu \overline{N}^G - J\right) , \tag{II.22}$$

D'après (1) et (6), nous obtenons

$$D^2(\ln \Xi) = D\left[\frac{1}{\Xi}\sum_{(\ell)} A_\ell f_\ell\right] = -\frac{1}{\Xi^2}(D\Xi).\sum_{(\ell)} A_\ell f_\ell + \frac{1}{\Xi}\sum_{(\ell)} A_\ell D f_\ell , \tag{9}$$

car $DA_\ell = 0$. On peut utiliser à nouveau la formule (1) pour transformer le premier terme, et la formule (7) dans le second, ce qui donne finalement

$$D^2(\ln \Xi) = \overline{A^2}^G - (\overline{A}^G)^2 = (\Delta A^G)^2 . \tag{10}$$

Ce résultat se généralise à l'ordre 3 :

$$D^3(\ln \Xi) = \overline{(A - \overline{A}^G)^3}^G , \tag{11}$$

mais pas directement aux ordres supérieurs.

On remarquera que des arguments semblables peuvent être appliqués à la formule III (II.8), donnant la valeur moyenne canonique $\overline{E}^c$ à partir de la fonction de partition Z; ils fournissent alors directement la formule III (II.9) et ses généralisations (III, note 28, formule (4)).

12. $$S^G = -k\sum_{(\ell)} P_\ell^G \ln P_\ell^G = -k\frac{1}{\Xi}\sum_{(\ell)} e^{-(E_\ell - \mu N_\ell)/kT}\left[-\frac{1}{kT}(E_\ell - \mu N_\ell) - \ln \Xi\right]$$
$$= \frac{1}{T}\frac{1}{\Xi}\sum_{(\ell)}(E_\ell - \mu N_\ell)\, e^{-(E_\ell - \mu N_\ell)/kT} + k \ln \Xi = \frac{1}{T}\left(\overline{E}^G - \mu \overline{N}^G\right) - \frac{1}{T} J .$$

que nous mettrons plutôt sous la forme

$$\boxed{J = \overline{E}^G - \mu \overline{N}^G - TS^G} \quad . \tag{II.23}$$

L'entropie grand-canonique s'exprime aussi très simplement[13] comme

$$\boxed{S^G = -\frac{\partial J}{\partial T}} \tag{II.24}$$

(la dérivée étant prise à μ, V, ... constants).

4. PRESSION GRAND-CANONIQUE

La définition de la pression grand-canonique p^G est analogue à celle de la pression microcanonique (II, § III.B.1) et de la pression canonique (III, § II.B.4) :

$$\boxed{p^G = -\frac{\partial J}{\partial V}} \quad . \tag{II.25}$$

Nous montrerons au paragraphe III.A.2 que, à la limite thermodynamique, p^G coïncide avec la pression canonique p^c et la pression microcanonique p^*, de même que S^G coïncide avec S^c et S^*.

C. Système constitué de sous-systèmes indépendants

1. FACTORISATION DE LA GRANDE FONCTION DE PARTITION ET ADDITIVITÉ DU GRAND POTENTIEL

Considérons un système $\mathcal{S}$ en équilibre avec un réservoir de particules et d'énergie, et supposons qu'il soit constitué de deux sous-systèmes $\mathcal{S}_1$ et $\mathcal{S}_2$ *indépendants*. Il peut s'agir de systèmes de *taille microscopique* dont le nombre de particules n'est pas fixé : par exemple, les molécules d'un gaz adsorbées sur les parois du récipient le sont souvent en des sites localisés pouvant accueillir zéro ou une particule (complément V.A); ou les « pièges » à électrons que constituent dans un semi-conducteur certaines impuretés contiennent de même zéro ou un électron, et dans

13. En effet, la dérivée de J par rapport à T vaut, d'après la définition (II.7),

$$\frac{\partial J}{\partial T} = -k \ln \Xi - kT \frac{\partial}{\partial T} \ln \Xi = -k \ln \Xi + k\beta \frac{\partial}{\partial \beta} \ln \Xi \, . \tag{1}$$

Mais (cf. (II.12))

$$\frac{\partial}{\partial \beta} (\ln \Xi) = -\overline{E}^G + \mu \overline{N}^G \, , \tag{2}$$

de sorte que

$$\frac{\partial J}{\partial T} = -k \ln \Xi + \frac{1}{T} \left(-\overline{E}^G + \mu \overline{N}^G \right) . \tag{3}$$

On reconnaît au second membre l'expression (II.22) de S^G.

certains cas jusqu'à deux (complément VI.A, § III). Si $\mathcal{S}_1$ et $\mathcal{S}_2$ sont *macroscopiques* et séparés par une cloison, celle-ci est fixe (pas d'échanges de volume); qu'elle soit diatherme ou adiabatique, et perméable ou imperméable aux particules, est sans importance, puisque $\mathcal{S}_1$ et $\mathcal{S}_2$ échangent de toute façon des particules et de la chaleur à travers le réservoir, qui les maintient tous deux à la même température T et au même potentiel chimique μ. Dans ces conditions, la grande fonction de partition du système $\mathcal{S}$ est le *produit des grandes fonctions de partition des sous-systèmes* $\mathcal{S}_1$ et $\mathcal{S}_2$, et son grand potentiel la *somme des grands potentiels* de $\mathcal{S}_1$ et $\mathcal{S}_2$.

En effet, un état microscopique (ℓ) quelconque du système $\mathcal{S}$ peut être caractérisé par la donnée des états microscopiques respectifs (ℓ_1) et (ℓ_2) des deux sous-systèmes :

$$(\ell)=(\ell_1, \ell_2)\ , \tag{II.26}$$

l'énergie E_ℓ et le nombre de particules N_ℓ correspondants étant simplement

$$E_\ell = E_{\ell_1} + E_{\ell_2}\ , \tag{II.27.a}$$

$$N_\ell = N_{\ell_1} + N_{\ell_2} \tag{II.27.b}$$

(E_{ℓ_1} et N_{ℓ_1} sont l'énergie et le nombre de particules associés à l'état (ℓ_1) de $\mathcal{S}_1$, E_{ℓ_2} et N_{ℓ_2} ceux de l'état (ℓ_2) de $\mathcal{S}_2$). Par conséquent, la grande fonction de partition du système $\mathcal{S}$,

$$\Xi = \sum_{(\ell_1, \ell_2)} \exp\{-\beta[E_{\ell_1} + E_{\ell_2} - \mu(N_{\ell_1} + N_{\ell_2})]\}\ , \tag{II.28}$$

peut s'écrire

$$\boxed{\Xi = \Xi_1 . \Xi_2}\ , \tag{II.29}$$

où $$\Xi_i = \sum_{(\ell_i)} e^{-\beta(E_{\ell_i} - \mu N_{\ell_i})}\ ; \quad i = 1, 2 \tag{II.30}$$

est la grande fonction de partition du sous-système $\mathcal{S}_i$. On en déduit aussitôt que le grand potentiel est additif :

$$\boxed{J = J_1 + J_2}\ , \tag{II.31}$$

où $$J_i = -kT \ln \Xi_i\ ; \quad i = 1, 2 \tag{II.32}$$

est le grand potentiel du sous-système $\mathcal{S}_i$.

Le raisonnement se généralise aussitôt à un nombre quelconque de sous-systèmes indépendants.

Dans le cas où $\mathcal{S}$ est un système *macroscopique,* il peut toujours être considéré comme la réunion d'un certain nombre de sous-systèmes macroscopiques : il suffit de diviser par la pensée son volume V en volumes V_i $(i = 1, 2, \ldots)$ plus petits mais toujours macroscopiques. *Si le couplage entre ces sous-systèmes est faible,* c'est-à-dire (complément II.A) si les forces entre particules sont à courte portée, les résultats précédents peuvent être appliqués; ils impliquent alors que *le grand potentiel J croît comme le volume V du système,* ou comme le nombre moyen $\overline{N}^G$ de particules qu'il contient[14].

14. En d'autres termes (complément II.A, § II), le grand potentiel est une grandeur *extensive,* de même que l'énergie moyenne, le nombre moyen de particules et l'entropie.

2. APPLICATION AU CAS PARTICULIER D'UN FLUIDE SIMPLE

Un fluide simple est un système *macroscopique* dont les propriétés dépendent de *trois variables seulement* qui, dans la situation grand-canonique, sont T, μ et le volume V :

$$J = J(T, \mu, V) . \tag{II.33}$$

Les gaz ou liquides ordinaires constitués d'un seul type de particules (c'est-à-dire purs) entrent dans cette catégorie.

Supposons que, T et μ étant maintenus constants, on multiplie par α (réel positif) le volume d'un tel système. Les résultats du paragraphe 1 indiquent que, si les fluctuations sont négligeables (limite thermodynamique), le grand potentiel J est multiplié par le même facteur, c'est-à-dire que

$$J(T, \mu, \alpha V) = \alpha J(T, \mu, V) . \tag{II.34}$$

En revanche, la pression grand-canonique p^G reste inchangée, d'après sa définition (II.25)[15] :

$$p^G(T, \mu, \alpha V) = p^G(T, \mu, V) . \tag{II.35}$$

Comme cette égalité doit être vraie pour tout α, *la pression grand-canonique dépend de T et μ mais pas de V :*

$$p^G = p^G(T, \mu) . \tag{II.36}$$

La dérivée de J par rapport à V, qui est égale à $-p^G$ d'après (II.25), est ici indépendante de V. Le grand potentiel s'écrit donc simplement

$$\boxed{J = -p^G V} . \tag{II.37}$$

Soulignons que les résultats (II.36) et (II.37) cesseraient d'être valables si les propriétés macroscopiques du système dépendaient de plus de trois variables[16], ou si les fluctuations n'étaient pas négligeables (voir paragraphe III.A.1 ci-dessous).

---

15. La formule (II.35) s'obtient immédiatement si l'on dérive (II.34) par rapport à V :

$$\alpha \frac{\partial J}{\partial V}(T, \mu, \alpha V) = \alpha \frac{\partial J}{\partial V}(T, \mu, V) . \tag{1}$$

On peut d'ailleurs démontrer directement le résultat (II.37) en dérivant (II.34) par rapport à α :

$$V \frac{\partial J}{\partial V}(T, \mu, \alpha V) = J(T, \mu, V) , \tag{2}$$

et $\partial J / \partial V$ n'est autre que $-p^G$.

16. Supposons que J dépende d'au moins une autre variable extensive comme le volume, c'est-à-dire comme lui proportionnelle à la taille du système (nous noterons x cette quatrième variable). La relation (II.34), obtenue en multipliant par α la taille du système, est ici remplacée par

$$J(T, \mu, \alpha V, \alpha x) = \alpha J(T, \mu, V, x) , \tag{1}$$

et la relation (II.35) par

$$p^G(T, \mu, \alpha V, \alpha x) = p^G(T, \mu, V, x) . \tag{2}$$

Nous en déduisons que p^G dépend alors de T et μ, mais aussi du rapport V/x (inchangé dans un changement de taille du système) :

$$p^G = p^G(T, \mu, V/x) . \tag{3}$$

D. Évolution spontanée du système après relâchement d'une contrainte

Les considérations du chapitre III (§ II.D) se généralisent facilement au cas qui nous occupe ici d'un système $\mathcal{S}$ en contact avec un réservoir $\mathcal{R}$ de particules et d'énergie. La définition du réservoir peut être précisée de la manière suivante : d'une part, la probabilité $P_{\ell, L}$ pour que le système global $\mathcal{S} \cup \mathcal{R}$ se trouve dans l'état (ℓ, L) est donnée par

$$P_{\ell, L} = P^*_L(E_{\text{tot}} - E_\ell, N_{\text{tot}} - N_\ell, \ldots) . P_\ell , \tag{II.38}$$

où $P^*_L(E_L, N_L, \ldots)$ est la distribution microcanonique pour le réservoir; d'autre part, la température et le potentiel chimique du réservoir restant invariables, son entropie microcanonique s'écrit (formule (I.14)).

$$S^*_{\mathcal{R}}(E_{\text{tot}} - E_\ell, N_{\text{tot}} - N_\ell, \ldots) = S^*_{\mathcal{R}}(E_{\text{tot}}, N_{\text{tot}}, \ldots) - \frac{E_\ell}{T} + \frac{\mu}{T} N_\ell . \tag{II.39}$$

L'entropie statistique associée à la distribution (II.38) se calcule facilement (cf. III, note 45) :

$$S_{\mathcal{S} \cup \mathcal{R}} = S + \sum_{(\ell)} P_\ell S^*_{\mathcal{R}}(E_{\text{tot}} - E_\ell, N_{\text{tot}} - N_\ell, \ldots) , \tag{II.40}$$

où S est l'entropie statistique associée à la distribution $\{P_\ell\}$ des états (ℓ) du système $\mathcal{S}$. Compte tenu de (II.39), on trouve (cf. III, note 46)

$$S_{\mathcal{S} \cup \mathcal{R}} = S + S^*_{\mathcal{R}}(E_{\text{tot}}, N_{\text{tot}}, \ldots) - \frac{1}{T} \overline{E} + \frac{\mu}{T} \overline{N} . \tag{II.41}$$

C'est la situation d'équilibre (microcanonique pour $\mathcal{S} \cup \mathcal{R}$, grand-canonique pour $\mathcal{S}$) qui nous intéresse ici. La formule (II.41) s'écrit dans ce cas, si l'on utilise l'expression (II.22) de l'entropie grand-canonique,

$$S^*_{\mathcal{S} \cup \mathcal{R}} = S^*_{\mathcal{R}}(E_{\text{tot}}, N_{\text{tot}}, \ldots) - \frac{1}{T} J . \tag{II.42}$$

La température T étant toujours positive et $S^*_{\mathcal{R}}(E_{\text{tot}}, N_{\text{tot}}, \ldots)$ étant invariable dans la situation envisagée, *à une augmentation de $S^*_{\mathcal{S} \cup \mathcal{R}}$ correspond une diminution de* J. Ainsi,

> *l'évolution spontanée d'un système maintenu en contact avec un réservoir de particules et d'énergie s'accompagne d'une diminution de son grand potentiel* J; le nouvel *état d'équilibre* atteint par le système après relâchement d'une contrainte correspond au *minimum de son grand potentiel* compatible avec les contraintes restantes.

On peut alors analyser les *échanges de volume* entre deux systèmes macroscopiques $\mathcal{S}_1$ et $\mathcal{S}_2$ maintenus par un réservoir à la même température T et au même potentiel chimique μ. On trouve en premier lieu que la *condition d'équilibre* s'écrit simplement

$$p^G_1(T, \mu, V_{1m}, \ldots) = p^G_2(T, \mu, V_2 = V - V_{1m}, \ldots) , \tag{II.43}$$

où V est le volume total, fixe, du système $\mathcal{S}_1 \cup \mathcal{S}_2$, V_{1m} la valeur la plus probable du volume de $\mathcal{S}_1$, p^G_i la pression grand-canonique du système $\mathcal{S}_i$ (formule (II.25)). Hors d'équilibre, le sens des échanges de volume est tel que *le système dont la pression est la plus forte augmente de volume.* Enfin, la *stabilité de l'équilibre* implique

$$\frac{\partial p^G}{\partial V} < 0 \quad \text{à } T, \mu, \ldots \text{ constants}, \tag{II.44}$$

c'est-à-dire, d'après (II.25),

$$\frac{\partial^2 J}{\partial V^2} > 0 . \tag{II.45}$$

Remarquons cependant que, pour un fluide simple (§ C.2), l'équation (II.43) est vérifiée pour toute valeur de V_{1m} puisque alors p^G ne dépend que de T et μ; les deux sous-systèmes sont dans ce cas en équilibre pour toute valeur de V_1.

III. LE FORMALISME GRAND CANONIQUE A LA LIMITE THERMODYNAMIQUE

Le formalisme grand-canonique que nous développons dans ce chapitre s'applique à des systèmes de taille quelconque, microscopique ou macroscopique. Dans ce dernier cas, le nombre de particules est constamment grand devant 1, mais ceci n'exclut pas toujours l'existence de fluctuations. Nous allons examiner ici comment se comporte la description grand-canonique dans la limite où ces fluctuations deviennent négligeables (limite thermodynamique; cf. III, § III.A), et la comparer aux descriptions microcanonique et canonique.

A. Limite thermodynamique de la description grand-canonique

1. DESCRIPTION GRAND-CANONIQUE A L'APPROXIMATION GAUSSIENNE

Nous avons explicité plus haut (§ I.C.1) la distribution statistique du nombre de particules dans un système en équilibre avec un réservoir de particules et d'énergie. Pour un système de taille suffisante, cette distribution a pratiquement la forme d'une gaussienne :

$$p^G(N) \simeq \frac{1}{\sqrt{2\pi(\Delta N)^2}} \exp\left[-\frac{(N-N_m)^2}{2(\Delta N)^2}\right], \tag{III.1}$$

où N_m et ΔN, donnés par (I.25) et (I.29), vérifient

$$N_m \gg 1 , \tag{III.2.a}$$

$$\frac{\Delta N}{N_m} \sim \frac{1}{\sqrt{N_m}} . \tag{III.2.b}$$

Calculons, dans le cadre de cette approximation, la *grande fonction de partition* Ξ du système. Nous pouvons pour cela utiliser l'expression (II.5), en y remplaçant Z par $\exp[-F/kT]$ et en transformant en intégrale la somme sur N :

$$\Xi = \int_0^{+\infty} \mathrm{d}N \, \exp\{-[F(T, N, V, \ldots) - \mu N]/kT\} . \tag{III.3}$$

La fonction à intégrer est précisément celle qu'il nous a fallu développer, au paragraphe I.C.1.b, pour obtenir l'expression approchée (III.1) de $P^G(N)$: l'approximation gaussienne consiste à prendre

$$\exp\{-[F(T, N, V, \ldots) - \mu N]/kT\} \simeq \exp\{-[F(T, N_m, V, \ldots) - \mu N_m]/kT\} \times \exp[-(N-N_m)^2/2(\Delta N)^2] . \tag{III.4}$$

En reportant cette expression approchée dans (III.3), on trouve sans difficulté

$$\Xi \simeq \exp\{-[F(T, N_m, V, \ldots) - \mu N_m]/kT\} \cdot \int_0^{+\infty} \mathrm{d}N \, \mathrm{e}^{-(N-N_m)^2/2(\Delta N)^2}$$
$$\simeq \sqrt{2\pi(\Delta N)^2} \exp\{-[F(T, N_m, V, \ldots) - \mu N_m]/kT\} . \tag{III.5}$$

On déduit aussitôt de cette formule le *grand potentiel* :

$$J(T, \mu, V, \ldots) = -kT \ln \Xi \simeq F(T, N_m, V, \ldots) - \mu N_m - \frac{kT}{2} \ln[2\pi(\Delta N)^2] . \tag{III.6}$$

Dans cette expression, N_m doit être considéré comme une fonction des paramètres extérieurs grand-canoniques[17] :

$$N_m = N_m(T, \mu, V, \ldots) . \tag{III.7}$$

On peut à partir de là calculer les autres caractéristiques du système, selon les formules du paragraphe II.B.

2. PASSAGE A LA LIMITE THERMODYNAMIQUE

La taille du système étudié peut ici être caractérisée par son volume V. Lorsque (la température T et le potentiel chimique μ restant constants) le volume V devient grand, le grand potentiel J du système croît comme V (§ II.C). Il en est de même, d'après (II.10), pour le nombre moyen de particules $\overline{N}^G$, et donc pour N_m qui coïncide avec lui à l'approximation gaussienne. L'écart quadratique moyen ΔN, lui, croît seulement comme $\sqrt{N_m}$ (cf. § I.C.1.b), c'est-à-dire comme $\sqrt{V}$. A la limite thermodynamique, ΔN devient inappréciable devant N_m, et tout se passe donc comme si le nombre de particules du système prenait la seule valeur N_m.

Dans la formule (III.6), les deux premiers termes croissent comme N_m ou V; le troisième, d'ordre $\ln N_m$, devient rapidement négligeable devant eux dès que N_m et V prennent des valeurs macroscopiques. A la limite thermodynamique, on a donc

$$J(T, \mu, V, \ldots) = F(T, N_m, V, \ldots) - \mu N_m \quad \text{(lim. thermo.)}. \tag{III.8}$$

L'expression (III.8) fournit celle de toutes les autres grandeurs physiques à la limite thermodynamique : l'énergie moyenne $\overline{E}^G$ (qui est, dans cette limite, *la* valeur de l'énergie) s'écrit[18]

$$\overline{E}^G = F(T, N_m, V, \ldots) - T\frac{\partial}{\partial T} F(T, N_m, V, \ldots) \quad \text{(lim. thermo.)}, \tag{III.9}$$

---

17. Dans l'équation (I.25) qui détermine N_m, le paramètre μ apparaît explicitement au second membre, et le potentiel chimique canonique μ^c est une fonction des autres paramètres extérieurs T (ou β), V, ... (outre sa dépendance par rapport à N_m). La racine N_m de cette équation dépend donc de T, μ, V, ...

18. Partons de la formule (II.14), en explicitant les variables maintenues constantes dans les dérivations partielles :

$$\overline{E}^G = J + \beta\left(\frac{\partial J}{\partial \beta}\right)_{\mu, V} - \mu\left(\frac{\partial J}{\partial \mu}\right)_{\beta, V} , \tag{1}$$

et calculons successivement les deux derniers termes en dérivant l'expression (III.8). D'une part,

$$\beta\left(\frac{\partial J}{\partial \beta}\right)_{\mu, V} = -T\left(\frac{\partial J}{\partial T}\right)_{\mu, V} = -T\left(\frac{\partial F}{\partial T}\right)_{N, V} - T\left(\frac{\partial F}{\partial N}\right)_{T, V}\left(\frac{\partial N_m}{\partial T}\right)_{\mu, V} + T\mu\left(\frac{\partial N_m}{\partial T}\right)_{\mu, V} . \tag{2}$$

D'autre part,

$$\left(\frac{\partial J}{\partial \mu}\right)_{\beta, V} = \left(\frac{\partial F}{\partial N}\right)_{T, V}\left(\frac{\partial N_m}{\partial \mu}\right)_{T, V} - N_m - \mu\left(\frac{\partial N_m}{\partial \mu}\right)_{T, V} . \tag{3}$$

Mais, d'après la définition III (II.30.b),

$$\left(\frac{\partial F}{\partial N}\right)_{T, V}(T, N_m, V, \ldots) = \mu^c(T, N_m, V, \ldots) . \tag{4}$$

L'égalité (I.25) montre alors que les résultats (2) et (3) se simplifient en

$$\beta\left(\frac{\partial J}{\partial \beta}\right)_{\mu, V} = -T\left(\frac{\partial F}{\partial T}\right)_{N, V}(T, N_m, V, \ldots) . \tag{5.a}$$

$$\left(\frac{\partial J}{\partial \mu}\right)_{\beta, V} = -N_m . \tag{5.b}$$

Il suffit de les reporter dans (1) pour obtenir la formule (III.9) :

$$\overline{E}^G = F - \mu N_m - T\left(\frac{\partial F}{\partial T}\right)_{N, V} + \mu N_m = F - T\left(\frac{\partial F}{\partial T}\right)_{N, V} . \tag{6}$$

l'entropie grand-canonique[19]

$$S^G = -\frac{\partial}{\partial T} F(T, N_m, V, \ldots) \quad \text{(lim. thermo.)} \tag{III.10}$$

et la pression grand-canonique[19]

$$p^G = -\frac{\partial}{\partial V} F(T, N_m, V, \ldots) \quad \text{(lim. thermo.)}. \tag{III.11}$$

Attention : dans les formules (III.9), (III.10) et (III.11), les dérivées partielles sont prises par rapport au système de variables, T, N_m, V, ... (et non plus T, μ, V, ...), c'est-à-dire que N_m *est maintenu constant.*

B. Équivalence des trois descriptions « canoniques » dans la limite thermodynamique

Considérons un système thermodynamique, c'est-à-dire de taille suffisante pour que les fluctuations y soient négligeables (III, § III.A).

Lorsque ce système est isolé et à l'équilibre (situation microcanonique), son état macroscopique est caractérisé par les paramètres extérieurs E, N, V, ...; toutes les autres grandeurs macroscopiques sont des fonctions des paramètres extérieurs : la température $T^*(E, N, V, \ldots)$, le potentiel chimique $\mu^*(E, N, V, \ldots)$, la pression $p^*(E, N, V, \ldots)$, mais aussi toutes les variables internes y, qui prennent chacune une valeur unique $y_m^*(E, N, V, \ldots)$ puisqu'il s'agit d'un système thermodynamique. Nous avons montré au chapitre III (§ III.C) que, si on le met en contact avec un thermostat dont la température T vaut précisément.

$$T = T^*(E, N, V, \ldots)\,, \tag{III.12}$$

les propriétés du système ne sont pas modifiées (équivalence entre les descriptions canonique et microcanonique d'un système thermodynamique). Nous avons vu en particulier que l'égalité (III.12) implique

$$\mu^c(T, N, V, \ldots) = \mu^*(E, N, V, \ldots) \quad \text{(lim. thermo.)}. \tag{III.13}$$

Supposons maintenant que l'on couple ce système thermodynamique avec un réservoir d'énergie et de particules, dont la température T vérifie (III.12) et dont le potentiel chimique μ soit tel que

$$\mu = \mu^*(E, N, V, \ldots)\,. \tag{III.14}$$

Les propriétés du système restent ici encore inchangées. En effet, dans la nouvelle situation (grand-canonique), le nombre de particules du système ne s'écarte pas de sa valeur la plus probable N_m puisque la taille du système rend les fluctuations négligeables; de plus, cette valeur est, d'après (I.25), (III.13) et (III.14), telle que

$$N_m = N\,. \tag{III.15}$$

---

19. La démonstration des formules (III.10) et (III.11) est tout à fait analogue à celle de la note précédente. Les compensations entre termes qu'on y constate ne sont pas fortuites : elles proviennent des propriétés de la correspondance entre $F(T, N, V, \ldots)$ et $J(T, \mu, V, \ldots)$:

$$J(T, \mu, V, \ldots) = F(T, N, V, \ldots) - \mu N$$

(formule (III.8)), que l'on appelle en mathématiques la *transformation de Legendre.*

Si l'on compare alors les résultats (III.9), (III.10) et (III.11) du paragraphe précédent avec les formules correspondantes du chapitre III, (II.10), (II.28) et (II.30.a), on constate que

$$\left.\begin{aligned} \overline{E}^G(T, \mu, V, \ldots) &= \overline{E}^c(T, N, V, \ldots) \\ S^G(T, \mu, V, \ldots) &= S^c(T, N, V, \ldots) \\ p^G(T, \mu, V, \ldots) &= p^c(T, N, V, \ldots) \end{aligned}\right\} \quad \text{(lim. thermo.)}, \tag{III.16}$$

Un raisonnement analogue à celui de la note 59 du chapitre III permet de montrer également que, pour une variable interne y quelconque,

$$y_m^G(T, \mu, V, \ldots) = y_m^c(T, N, V, \ldots) \quad \text{(lim. thermo.)}. \tag{III.17}$$

Compte tenu de (III.15) et des résultats du chapitre III (§ III.C), on peut écrire finalement

$$\left.\begin{aligned} \overline{N}^G(T, \mu, V, \ldots) &= N \\ \overline{E}^G(T, \mu, V, \ldots) &= \overline{E}^c(T, N, V, \ldots) = E \\ S^G(T, \mu, V, \ldots) &= S^c(T, N, V, \ldots) \\ &= S^*(E, N, V, \ldots) = S \\ p^G(T, \mu, V, \ldots) &= p^c(T, N, V, \ldots) \\ &= p^*(E, N, V, \ldots) = p \end{aligned}\right\} \quad \text{(lim. thermo.)}, \tag{III.18}$$

et, pour une variable interne y quelconque,

$$\begin{aligned} y_m^G(T, \mu, V, \ldots) &= y_m^c(T, N, V, \ldots) \\ &= y_m^*(E, N, V, \ldots) = y \quad \text{(lim. thermo.)}. \end{aligned} \tag{III.19}$$

Le nombre de particules N, l'énergie E, l'entropie S et la pression p, ainsi que la valeur y de chaque variable interne, sont ici des *grandeurs thermodynamiques.* On peut également définir à partir de (III.8) le *grand potentiel thermodynamique* J du système[20] :

$$J \widehat{=} F - \mu N \quad \text{(lim. thermo.)}. \tag{III.20}$$

Nous dirons pour résumer que, *dans un système thermodynamique à l'équilibre, chaque grandeur prend une valeur unique,* bien déterminée; *cette valeur est la même que le système soit isolé, ou en équilibre avec un thermostat qui fixe sa température* (pourvu que la condition (III.12) soit vérifiée), *ou en équilibre avec un réservoir qui fixe sa température et son potentiel chimique* (pourvu que les conditions (III.12) et (III.14) soient vérifiées). Pour décrire un tel système, *on a donc le choix* : on peut décider que sont fixés (outre les autres paramètres extérieurs tels que le volume V) l'énergie E et le nombre de particules N (description microcanonique), ou bien plutôt la température T et le nombre de particules N (description canonique), ou bien encore la température T et le potentiel chimique μ (description grand-canonique). *Ces trois descriptions sont équivalentes*[21].

---

20. Pour un *fluide simple,* le grand potentiel thermodynamique peut être mis sous la forme (cf. (II.37))

$$J = -pV \quad \text{(lim. thermo. – fluide simple)}.$$

21. On peut envisager également de mettre le système étudié en contact avec un *réservoir de volume,* qui maintient une *pression p constante.* L'équivalence, à la limite thermodynamique, entre les diverses descriptions statistiques du système englobe également ce nouveau cas (cf. complément III.D).

Soulignons cependant que cette équivalence ne vaut que *dans la limite thermodynamique,* et pour un système se trouvant *à l'équilibre macroscopique.* Si en effet on relâche une contrainte dans un système thermodynamique, le nouvel état d'équilibre qu'il atteint après évolution n'est pas le même si c'est l'énergie et le nombre de particules qui sont maintenus constants, ou bien la température et le nombre de particules, ou encore la température et le potentiel chimique.

IV. DESCRIPTION GRAND-CANONIQUE D'UN SYSTÈME DE PARTICULES INDÉPENDANTES IDENTIQUES ET INDISCERNABLES

Reprenons, dans le cadre grand-canonique, le problème des systèmes de particules *indépendantes* dont nous avons présenté la description canonique au chapitre III. Il est important de souligner d'emblée que nous nous limitons ici aux systèmes de particules *identiques et indiscernables*[22].

A. Formules générales

Nous utiliserons les mêmes notations qu'au chapitre III, paragraphe IV : les états individuels possibles pour chacune des particules du système seront désignés par (λ), l'énergie correspondante étant ε_λ ; chaque état microscopique (ℓ) du système sera repéré par la *configuration* $\{N_\lambda\}$ qui le caractérise, c'est-à-dire l'ensemble des *nombres d'occupation* N_λ des divers états individuels (λ) :

$$(\ell) \iff \{N_\lambda\} . \tag{IV.1}$$

Il sera commode de considérer que, dans chaque configuration, est spécifié un nombre d'occupation pour chaque état individuel, N_λ étant nul si aucune particule ne se trouve dans l'état (λ).

1. FACTORISATION DE LA GRANDE FONCTION DE PARTITION

Avec ces notations, l'énergie E_ℓ et le nombre de particules N_ℓ de l'état microscopique (ℓ) s'écrivent

$$E_\ell = \sum_{(\lambda)} N_\lambda \varepsilon_\lambda , \tag{IV.2.a}$$

$$N_\ell = \sum_{(\lambda)} N_\lambda , \tag{IV.2.b}$$

la sommation portant toujours sur l'ensemble de tous les états individuels possibles. La grande fonction de partition du système a donc pour expression

$$\Xi(\beta, \mu, V, \ldots) = \sum_{\{N_\lambda\}} \exp\left[-\beta \sum_{(\lambda)} N_\lambda(\varepsilon_\lambda - \mu)\right] . \tag{IV.3}$$

22. Rappelons que nous avons qualifié de « discernables », au chapitre III (§ IV.A.2), des particules qui étaient en réalité indiscernables : c'étaient les sites où elles étaient placées qui étaient discernables.

Il est alors facile de mettre Ξ sous forme d'un produit[23] :

$$\boxed{\Xi = \prod_{(\lambda)} \xi_\lambda} \quad , \tag{IV.4}$$

avec

$$\boxed{\xi_\lambda(\beta, \mu, V, \ldots) \widehat{=} \sum_{N_\lambda} \exp\left[-\beta N_\lambda(\varepsilon_\lambda - \mu)\right]} \tag{IV.5}$$

(la sommation porte ici simplement sur toutes les valeurs possibles pour le nombre d'occupation N_λ). Notons que, dans l'expression factorisée (IV.4), *chaque facteur est associé à un état individuel* (λ), et non pas, comme dans la fonction de partition d'un système de particules indépendantes discernables, à une particule du système : ξ_λ est en quelque sorte la *grande fonction de partition pour un état individuel.*

Soulignons également que la factorisation que nous venons de démontrer repose sur l'*indépendance des divers nombres d'occupation* N_λ (cf. note 23). Or cette indépendance disparaîtrait si le nombre total de particules était fixé à N, c'est-à-dire si, comme dans la description canonique, on imposait la contrainte $\sum_{(\lambda)} N_\lambda = N$. C'est en éliminant cette contrainte que le formalisme grand-canonique parvient à traiter exactement les systèmes de particules indépendantes identiques et indiscernables : la présence du réservoir de particules permet à chaque état individuel (λ) de « se peupler » librement, selon (IV.5), sans avoir à tenir compte des autres états possibles pour partager avec eux un nombre de particules N imposé.

2. PROPRIÉTÉS DU SYSTÈME

La connaissance de l'ensemble des états individuels (λ) possibles et de la fonction $\xi_\lambda(\beta, \mu, V, \ldots)$ associée à chacun d'eux (voir paragraphes B et C) suffit pour déterminer les propriétés du système. Pour simplifier les notations, nous supprimons

---

23. La sommation sur toutes les configurations $\{N_\lambda\}$ possibles revient à une sommation sur toutes les valeurs possibles pour le nombre d'occupation N_1 du premier état individuel (1), suivie d'une sommation *indépendante* sur le nombre d'occupation N_2 du deuxième état individuel (2), ... et ainsi et suite :

$$\sum_{\{N_\lambda\}} \equiv \sum_{N_1} \sum_{N_2} \ldots \sum_{N_\lambda} \ldots \, . \tag{1}$$

Or la fonction à sommer est elle-même un produit de facteurs, un pour chaque état individuel :

$$\exp\left[-\beta \sum_{(\lambda)} N_\lambda(\varepsilon_\lambda - \mu)\right] = e^{-\beta N_1(\varepsilon_1 - \mu)} . e^{-\beta N_2(\varepsilon_2 - \mu)} \ldots . e^{-\beta N_\lambda(\varepsilon_\lambda - \mu)} \ldots \tag{2}$$

(à un état individuel non occupé correspond un facteur égal à 1). Comme N_1 ne figure que dans le premier de ces facteurs, N_2 que dans le deuxième, etc., l'expression (IV.3) devient simplement

$$\Xi = \left[\sum_{N_1} e^{-\beta N_1(\varepsilon_1 - \mu)}\right] . \left[\sum_{N_2} e^{-\beta N_2(\varepsilon_2 - \mu)}\right] \ldots . \left[\sum_{N_\lambda} e^{-\beta N_\lambda(\varepsilon_\lambda - \mu)}\right] \ldots , \tag{3}$$

formule que l'on peut écrire sous la forme condensée (IV.4)-(IV.5).

dans la suite l'indice G qui distingue les grandeurs grand-canoniques; si le système est assez grand pour se situer dans la limite thermodynamique, on peut également supprimer les symboles de valeur moyenne sur $\overline{E}$ et $\overline{N}$ (cf. § III.B).

Le grand potentiel J se calcule par

$$J(T, \mu, V, \ldots) = -kT \sum_{(\lambda)} \ln \xi_\lambda(T, \mu, V, \ldots) , \tag{IV.6}$$

d'où l'on peut tirer les autres grandeurs. La formule (II.10) donne immédiatement le nombre moyen de particules sous la forme

$$\overline{N} = \sum_{(\lambda)} \overline{N}_\lambda , \tag{IV.7}$$

où le *nombre d'occupation moyen* $\overline{N}_\lambda$ de l'état (λ) vaut[24]

$$\overline{N}_\lambda = kT \frac{\partial}{\partial \mu} (\ln \xi_\lambda) . \tag{IV.8}$$

Quant à l'énergie moyenne, on peut l'obtenir à partir de (II.12) ou (II.14). On trouve bien sûr[25]

$$\overline{E} = \sum_{(\lambda)} \overline{N}_\lambda \varepsilon_\lambda \ ; \tag{IV.9}$$

les divers états individuels (λ) contribuent de façon indépendante, l'énergie de chacun d'eux étant multipliée par le nombre moyen $\overline{N}_\lambda$ de particules qui l'occupent. Enfin, l'entropie S et la pression p se déduisent facilement des formules (II.24) et (II.25) :

$$S = k \sum_{(\lambda)} \left[\ln \xi_\lambda + T \frac{\partial}{\partial T} \ln \xi_\lambda \right] , \tag{IV.10}$$

$$p = kT \sum_{(\lambda)} \frac{\partial}{\partial V} (\ln \xi_\lambda) . \tag{IV.11}$$

---

24. Notons que $\overline{N}_\lambda$, nombre d'occupation moyen d'un état individuel (λ) particulier, peut également s'exprimer à partir de la grande fonction de partition globale Ξ :

$$\overline{N}_\lambda = -kT \frac{\partial}{\partial \varepsilon_\lambda} \ln \Xi . \tag{1}$$

En effet les formules (IV.4) et (IV.5) impliquent

$$\frac{\partial}{\partial \varepsilon_\lambda} \ln \Xi = \frac{\partial}{\partial \varepsilon_\lambda} \ln \xi_\lambda = -\frac{\partial}{\partial \mu} \ln \xi_\lambda . \tag{2}$$

25. La formule (II.12) s'écrit ici

$$\overline{E} = \mu \overline{N} - \frac{\partial}{\partial \beta} \ln \Xi = \sum_{(\lambda)} \left[\mu \overline{N}_\lambda - \frac{\partial}{\partial \beta} \ln \xi_\lambda \right] . \tag{1}$$

Or la définition (IV.5) implique une relation simple entre les dérivées de ξ_λ par rapport à β et par rapport à μ :

$$\frac{\partial}{\partial \beta} (\ln \xi_\lambda) = -\frac{1}{\xi_\lambda} \sum_{N_\lambda} N_\lambda (\varepsilon_\lambda - \mu) \, \mathrm{e}^{-\beta N_\lambda (\varepsilon_\lambda - \mu)} = -(\varepsilon_\lambda - \mu) \frac{1}{\beta} \frac{\partial}{\partial \mu} (\ln \xi_\lambda) . \tag{2}$$

Compte tenu de (IV.8), l'expression (1) de $\overline{E}$ devient donc

$$\overline{E} = \sum_{(\lambda)} [\mu \overline{N}_\lambda + (\varepsilon_\lambda - \mu) \overline{N}_\lambda] = \sum_{(\lambda)} \overline{N}_\lambda \varepsilon_\lambda . \tag{3}$$

B. Les statistiques quantiques

La fonction $\xi_\lambda(T, \mu, V, \ldots)$ est facile à déterminer explicitement, car, dans sa définition (IV.5), la dépendance en N_λ de l'expression à sommer est très simple. Il faut évidemment distinguer deux cas, car l'ensemble des valeurs possibles pour N_λ n'est pas le même suivant que les particules considérées sont des fermions ou des bosons.

1. STATISTIQUE DE FERMI-DIRAC

S'il s'agit de fermions, le nombre d'occupation N_λ d'un état donné ne peut prendre que les valeurs 0 et 1; dans (IV.5), la somme se réduit alors à deux termes :

$$\xi_\lambda^F = 1 + e^{-\beta(\varepsilon_\lambda - \mu)} . \tag{IV.12}$$

Le nombre d'occupation moyen correspondant s'en déduit aussitôt (formule (IV.8))[26] :

$$\boxed{\overline{N}_\lambda^F = \frac{1}{e^{+\beta(\varepsilon_\lambda - \mu)} + 1}} . \tag{IV.13}$$

L'expression (IV.13) caractérise ce que l'on appelle la *distribution de Fermi.* Nous étudierons en détail sa forme et ses implications dans le chapitre suivant (VI, § I). Notons que

$$\xi_\lambda^F = \frac{1}{1 - \overline{N}_\lambda^F} , \tag{IV.14}$$

de sorte que toutes les propriétés du système peuvent être calculées à partir de (IV.13).

2. STATISTIQUE DE BOSE-EINSTEIN

Pour des *bosons,* le nombre d'occupation N_λ peut prendre toutes les valeurs entières positives ou nulles, de sorte que la somme de la définition (IV.5) est celle d'une série :

$$\xi_\lambda^B = \sum_{N_\lambda = 0}^{+\infty} e^{-\beta N_\lambda(\varepsilon_\lambda - \mu)} . \tag{IV.15}$$

Mais le calcul est encore simple, car il s'agit d'une *série géométrique :*

$$\xi_\lambda^B = \sum_{N_\lambda = 0}^{+\infty} \left[e^{-\beta(\varepsilon_\lambda - \mu)}\right]^{N_\lambda} . \tag{IV.16}$$

a. Condition sur le potentiel chimique

Il est bien connu qu'une série géométrique converge si et seulement si sa raison est inférieure à 1; la fonction ξ_λ^B n'existe donc que si

$$e^{-\beta(\varepsilon_\lambda - \mu)} < 1 , \tag{IV.17}$$

c'est-à-dire, β étant toujours positif,

$$\varepsilon_\lambda - \mu > 0 . \tag{IV.18}$$

Cette condition doit bien entendu être vérifiée pour tous les états individuels (λ); si ε_0 est l'énergie individuelle la plus basse :

$$\varepsilon_\lambda \geqslant \varepsilon_0 , \tag{IV.19}$$

---

26. $\overline{N}_\lambda^F = kT \dfrac{\partial}{\partial \mu} (\ln \xi_\lambda^F) = \dfrac{1}{\beta} \dfrac{\beta\, e^{-\beta(\varepsilon_\lambda - \mu)}}{1 + e^{-\beta(\varepsilon_\lambda - \mu)}} = \dfrac{1}{e^{+\beta(\varepsilon_\lambda - \mu)} + 1}$.

il suffit pour cela que

$$\mu < \varepsilon_0 \,. \tag{IV.20}$$

Pour un système de *bosons* indépendants, l'équilibre n'est possible que si *le potentiel chimique* imposé par le réservoir *est inférieur au niveau fondamental individuel.*

b. Distribution de Bose

La condition précédente étant remplie, la somme de la série géométrique (IV.16) est également bien connue :

$$\xi_\lambda^B = \frac{1}{1 - e^{-\beta(\varepsilon_\lambda - \mu)}} \,. \tag{IV.21}$$

Comme dans le cas des fermions, on utilise plutôt le nombre d'occupation moyen $\overline{N}_\lambda^B$, qui se déduit de (IV.21) selon la formule (IV.8) (le calcul est analogue à celui de la note 26) :

$$\boxed{\overline{N}_\lambda^B = \frac{1}{e^{+\beta(\varepsilon_\lambda - \mu)} - 1} \,; \quad \mu < \varepsilon_0} \tag{IV.22}$$

Les résultat du paragraphe A.2 donnent alors les propriétés du système à partir de cette expression, puisque

$$\xi_\lambda^B = 1 + \overline{N}_\lambda^B \,. \tag{IV.23}$$

La formule (IV.22) caractérise la *distribution de Bose,* que nous analyserons au chapitre VI, paragraphe II.

C. La statistique classique de Maxwell-Boltzmann

1. LIMITE CLASSIQUE DES STATISTIQUES QUANTIQUES

Examinons ce que deviennent les statistiques quantiques du paragraphe B lorsque *les nombres d'occupation moyens des divers états individuels sont petits devant* 1. Les formules (IV.13) et (IV.22) montrent que ceci se produit si

$$e^{+\beta(\varepsilon_\lambda - \mu)} \gg 1 \,. \tag{IV.24}$$

$\overline{N}_\lambda^F$ et $\overline{N}_\lambda^B$ tendent alors vers la *même expression limite*[27], que l'on appelle la *distribution de Maxwell-Boltzmann :*

$$\overline{N}_\lambda^F \simeq \overline{N}_\lambda^B \simeq \overline{N}_\lambda^{\text{M.B.}} \,, \tag{IV.25}$$

avec

$$\boxed{\overline{N}_\lambda^{\text{M.B.}} = e^{-\beta(\varepsilon_\lambda - \mu)}} \,. \tag{IV.26}$$

---

27. Bien entendu, les termes correctifs que l'on obtient en développant $\overline{N}_\lambda^F$ et $\overline{N}_\lambda^B$ jusqu'à l'ordre suivant en $\exp[-\beta(\varepsilon_\lambda - \mu)]$ sont différents pour les deux statistiques quantiques :

$$\frac{1}{e^{\beta(\varepsilon_\lambda - \mu)} \pm 1} = e^{-\beta(\varepsilon_\lambda - \mu)} [1 \mp e^{-\beta(\varepsilon_\lambda - \mu)} + \ldots] \,.$$

La distribution de Maxwell-Boltzmann (IV.26) est donc en quelque sorte intermédiaire entre les deux distributions quantiques, qui tendent vers elle l'une par valeurs inférieures (fermions), l'autre par valeurs supérieures (bosons).

Il est facile de montrer que cette limite correspond effectivement à l'approximation de Maxwell-Boltzmann que nous avons introduite et discutée au chapitre III (§ IV.A.3.c). En effet, les formules (IV.12) et (IV.21) deviennent ici

$$\ln \xi_\lambda^F \simeq \ln \xi_\lambda^B \simeq e^{-\beta(\epsilon_\lambda - \mu)} . \tag{IV.27}$$

La grande fonction de partition du système est donc telle que

$$\ln \Xi^{M.B.} = \sum_{(\lambda)} e^{-\beta(\epsilon_\lambda - \mu)} = e^{\beta\mu} \sum_{(\lambda)} e^{-\beta\epsilon_\lambda} , \tag{IV.28}$$

où l'on reconnaît la fonction de partition z d'une particule :

$$\Xi^{M.B.} = \exp\left[z\, e^{\beta\mu}\right] . \tag{IV.29}$$

En développant l'exponentielle, on peut mettre cette expression sous la forme

$$\Xi^{M.B.}(\beta, \mu, V, \ldots) = \sum_{N=0}^{+\infty} e^{\beta\mu N} \frac{z^N}{N!} . \tag{IV.30}$$

Si l'on identifie avec la formule générale (II.5)[28], on en déduit la *fonction de partition* du système *dans la limite de Maxwell-Boltzmann :*

$$Z^{M.B.}(\beta, N, V, \ldots) = \frac{1}{N!}\left[z(\beta, V, \ldots)\right]^N . \tag{IV.31}$$

On retrouve bien le résultat III (IV.19.a).

Ainsi, *l'approximation de Maxwell-Boltzmann apparait comme la limite commune des statistiques quantiques* de Fermi-Dirac et de Bose-Einstein *lorsque tous les nombres d'occupation deviennent petits devant* 1[29].

---

28. Les deux expressions doivent bien sûr coïncider pour toutes les valeurs permises de μ.

29. La condition (IV.24) est différente de celle du chapitre III (§ IV.A.3.c.β). Ceci est simplement dû au fait que c'est ici le potentiel chimique μ qui est fixé, et non plus le nombre de particules N.

Mais on retrouve facilement la condition III (IV.23). Le nombre moyen de particules vaut en effet

$$\overline{N}^{M.B.} = -\frac{\partial J}{\partial \mu} = kT \frac{\partial}{\partial \mu} \ln \Xi^{M.B.} = z\, e^{\beta\mu} , \tag{1}$$

de sorte que (IV.24) peut s'écrire

$$e^{\beta\epsilon_\lambda} \cdot \frac{z}{\overline{N}} \gg 1 , \tag{2}$$

c'est-à-dire ($\beta > 0$)

$$e^{\beta\epsilon_0} \cdot z \gg \overline{N} . \tag{3}$$

Cette inégalité, qui diffère de III (IV.23) seulement en ce que $\overline{N}$ remplace N, lui est équivalente pour les valeurs de N suffisamment grandes.

2. DESCRIPTION GRAND-CANONIQUE DES GAZ PARFAITS CLASSIQUES MONOATOMIQUES

Nous connaissons la fonction de partition z pour une particule ponctuelle libre (III, formule (IV.30)) :

$$z = (2s+1)\frac{V}{\Lambda^3}, \tag{IV.32.a}$$

avec $$\Lambda = \left(\frac{2\pi\hbar^2}{m}\beta\right)^{1/2}. \tag{IV.32.b}$$

Si la condition (IV.24) est satisfaite et si le nombre (moyen) de particules est suffisamment grand pour que la limite thermodynamique soit valable, le système considéré constitue un gaz parfait classique monoatomique, tel que nous l'avons défini au chapitre III (§ IV.B.1). On peut en effet retrouver alors à partir de (IV.29) et (IV.32), en utilisant les formules du paragraphe II.B, tous les résultats du chapitre III (§ IV.B.3).

Écrivons tout d'abord le grand potentiel :

$$J = -kTz\, e^{\beta\mu}, \tag{IV.33}$$

c'est-à-dire

$$J(\beta, \mu, V, \ldots) = -(2s+1)\frac{1}{\beta}\frac{V}{\Lambda^3}e^{\beta\mu}. \tag{IV.34}$$

Nous constatons que, dans ce cas particulier, les dépendances en V et μ sont particulièrement simples : *J est proportionnel à V et à la fugacité φ.*

Calculons alors le *nombre de particules,* qui coïncide ici avec sa valeur moyenne :

$$N = -\frac{\partial J}{\partial \mu} = (2s+1)\frac{V}{\Lambda^3}e^{\beta\mu}. \tag{IV.35}$$

Cette relation peut être inversée pour exprimer le potentiel chimique μ en fonction de N, V et T.

La pression du gaz est donnée par

$$p = -\frac{\partial J}{\partial V} = (2s+1)\frac{1}{\beta}\frac{1}{\Lambda^3}e^{\beta\mu}. \tag{IV.36}$$

On aboutit ainsi directement à la relation III (IV.50) de *proportionnalité entre la fugacité et la pression.* Si l'on élimine μ au profit de N grâce à (IV.35), la formule (IV.36) s'écrit

$$p = \frac{N}{V}kT, \tag{IV.37}$$

où l'on reconnaît l'*équation d'état des gaz parfaits.*

De même, l'*entropie* s'obtient par

$$S = -\frac{\partial J}{\partial T}. \tag{IV.38}$$

Un calcul simple[30] redonne la formule de Sackur-Tétrode (III (IV.43)).

Enfin, l'*énergie moyenne* du gaz peut être déduite de la formule (II.23), et l'on retrouve[31]

$$\overline{E}=\frac{3}{2}\,NkT\,. \tag{IV.39}$$

30. $$S=\frac{\partial}{\partial T}\left\{(2s+1)kTV\left(\frac{mkT}{2\pi\hbar^2}\right)^{3/2}\mathrm{e}^{\mu/kT}\right\}$$

$$=(2s+1)\,k\,\frac{V}{\Lambda^3}\,\mathrm{e}^{\mu/kT}+(2s+1)\,kT\,\frac{V}{\Lambda^3}\times\frac{3}{2T}\,\mathrm{e}^{\mu/kT}+(2s+1)\,kT\,\frac{V}{\Lambda^3}\left(-\frac{\mu}{kT^2}\right)\mathrm{e}^{\mu/kT}\,. \tag{1}$$

La relation (IV.35) permet d'écrire ce résultat sous la forme

$$S=Nk+\frac{3}{2}\,Nk-N\frac{\mu}{T}\,. \tag{2}$$

Il suffit alors de remplacer μ en fonction de N, V et T, en utilisant (IV.35) :

$$\mu=kT\,\ln\left[\frac{N}{V}\,\frac{\Lambda^3}{2s+1}\right], \tag{3}$$

pour obtenir finalement

$$S=NK\left[\frac{5}{2}-\ln\frac{N}{V}-\ln\,\Lambda^3+\ln(2s+1)\right], \tag{4}$$

qui est bien la formule de Sackur-Tétrode.

31. $$\overline{E}=J+\mu N+TS\,. \tag{1}$$

Si l'on utilise pour S l'expression donnée par la formule (2) de la note précédente, le terme μN disparaît;

$$\overline{E}=-(2s+1)\,\frac{1}{\beta}\,\frac{V}{\Lambda^3}\,\mathrm{e}^{\beta\mu}+\mu N+\frac{5}{2}\,NkT-N\mu\,. \tag{2}$$

On élimine ensuite la fugacité à l'aide de la relation (IV.35), pour obtenir finalement

$$\overline{E}=-NkT+\frac{5}{2}\,NkT=\frac{3}{2}\,NkT\,. \tag{3}$$

COMPLÉMENTS DU CHAPITRE V

V.A. Adsorption	Application simple des notions et résultats du chapitre. Le début (§ I) est très facile, la suite à peine plus difficile.
V.B. Équilibre d'un système macroscopique dans un champ de forces extérieur	Problème physique important, mais un peu délicat à bien comprendre. L'application aux fluides (§ II.1) est très facile, l'étude des conducteurs (§ II.2) nettement plus difficile.
V.C. Les gaz réels et le développement du viriel **V.D. Le postulat fondamental sous sa forme «généralisée»** **V.E. Utilisation du formalisme des opérateurs-densités**	Sujets essentiellement théoriques. Celui que traite V.C est particulièrement important. V.C. : prolonge III.G (équation de van der Waals), dont la connaissance détaillée n'est cependant pas nécessaire; les paragraphes I et II sont facilement accessibles, le paragraphe III plus complexe techniquement. V.D. : Énonce et discute le «principe du désordre maximum»; apporte un éclairage nouveau sur la signification des distributions «canoniques»; pas difficile à suivre, mais l'interprétation des résultats (§ III) est plus délicate. V.E. : Suppose connu l'opérateur-densité (complément I.H); le paragraphe I reprend dans ce cadre les diverses distributions introduites aux chapitres II, III et V; il est facile si I.H est bien assimilé; l'application du formalisme au développement de perturbation (§ II.1) est difficile, mais la méthode variationnelle (§ II.2) est assez facile à comprendre.
V.F. Fluctuations dans un système à l'équilibre statistique	Aborde le problème des écarts par rapport aux valeurs moyennes, c'est-à-dire par rapport à la limite thermodynamique dont on se contente le plus souvent. Les raisonnements sont faciles à suivre si l'on a bien compris la différence entre les diverses descriptions «canoniques» d'un système macroscopique; les détails de calcul sont souvent assez complexes techniquement.

COMPLÉMENT V.A

Adsorption

I. Étude d'un modèle très simple
  1. Le modèle
  2. Analyse grand-canonique
  3. Propriétés thermodynamiques
    a. Isothermes d'adsorption
    b. Énergie et capacité calorifique
    c. Chaleur d'adsorption

II. Généralisation
  1. Analyse grand-canonique
  2. Propriétés thermodynamiques
    a. Isothermes d'adsorption
    b. Énergie et capacité calorifique
    c. Chaleur d'adsorption

III. Cas ou l'adsorption s'accompagne d'une dissociation de la molécule
  1. Condition d'équilibre
  2. Propriétés du système

Les molécules d'un gaz sont souvent attirées par les solides avec lesquels elles sont en contact. Elles peuvent même rester « piégées » à la surface du solide : c'est le phénomène de l'*adsorption*[1]. Le solide s'appelle alors le *substrat* (ou adsorbant); les molécules du gaz piégées à la surface du substrat forment la *phase adsorbée* (ou adsorbat).

Nous allons analyser ici, à l'aide des techniques grand-canoniques, l'équilibre entre les molécules du gaz et les molécules adsorbées.

I. ÉTUDE D'UN MODÈLE TRÈS SIMPLE

1. LE MODÈLE

Un récipient de volume V contient un gaz, supposé parfait, de particules monoatomiques de spin nul. Le gaz est en contact avec un substrat dont la surface présente A pièges, susceptibles d'adsorber chacun *un seul atome,* dans un *état unique,* d'énergie $-\varepsilon_0$. Le paramètre $\varepsilon_0 > 0$ représente l'énergie de liaison de l'atome dans le piège. L'ensemble est maintenu à la température T (fig. 1). Nous supposons d'autre part que le nombre d'atomes (dans le gaz et dans la phase adsorbée) est suffisamment grand pour qu'on puisse considérer que la limite thermodynamique est atteinte.

Ce modèle est évidemment très simple; il retient cependant les ingrédients essentiels à une première compréhension du phénomène de l'adsorption et son étude constitue un bon exercice pour se familiariser avec les techniques grand-canoniques.

1. On distingue deux types d'adsorption : la physisorption et la chimisorption. Dans la *physisorption* ce sont les forces de van der Waals qui attirent et lient l'adsorbat sur le substrat; les énergies mises en jeu sont relativement faibles (inférieures à 0,5 eV), et les atomes adsorbés se déplacent facilement sur la surface du substrat. La *chimisorption,* au contraire, est une véritable réaction chimique entre les atomes adsorbés et ceux du substrat; les énergies correspondantes ont les valeurs typiques des énergies des liaisons chimiques (1 à 5 eV) et les atomes adsorbés restent fixés sur les sites du substrat auxquels ils sont liés. C'est le phénomène de chimisorption qui nous intéresse principalement dans ce complément.

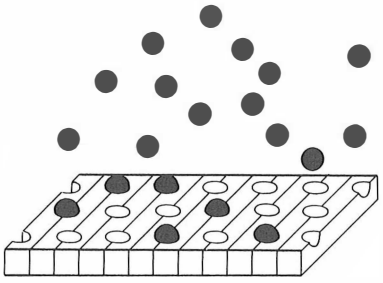


FIGURE 1
Représentation schématique du phénomène d'adsorption. Le substrat (hachuré) peut piéger à sa surface certaines des molécules du gaz avec lequel il est en contact.

2. ANALYSE GRAND-CANONIQUE

Considérons le *système* constitué par les *atomes adsorbés.* Son nombre de particules n'est pas fixé puisqu'il peut échanger des atomes avec le gaz environnant. Le *gaz* constitue ainsi un *réservoir* de particules et d'énergie qui impose au système sa *température* T et son *potentiel chimique* μ[2].

Les propriétés d'un tel système sont déterminées dès que l'on connaît sa grande fontion de partition Ξ. Les atomes étant indépendants et indiscernables, Ξ se factorise, à raison d'un facteur par piège (cf. V, § II.C.1) :

$$\Xi = \xi^A , \tag{A.1}$$

puisque tous les pièges ont la même grande fonction de partition ξ. Un piège peut contenir zéro ou un atome, et donc

$$\xi = 1 + e^{\beta(\varepsilon_0+\mu)} , \quad \beta = 1/kT . \tag{A.2}$$

On en déduit le nombre moyen d'atomes adsorbés :

$$N_a = kT \frac{\partial \ln \Xi}{\partial \mu} = A \frac{e^{\beta(\varepsilon_0+\mu)}}{1+e^{\beta(\varepsilon_0+\mu)}} . \tag{A.3}$$

Le rapport

$$\theta \hat{=} \frac{N_a}{A} = \frac{1}{1+e^{-\beta(\varepsilon_0+\mu)}} , \tag{A.4}$$

qui donne la probabilité pour qu'un piège soit rempli, s'appelle le *« taux d'adsorption »*.

L'énergie totale E_a des particules adsorbées s'obtient à partir de

$$E_a - \mu N_a = -\frac{\partial \ln \Xi}{\partial \beta} ; \tag{A.5}$$

on trouve

$$E_a = \frac{-\varepsilon_0 A}{1+e^{-\beta(\varepsilon_0+\mu)}} = -\varepsilon_0 N_a , \tag{A.6}$$

résultat attendu, puisque chaque atome adsorbé n'a qu'une seule énergie possible ($-\varepsilon_0$).

2. Le gaz constitue réellement un réservoir si le nombre d'atomes à l'état gazeux est bien supérieur à celui des atomes adsorbés, ce qui est en général le cas. En fait, on peut toujours considérer que le système des particules adsorbées est en contact avec un vrai réservoir, dont la température T et le potentiel chimique μ sont ceux du gaz à l'équilibre : à la limite thermodynamique, toutes les descriptions sont équivalentes.

Enfin, l'entropie S_a se calcule directement lorsqu'on connaît le grand potentiel (formule V (II.23)) :

$$J \mathrel{\widehat{=}} -kT \ln \Xi = E_a - \mu N_a - TS_a \ ; \tag{A.7}$$

il vient

$$S_a = kA \ln(1 + e^{\beta(\varepsilon_0 + \mu)}) - \frac{\varepsilon_0 + \mu}{T} N_a \ . \tag{A.8}$$

De la relation (A.3) on tire aisément

$$\frac{\varepsilon_0 + \mu}{kT} = \ln \frac{N_a}{A - N_a} \ ; \tag{A.9}$$

l'entropie S_a s'exprime donc en fonction du nombre de particules adsorbées N_a, sous la forme[3]

$$S_a = k\left[A \ln A - N_a \ln N_a - (A - N_a)\ln(A - N_a)\right] . \tag{A.10}$$

3. PROPRIÉTÉS THERMODYNAMIQUES

Les propriétés du système des particules adsorbées sont donc connues dès que la température T et le potentiel chimique μ le sont. Or ce dernier est fixé par le gaz qui sert de réservoir. Pour un gaz parfait monoatomique, μ est relié à la pression p et à la température T par (III, § IV.B.3.c)

$$e^{\mu/kT} = \frac{p}{kT}\left(\frac{2\pi\hbar^2}{mkT}\right)^{3/2} . \tag{A.11}$$

En remplaçant μ par son expression dans les formules du paragraphe précédent, on obtient les propriétés du système en fonction des paramètres T et p, plus directement accessibles à l'expérience.

a. Isothermes d'adsorption

Il est commode de poser

$$p_0(T) \mathrel{\widehat{=}} kT\left(\frac{mkT}{2\pi\hbar^2}\right)^{3/2} e^{-\varepsilon_0/kT} . \tag{A.12}$$

En reportant (A.11) et (A.12) dans (A.4), on écrit le taux d'adsorption sous la forme

$$\theta \mathrel{\widehat{=}} \frac{N_a}{A} = \frac{p}{p + p_0(T)} . \tag{A.13}$$

---

3. La formule (A.10) n'est autre que

$$S_a = k \ln \Omega_a \ , \tag{1}$$

où Ω_a est ici le nombre de façons de ranger les N_a atomes adsorbés dans les A pièges :

$$\Omega_a = \frac{A!}{N_a!(A - N_a)!} . \tag{2}$$

Pour chaque valeur de T on peut alors tracer θ en fonction de la pression (fig. 2); on obtient ainsi une série de courbes connues sous le nom d'*isothermes d'adsorption de Langmuir*[4]. Aux faibles pressions, θ est proportionnel à p :

$$\theta \simeq \frac{p}{p_0(T)} \quad \text{si} \quad p \ll p_0(T) ; \tag{A.14}$$

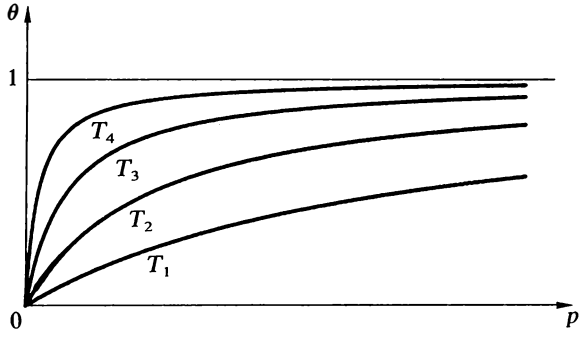


FIGURE 2
Isothermes de Langmuir : taux d'adsorption θ en fonction de la pression p pour différentes valeurs de la température ($T_1 > T_2 > T_3 > T_4$).

aux fortes pressions, on observe une saturation; tous les sites sont occupés :

$$\theta \simeq 1 \quad \text{si} \quad p \gg p_0(T) . \tag{A.15}$$

$p_0(T)$ est une fonction croissante de la température; à p fixée, θ décroît donc quand la température augmente : plus la température est élevée, plus l'agitation thermique a tendance à arracher les atomes des pièges pour les transférer dans le gaz. Si l'on veut « dégazer » un solide des atomes adsorbés qu'il contient, il faut donc le chauffer dans une enceinte à basse pression.

b. Énergie et capacité calorifique

L'énergie moyenne E_T du *système global* constitué par *le gaz et l'adsorbat* est simplement la somme de deux termes : à l'énergie (A.6) des atomes adsorbés s'ajoute celle des atomes à l'état gazeux. En moyenne, chaque atome du gaz a l'énergie $3kT/2$; si N_g est leur nombre moyen, on a

$$E_T = E_g + E_a = \frac{3}{2} N_g kT - N_a \varepsilon_0 . \tag{A.16}$$

Pour calculer la *capacité calorifique du système global,* à l'équilibre dans le récipient de volume V, on utilise la conservation du nombre total d'atomes :

$$dN_g + dN_a = 0 . \tag{A.17}$$

4. Les premières études expérimentales ont été réalisées par I. Langmuir pour l'adsorption de l'azote sur le mica (*J. Chem. Soc. 40,* 1361 (1918)).

On trouve ainsi

$$C_V \widehat{=} \frac{dE_T}{dT} = \frac{3}{2} N_g k + \left(\frac{3}{2} kT + \varepsilon_0\right)\left(\frac{dN_g}{dT}\right)_{eq} . \tag{A.18}$$

La notation $(dN_g/dT)_{eq}$ signifie que la dérivée de N_g par rapport à la température doit être prise le long de la courbe d'équilibre, c'est-à-dire avec N_a donné par la relation (A.13), dans laquelle p dépend de T par l'intermédiaire de l'équation d'état des gaz parfaits[5] :

$$p = \frac{N_g kT}{V} . \tag{A.19}$$

A la capacité calorifique propre des atomes du gaz $3N_g k/2$ (celle des atomes adsorbés est ici nulle puisque, à N_a fixé, leur énergie ne dépend pas de la température) s'ajoute donc un terme supplémentaire car une augmentation de température fait passer un certain nombre d'atomes de la phase adsorbée à la phase gazeuse. Autrement-dit, lorsqu'on fournit de la chaleur au système global, la température augmente moins que si le passage des atomes d'une phase à l'autre était interdit[6].

c. Chaleur d'absorption

Lorsque δn atomes passent de la phase gazeuse à la phase adsorbée, à température et volume constants (par fluctuation autour de l'équilibre, ou au cours de son établissement), l'énergie du système global varie de δE_T. Cette énergie est récupérée par le thermostat. Le rapport

$$Q_V \widehat{=} \frac{\delta E_T}{\delta n} \tag{A.20}$$

est appelé *chaleur d'adsorption à volume constant.* δE_T étant négative pour un δn positif, Q_V est négative : l'adsorption s'accompagne d'un dégagement de chaleur.

Dans le modèle simple que nous considérons ici,

$$\delta E_T = (-\varepsilon_0)\,\delta n - \frac{3}{2} kT \delta n , \tag{A.21}$$

et

$$Q_V = -\left(\varepsilon_0 + \frac{3}{2} kT\right) . \tag{A.22}$$

---

5. On peut exprimer $(dN_g/dT)_{eq}$ en fonction de p et T. Des deux formules

$$\frac{N_a}{A} = \frac{p}{p + p_0(T)} , \tag{1}$$

$$pV = N_g kT , \tag{2}$$

on tire

$$\frac{p_0 V}{kT} \widehat{=} Vf(T) = \frac{(A - N_a) N_g}{N_a} . \tag{3}$$

En différentiant les deux membres, et en utilisant (A.17), on obtient

$$\left(\frac{dN_g}{dT}\right)_{eq} = \frac{Vf'(T)}{\frac{A}{N_a}\left(1 + \frac{N_g}{N_a}\right) - 1} , \tag{4}$$

ou encore, en remplaçant N_a et N_g par leur expression tirée de (1) et (2),

$$\left(\frac{dN_g}{dT}\right)_{eq} = \frac{pVf'(T)}{(p + p_0)^2 \frac{V}{AkT} + p_0} . \tag{5}$$

6. C'est là un exemple du fameux « principe de modération » de Le Châtelier : le système réagit de façon à atténuer l'augmentation de température qu'on cherche à lui imposer.

II. GÉNÉRALISATION

Le modèle précédent se généralise aisément. Nous supposons encore que le récipient contient un gaz parfait, mais constitué cette fois de molécules possédant éventuellement plusieurs degrés de liberté internes (rotation, vibration, spin, ...) tels que nous les avons étudiés au complément III.B. La surface du substrat présente toujours A pièges susceptibles d'adsorber chacun *une seule molécule*[7], mais il existe pour chaque molécule adsorbée une série d'états (λ) possibles, d'énergie ε_λ.

1. ANALYSE GRAND-CANONIQUE

L'analyse grand-canonique du cas général suit la même démarche qu'au paragraphe I. On considère donc le système constitué par les particules adsorbées, en équilibre avec un réservoir de température T et de potentiel chimique μ.

Chaque piège peut toujours contenir zéro ou une particule (mais dans toute la série possible des états (λ)); la grande fonction de partition d'*un piège* vaut donc

$$\xi = 1 + \sum_{(\lambda)} e^{-\beta(\varepsilon_\lambda - \mu)} , \tag{A.23}$$

qu'on peut encore écrire

$$\xi = 1 + z_a(\beta)\, e^{\beta\mu} , \tag{A.24}$$

où $$z_a(\beta) \mathrel{\widehat{=}} \sum_{(\lambda)} e^{-\beta\varepsilon_\lambda} \tag{A.25}$$

est la fonction de partition (canonique) d'*une particule piégée.*

On en déduit immédiatement la grande fonction de partition du système des particules adsorbées :

$$\Xi = \xi^A = \left(1 + z_a(\beta)\, e^{\beta\mu}\right)^A , \tag{A.26}$$

le nombre moyen de molécules piégées :

$$N_a = kT \frac{\partial \ln \Xi}{\partial \mu} = A \frac{z_a(\beta)\, e^{\beta\mu}}{1 + z_a(\beta)\, e^{\beta\mu}} , \tag{A.27}$$

le taux d'adsorption :

$$\theta \mathrel{\widehat{=}} \frac{N_a}{A} = \frac{z_a(\beta)}{z_a(\beta) + e^{-\beta\mu}} , \tag{A.28}$$

et l'énergie moyenne des particules adsorbées :

$$E_a = \mu N_a - \frac{\partial \ln \Xi}{\partial \beta} = -A \frac{\dfrac{dz_a}{d\beta} e^{\beta\mu}}{1 + z_a(\beta)\, e^{\beta\mu}} , \tag{A.29}$$

qu'on peut, en faisant intervenir l'énergie moyenne d'une particule piégée

$$\bar{\varepsilon}_a = -\frac{d \ln z_a}{d\beta} , \tag{A.30}$$

mettre sous la forme

$$E_a = N_a \bar{\varepsilon}_a . \tag{A.31}$$

---

7. Nous supposons ici que la molécule piégée reste liée à un site particulier de la surface du substrat, c'est-à-dire que nous ne considérons que le phénomène de chimisorption (cf. note 1). Dans le cas de la physisorption, les molécules adsorbées peuvent se déplacer plus ou moins librement à la surface du solide; les propriétés thermodynamiques sont alors un peu différentes (voir exercice V.6).

Quant à l'entropie

$$S_a = kA \ln[1 + z_a \, e^{\beta\mu}] + \frac{N_a}{T}(\bar{\varepsilon}_a - \mu) , \tag{A.32}$$

on vérifie facilement, en tirant μ de (A.27), qu'elle peut s'écrire

$$S_a = k[A \ln A - N_a \ln N_a - (A - N_a) \ln(A - N_a)] + N_a s_a , \tag{A.33}$$

où $$s_a = k \frac{\mathrm{d}}{\mathrm{d}T}(T \ln z_a) = k \ln z_a + \frac{\bar{\varepsilon}_a}{T} \tag{A.34}$$

est l'entropie d'une molécule adsorbée sur un piège donné : à l'entropie « configurationnelle » (A.10) (cf. note 3) s'ajoute l'entropie propre à chaque molécule dans son piège.

2. PROPRIÉTÉS THERMODYNAMIQUES

Comme au paragraphe I, les propriétés thermodynamiques s'expriment en fonction de la température T et de la pression p du gaz en remplaçant μ, dans les formules précédentes, par le potentiel chimique du gaz. Pour un gaz parfait quelconque, μ est relié à p et T par (complément III.B, § I.3)

$$e^{\beta\mu} = \frac{p}{kT} \frac{V}{z_g} = \frac{p}{kT} \left(\frac{2\pi\hbar^2}{mkT}\right)^{3/2} \frac{1}{\zeta(T)} , \tag{A.35}$$

où z_g est la fonction de partition d'une molécule du gaz et $\zeta(T)$ sa « fonction de partition interne ». Celle-ci prend en compte les degrés de liberté de rotation, de vibration et de spin de la molécule; elle ne dépend que de la température.

a. Isothermes d'adsorption

Si on pose

$$p_0(T) \mathrel{\widehat{=}} \frac{kT}{V} \frac{z_g}{z_a} = kT \left(\frac{mkT}{2\pi\hbar^2}\right)^{3/2} \frac{\zeta(T)}{z_a(T)} , \tag{A.36}$$

la formule (A.28) s'écrit encore

$$\theta \mathrel{\widehat{=}} \frac{N_a}{A} = \frac{p}{p + p_0(T)} \tag{A.37}$$

comme au paragraphe I.3.a, avec simplement une expression de $p_0(T)$ plus générale. Les conclusions que nous avons tirées plus haut de la formule (A.13) restent donc valables ici[8].

b. Énergie et capacité calorifique

Chaque molécule du gaz possède en moyenne l'énergie

$$\bar{\varepsilon}_g = -\frac{\partial \ln z_g}{\partial\beta} = \frac{3}{2} kT - \frac{\partial \ln \zeta}{\partial\beta} . \tag{A.38}$$

8. Cela suppose évidemment que $p_0(T)$ reste une fonction croissante de la température comme au paragraphe I. C'est en général le cas, la dépendance en température du rapport $\zeta(T)/z_a(T)$ étant dominée par le facteur $\exp(-\varepsilon_0/kT)$ où ε_0 (>0) est l'énergie de liaison de la molécule sur le substrat.

Si N_g est le nombre moyen de molécules à l'état gazeux, l'énergie du *système global,* gaz + molécules adsorbées, s'écrit

$$E_T = \bar{\varepsilon}_g N_g + \bar{\varepsilon}_a N_a \,, \tag{A.39}$$

et sa capacité calorifique

$$C_V = \frac{\mathrm{d}E_T}{\mathrm{d}T} = \frac{\mathrm{d}\bar{\varepsilon}_g}{\mathrm{d}T} N_g + \frac{\mathrm{d}\bar{\varepsilon}_a}{\mathrm{d}T} N_a + (\bar{\varepsilon}_g - \bar{\varepsilon}_a)\left(\frac{\mathrm{d}N_g}{\mathrm{d}T}\right)_{\mathrm{eq}} . \tag{A.40}$$

A la capacité calorifique

$$C_V^{(g)} = N_g \frac{\mathrm{d}\bar{\varepsilon}_g}{\mathrm{d}T} \tag{A.41}$$

des N_g molécules de gaz, il faut ici ajouter celle

$$C^{(a)} = N_a \frac{\mathrm{d}\bar{\varepsilon}_a}{\mathrm{d}T} \tag{A.42}$$

des N_a molécules adsorbées, puis un terme correctif qui traduit, là encore, le fait qu'une augmentation de la température transfère un certain nombre de particules de la phase adsorbée à la phase gazeuse. En utilisant (A.30), (A.38) et (A.37) on peut exprimer C_V en fonction de p et T sous la forme[9]

$$C_V = C_V^{(g)} + C^{(a)} + kT\left[T\left(\frac{\partial \ln p}{\partial T}\right)_{N_a} - 1\right]\left(\frac{\mathrm{d}N_g}{\mathrm{d}T}\right)_{\mathrm{eq}} . \tag{A.43}$$

Si l'on connaît de façon précise les isothermes de Langmuir, une mesure de C_V permet donc de remonter à la capacité calorifique $C^{(a)}$ propre aux particules adsorbées[10].

---

9. De la définition (A.36) on déduit

$$\frac{\mathrm{d}\ln p_0}{\mathrm{d}T} = -\frac{1}{kT^2}\frac{\mathrm{d}\ln p_0}{\mathrm{d}\beta} = -\frac{1}{kT^2}\left[-kT + \frac{\mathrm{d}\ln z_g}{\mathrm{d}\beta} - \frac{\mathrm{d}\ln z_a}{\mathrm{d}\beta}\right], \tag{1}$$

c'est-à-dire, d'après (A.30) et (A.38),

$$\bar{\varepsilon}_g - \bar{\varepsilon}_a = kT^2 \frac{\mathrm{d}\ln p_0}{\mathrm{d}T} - kT \,. \tag{2}$$

D'autre part, on tire de (A.37)

$$p = \frac{p_0 N_a / A}{1 - N_a/A} \,, \tag{3}$$

d'où $$\left(\frac{\partial \ln p}{\partial T}\right)_{N_a} = \frac{\mathrm{d}\ln p_0}{\mathrm{d}T} \,. \tag{4}$$

En reportant (4) dans (2), on exprime la différence des énergies moyennes par particule sous la forme

$$\bar{\varepsilon}_g - \bar{\varepsilon}_a = kT\left[T\left(\frac{\partial \ln p}{\partial T}\right)_{N_a} - 1\right]. \tag{5}$$

10. La dérivée $(\mathrm{d}N_g/\mathrm{d}T)_{\mathrm{eq}}$ s'exprime aussi à partir des propriétés directement accessibles à l'expérience. Le calcul de la note 5 reste en effet valable, et on peut écrire son résultat en termes de la fonction $p(T, N_a)$: on vérifiera que la relation (5) de la note 5 se met sous la forme

$$\left(\frac{\mathrm{d}N_g}{\mathrm{d}T}\right)_{\mathrm{eq}} = \frac{pV}{kT^2}\left[T\left(\frac{\partial \ln p}{\partial T}\right)_{N_a} - 1\right] \times \left[1 + \frac{pV}{kT}\left(\frac{\partial \ln p}{\partial N_a}\right)_T\right]^{-1} .$$

c. Chaleur d'adsorption

La chaleur d'adsorption à volume constant (§ I.3.c) vaut ici

$$Q_V \widehat{=} \frac{\delta E_T}{\delta n} = \bar{\varepsilon}_a - \bar{\varepsilon}_g \,, \tag{A.44}$$

c'est-à-dire (cf. note 9)

$$Q_V = -kT\left[T\left(\frac{\partial \ln p}{\partial T}\right)_{N_a} - 1\right] . \tag{A.45}$$

En résumé, *toutes les propriétés macroscopiques du système sont connues dès que la fonction* $p_0(T)$ *l'est.* Celle-ci dépend évidemment du modèle microscopique qu'on se donne, tant pour les molécules du gaz (*via* la fonction de partition $\zeta(T)$) que pour la particule piégée sur un site de la surface du substrat (*via* la fonction de partition $z_a(T)$).

III. CAS OU L'ADSORPTION S'ACCOMPAGNE D'UNE DISSOCIATION DE LA MOLÉCULE

Il arrive parfois que les molécules du gaz se dissocient lorsqu'elles s'adsorbent, chaque atome constitutif allant se placer sur un piège différent. Pour simplifier, nous ne traiterons ici que le cas des *molécules diatomiques* formées de *deux atomes identiques;* plus explicitement, nous prendrons l'exemple de la molécule d'hydrogène H_2 qui s'adsorbe sur le platine ou le tungstène sous forme d'atomes H.

1. CONDITION D'ÉQUILIBRE

La différence essentielle avec les deux paragraphes précédents est que les particules de la phase gazeuse (molécules H_2) ne sont pas les mêmes que celles de la phase adsorbée (atomes H). On ne peut donc plus égaler le potentiel chimique des particules adsorbées à celui du gaz.

Pour trouver la nouvelle condition d'équilibre, il faut considérer le système global gaz + atomes adsorbés et se placer, soit dans l'ensemble microcanonique (système isolé) soit dans l'ensemble canonique (système en équilibre avec un thermostat). Choisissons par exemple la deuxième méthode[11].

Soit $F_g(T, N_g)$ l'énergie libre du gaz à la température T lorsqu'il contient N_g molécules H_2, et $F_a(T, N_a)$ celle de la phase adsorbée lorsqu'elle comporte N_a atomes piégés. L'énergie libre totale, pour N_g et N_a fixés, est la somme des énergies libres des deux phases :

$$F(T, N_g, N_a) = F_g(T, N_g) + F_a(T, N_a) \,. \tag{A.46}$$

Mais N_g et N_a sont des variables internes, libres de fluctuer, à condition de conserver le nombre total d'atomes d'hydrogène enfermés dans le récipient, c'est-à-dire de vérifier

$$2N_g + N_a = \text{constante} \,. \tag{A.47}$$

---

11. La méthode la plus rapide est une troisième : si on raisonne dans l'ensemble $T-p$ (complément III.D), la fonction à minimiser par rapport à N_g est l'enthalpie libre, qui s'écrit simplement (formule (D.26) du complément III.D)

$$G(T, p, N_g, N_a) = N_g \mu_{H_2}(T, p) + N_a \mu_H(T, p) \,.$$

Compte tenu de (A.47), (A.51) s'en déduit immédiatement.

A l'équilibre, $F(T, N_g, N_a)$ doit être minimum par rapport aux variations de N_g (voir III, § II.D.2); on a donc

$$\frac{\partial F}{\partial N_g} = \frac{\partial F_g}{\partial N_g} + \frac{\partial F_a}{\partial N_a}\frac{\mathrm{d}N_a}{\mathrm{d}N_g} = 0 \,. \tag{A.48}$$

Or
$$\frac{\partial F_g}{\partial N_g} = \mu_{\mathrm{H}_2} \,, \tag{A.49}$$

potentiel chimique des molécules d'hydrogène dans la phase gazeuse, et

$$\frac{\partial F_a}{\partial N_a} = \mu_{\mathrm{H}} \,, \tag{A.50}$$

potentiel chimique des atomes d'hydrogène adsorbés. Lorsqu'on tient compte de la contrainte (A.47), la condition d'équilibre (A.48) s'écrit donc

$$\mu_{\mathrm{H}_2} = 2\mu_{\mathrm{H}} \,. \tag{A.51}$$

2. PROPRIÉTÉS DU SYSTÈME

Si l'on suppose les atomes adsorbés en équilibre avec un réservoir à la température T et au potentiel chimique μ_{H}, les relations du paragraphe II.1 restent valables; en particulier, le taux d'adsorption a pour expression

$$\theta \mathrel{\widehat{=}} \frac{N_a}{A} = \frac{z_a(\beta)}{z_a(\beta) + \mathrm{e}^{-\beta\mu_{\mathrm{H}}}} \,. \tag{A.52}$$

D'autre part, le potentiel chimique des molécules du gaz est relié à la pression p par

$$\exp(\beta\mu_{\mathrm{H}_2}) = \frac{p}{kT}\left(\frac{2\pi\hbar^2}{m_{\mathrm{H}_2}kT}\right)^{3/2}\frac{1}{\zeta_{\mathrm{H}_2}(T)} \,. \tag{A.53}$$

La condition d'équilibre (A.51) donne donc

$$\theta \mathrel{\widehat{=}} \frac{N_a}{A} = \frac{p^{1/2}}{p^{1/2} + \tilde{p}_0(T)^{1/2}} \,, \tag{A.54}$$

où nous avons posé

$$\tilde{p}_0(T) \mathrel{\widehat{=}} kT\left(\frac{m_{\mathrm{H}_2}kT}{2\pi\hbar^2}\right)^{3/2}\frac{\zeta_{\mathrm{H}_2}(T)}{[z_a(T)]^2} \,. \tag{A.55}$$

La différence essentielle avec les cas précédents est donc le comportement en $p^{1/2}$ à basse pression ($p \ll \tilde{p}_0$); les isothermes ont l'allure indiquée sur la figure 3.

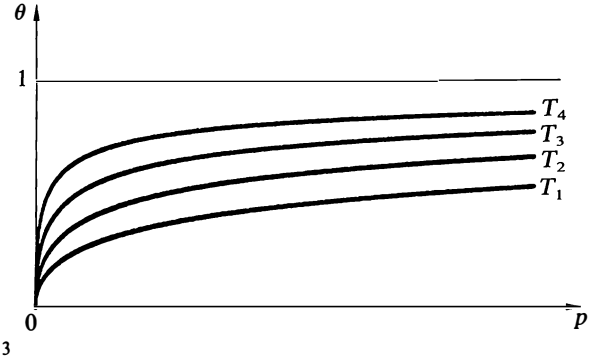


FIGURE 3

Isothermes d'adsorption de l'hydrogène atomique en fonction de la pression du gaz d'hydrogène moléculaire ($T_1 > T_2 > T_3 > T_4$).

Dans le cas particulier de l'adsorption d'atomes d'hydrogène, on peut expliciter la dépendance de $\tilde{p}_0$ avec T. Au voisinage de la température ordinaire et au-dessus, $\zeta_{H_2}(T)$ est donnée par la formule (B.69) du complément III.B :

$$\zeta_{H_2}(T)=\frac{2IkT\,e^{\varepsilon_L/kT}}{\hbar^2\,\mathrm{sh}\,\dfrac{\hbar\omega}{2kT}}\,, \tag{A.56}$$

où I est le moment d'inertie de la molécule H_2, ω sa pulsation propre de vibration, et ε_L son énergie de liaison.

Quant aux atomes adsorbés, ils peuvent vibrer à l'intérieur du « piège » où ils sont localisés. On peut supposer que les vibrations parallèles à la surface du substrat ont toutes la même pulsation propre $\omega_\parallel$, alors que, dans la direction perpendiculaire, les vibrations sont caractérisées par une pulsation $\omega_\perp$ différente. Les niveaux d'énergie sont alors de la forme

$$\varepsilon(n_1,\,n_2,\,n_3)=-\varepsilon_0+(n_1+n_2+1)\hbar\omega_\parallel+\left(n_3+\frac{1}{2}\right)\hbar\omega_\perp\,, \tag{A.57}$$

où n_1, n_2, n_3 peuvent prendre toutes les valeurs entières positives ou nulles; ε_0 (>0) est l'énergie de liaison de l'atome sur le piège (compte non tenu de l'énergie de vibration de point zéro). Enfin, le spin nucléaire de l'atome peut prendre deux valeurs[12]; il vient alors[13]

$$z_a(T)=\frac{e^{\varepsilon_0/kT}}{4\,\mathrm{sh}^2\,\dfrac{\hbar\omega_\parallel}{2kT}\,\mathrm{sh}\,\dfrac{\hbar\omega_\perp}{2kT}}\,. \tag{A.58}$$

Les expressions (A.56) et (A.58) permettent donc de préciser la dépendance de $\tilde{p}_0$ avec la température (formule (A.55)). On obtient (m_p étant la masse de l'atome d'hydrogène, $2m_p$ celle de la molécule)

$$\tilde{p}_0(T)=\frac{32m_P^{3/2}(kT)^{7/2}I}{\pi^{3/2}\hbar^5}\,\frac{\mathrm{sh}^4\,\dfrac{\hbar\omega_\parallel}{2kT}\,\mathrm{sh}^2\,\dfrac{\hbar\omega_\perp}{2kT}}{\mathrm{sh}\,\dfrac{\hbar\omega}{2kT}}\,\exp\frac{\varepsilon_L-2\varepsilon_0}{kT}\,. \tag{A.59}$$

---

12. L'atome d'hydrogène forme avec le substrat une liaison chimique; les spins des deux électrons mis en jeu dans cette liaison doivent être antiparallèles, comme dans la molécule H_2. Ceci fixe l'état de spin pour l'électron de l'atome H et ne laisse libre que celui du noyau.

13. Le calcul est connu (cf. complément III.B, § II.2) :

$$z_a\triangleq\sum_{(\lambda)}e^{-\beta\varepsilon_\lambda}=2\sum_{n_1=0}^{\infty}\sum_{n_2=0}^{\infty}\sum_{n_3=0}^{\infty}\exp-\beta\left[-\varepsilon_0+(n_1+n_2+1)\,\hbar\omega_\parallel+\left(n_3+\frac{1}{2}\right)\hbar\omega_\perp\right] \tag{1}$$

(le facteur 2 provient du spin nucléaire). On a donc

$$z_a=2\,e^{\beta\varepsilon_0}\left[\sum_{n=0}^{\infty}e^{-\beta(n+\frac{1}{2})\hbar\omega_\parallel}\right]^2\left[\sum_{n=0}^{\infty}e^{-\beta(n+\frac{1}{2})\hbar\omega_\perp}\right]. \tag{2}$$

Les sommes à effectuer correspondent à des séries géométriques. Par exemple,

$$\sum_{n=0}^{\infty}e^{-\beta(n+\frac{1}{2})\hbar\omega_\perp}=\frac{e^{-\beta\hbar\omega_\perp/2}}{1-e^{-\beta\hbar\omega_\perp}}=\frac{1}{2\,\mathrm{sh}\,\dfrac{\beta\hbar\omega_\perp}{2}}\,. \tag{3}$$

En reportant cette expression dans (2), on trouve finalement le résultat (A.58).

Pour la molécule H_2, $\hbar\omega/k \approx 6000$ K. Donc, pour des températures ne dépassant pas 1000 K, on a

$$\frac{\hbar\omega}{kT} \gg 1 \,. \tag{A.60}$$

Si l'on admet qu'une inégalité analogue est valable pour les fréquences de vibration des atomes adsorbés, l'expression (A.59) se simplifie quelque peu :

$$\tilde{p}_0(T) \simeq \frac{m_P^{3/2}(kT)^{7/2}I}{\pi^{3/2}\hbar^5} \exp -\frac{1}{kT}\left[2\varepsilon_0 - \varepsilon_L + \hbar\left(\frac{\omega}{2} - \omega_\perp - 2\omega_\parallel\right)\right] ,$$

$$\text{pour } \frac{kT}{\hbar} \ll \min(\omega, \omega_\perp, \omega_\parallel) \,. \tag{A.61}$$

La connaissance de la fonction $\tilde{p}_0(T)$ permet de déterminer complètement les propriétés macroscopiques du système constitué par le gaz H_2 en équilibre avec les atomes H adsorbés[14].

14. Là encore $\tilde{p}_0(T)$ est une fonction croissante de la température; le phénomène d'adsorption n'est en effet important que parce qu'il est énergétiquement favorable : l'énergie de deux atomes piégés est inférieure à celle d'une molécule :

$$2\varepsilon_0 - \varepsilon_L + \hbar\left(\frac{\omega}{2} - \omega_\perp - 2\omega_\parallel\right) > 0 \,.$$

COMPLÉMENT V.B

Équilibre d'un système macroscopique dans un champ de forces extérieur

I. Conditions d'équilibre en présence du champ extérieur
   1. Influence du champ extérieur sur les propriétés locales du système
   2. Expression locale de la condition d'équilibre global
   3. Interprétation canonique du potentiel chimique local

II. Exemples d'application
   1. Équilibre d'un fluide dans un champ extérieur
      a. Répartition de la pression dans le fluide
      b. Exemple d'un gaz parfait dans le champ de pesanteur
   2. Équilibre électrique d'un conducteur métallique
      a. Position du problème et méthode de solution
      b. Écrantage des charges extérieures
      c. Différence de potentiel de contact entre deux métaux

Nous considérons un système macroscopique constitué de particules identiques, soumises à une *force extérieure.* Nous entendons par là que ces particules interagissent avec des corps ne faisant pas partie du système : par exemple, les molécules d'un gaz subissent l'attraction de la Terre (force de pesanteur), ou les électrons d'un métal peuvent être plongés dans un champ électrique créé par des sources extérieures au métal.

Nous poserons les *hypothèses* suivantes :

(i) Le champ de forces extérieur est *indépendant du temps* et dérive d'une *énergie potentielle* $\Phi(\vec{r})$.

(ii) Sa présence ne modifie pas les interactions entre les particules du système[1].

(iii) *La fonction potentiel* $\Phi(\vec{r})$ *ne varie de façon appréciable qu'à l'échelle macroscopique,* c'est-à-dire sur des distances très supérieures à celles qui séparent en moyenne des particules voisines.

I. CONDITIONS D'ÉQUILIBRE EN PRÉSENCE DU CHAMP EXTÉRIEUR

1. INFLUENCE DU CHAMP EXTÉRIEUR SUR LES PROPRIÉTÉS LOCALES DU SYSTÈME

Découpons le système étudié en *sous-systèmes* $\delta\mathcal{S}$ *fixes,* ayant tous (pour simplifier) même volume δV et des dimensions petites à l'échelle macroscopique caractéristique des variations de Φ, mais grandes devant les distances intermoléculaires.

1. Cette hypothèse n'est pas vérifiée par exemple si l'on applique un champ électrique intense à un système constitué de molécules polarisables : le champ électrique induit dans chacune des molécules un moment dipolaire, ce qui affecte profondément leurs interactions.

Considérons alors un sous-système $\delta\mathcal{S}$ particulier, situé autour du point $\vec{r}$. Il contient en moyenne un grand nombre de particules, qui ont *toutes pratiquement la même énergie potentielle* $\Phi(\vec{r})$ dans le champ extérieur. Ce sous-système, de volume δV invariable, échange de la chaleur et des particules avec le reste du système, qui joue pour lui le rôle d'un réservoir-thermostat lui imposant la température T et le potentiel chimique μ qui règnent dans le système à l'équilibre (et que l'on peut supposer fixés par un réservoir-thermostat extérieur).

En l'absence de champ extérieur, l'hamiltonien H décrivant l'ensemble des particules contenues dans $\delta\mathcal{S}$ (dont le nombre fluctue) a pour états propres les états (ℓ), auxquels correspondent le nombre de particules N_ℓ et l'énergie E_ℓ. Une grandeur physique A quelconque du sous-système $\delta\mathcal{S}$ prend dans l'état (ℓ) la valeur

$$A_\ell = \langle \ell | A | \ell \rangle , \tag{B.1}$$

et sa valeur moyenne grand-canonique $\overline{A}^G$ se calcule par la formule

$$\overline{A}^G(\beta, \mu) = \sum_{(\ell)} A_\ell P_\ell^G(\beta, \mu) = \frac{1}{\Xi(\beta, \mu)} \sum_{(\ell)} A_\ell \, \mathrm{e}^{-\beta(E_\ell - \mu N_\ell)} , \tag{B.2}$$

où la grande fonction de partition $\Xi(\beta, \mu)$ du sous-système $\delta\mathcal{S}$ est par définition

$$\Xi(\beta, \mu) \widehat{=} \sum_{(\ell)} \mathrm{e}^{-\beta(E_\ell - \mu N_\ell)} . \tag{B.3}$$

En présence du champ extérieur, toutes les énergies des états du sous-système $\delta\mathcal{S}$ *sont décalées* d'une quantité égale au produit de l'énergie potentielle $\Phi(\vec{r})$ par le nombre de particules : l'état (ℓ), qui décrit toujours N_ℓ particules, a maintenant l'énergie[2]

$$E'_\ell = E_\ell + N_\ell \Phi(\vec{r}) . \tag{B.4}$$

Les autres propriétés de l'état (ℓ) *restent inchangées* : la valeur A_ℓ prise dans l'état (ℓ) par une grandeur A (autre que l'énergie) est la même qu'en l'absence de champ extérieur (formule (B.1)). Mais, en présence du champ extérieur, *la probabilité* P'^G_ℓ pour que le sous-système $\delta\mathcal{S}$ se trouve dans l'état (ℓ) *est changée :*

$$P'^G_\ell \propto \exp\left[-\beta(E'_\ell - \mu N_\ell)\right] \neq \exp\left[-\beta(E_\ell - \mu N_\ell)\right] . \tag{B.5}$$

Il en est donc *de même pour les propriétés statistiques du sous-système* $\delta\mathcal{S}$ *:*

$$\overline{A}'^G \neq \overline{A}^G . \tag{B.6}$$

En outre, *ces modifications dépendent du sous-système choisi,* puisque les énergies E'_ℓ (en présence du champ extérieur) dépendent de $\vec{r}$ selon la formule (B.4). Donc, *lorsqu'un système est à l'équilibre dans un champ de forces extérieur, ses propriétés statistiques varient d'un point à l'autre.* Cette variation est régie par les formules suivantes : d'après (B.4), la probabilité P'^G_ℓ pour que le sous-système $\delta\mathcal{S}$ situé au point $\vec{r}$ se trouve dans l'état microscopique (ℓ) est donnée par

$$P'^G_\ell(\beta, \mu; \vec{r}) = \frac{1}{\Xi'(\beta, \mu; \vec{r})} \exp\left\{-\beta\left[E_\ell + N_\ell \Phi(\vec{r}) - \mu N_\ell\right]\right\} , \tag{B.7}$$

où

$$\Xi'(\beta, \mu; \vec{r}) = \sum_{(\ell)} \mathrm{e}^{-\beta\left[E_\ell + N_\ell \Phi(\vec{r}) - \mu N_\ell\right]} \tag{B.8}$$

est la grande fonction de partition du sous-système $\delta\mathcal{S}$ en présence du champ extérieur.

---

2. L'hamiltonien H' décrivant (en présence du champ extérieur) les particules contenues dans $\delta\mathcal{S}$ s'obtient très simplement à partir de H :

$$H' = H + \Phi(\vec{r}) \widehat{N} ,$$

où $\widehat{N}$ est l'opérateur nombre de particules. Comme les états (ℓ) sont états propres communs à H et $\widehat{N}$ (avec les valeurs propres respectives E_ℓ et N_ℓ), *ils restent états propres du nouvel hamiltonien* H', avec les valeurs propres E'_ℓ données par (B.4).

La valeur moyenne d'une grandeur physique A quelconque s'en déduit aussitôt :

$$\overline{A}'^{G}(\beta, \mu; \vec{r}) = \sum_{(\ell)} A_\ell P'^{G}_\ell(\beta, \mu; \vec{r}) . \tag{B.9}$$

Dans les formules (B.7)-(B.9), les quantités pour lesquelles la dépendance en $\vec{r}$ a été indiquée sont les seules à varier d'un point à l'autre; en particulier, si le système dans son ensemble est à l'équilibre, *β et μ sont indépendants de $\vec{r}$* même lorsque le champ extérieur est présent, puisque les échanges d'énergie et de particules entre les divers sous-systèmes égalisent leurs températures et leurs potentiels chimiques.

2. EXPRESSION LOCALE DE LA CONDITION D'ÉQUILIBRE GLOBAL

Comparons les formules (B.7)-(B.9) aux expressions grand-canoniques habituelles, en l'absence de champ extérieur, rappelées en (B.2)-(B.3).

La grande fonction de partition du sous-système situé en $\vec{r}$, en présence du champ extérieur, peut être écrite sous la forme

$$\Xi'(\beta, \mu; \vec{r}) = \sum_{(\ell)} \mathrm{e}^{-\beta[E_\ell - [\mu - \Phi(\vec{r})]N_\ell]} = \Xi(\beta, \tilde{\mu}(\vec{r})) , \tag{B.10}$$

où $\Xi(\beta, \mu)$ est la grande fonction de partition en l'absence de champ extérieur et

$$\tilde{\mu}(\vec{r}) \widehat{=} \mu - \Phi(\vec{r}) . \tag{B.11}$$

La probabilité P'^{G}_ℓ en présence du champ extérieur s'exprime de façon analogue à partir de la probabilité $P^{G}_\ell(\beta, \mu)$ en l'absence de champ extérieur :

$$P'^{G}_\ell(\beta, \mu; \vec{r}) = P^{G}_\ell(\beta, \tilde{\mu}(\vec{r})) . \tag{B.12}$$

Comme les valeurs microscopiques A_ℓ d'une grandeur quelconque A (autre que l'énergie) ne sont pas modifiées par le champ extérieur, on peut écrire de même (voir (B.9) et (B.2))

$$\overline{A}'^{G}(\beta, \mu; \vec{r}) = \overline{A}^{G}(\beta, \tilde{\mu}(\vec{r})) . \tag{B.13}$$

En particulier, le nombre moyen de particules des sous-systèmes $\delta\mathcal{S}$ (ou leur densité puisqu'ils ont tous même volume δV) varie avec $\vec{r}$ selon la formule[3]

$$\overline{N}'^{G}(\beta, \mu; \vec{r}) = \frac{1}{\beta} \frac{\partial}{\partial \mu} \ln \Xi'(\beta, \mu; \vec{r}) = \frac{1}{\beta} \frac{\partial}{\partial \tilde{\mu}} \ln \Xi(\beta, \tilde{\mu}) , \tag{B.14}$$

dans laquelle on remplace ensuite $\tilde{\mu}$ par $\mu - \Phi(\vec{r})$.

Bien entendu, *l'énergie fait exception* à la règle qui se dégage des expressions (B.10)-(B.14), car ses valeurs possibles $E'_\ell(\vec{r})$ dans le champ extérieur sont différentes des valeurs E_ℓ sans champ extérieur. Cependant, la modification est très simple : on

---

3. Selon sa définition, le nombre moyen de particules du sous-système $\delta\mathcal{S}$ situé au point $\vec{r}$ est donné par

$$\overline{N}'^{G}(\beta, \mu; \vec{r}) = \sum_{(\ell)} N_\ell P'^{G}_\ell(\beta, \mu; \vec{r}) . \tag{1}$$

Il vaut d'après (B.12)

$$\overline{N}'^{G}(\beta, \mu; \vec{r}) = \sum_{(\ell)} N_\ell P^{G}_\ell(\beta, \tilde{\mu}(\vec{r})) , \tag{2}$$

que l'on peut écrire, compte tenu de (B.10), sous la forme (B.14).

ajoute $\Phi(\vec{r})$ par particule. La valeur moyenne de l'énergie du sous-système $\delta\mathcal{S}$ dans le champ extérieur est effectivement donnée par[4]

$$\overline{E}'^{G}(\beta, \mu; \vec{r}) = \overline{E}^{G}(\beta, \tilde{\mu}(\vec{r})) + \Phi(\vec{r})\,\overline{N}^{G}(\beta, \tilde{\mu}(\vec{r}))\,. \tag{B.15}$$

Nous venons de montrer que les *propriétés locales du système en présence du champ extérieur* peuvent être déduites des *formules grand-canoniques sans champ extérieur* dans lesquelles on a remplacé le paramètre μ par le *«potentiel chimique local»* $\tilde{\mu}(\vec{r})$ défini par (B.11). Mais l'équilibre du système dans son ensemble reste caractérisé par *l'égalité de la température T et du potentiel chimique* μ *entre les divers sous-systèmes* $\delta\mathcal{S}$, qui échangent de la chaleur et des particules; le «potentiel chimique local» $\tilde{\mu}(\vec{r})$ doit donc vérifier la *condition d'équilibre*

$$\boxed{\tilde{\mu}(\vec{r}) + \Phi(\vec{r}) = \mu = \text{c}^{\text{te}}}\,, \tag{B.16}$$

complétée par celle qui concerne la température :

$$T = \text{c}^{\text{te}} \tag{B.17}$$

(dans ces deux égalités, les «constantes» sont indépendantes de $\vec{r}$).

3. INTERPRÉTATION CANONIQUE DU POTENTIEL CHIMIQUE LOCAL

Nous avons jusqu'ici raisonné exclusivement dans le cadre grand-canonique : le système considéré était en équilibre avec un réservoir-thermostat qui lui imposait sa température T et son potentiel chimique μ, qu'il imposait à son tour à chacun de ses sous-systèmes $\delta\mathcal{S}$. Nous allons montrer maintenant que le «potentiel chimique local» $\tilde{\mu}(\vec{r})$, introduit de façon formelle au paragraphe 2, acquiert dans le cadre *canonique* une interprétation simple.

Prenons pour cela un sous-système $\delta\mathcal{S}$ déterminé, repéré par le vecteur $\vec{r}$, et imaginons (même si ce n'est pas réaliste) qu'il se trouve dans la situation canonique : il est maintenu à la température T, mais *son nombre de particules est ici fixé* à N. Calculons alors sa fonction de partition $Z'(T, \delta V, N, \ldots)$ *en présence du champ extérieur* :

$$Z'(T, \delta V, N, \ldots) = \sum_{(\ell)\,:\,N_\ell = N} \text{e}^{-\beta E'_\ell}\,. \tag{B.18}$$

---

4. La valeur moyenne de l'énergie en présence du champ extérieur se déduit de la grande fonction de partition correspondante $\Xi'(\beta, \mu; \vec{r})$ par la formule habituelle :

$$\overline{E}'^{G}(\beta, \mu; \vec{r}) = -\frac{\partial}{\partial\beta} \ln \Xi' + \mu \overline{N}'^{G}\,. \tag{1}$$

Il suffit de remplacer Ξ' et $\overline{N}'^{G}$ par leurs expressions (B.10) et (B.14) pour aboutir à

$$\overline{E}'^{G}(\beta, \mu; \vec{r}) = -\frac{\partial}{\partial\beta} \ln \Xi(\beta, \tilde{\mu}(\vec{r})) + \mu \overline{N}^{G}(\beta, \tilde{\mu}(\vec{r}))\,. \tag{2}$$

Attention : dans cette formule, c'est le potentiel chimique local $\tilde{\mu}(\vec{r})$ qui apparaît dans Ξ et $\overline{N}^{G}$, mais c'est le véritable potentiel chimique μ qui multiplie $\overline{N}^{G}$ dans le second terme. Compte tenu de la relation (B.11) entre $\tilde{\mu}$ et μ, cette formule donne bien (B.15).

On trouve facilement[5]

$$Z'(T, \delta V, N, \ldots) = e^{-\beta N \Phi(\vec{r})} Z(T, \delta V, N, \ldots) , \tag{B.19}$$

où Z est la fonction de partition en l'absence du champ extérieur. On en déduit[6] le *potentiel chimique canonique* μ'^c *en présence du champ extérieur* en fonction du potentiel chimique canonique μ^c sans champ extérieur et de l'énergie potentielle $\Phi(\vec{r})$ de chaque particule :

$$\mu'^c(T, \delta V, N, \ldots) = \Phi(\vec{r}) + \mu^c(T, \delta V, N, \ldots) . \tag{B.20}$$

Le sous-système $\delta\mathcal{S}$ ne se trouve pas en réalité dans la situation canonique, mais dans la situation grand-canonique : le reste du système joue pour lui le rôle d'un réservoir, qui lui impose non seulement sa température T mais son potentiel chimique μ. Le nombre N de particules du sous-système $\delta\mathcal{S}$ fluctue donc. Sa *valeur la plus probable* $N'(\vec{r})$, *en présence du champ extérieur,* est donnée par (cf. V, § I.C.1.b)

$$\mu'^c(T, \delta V, N'(\vec{r}), \ldots) = \mu . \tag{B.21}$$

Si l'on rapproche cette égalité, qui détermine $N'(\vec{r})$, de la condition d'équilibre (B.16) et que l'on tient compte de l'expression (B.20) de μ'^c, on constate que

$$\tilde{\mu}(\vec{r}) = \mu^c(T, \delta V, N'(\vec{r}), \ldots) : \tag{B.22}$$

le «*potentiel chimique local*» $\tilde{\mu}(\vec{r})$ n'est autre que le *potentiel chimique canonique* qu'aurait le sous-système $\delta\mathcal{S}$ (situé au point $\vec{r}$) *en l'absence de champ extérieur, mais pour un nombre de particules fixé à sa valeur la plus probable* $N'(\vec{r})$ *en présence du champ.* Cette valeur $N'(\vec{r})$ est déterminée par la condition d'équilibre (B.16), qui s'écrit en utilisant (B.22) sous la forme

$$\mu^c(T, \delta V, N'(\vec{r}), \ldots) + \Phi(\vec{r}) = \mu = c^{te} . \tag{B.23}$$

Si la taille du sous-système $\delta\mathcal{S}$ est suffisante pour que les fluctuations du nombre de particules qu'il contient soient négligeables (*limite thermodynamique*, voir V, § III), on peut raisonner comme si ce nombre de particules était effectivement fixé à $N'(\vec{r})$. En outre, le potentiel chimique canonique devient une grandeur thermodynamique, que l'on peut considérer comme fonction de n'importe quel autre ensemble de variables caractérisant l'état macroscopique du système.

Le point de vue présenté dans ce paragraphe 3 permet une compréhension intuitive simple du comportement d'un système soumis à un potentiel extérieur. Prenons en effet deux sous-systèmes contigus $\delta\mathcal{S}_1$ et $\delta\mathcal{S}_2$ et supposons par exemple que l'énergie potentielle Φ du champ extérieur décroisse de $\delta\mathcal{S}_1$ à $\delta\mathcal{S}_2$:

$$\Phi(\vec{r}_1) > \Phi(\vec{r}_2) . \tag{B.24}$$

---

5. La fonction de partition Z' s'écrit en effet d'après (B.4)

$$Z' = \sum_{(\ell) : N_\ell = N} e^{-\beta[E_\ell + N_\ell \Phi(\vec{r})]} . \tag{1}$$

Comme la somme ne porte que sur les états (ℓ) dont le nombre N_ℓ de particules est égal à N, Z' se factorise :

$$Z' = e^{-\beta N \Phi(\vec{r})} \sum_{(\ell) : N_\ell = N} e^{-\beta E_\ell} . \tag{2}$$

La somme qui reste à effectuer n'est autre que la fonction de partition Z en l'absence du champ extérieur.

6. La formule (B.19) donne aussitôt l'énergie libre F' en fonction de F et $\Phi(\vec{r})$:

$$F' = -kT \ln Z' = N\Phi(\vec{r}) + F . \tag{1}$$

Le potentiel chimique canonique s'obtient en dérivant par rapport au nombre N de particules :

$$\mu'^c = \frac{\partial F'}{\partial N} = \Phi(\vec{r}) + \mu^c . \tag{2}$$

La force qu'exerce ce champ sur chacune des particules (égale à l'opposé du gradient de Φ) les pousse alors de $\delta\mathcal{S}_1$ vers $\delta\mathcal{S}_2$. Mais la condition d'équilibre (B.16) implique que, dans ce cas, le potentiel chimique local croît de $\delta\mathcal{S}_1$ à $\delta\mathcal{S}_2$:

$$\tilde{\mu}(\vec{r}_1) < \tilde{\mu}(\vec{r}_2) . \tag{B.25}$$

Sous l'influence de cette différence entre les potentiels chimiques locaux, les particules ont tendance, du point de vue statistique, à passer au contraire de $\delta\mathcal{S}_2$ à $\delta\mathcal{S}_1$. L'équilibre résulte alors de la compensation entre ces deux effets opposés.

II. EXEMPLES D'APPLICATION

1. ÉQUILIBRE D'UN FLUIDE DANS UN CHAMP EXTÉRIEUR

Appliquons les résultats du paragraphe I au système le plus simple possible : un gaz ou un liquide pur. Les variables d'état macroscopiques sont dans ce cas au nombre de trois seulement. Nous supposons que les distances caractéristiques des variations du potentiel extérieur $\Phi(\vec{r})$ permettent de choisir des sous-systèmes $\delta\mathcal{S}$ dans lesquels les fluctuations sont négligeables (limite thermodynamique).

a. Répartition de la pression dans le fluide

Il est commode de choisir pour variables la température T, la pression p et le nombre N de particules. En effet, avec ce jeu de variables, le potentiel chimique thermodynamique du système, grandeur intensive, est indépendant de N, seule variable extensive (cf. complément III.D, § I.3.c); en l'absence de champ, c'est une fonction de T et p seulement, que nous noterons $\mu_0(T, p)$.

Si le fluide est soumis à un potentiel extérieur $\Phi(\vec{r})$, la température T reste indépendante de $\vec{r}$ à l'équilibre (condition (B.17)); par contre, le « potentiel chimique local » $\tilde{\mu}$ doit dépendre de $\vec{r}$ si le potentiel Φ n'est pas uniforme; comme $\tilde{\mu}(\vec{r})$ s'identifie au potentiel chimique thermodynamique en l'absence de champ (§ I.3), *la pression p varie d'un point à l'autre :*

$$\tilde{\mu}(\vec{r}) = \mu_0(T, p(\vec{r})) . \tag{B.26}$$

La loi de répartition de la pression $p(\vec{r})$ dans le fluide est alors donnée par la condition d'équilibre (B.16) :

$$\mu_0(T, p(\vec{r})) + \Phi(\vec{r}) = c^{te} . \tag{B.27}$$

Cette équation permet de déterminer la dépendance de la pression p par rapport à $\vec{r}$ si l'on connaît la fonction $\mu_0(T, p)$ et le potentiel extérieur $\Phi(\vec{r})$. On peut ensuite en déduire, grâce à l'équation d'état du fluide, la densité locale n en fonction de $\vec{r}$[7].

---

7. Il n'est en réalité pas nécessaire de connaître explicitement la fonction $\mu_0(T, p)$; l'équation d'état suffit. On sait en effet (complément III.F, § I.1) que

$$\left(\frac{\partial \mu_0}{\partial p}\right)_T = \frac{V}{N} = \frac{1}{n} . \tag{1}$$

On peut dès lors expliciter la condition d'équilibre (B.27) en prenant le gradient des deux membres, ce qui donne l'équation

$$\frac{1}{n(\vec{r})} \vec{\nabla} p(\vec{r}) + \vec{\nabla}\Phi(\vec{r}) = \vec{0} , \tag{2}$$

où $n(\vec{r})$ et $p(\vec{r})$ sont reliés par l'équation d'état (voir l'exemple de la note 9).

Notons que l'équation (2) n'est rien d'autre que la condition d'équilibre mécanique du fluide en présence du champ extérieur : un volume δV du fluide, situé au point $\vec{r}$, contient $n(\vec{r}).\delta V$ molécules; il est en équilibre sous l'action simultanée de la force de pression $-\delta V \times \vec{\nabla} p(\vec{r})$ et de la force extérieure $n(\vec{r})\delta V \times [-\vec{\nabla}\Phi(\vec{r})]$.

b. Exemple d'un gaz parfait dans le champ de pesanteur

Pour un gaz parfait (III, § IV.B.3.c et complément III.B, § I.3), le potentiel chimique μ_0 est tel que

$$e^{\mu_0/kT} = p\alpha(T) . \tag{B.28}$$

D'autre part, dans le champ de pesanteur (d'accélération g), une molécule de masse m a pour énergie potentielle

$$\Phi(x, y, z) = mgz . \tag{B.29}$$

La condition d'équilibre (B.27) donne alors l'équation différentielle[8], [9]

$$kT \frac{1}{p} \frac{dp}{dz} + mg = 0 , \tag{B.30}$$

dont la solution est immédiate :

$$p(z) = p(z_0)\, e^{-mg(z-z_0)/kT} . \tag{B.31}$$

La densité n, que l'on déduit de l'équation d'état

$$p = nkT , \tag{B.32}$$

varie de la même façon que p :

$$n(z) = n(z_0)\, e^{-mg(z-z_0)/kT} . \tag{B.33}$$

Le résultat (B.31) ou (B.33) (que l'on peut obtenir par d'autres méthodes) est connu sous le nom de *« formule du nivellement barométrique »* : la pression et la densité du gaz décroissent exponentiellement lorsque l'altitude croît. L'ordre de grandeur de cet effet est cependant très faible dans des récipients de dimensions raisonnables : pour l'air à température ordinaire, la diminution relative de pression ou de densité dépasse à peine un pour mille pour une élévation de 10 mètres[10].

2. ÉQUILIBRE ÉLECTRIQUE D'UN CONDUCTEUR MÉTALLIQUE

a. Position du problème et méthode de solution

La structure d'un métal peut en première approximation être schématisée comme suit (complément VI.A, § II.1) : un réseau d'ions positifs fixes (constitués par les noyaux atomiques entourés des électrons des couches profondes) est baigné par un

---

8. D'après (B.28) et (B.29), la condition d'équilibre (B.27) s'écrit ici

$$kT\left[\ln p\left(\vec{r}\right) + \ln \alpha(T)\right] + mgz = c^{te} .$$

Il suffit de dériver par rapport à z pour obtenir l'équation (B.30). Noter que ceci élimine la fonction $\alpha(T)$, qui dépend du gaz considéré, et donne une loi identique pour tous les gaz parfaits (cf. notes 7 et 9).

9. L'équation différentielle (B.30) peut être déduite directement de l'équation d'état (B.32) des gaz parfaits, comme l'indique la note 7. La formule (2) de cette note s'écrit en effet ici

$$\frac{1}{n(z)} \frac{dp}{dz} + mg = 0 ,$$

d'où découle aussitôt (B.30) si l'on remplace n en fonction de T et p selon (B.32).

10. Sur des différences d'altitude nettement supérieures, l'atmosphère terrestre n'est pas à l'équilibre thermique (dans la basse atmosphère, la température décroît d'environ 6,5° par km d'altitude); on ne peut donc pas lui appliquer les résultats obtenus ici.

Toutefois la formule (B.30), qui peut être déduite de la condition d'équilibre mécanique (notes 7 et 9), reste vraie, même si T dépend de z. Pour intégrer cette équation différentielle, il faut alors connaître T en fonction de z, ou la loi qui régit la dépendance de la température par rapport à l'altitude.

gaz d'électrons de conduction pratiquement libres et indépendants les uns des autres; ce gaz est fortement «dégénéré» (VI, § I.B.3), c'est-à-dire que les effets quantiques y sont très importants. En l'absence de champ extérieur, la densité volumique de charge est partout nulle à l'intérieur du métal : si d^3r est un volume infiniment petit à l'échelle macroscopique mais grand à l'échelle microscopique, les $n d^3r$ électrons de conduction présents dans ce volume compensent constamment (à des fluctuations négligeables près) la charge des ions qu'il contient. Le nombre n d'électrons par unité de volume est le même en tous les points du métal (sauf peut-être au voisinage immédiat de sa surface), et le champ électrique y est partout nul (à l'échelle macroscopique).

Supposons que l'on impose au métal un *champ électrique extérieur* indépendant du temps : ce champ est créé par des charges (autres que celles du métal) dont nous noterons $\rho_{ex}(\vec{r})$ la densité volumique; il dérive du potentiel électrostatique $\varphi_{ex}(\vec{r})$. Sous son influence, *les électrons de conduction du métal vont se redistribuer*; il apparaîtra ainsi une densité volumique de *charge induite* $\rho_{in}(\vec{r})$ qui va elle aussi créer un champ électrique, dérivant du potentiel $\varphi_{in}(\vec{r})$. Finalement, le potentiel électrostatique $\varphi(\vec{r})$ régnant, à l'équilibre, en chaque point de l'espace va être la somme

$$\varphi(\vec{r}) = \varphi_{ex}(\vec{r}) + \varphi_{in}(\vec{r}) \tag{B.34}$$

des deux contributions précédentes.

Si la densité $\rho_{ex}(\vec{r})$ des charges extérieures est donnée, on sait calculer le potentiel φ_{ex} qu'elles créent :

$$\varphi_{ex}(\vec{r}) = \frac{1}{4\pi\varepsilon_0} \int d^3r' \, \frac{\rho_{ex}(\vec{r}')}{|\vec{r}' - \vec{r}|} \, . \tag{B.35}$$

A partir de là, la détermination de ρ_{in} et φ_{in} pose un *«problème d'autocohérence»*. D'une part, en effet, chaque particule ressent non seulement φ_{ex} mais aussi φ_{in}; c'est donc le *potentiel total* φ qui doit figurer dans la condition d'équilibre (B.16) :

$$\tilde{\mu}(\vec{r}) - e\varphi(\vec{r}) = c^{te} \, , \tag{B.36}$$

où $-e$ est la charge (négative) de l'électron. Or le potentiel chimique local $\tilde{\mu}$ du gaz d'électrons en présence du champ extérieur dépend du point $\vec{r}$ par l'intermédiaire du nombre $n'(\vec{r})$ d'électrons par unité de volume (§ I.3) :

$$\tilde{\mu}(\vec{r}) = \mu_0(T, n'(\vec{r})) \, , \tag{B.37}$$

où $\mu_0(T, n)$ est le potentiel chimique du gaz en l'absence de champ extérieur[11]. La condition (B.36) relie donc $n'(\vec{r})$ et $\varphi(\vec{r})$. Mais d'autre part l'excès (algébrique) d'électrons au point $\vec{r}$ donne directement la *densité de charge induite :*

$$\rho_{in}(\vec{r}) = -e[n'(\vec{r}) - n] \, , \tag{B.38}$$

et $\varphi_{in}(\vec{r})$ s'obtient à partir de ρ_{in} par une relation analogue à (B.35). On a donc à résoudre un *système d'équations couplées* : le potentiel électrostatique $\varphi(\vec{r})$ doit vérifier à la fois l'équation de Poisson

$$\Delta\varphi + \frac{1}{\varepsilon_0}(\rho_{in} + \rho_{ex}) = 0 \, , \tag{B.39}$$

où ρ_{ex} est connue et ρ_{in} est donnée en fonction de $n'(\vec{r})$ par (B.38), et la condition d'équilibre

$$\mu_0(T, n'(\vec{r})) - e\varphi(\vec{r}) = c^{te} \, . \tag{B.40}$$

---

11. Nous choisissons ici les variables canoniques T, V et N. Le potentiel chimique μ_0, grandeur intensive, ne peut dépendre des variables extensives N et V qu'à travers leur rapport $n = N/V$:

$$\mu_0(T, V, N) = \mu_0(T, n)$$

(le raisonnement est analogue à celui du complément III.F, note 10).

En pratique, la différence $n'(\vec{r})-n$ est suffisamment petite devant n pour qu'on puisse la traiter au premier ordre (voir plus loin la note 19) :

$$\mu_0(T, n'(\vec{r})) \simeq \mu_0(T, n)+[n'(\vec{r})-n]\frac{\partial \mu_0}{\partial n}(T, n) . \tag{B.41}$$

L'équation (B.40) implique dans ce cas, compte tenu de (B.38), que le potentiel total φ est simplement linéaire par rapport à ρ_{in} :

$$\varphi(\vec{r})=\varphi_0-\frac{1}{e^2}\frac{\partial \mu_0}{\partial n}\rho_{in}(\vec{r}) , \tag{B.42}$$

où φ_0 est indépendant de $\vec{r}$. L'équation de Poisson (B.39) permet de déterminer ρ_{in} à partir de ρ_{ex} (et des conditions aux limites qui restent à préciser) :

$$\Delta\rho_{in}-k_0^2(\rho_{in}+\rho_{ex})=0 , \tag{B.43}$$

où nous avons posé[12]

$$k_0 \widehat{=} \sqrt{\frac{e^2}{\varepsilon_0 \partial\mu_0/\partial n}} . \tag{B.44}$$

b. Écrantage des charges extérieures

Considérons le cas le plus simple, celui où le champ extérieur est créé par une *charge ponctuelle* e_0. Supposons d'abord que cette charge est située *à l'intérieur du métal*[13], en un point que nous prenons pour origine O des coordonnées. On a alors

$$\rho_{ex}(\vec{r})=e_0\delta(\vec{r}) , \tag{B.45}$$

où δ est la «fonction» de Dirac à trois dimensions, de sorte que l'équation (B.43) donnant ρ_{in} s'écrit

$$(\Delta-k_0^2)\rho_{in}(\vec{r})=k_0^2 e_0\delta(\vec{r}) . \tag{B.46}$$

Nous supposons le point O suffisamment éloigné de la surface du métal pour que la solution que nous recherchons présente la symétrie sphérique autour de O : la densité de charge induite ρ_{in} dépend seulement de la distance r au point O. On trouve alors[14]

$$\rho_{in}(r)=-\frac{k_0^2 e_0}{4\pi}\frac{1}{r}e^{-k_0 r} . \tag{B.47}$$

---

12. On sait de façon générale (III, § II.D.3.b) que la stabilité du système exige que la dérivée $\partial\mu_0/\partial n$ soit positive.

13. Dans la réalité, une telle situation se présente lorsqu'un ion du métal est remplacé par un atome d'impureté, de charge différente.

14. La fonction $\rho_{in}(r)$ doit d'abord être telle que

$$(\Delta-k_0^2)\rho_{in}(r)=\frac{1}{r}\frac{d^2}{dr^2}[r\rho_{in}(r)]-k_0^2\rho_{in}(r)=0 \text{ pour tout } r\neq 0, \tag{1}$$

ce qui lui donne la forme

$$\rho_{in}(r)=C\frac{1}{r}e^{-k_0 r} \tag{2}$$

(la solution en $\exp(+k_0 r)$ est rejetée comme non physique, car elle donnerait une charge induite croissant indéfiniment avec la distance au point 0). La constante C est déterminée par la condition

$$(\Delta-k_0^2)\rho_{in}(r) \underset{r\to 0}{\sim} k_0^2 e_0\delta(\vec{r}) . \tag{3}$$

.../...

On en déduit aussitôt, grâce à (B.42) et (B.44), le potentiel électrostatique total (nous choisissons la constante de façon qu'il s'annule à l'infini) :

$$\varphi(r)=\frac{e_0}{4\pi\varepsilon_0}\frac{1}{r}\,e^{-k_0 r}\,. \tag{B.48}$$

Cette expression est celle d'un *«potentiel écranté»*[15] : $\varphi(r)$ décroît beaucoup plus vite (en module), quand r croît, que le potentiel $\varphi_{ex}(r)$ de la charge ponctuelle e_0, donné par

$$\varphi_{ex}(r)=\frac{e_0}{4\pi\varepsilon_0}\frac{1}{r}\,. \tag{B.49}$$

Mais la forme (2) de ρ_{in} implique

$$\Delta\rho_{in}=C\left\{e^{-k_0 r}\Delta\left(\frac{1}{r}\right)+2\,\vec{\nabla}\left(\frac{1}{r}\right).\vec{\nabla}\,e^{-k_0 r}+\frac{1}{r}\,\Delta\left(e^{-k_0 r}\right)\right\}$$

$$=C\left\{e^{-k_0 r}\Delta\left(\frac{1}{r}\right)+k_0^2\frac{1}{r}\,e^{-k_0 r}\right\}, \tag{4}$$

de sorte que la condition (3) s'écrit

$$\Delta\rho_{in}-k_0^2\rho_{in}\underset{r\to 0}{\sim}C\Delta\left(\frac{1}{r}\right)=k_0^2 e_0\delta(\vec{r})\,. \tag{5}$$

Or on sait d'après l'expression du potentiel créé par une charge ponctuelle que

$$\Delta\left(\frac{1}{4\pi r}\right)=-\delta(\vec{r})\,. \tag{6}$$

On en déduit la valeur de C :

$$C=-\frac{k_0^2 e_0}{4\pi}\,. \tag{7}$$

15. Un résultat analogue a été obtenu au complément III.H, paragraphe II.2. Pour préciser cette analogie, calculons la dérivée $\partial\mu_0/\partial n$ pour un gaz parfait classique. On a dans ce cas (cf. formule (B.28))

$$e^{\mu_0/kT}=p\alpha(T)=nkT\alpha(T)\,, \tag{1}$$

ce qui donne aussitôt

$$\frac{\partial\mu_0}{\partial n}=\frac{kT}{n}\,. \tag{2}$$

Si l'on remplace dans (B.44) $\partial\mu_0/\partial n$ par cette expression, on obtient

$$k_0=\left[\frac{ne^2}{\varepsilon_0 kT}\right]^{1/2}\,. \tag{3}$$

Ce résultat peut effectivement être comparé à la formule (H.13) du complément III.H donnant la *longueur de Debye* dans un électrolyte :

$$\ell_D=\left[\frac{\varepsilon VkT}{\sum_{(a)}q_a^2 N_a}\right]^{1/2}\,. \tag{4}$$

Il y a ici une seule espèce de charges mobiles, les électrons, de sorte que $\sum_{(a)}q_a^2 N_a/V$ doit être identifié à $e^2N/V=ne^2$, et la permittivité ε de l'électrolyte doit être remplacée par celle du vide ε_0; ℓ_D et $1/k_0$ coïncident alors.

Mais, pour les électrons de conduction d'un métal, les effets quantiques sont prépondérants et le potentiel chimique μ_0 n'est pas celui d'un gaz parfait classique : il est déterminé par le principe de Pauli (voir note 16). La longueur $1/k_0$ correspondante est appelée *«longueur de Thomas-Fermi»*.

C'est évidemment la charge induite dans le métal qui produit cette modification importante : suivant son signe, la charge e_0 attire ou repousse les électrons de conduction, de sorte qu'elle se trouve entourée par une charge de signe opposé au sien qui atténue ses effets à grande distance. L'échelle de longueur caractéristique de cet écrantage est donnée par

$$r_0 = \frac{1}{k_0} : \tag{B.50}$$

aux points dont la distance r à la charge ponctuelle est grande devant r_0, le métal conserve les propriétés qu'il avait en l'absence de champ extérieur : densité volumique de charge et champ électrique nuls à l'échelle macroscopique. Numériquement, on peut montrer[16] que r_0 *est de l'ordre de grandeur des distances entre ions voisins*;

16. La valeur de r_0 est déterminée par celle de la dérivée $\partial\mu_0/\partial n$:

$$r_0 = \left[\frac{\varepsilon_0}{e^2}\frac{\partial\mu_0}{\partial n}\right]^{1/2} . \tag{1}$$

Or, dans tout le domaine où le métal existe en tant que tel, la température T est négligeable devant la « température de Fermi » T_F du gaz d'électrons (cf. VI, § B.1.a). La répartition des électrons de conduction sur leurs états quantiques individuels est donc pratiquement celle du zéro absolu, qui est donnée de façon simple par le *principe de Pauli* : tous les états dont l'énergie est inférieure au potentiel chimique μ_0 sont occupés, tous ceux d'énergie supérieure à μ_0 sont vides. Connaissant la densité d'états individuels (I, § III.B.2.a), on en déduit facilement l'expression de μ_0 en fonction de n (voir VI, paragraphe I.B.1.a) :

$$\mu_0 = \eta_0 + \frac{\hbar^2}{2m}(3\pi^2 n)^{2/3} , \tag{2}$$

où la constante η_0 repère le fond de la « bande de conduction », énergie minimale des électrons du gaz. D'où

$$r_0 = \left[\frac{\varepsilon_0\hbar^2}{3me^2}(3\pi^2)^{2/3}n^{-1/3}\right]^{1/2} . \tag{3}$$

Si l'on fait apparaître le rayon de Bohr

$$a_0 \widehat{=} \frac{4\pi\varepsilon_0\hbar^2}{me^2} \simeq 0{,}53 \text{ Å} , \tag{4}$$

on obtient l'expression simple

$$r_0 = \frac{(3\pi^2)^{1/3}}{(12\pi)^{1/2}}[a_0 n^{-1/3}]^{1/2} \simeq 0{,}504\,[a_0 n^{-1/3}]^{1/2} . \tag{5}$$

Pour le cuivre par exemple, le nombre d'électrons de conduction par unité de volume vaut environ

$$n \simeq 8{,}5 \times 10^{28} \text{ électrons/m}^3 , \tag{6}$$

de sorte que

$$n^{-1/3} \simeq 2{,}3 \text{ Å} \tag{7}$$

et finalement

$$r_0 \simeq 0{,}55 \text{ Å} . \tag{8}$$

L'ordre de grandeur reste le même pour tous les métaux (on trouvera quelques valeurs de n au complément VI.A, tableau I).

l'écrantage réalisé par le gaz d'électrons de conduction est donc extrêmement efficace[17].

Supposons maintenant que la charge ponctuelle e_0 soit placée *à l'extérieur du métal.* Les électrons de conduction se répartissent à nouveau de façon à écranter le potentiel qu'elle crée. Mais, ne pouvant sortir du métal[18], ils réalisent cet écrantage en se redistribuant au voisinage immédiat de la surface du conducteur, dans la région située en regard de la charge extérieure : ce phénomène est connu en électrostatique sous le nom d'*« influence électrique »*. La répartition précise des charges sur le bloc de métal dépend de sa forme et de la position de la charge extérieure. Cette répartition peut être considérée, à l'échelle macroscopique, comme *superficielle,* puisqu'elle est concentrée sur une épaisseur de l'ordre de r_0, c'est-à-dire sur quelques distances interatomiques au plus; elle a pour effet d'*annuler le champ électrique en tous les points intérieurs au métal,* dès qu'ils se trouvent à une distance appréciable de la surface.

Un phénomène analogue se produit lorsqu'on soumet le métal à un champ électrique créé par une distribution quelconque de charges placées à l'extérieur : les électrons de conduction se répartissent de façon que *le potentiel électrostatique résultant soit constant dans le volume du métal,* sauf sur une couche d'épaisseur microscopique au voisinage immédiat de la surface[19].

c. Différence de potentiel de contact entre deux métaux

Prenons deux échantillons de métaux M_1 et M_2 de nature différente. Chacun d'eux contient un gaz d'électrons de conduction quasi-libres; les potentiels chimiques μ_1 et μ_2 de ces deux gaz sont en général différents.

---

17. Il est tellement efficace que le potentiel résultant $\varphi(\vec{r})$ varie sur des distances d'ordre microscopique, mettant en défaut l'hypothèse (iii) posée au début de ce complément : on ne peut parler de densité volumique de charge à cette échelle. Le *« modèle de Thomas-Fermi »* que nous développons ici ne peut donc donner qu'une idée qualitative du phénomène d'écrantage (voir référence 1, p. 343).

18. Il est en réalité possible, à l'aide d'un champ électrique intense, d'arracher des électrons d'un métal : c'est le phénomène d'*« émission de champ »*. Nous supposons ici que la force exercée par le champ extérieur sur les électrons de conduction n'est pas suffisante pour que ce phénomène se produise.

19. On montre facilement que le nombre d'électrons nécessaire pour réaliser l'écrantage est extrêmement faible devant le nombre total d'électrons de conduction, ce qui justifie l'approximation (B.41) et assure que les propriétés du gaz restent inchangées à l'intérieur du métal. Pour obtenir un ordre de grandeur, on peut par exemple calculer la charge totale Q d'une boule de cuivre de rayon $R = 10$ cm portée à un potentiel φ de 10 000 volts : la formule bien connue (théorème de Gauss)

$$\varphi = \frac{Q}{4\pi\varepsilon_0}\frac{1}{R}$$

donne $Q \simeq 10^{-7}$ coulomb, ce qui correspond à un excès ou un déficit d'environ 10^{12} électrons au voisinage de la surface de la boule. Ceux-ci sont à prendre sur les

$$N \simeq 8{,}5 \times 10^{28} \times \frac{4\pi}{3} \times 10^{-3} \simeq 3{,}5 \times 10^{26}$$

électrons de conduction que contient la boule. La proportion d'électrons concernés est effectivement infime.

Si l'on met M_1 et M_2 en contact électrique, ils pourront échanger des électrons; ceux-ci, selon la règle générale (III, § II.D.3.b), passeront *du métal ayant le potentiel chimique le plus élevé à l'autre,* disons pour fixer les idées de M_1 à M_2 si $\mu_1 > \mu_2$.

Mais les électrons portent une *charge électrique,* de sorte que ceux qui sont transférés de M_1 à M_2 laissent derrière eux dans M_1 des charges positives non compensées et chargent M_2 négativement (nous supposons que M_1 et M_2 étaient globalement neutres avant l'établissement du contact électrique). *Le mouvement de transfert d'électrons est donc freiné par la distribution de charge qu'ils créent.*

Ces deux effets antagonistes aboutissent rapidement à une *situation d'équilibre* dans laquelle ils se compensent statistiquement. A l'échelle macroscopique, chacun des deux blocs M_1 et M_2 est dans cette situation un conducteur en équilibre électrostatique, portant ici une charge globale non nulle. Par suite du phénomène d'écrantage décrit au paragraphe b, *les charges non compensées se situent dans une couche de quelques angströms d'épaisseur à la surface de M_1 et M_2, et le potentiel résultant $\varphi(\vec{r})$ qu'elles créent a même valeur en tous les points d'un même bloc* (sauf au voisinage immédiat de la surface)[20], φ_1 pour M_1 et φ_2 pour M_2.

La zone de contact entre les deux blocs présente évidemment, sur une épaisseur d'une dizaine d'angströms, une structure électrique compliquée : le potentiel $\varphi(\vec{r})$ doit y passer de la valeur φ_1 à la valeur φ_2 sur une faible distance, ce qui se traduit par un champ électrique non nul, et même très intense[21]; la répartition des électrons y est fortement inhomogène, donnant lieu à un excès de charges positives à certains endroits, négatives à d'autres. Mais cette zone est très mince, et les phénomènes redeviennent simples dès qu'on la quitte; ce sont ces phénomènes simples que nous allons analyser maintenant.

A l'intérieur de chacun des blocs, la distribution des ions positifs n'a pas été modifiée, puisqu'elle est rigide; comme la densité globale de charge est nulle en tous les points intérieurs, le nombre d'électrons par unité de volume est le même qu'avant établissement du contact, et donc aussi leur potentiel chimique local. La condition d'équilibre (B.16) est donc aisée à écrire dans ce cas. D'une part, en effet, l'énergie potentielle supplémentaire $-e\varphi(\vec{r})$ des électrons est constante à l'intérieur de chacun des deux blocs, mais différente de l'un à l'autre : $-e\varphi_1$ dans M_1, $-e\varphi_2$ dans M_2. D'autre part, leur « potentiel chimique local » $\tilde{\mu}(\vec{r})$ est égal à μ_1 en tous les points de M_1 et à μ_2 en tous les points de M_2, μ_1 et μ_2 ayant les mêmes valeurs que lorsque les deux blocs étaient isolés l'un de l'autre. On a donc simplement[22]

$$\mu_1 - e\varphi_1 = \mu_2 - e\varphi_2 \ . \tag{B.51}$$

---

20. Pour le conducteur M_1, par exemple, la charge portée par M_2 est une charge extérieure. Les électrons de conduction de M_1 se répartissent alors au voisinage de sa surface de façon que le champ qu'ils produisent eux-mêmes compense le champ extérieur en tout point intérieur de M_1; le potentiel électrostatique résultant est alors constant dans le volume de M_1. Un phénomène analogue se produit évidemment dans M_2. Notons cependant qu'ici les charges extérieures ne sont pas données à l'avance comme nous l'avons supposé dans les paragraphes a et b : la charge de l'un des conducteurs, et sa répartition superficielle, dépendent de la présence et des propriétés de l'autre; il s'agit à nouveau d'un « problème d'autocohérence ». Finalement, les charges superficielles des deux blocs M_1 et M_2 s'arrangent à l'équilibre de façon à annuler le champ électrique total en tous les points intérieurs de M_1 et de M_2.

21. Les ordres de grandeur sont de quelques dixièmes de volt pour la d.d.p. $|\varphi_2 - \varphi_1|$ et une dizaine d'angströms pour la distance sur laquelle elle s'établit. Ceci correspond à un champ électrique de l'ordre de quelques 10^8 V/m, valeur considérable à l'échelle macroscopique, à peine 500 fois plus faible que la valeur typique des champs intra-atomiques.

22. La somme $\mu - e\varphi$, que l'on appelle parfois « potentiel électrochimique », est ici le véritable potentiel chimique du gaz d'électrons commun à M_1 et M_2, celui dont la valeur est la même dans l'ensemble du système (voir la condition d'équilibre (B.16) et plus loin la figure 4).

La *différence de potentiel de contact* $\varphi_2 - \varphi_1$ entre les métaux M_2 et M_1 est donc donnée par la différence entre les potentiels chimiques μ_2 et μ_1 de leurs électrons de conduction avant contact :

$$\varphi_2 - \varphi_1 = \frac{1}{e}(\mu_2 - \mu_1) \,. \tag{B.52}$$

Cette différence de potentiel est typiquement de l'ordre de quelques dixièmes de volt.

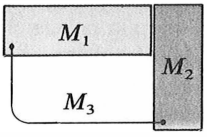


FIGURE 1
Aucun courant permanent ne circule dans un circuit fermé constitué de métaux différents.

Le système ne constitue cependant pas en lui-même un générateur, car aucune source d'énergie ne lui est associée. Si l'on construit un circuit fermé en reliant par un fil conducteur deux blocs M_1 et M_2 de métaux différents qui sont par ailleurs en contact (fig. 1), aucun courant permanent ne circule[23]. La d.d.p. de contact ne peut donc pas être mesurée par un voltmètre ordinaire[24], [25].

23. Les électrons se déplacent évidemment dans le circuit pendant la période transitoire d'établissement de l'équilibre, qui ne dure pas plus de 10^{-12} s.

24. Un voltmètre est sensible, non pas à la différence de potentiel électrique entre ses deux bornes, mais à la différence de « potentiel électrochimique » (note 22) : un courant ne le parcourt que si le potentiel chimique véritable des électrons n'est pas le même aux deux points où on le branche.

25. Pour mesurer la d.d.p. de contact entre deux métaux M_1 et M_2, on peut réaliser un condensateur dont une armature est faite de M_1, l'autre de M_2 (fig. 2). La d.d.p. entre les deux armatures, que l'on connecte par un fil, s'établit alors à $\varphi_2 - \varphi_1$. La charge Q du condensateur dépend de sa capacité C :

$$Q = C(\varphi_2 - \varphi_1) \,.$$

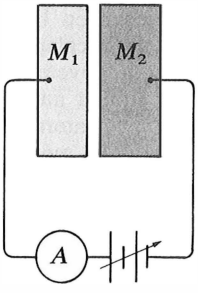


FIGURE 2
Mesure de la d.d.p. de contact entre deux métaux M_1 et M_2 à l'aide d'un générateur ajustable monté en opposition sur un condensateur construit avec ces deux métaux et que l'on peut faire vibrer.

Si l'on fait varier cette capacité (par exemple en modifiant la distance d entre les armatures), la charge Q varie, ce qui donne lieu à un courant dans le fil. En pratique, on monte un générateur ajustable en opposition avec la d.d.p. de contact, et on le règle pour qu'aucun courant ne parcoure le fil lorsqu'on écarte et rapproche alternativement les armatures; la d.d.p. du générateur est alors exactement opposée à la d.d.p. de contact.

Pour préciser les phénomènes à l'échelle microscopique, nous avons représenté sur la figure 3 l'énergie potentielle d'un électron dans chacun des deux blocs au voisinage de leur surface, lorsqu'ils sont neutres, isolés l'un de l'autre, et qu'ils ne sont soumis à aucun champ extérieur. Cette énergie potentielle est due à la fois aux ions positifs et à l'ensemble des électrons autres que celui que l'on considère (complément VI.A, note 4); à l'intérieur du métal, elle a la périodicité du réseau d'ions. Nous avons choisi l'origine de telle sorte que l'électron ait une énergie nulle lorsqu'il est au repos à l'extérieur du métal (suffisamment loin de la surface). Dans ce cas l'énergie potentielle est partout négative à l'intérieur du bloc, et le potentiel chimique, qui est pratiquement égal à l'énergie maximale d'un électron dans le métal (note 16), a lui aussi une valeur négative. Lorsqu'on atteint la surface, l'énergie potentielle croît rapidement, sur une distance de l'ordre de celle qui sépare deux ions voisins : l'électron qui tenterait de quitter le métal est rappelé vers l'intérieur par la charge positive qu'il laisse derrière lui[26]. L'énergie qu'il faut fournir à un électron pour l'arracher du métal (avec nos conventions, elle est égale à l'opposé du potentiel chimique) est appelée le *«travail d'extraction»;* sa valeur est de l'ordre de quelques volts.

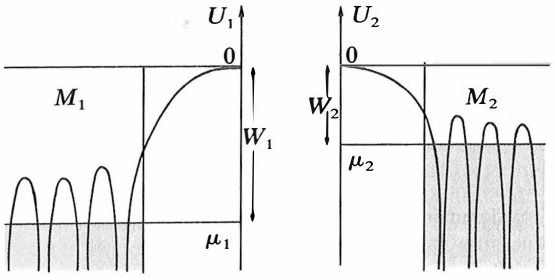


FIGURE 3

Énergie potentielle d'un électron dans chacun des deux blocs métalliques M_1 et M_2 au voisinage de leur surface, lorsqu'ils sont neutres et isolés. Cette énergie potentielle, U_1 ou U_2, est représentée dans les deux cas le long d'une ligne d'ions positifs. Avec l'origine choisie, elle est partout négative. Elle devient très grande, en valeur absolue, à la position de chacun des ions. Périodique à l'intérieur du métal, elle se déforme à l'approche de la surface puis croît rapidement vers zéro après le dernier ion. La valeur μ_1 ou μ_2 du potentiel chimique des électrons est également indiquée pour chacun des deux métaux, ainsi que le travail d'extraction W_1 ou W_2.

Établissons maintenant le contact entre les deux métaux M_1 et M_2. A l'équilibre, toutes les énergies électroniques de l'échantillon M_1 sont déplacées en bloc de la quantité constante $-e\varphi_1$, toutes celles de M_2 de $-e\varphi_2$. Ce double décalage est tel, d'après la condition (B.51), que les deux énergies maximales sont devenues égales (fig. 4). Les charges globales non nulles des deux blocs, qui se situent au voisinage immédiat de leur surface (sur une épaisseur de l'ordre de la distance interatomique), ne modifient pas la forme de l'énergie potentielle à l'intérieur des deux métaux; elles créent évidemment un champ électrique entre les deux conducteurs, correspondant à la d.d.p. $\varphi_2 - \varphi_1$ (voir note 21).

26. Les effets électriques qui se produisent au voisinage de la surface d'un cristal métallique sont décrits plus en détail dans la référence 1, chapitre 18.

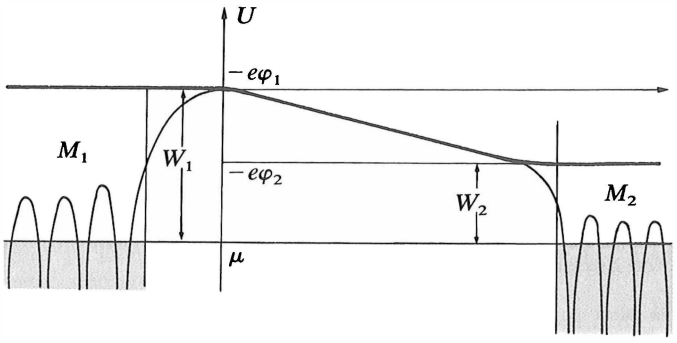


FIGURE 4

Énergie potentielle d'un électron dans l'ensemble des deux blocs M_1 et M_2 lorsqu'ils sont au contact. Cette figure est construite en raccordant les courbes de la figure 3 de manière que le potentiel chimique des électrons soit le même dans les deux blocs, conformément à la condition d'équilibre (B.51) (nous avons choisi ici pour origine des potentiels celui du bloc M_1). Entre les deux blocs, dans la zone de contact, le potentiel électrostatique varie de $\varphi_1 = 0$ à φ_2 (trait gras).

Référence

1. N.W. Ashcroft et N.D. Mermin, *Solid State Physics*, Holt Saunders, Tokyo (1981).

COMPLÉMENT V.C

Les gaz réels et le développement du viriel

I. Le théorème du viriel
  1. Définition du viriel
  2. Démonstration du théorème
  3. Application à l'équation d'état d'un fluide
    a. Séparation du viriel intérieur et du viriel extérieur
    b. Expression de la valeur moyenne du viriel extérieur
    c. Le viriel intérieur et l'équation d'état

II. Évaluation directe de la moyenne du viriel
  1. Introduction de la fonction de corrélation à deux particules
  2. Expression approchée de la fonction de corrélation
  3. Équation d'état approchée

III. Développement du viriel en puissances de la densité
  1. Principe de la méthode
    a. L'équation d'état et les coefficients du viriel
    b. La technique de développement
  2. Développement de la pression et du nombre de particules en puissances du paramètre λ
  3. Le développement du viriel dans l'équation d'état
  4. Conclusion

Appliquée aux gaz réels, l'équation d'état des gaz parfaits ne peut être qu'une première approximation. On sait cependant que cette approximation est d'autant meilleure que le gaz est plus dilué. Nous allons montrer ici comment on peut obtenir de façon systématique les corrections à l'équation d'état du gaz parfait sous la forme d'un développement en série entière de la densité; c'est ce qu'on appelle le «développement du viriel».

I. LE THÉORÈME DU VIRIEL

Nous considérons un gaz enfermé dans un récipient de volume V et constitué de N particules *sans structure interne.* Nous utiliserons le formalisme de la *mécanique classique* et supposerons N et V suffisamment grands pour que soit atteinte la *limite thermodynamique.*

1. DÉFINITION DU VIRIEL

Soit $\vec{r}_i$ le vecteur position, à l'instant t, de la particule (i), et $\vec{f}_i$ la résultante des forces auxquelles elle est soumise. On appelle *viriel du système* la quantité

$$\mathcal{V} \widehat{=} \sum_{i=1}^{N} \vec{r}_i . \vec{f}_i \,. \tag{C.1}$$

Cette grandeur, quoique peu utilisée dans les cours de mécanique classique, joue un rôle important en mécanique statistique car sa valeur moyenne est reliée à celle de l'énergie cinétique par le théorème du viriel.

Nous supposons que le *centre d'inertie* du système est *au repos* dans le référentiel galiléen utilisé. La résultante globale des forces est donc nulle :

$$\sum_{i=1}^{N} \vec{f}_i = \vec{0} \,. \tag{C.2}$$

Dans ce cas, le viriel $\mathcal{V}$ est *indépendant de l'origine* choisie pour les coordonnées[1].

2. DÉMONSTRATION DU THÉORÈME

Soit $\vec{p}_i$ la quantité de mouvement de la particule (i). Les équations de son mouvement s'écrivent

$$\frac{\mathrm{d}\vec{r}_i}{\mathrm{d}t} = \frac{\vec{p}_i}{m_i} \,, \tag{C.3}$$

$$\frac{\mathrm{d}\vec{p}_i}{\mathrm{d}t} = \vec{f}_i \,. \tag{C.4}$$

On en déduit

$$\frac{\mathrm{d}}{\mathrm{d}t}(\vec{r}_i \cdot \vec{p}_i) = \vec{r}_i \cdot \vec{f}_i + \frac{\vec{p}_i^{\,2}}{m_i} \,. \tag{C.5}$$

Sommée sur l'ensemble des N particules, l'égalité (C.5) donne

$$\frac{\mathrm{d}}{\mathrm{d}t} \sum_{i=1}^{N} \vec{r}_i \cdot \vec{p}_i = \mathcal{V} + 2E_c \,, \tag{C.6}$$

où
$$E_c \mathrel{\widehat{=}} \sum_{i=1}^{N} \frac{\vec{p}_i^{\,2}}{2m_i} \tag{C.7}$$

est l'énergie cinétique totale du système. Prenons alors la *valeur moyenne dans le temps* des deux membres de (C.6); celle du premier membre est nulle[2], de sorte que

$$\widehat{E}_c = -\frac{1}{2}\,\widehat{\mathcal{V}} \,. \tag{C.8}$$

---

1. Si M_i est la position de la particule (i), le vecteur $\vec{r}_i$ qui la repère par rapport à l'origine O est par définition

$$\vec{r}_i \mathrel{\widehat{=}} \overrightarrow{OM}_i \,. \tag{1}$$

Lorsqu'on choisit une autre origine O', le viriel devient

$$\mathcal{V}' = \sum_{i=1}^{N} \vec{r}\,'_i \cdot \vec{f}_i \,, \tag{2}$$

avec
$$\vec{r}\,'_i \mathrel{\widehat{=}} \overrightarrow{O'M}_i = \overrightarrow{O'O} + \vec{r}_i \,. \tag{3}$$

On a donc

$$\mathcal{V}' = \mathcal{V} + \overrightarrow{O'O} \cdot \sum_{i=1}^{N} \vec{f}_i \,. \tag{4}$$

2. On a en effet à calculer une expression de la forme

$$\frac{\widehat{\mathrm{d}g}}{\mathrm{d}t} = \lim_{\tau \to \infty} \frac{1}{\tau} \int_0^{\tau} \frac{\mathrm{d}g}{\mathrm{d}t}\,\mathrm{d}t \,, \tag{1}$$

ce qui donne aussitôt

$$\frac{\widehat{\mathrm{d}g}}{\mathrm{d}t} = \lim_{\tau \to \infty} \frac{1}{\tau}\,[g(\tau) - g(0)] \,. \tag{2}$$

.../...

Si, d'autre part, le système a atteint l'équilibre macroscopique, moyennes dans le temps et moyennes d'ensemble coïncident (cf. principe ergodique, I, § IV.B) :

$$\overline{E_c} = -\frac{1}{2}\,\overline{\mathcal{U}}\,. \tag{C.9}$$

A l'équilibre macroscopique, la valeur moyenne de l'énergie cinétique du système est égale à la moyenne de son viriel multipliée par $-1/2$[3],[4].

3. APPLICATION A L'ÉQUATION D'ÉTAT D'UN FLUIDE

a. Séparation du viriel intérieur et du viriel extérieur

Les forces qui agissent sur la particule (i) sont de deux types :

— *les forces extérieures :* si nous négligeons la pesanteur (et supposons qu'il n'existe aucun autre champ extérieur), elles se réduisent aux interactions avec les parois du récipient. Nous appellerons $\vec{\varphi}_i$ la force qu'exercent les parois sur la particule (i). Cette force dépend évidemment de la position $\vec{r}_i$ de la particule.

— *les forces intérieures,* dues à l'interaction avec les autres particules du système. Nous supposerons (cf. complément III.G, § I.2) que l'on peut négliger les interactions à trois corps et plus, et ne considérer que des interactions à deux corps : la force $\vec{f}_{ij}$ exercée par la particule (j) sur la particule (i) ne dépend alors que de leur position relative $(\vec{r}_i - \vec{r}_j)$ et pas des positions des $(N-2)$ autres particules. Les forces intérieures vérifient le « principe des actions réciproques » :

$$\vec{f}_{ji} = -\vec{f}_{ij}\,. \tag{C.10}$$

On a donc

$$\vec{f}_i = \vec{\varphi}_i + \sum_{j \neq i} \vec{f}_{ij}\,. \tag{C.11}$$

Compte tenu de (C.10), le viriel du système peut se mettre sous la forme

$$\mathcal{U} = \sum_{i=1}^{N} \vec{r}_i \cdot \vec{\varphi}_i + \frac{1}{2} \sum_{i=1}^{N} \sum_{j \neq i} (\vec{r}_i - \vec{r}_j) \cdot \vec{f}_{ij}\,; \tag{C.12}$$

$\mathcal{U}$ est ainsi la somme de deux termes, que nous appellerons respectivement le *« viriel extérieur »* et le *« viriel intérieur »* (le facteur 1/2 dans (C.12) évite de compter deux fois l'interaction entre deux particules).

Si g reste bornée, la limite est nulle. Or on a ici

$$g = \sum_{i=1}^{N} \vec{r}_i \cdot \vec{p}_i\,. \tag{3}$$

Chaque variable de position $\vec{r}_i$ est limitée à l'intérieur du volume V. D'autre part, pour un système *isolé,* l'énergie totale est fixée, et l'énergie potentielle globale est bornée inférieurement; par conséquent, l'énergie cinétique E_c est bornée supérieurement; d'après (C.7), chacune des quantités de mouvement $\vec{p}_i$ est donc elle-même bornée en module. L'expression (3) reste donc bien bornée et la valeur moyenne (2) est nulle.

3. A la limite thermodynamique, les fluctuations relatives autour de la valeur moyenne sont négligeables et la relation (C.9) peut, en pratique, s'écrire sans le symbole de valeur moyenne.

4. Le théorème du viriel est également valable en mécanique quantique pour les valeurs moyennes des opérateurs associés à l'énergie cinétique et au viriel, lorsque le système est dans un état propre de l'hamiltonien ou un mélange statistique de tels états (voir C.D.L., complément G.XI, § 5).

b. Expression de la valeur moyenne du viriel extérieur

Les interactions entre les particules et les parois du récipient sont à très courte portée (quelques Å); les seules particules qui contribuent au viriel extérieur sont donc celles qui se trouvent au voisinage immédiat des parois.

Plutôt que d'étudier les $\vec{\varphi}_i$ dans le détail, nous allons utiliser le principe des actions réciproques. Si p est la pression du gaz, les particules exercent *en moyenne* sur un élément $d\vec{s}$ de la paroi (orienté vers l'extérieur) la force

$$d\vec{f} = p\, d\vec{s}\,. \tag{C.13}$$

Inversement, cet élément de paroi exerce sur les particules de son voisinage immédiat (dont les positions $\vec{r}$ se confondent pratiquement avec celle de l'élément de surface $d\vec{s}$) une force (moyenne) égale et opposée. La moyenne du viriel extérieur vaut donc

$$\overline{\mathcal{V}}_{ex} = -\int_s \vec{r}\,.p\, d\vec{s}\,, \tag{C.14}$$

l'intégrale étant étendue à la surface s limitant le volume V du récipient. Le théorème de Green-Ostrogradsky permet de transformer cette intégrale de surface en une intégrale de volume, p étant la même en tous les points :

$$\overline{\mathcal{V}}_{ex} = -p \int_V \text{div}\, \vec{r}\; d^3r\,, \tag{C.15}$$

c'est-à-dire, puisque div $\vec{r} = 3$,

$$\overline{\mathcal{V}}_{ex} = -3pV\,. \tag{C.16}$$

c. Le viriel intérieur et l'équation d'état

L'énergie cinétique moyenne des particules (que nous avons supposées sans structure interne) est liée à la température par la relation (cf. III, § V.C)

$$\overline{E}_c = \frac{3}{2} NkT\,. \tag{C.17}$$

D'autre part, la somme $\sum_{j\neq i} \overline{(\vec{r}_i - \vec{r}_j).\vec{f}_{ij}}$ est indépendante de (i)[5]. L'expression de la valeur moyenne du viriel intérieur se simplifie donc quelque peu :

$$\overline{\mathcal{V}}_{int} = \frac{1}{2}\sum_{i=1}^{N}\sum_{j\neq i} \overline{(\vec{r}_i - \vec{r}_j).\vec{f}_{ij}} = \frac{N}{2}\sum_{j=2}^{N} \overline{(\vec{r}_1 - \vec{r}_j).\vec{f}_{1j}}\,. \tag{C.18}$$

En reportant les formules (C.16), (C.17) et (C.18) dans l'expression (C.9) du théorème du viriel, on obtient la relation

$$pV = N\left[kT + \frac{1}{6}\sum_{j=2}^{N} \overline{(\vec{r}_1 - \vec{r}_j).\vec{f}_{1j}}\right]. \tag{C.19}$$

Pour des particules indépendantes ($\vec{f}_{ij} = \vec{0}$), on retrouve évidemment l'équation d'état d'un gaz parfait. Lorsque les interactions ne sont pas négligeables, la relation (C.19) permet d'obtenir l'équation d'état du fluide considéré si l'on peut calculer, ou pour le moins évaluer, la moyenne du viriel intérieur.

---

5. La somme

$$\sum_{j\neq i} (\vec{r}_i - \vec{r}_j).\vec{f}_{ij}$$

est fonction de la position de la particule (i) choisie (et de celle des autres particules). Sa *valeur moyenne*, cependant, est indépendante de (i) puisque toutes les particules jouent des rôles identiques.

La relation (C.19) permet déjà quelques remarques qualitatives. Pour un *gaz dilué,* les distances entre particules sont grandes, de sorte que les forces qu'elles exercent les unes sur les autres sont principalement attractives (figure 1 du complément III.G). Le terme correctif à l'équation d'état du gaz parfait, dans (C.19), est donc négatif : la pression d'un gaz réel dilué est inférieure à celle d'un gaz parfait de même densité et de même température. C'est le contraire pour un liquide très dense : les forces répulsives deviennent alors prépondérantes et la pression est très supérieure à celle du gaz parfait correspondant[6].

Pour aller au-delà de ces remarques qualitatives, nous allons mettre en œuvre deux méthodes. Dans la première, nous tenterons une évaluation directe approchée de la valeur moyenne du viriel intérieur. La seconde, plus systématique, fournira un développement du viriel intérieur moyen en série entière de la densité N/V.

II. ÉVALUATION DIRECTE DE LA MOYENNE DU VIRIEL

1. INTRODUCTION DE LA FONCTION DE CORRÉLATION A DEUX PARTICULES

Pour calculer le second terme du membre de droite de (C.19), qui provient du viriel intérieur, nous allons traiter les $(N-1)$ particules qui agissent sur la particule (1) *comme un milieu continu* : nous noterons dn le nombre de ces particules qui se trouvent, en moyenne, dans le volume élémentaire d^3r repéré par le vecteur $\vec{r}$. Dans un gaz parfait, les molécules se meuvent librement et ce nombre vaut simplement[7]

$$dn_0 = \frac{N}{V} d^3r . \tag{C.20}$$

Dans un gaz réel, la particule (1) perturbe son voisinage immédiat : elle repousse (très fortement) les particules qui voudraient s'approcher à des distances $|\vec{r} - \vec{r}_1|$ inférieures au diamètre des molécules; celles qui se situent à des distances intermédiaires sont attirées (assez faiblement) vers le minimum du potentiel $u(\rho)$ dont dérive la force intermoléculaire $\vec{f}_{ij}$ (figure 1 du complément III.G); pour les grandes valeurs de $|\vec{r} - \vec{r}_1|$, l'influence de la particule (1) a pratiquement disparu, car le potentiel $u(\rho)$ est à courte portée. Nous écrirons

$$dn = \frac{N}{V} g(\vec{r} - \vec{r}_1) d^3r : \tag{C.21}$$

$g(\vec{r} - \vec{r}_1) d^3r/V$ *est la probabilité de trouver une particule dans le volume* d^3r *autour du point* $\vec{r}$*, sachant qu'il y a une particule au point* $\vec{r}_1$.

La fonction $g(\vec{r} - \vec{r}_1)$ est appelée *«fonction de corrélation à deux particules»*.

---

6. Dans ces conditions, les forces à trois corps (ou plus) jouent un rôle important et l'expression du viriel intérieur est plus compliquée; mais la conclusion reste qualitativement valable : il est très difficile de comprimer un liquide.

7. En toute rigueur,

$$dn_0 = \frac{N-1}{V} d^3r .$$

Dans toute la suite, puisque $N \gg 1$, nous remplacerons systématiquement $(N-1)$ par N.

D'après l'analyse précédente, on s'attend à ce que g s'annule aux courtes valeurs de $|\vec{r} - \vec{r}_1|$ et tende vers 1 ($dn \to dn_0$) aux grandes distances. Lorsqu'on peut la mesurer, on lui trouve l'allure indiquée sur la figure 1.

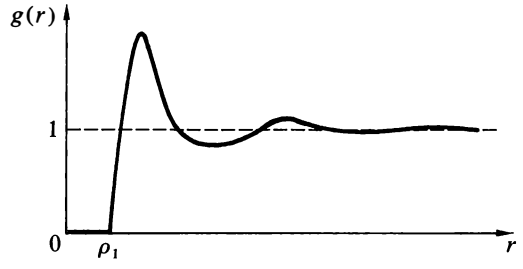


FIGURE 1

Allure de la fonction de corrélation à deux particules dans un fluide, en fonction de leur distance. Pour les faibles valeurs de $r(<\rho_1)$, la fonction est pratiquement nulle (forte répulsion). Du fait de la forme du potentiel d'interaction (figure 1 du complément III.G), elle passe ensuite par un maximum. La forte probabilité de présence des molécules au voisinage de ce maximum conduit, pour les plus grandes valeurs de r, à des oscillations qui s'amortissent rapidement. L'approximation du paragraphe II néglige ces oscillations et remplace $g(r)$ par $\exp -u(r)/kT$.

La moyenne du viriel intérieur s'exprime simplement à partir de la fonction de corrélation à deux particules; en prenant $\vec{r}_1$ pour origine, on obtient

$$\sum_{j=2}^{N} \overline{(\vec{r}_1 - \vec{r}_j).\vec{f}_{1j}} = \frac{N}{V} \int_V d^3r\, g(\vec{r})\, \vec{r}.\vec{f}(\vec{r}), \tag{C.22}$$

où $\vec{f}(\vec{r})$ désigne ici la force qu'exerce la particule (1) sur une particule située au point $\vec{r}$.

2. EXPRESSION APPROCHÉE DE LA FONCTION DE CORRÉLATION

La difficulté est évidemment de calculer la fonction de corrélation. Nous allons raisonner en première approximation comme si les $(N-1)$ particules autres que la particule (1) choisie comme référence étaient indépendantes les unes des autres. S'il en est ainsi, la probabilité pour que l'une d'elles se trouve au point $\vec{r}$ (sa quantité de mouvement n'étant pas précisée), où son énergie potentielle d'interaction avec la particule (1) vaut $u(\vec{r} - \vec{r}_1)$, est proportionnelle au facteur de Boltzmann[8] :

$$g(\vec{r} - \vec{r}_1) \propto \exp\left[-u(\vec{r} - \vec{r}_1)/kT\right]. \tag{C.23}$$

8. Si $U(\vec{r}_1, \vec{r}_2, \ldots, \vec{r}_N)$ est l'énergie potentielle d'interaction entre les N particules, la véritable fonction de corrélation à deux particules s'écrit

$$g(\vec{r}_1, \vec{r}_2) = \frac{V^2 \int d^3r_3\, d^3r_4 \ldots d^3r_N\, e^{-U/kT}}{\int d^3r_1\, d^3r_2\, d^3r_3 \ldots d^3r_N\, e^{-U/kT}}. \tag{1}$$

En effet, l'intégrale portant sur les N impulsions $\vec{p}_i$, en facteur au numérateur et au dénominateur, se simplifie; le coefficient V^2 provient de la normalisation de g : d'après (C.21), on doit avoir

$$\int_V d^3r \frac{N}{V} g(\vec{r} - \vec{r}_1) = N, \tag{2}$$

.../...

La condition aux grandes distances ($|\vec{r}-\vec{r}_1| \to \infty$, $u \to 0$, $g \to 1$) détermine la constante de proportionnalité[9]; on a donc

$$g(\vec{r}-\vec{r}_1)=\exp[-u(\vec{r}-\vec{r}_1)/kT]\,. \tag{C.24}$$

3. ÉQUATION D'ÉTAT APPROCHÉE

Le potentiel d'interaction $u(\vec{r}_i-\vec{r}_j)$ entre deux particules monoatomiques sans structure interne ne dépend que de la distance r_{ij} entre ces particules, de sorte que

$$\vec{f}_{ij}=-\frac{\mathrm{d}u}{\mathrm{d}r_{ij}}\frac{\vec{r}_i-\vec{r}_j}{r_{ij}}\,. \tag{C.25}$$

En reportant (C.25) et l'approximation (C.24) de la fonction de corrélation g dans (C.22), on obtient pour l'équation d'état (C.19)

$$pV=N\left\{kT-\frac{2\pi}{3}\frac{N}{V}\int_0^\infty \mathrm{d}r\,r^3\frac{\mathrm{d}u}{\mathrm{d}r}\exp[-u(r)/kT]\right\} \tag{C.26}$$

(l'intégrale peut sans inconvénient être prolongée jusqu'à l'infini car le potentiel $u(r)$ est à courte portée).

Si la fonction $u(r)$ est connue, une simple quadrature détermine, dans l'approximation considérée, l'équation d'état du gaz. Une intégration par parties[10] permet de transformer l'intégrale :

$$\int_0^\infty \mathrm{d}r\,r^3\frac{\mathrm{d}u}{\mathrm{d}r}\exp[-u(r)/kT]=-3kT\int_0^\infty \mathrm{d}r\,r^2[1-\exp(-u(r)/kT)]\,. \tag{C.27}$$

---

c'est-à-dire, de façon plus générale,

$$\int_V \mathrm{d}^3r_1\int_V \mathrm{d}^3r_2\,g(\vec{r}_1,\vec{r}_2)=V^2\,. \tag{3}$$

Si U est simplement la somme des potentiels $u(\vec{r}_i-\vec{r}_j)$ d'interaction des particules deux à deux, on peut mettre en facteur

$$\exp[-u(\vec{r}_1-\vec{r}_2)/kT]\,,$$

mais l'intégrale qui reste au numérateur dépend encore de $\vec{r}_1$ et $\vec{r}_2$ à travers les termes en $u(\vec{r}_i-\vec{r}_1)$ et $u(\vec{r}_j-\vec{r}_2)$ (pour i et j supérieurs à 2). L'approximation (C.23) consiste à négliger cette dépendance supplémentaire.

9. La constante de proportionnalité C peut aussi être obtenue par normalisation de la probabilité $g(\vec{r}-\vec{r}_1)\mathrm{d}^3r/V$. Prenant $\vec{r}_1$ comme origine, on écrit

$$\frac{C}{V}\int_V \mathrm{d}^3r\,\mathrm{e}^{-u(\vec{r})/kT}=1\,.$$

La fonction $u(\vec{r})$ étant à courte portée (quelques dizaines d'Å tout au plus), l'exponentielle vaut 1 dans presque tout le domaine d'intégration. Dans la limite thermodynamique ($V \to \infty$), on obtient donc $C=1$.

10. En effet

$$I\widehat{=}\int_0^\infty \mathrm{d}r\,r^3\frac{\mathrm{d}u}{\mathrm{d}r}\mathrm{e}^{-u(r)/kT}=\int_0^\infty \mathrm{d}r\,r^3kT\frac{\mathrm{d}}{\mathrm{d}r}[1-\mathrm{e}^{-u(r)/kT}]. \tag{1}$$

L'intégration par parties est alors très simple :

$$I=\{r^3kT[1-\mathrm{e}^{-u(r)/kT}]\}_{r=0}^{r=\infty}-3kT\int_0^\infty \mathrm{d}r\,r^2[1-\mathrm{e}^{-u(r)/kT}]\,. \tag{2}$$

Le terme tout intégré s'annule pour $r=0$ (à cause du facteur r^3) et pour $r=+\infty$ (puisque $u(r)\sim r^{-6}$ dans cette limite).

Sous cette forme, compte tenu de l'allure de $u(r)$, elle est facile à évaluer de façon approchée. On sait en effet (complément III.G, § I.2) que $u(r)$ est fortement répulsif pour des distances r inférieures à ρ_1, faiblement attractif aux distances intermédiaires et qu'il tend rapidement vers zéro aux grandes distances. On obtient donc une bonne approximation de (C.27) en prenant[11]

$$\begin{aligned} &u(r) \simeq +\infty \quad \text{pour} \quad 0 \leqslant r < \rho_1 , \\ &\frac{|u(r)|}{kT} \ll 1 \quad \text{pour} \quad r \geqslant \rho_1 , \end{aligned} \tag{C.28}$$

ce qui donne

$$\int_0^\infty \mathrm{d}r r^2 \left[1 - \mathrm{e}^{-u(r)/kT}\right] \simeq \frac{\rho_1^3}{3} + \int_{\rho_1}^\infty \mathrm{d}r r^2 \frac{u(r)}{kT} . \tag{C.29}$$

Si on pose, comme au complément III.G (formules (G.11) et (G.15)),

$$b \mathrel{\widehat{=}} \frac{2}{3} \pi \rho_1^3 \tag{C.30}$$

et

$$a \mathrel{\widehat{=}} -2\pi \int_{\rho_1}^\infty \mathrm{d}r r^2 u(r) , \tag{C.31}$$

l'équation d'état (C.26) prend la forme

$$pV = NkT \left[1 + \frac{N}{V}\left(b - \frac{a}{kT}\right)\right] . \tag{C.32}$$

On a ainsi calculé la première correction, proportionnelle à N/V, à l'équation d'état des gaz parfaits (ce résultat peut aussi être obtenu par la « théorie cinétique », voir complément III.C).

La relation (C.32) est très voisine de l'équation de van der Waals[12] (complément III.G)

$$\left(p + a \frac{N^2}{V^2}\right)(V - Nb) = NkT . \tag{C.33}$$

Pour obtenir cette dernière, il faut supposer

$$\frac{N}{V} b \ll 1 \tag{C.34}$$

et effectuer la substitution, valable dans cette limite,

$$1 + \frac{N}{V} b \longrightarrow \frac{1}{1 - bN/V} . \tag{C.35}$$

Cette ultime manipulation est parfaitement justifiée pour un gaz dilué : les formes (C.32) et (C.33) de l'équation d'état sont alors équivalentes. Mais nous avons montré dans le complément III.G que l'équation de van der Waals permet de prévoir la transition de phase gaz-liquide et de rendre compte de ses principales caractéristiques. *Cette propriété très importante n'est pas partagée par l'équation* (C.32) à laquelle nous a conduits le calcul développé ici. L'équation proposée en 1873 par van der Waals

11. Lorsqu'on parle de potentiel $u(r)$ faiblement attractif aux distances intermédiaires ($r \sim \rho_0$), c'est bien sûr par comparaison avec l'énergie thermique kT. La condition $|u(r)| \ll kT$ pour $r > \rho_1$ n'a donc de sens que pour un gaz; elle deviendrait fausse pour un liquide ($|u(\rho_0)| \sim kT$).

12. La différence avec le calcul du complément III.G provient du fait essentiel que l'approximation y était faite sur la fonction de partition, alors qu'ici elle est introduite directement sur l'équation d'état.

garde donc une grande partie de son mystère : on peut comprendre, par des arguments comme ceux de ce complément, qu'elle décrive les gaz dilués; mais elle est en réalité beaucoup plus puissante puisqu'elle permet de prévoir des phénomènes qui débordent largement du domaine où les raisonnements qui la « démontrent » sont vraiment valables.

III. DÉVELOPPEMENT DU VIRIEL EN PUISSANCES DE LA DENSITÉ

Dans cette troisième partie, nous allons exposer une méthode d'approximation plus systématique que celle du paragraphe II. Elle permet en principe de calculer autant de termes correctifs que l'on désire à l'équation d'état des gaz parfaits.

1. PRINCIPE DE LA MÉTHODE

a. L'équation d'état et les coefficients du viriel

De façon générale, l'approximation des gaz parfaits est valable, pour une température T donnée, si la densité N/V est suffisamment faible (complément I.A, § III.3). Pour un gaz quelconque (classique), les corrections à l'équation d'état des gaz parfaits doivent pouvoir s'écrire sous la forme d'un *développement en puissances de* N/V :

$$\frac{p}{kT}=\frac{N}{V}+B_2(T)\left(\frac{N}{V}\right)^2+B_3(T)\left(\frac{N}{V}\right)^3+\ldots+B_q(T)\left(\frac{N}{V}\right)^q+\ldots . \tag{C.36}$$

Le coefficient $B_q(T)$ s'appelle le q-ième *coefficient du viriel;* pour des raisons dimensionnelles (p, variable intensive, ne peut dépendre que des deux variables intensives T et N/V), les coefficients B_q ne dépendent que de la température. Au paragraphe précédent, nous avons calculé directement le deuxième coefficient du viriel $B_2(T)$. Il est cependant assez difficile d'améliorer la méthode du paragraphe II pour obtenir les coefficients suivants. Par contre, le formalisme grand-canonique fournit une méthode très élégante pour déterminer systématiquement les B_q. Nous nous contenterons ici d'exposer cette méthode et de déterminer les tout premiers coefficients.

b. La technique de développement

Si on suppose le système de particules en contact avec un réservoir de potentiel chimique μ, sa grande fonction de partition peut s'écrire (formule V(II.5))

$$\Xi=\sum_{n=0}^{\infty} Z_n(T,\,V)\,\mathrm{e}^{n\mu/kT}\,, \tag{C.37}$$

où Z_n est la fonction de partition d'un système de n particules :

$$Z_n=\frac{1}{n!}\frac{1}{h^{3n}}\int \mathrm{d}^3p_1\ldots\mathrm{d}^3p_n\int \mathrm{d}^3r_1\ldots\mathrm{d}^3r_n \exp\left\{-\frac{1}{kT}\left[\sum_{i=1}^{n}\frac{\vec{p}_i^{\,2}}{2m}+U_n(\vec{r}_1,\ldots,\vec{r}_n)\right]\right\}, \tag{C.38}$$

U_n représentant l'énergie potentielle d'interaction des n particules.

Il est commode de séparer dans la somme (C.37) le terme $n=0$ qui vaut 1 et d'effectuer dans (C.38) les intégrations sur les $\vec{p}_i$; ceci donne

$$\Xi=1+\sum_{n=1}^{\infty}\frac{1}{n!}\left(\frac{2\pi mkT}{h^2}\right)^{3n/2}\mathrm{e}^{n\mu/kT}Y_n(T,\,V)\,, \tag{C.39}$$

avec $$Y_n(T,\,V)\mathrel{\widehat{=}}\int \mathrm{d}^3r_1\ldots\mathrm{d}^3r_n\,\mathrm{e}^{-U_n/kT}\,. \tag{C.40}$$

Sous la forme (C.39), Ξ apparaît comme une série entière de puissances du paramètre

$$\lambda \widehat{=} \left(\frac{2\pi mkT}{h^2}\right)^{3/2} e^{\mu/kT} ; \tag{C.41}$$

les coefficients de la série, les $Y_n/n!$, sont, pour les premiers d'entre eux tout au moins, facilement calculables puisqu'ils ne font intervenir qu'un très petit nombre de particules.

L'idée de la méthode est alors la suivante : à partir de l'expression (C.39) de Ξ, on peut obtenir le développement en puissances de λ, pour la pression p d'une part, et pour le nombre moyen N de particules d'autre part. En éliminant λ entre ces deux séries, on aboutira à la relation (C.36) cherchée, qui exprime p en série de puissances du rapport N/V.

2. DÉVELOPPEMENT DE LA PRESSION ET DU NOMBRE DE PARTICULES EN PUISSANCES DU PARAMÈTRES λ

De la grande fonction de partition Ξ, on déduit le grand potentiel :

$$J \widehat{=} -kT \ln \Xi = -kT \ln \left[1 + \sum_{n=1}^{\infty} Y_n \frac{\lambda^n}{n!}\right] , \tag{C.42}$$

et la pression :

$$p = -\frac{J}{V} = \frac{kT}{V} \ln \left[1 + \sum_{n=1}^{\infty} Y_n \frac{\lambda^n}{n!}\right] . \tag{C.43}$$

En développant le logarithme dans cette formule, on obtient p comme une série de puissances de λ :

$$\frac{pV}{kT} = \sum_{n=1}^{\infty} a_n \lambda^n , \tag{C.44}$$

dont les coefficients a_n s'expriment en fonction des Y_n. Pour les premiers d'entre eux, on trouve[13]

$$a_1 = Y_1 , \tag{C.45.a}$$

$$a_2 = \frac{1}{2} [Y_2 - Y_1^2] , \tag{C.45.b}$$

$$a_3 = \frac{1}{3!} [Y_3 - 3Y_1Y_2 + 2Y_1^3] , \tag{C.45.c}$$

etc.

---

13. Il suffit d'écrire

$$\ln(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \ldots \tag{1}$$

et de remplacer x par

$$x = \sum_{n=1}^{\infty} Y_n \frac{\lambda^n}{n!} \tag{2}$$

pour aboutir au développement cherché. Par exemple, pour obtenir le développement jusqu'au troisième ordre, on détermine les coefficients de λ (pour a_1), λ^2 (pour a_2) et λ^3 (pour a_3) dans

$$\left(\lambda Y_1 + \frac{\lambda^2}{2} Y_2 + \frac{\lambda^3}{3!} Y_3\right) - \frac{1}{2}\left(\lambda Y_1 + \frac{\lambda^2}{2} Y_2\right)^2 + \frac{1}{3}(\lambda Y_1)^3 , \tag{3}$$

ce qui donne immédiatement les formules (C.45).

De la même façon, le nombre moyen N de particules est donné par

$$N = kT \frac{\partial}{\partial \mu} \ln \Xi = V \frac{\partial p}{\partial \mu} , \tag{C.46}$$

ou encore, si on utilise la relation (C.41) entre λ et μ,

$$N = V \frac{\partial p}{\partial \lambda} \frac{\lambda}{kT} . \tag{C.47}$$

Le développement de N se déduit donc de celui de la pression (C.44) :

$$N = \sum_{n=1}^{\infty} n a_n \lambda^n . \tag{C.48}$$

3. LE DÉVELOPPEMENT DU VIRIEL DANS L'ÉQUATION D'ÉTAT

Il suffit maintenant d'éliminer λ entre les relations (C.44) et (C.48) pour obtenir le développement du viriel (C.36). Le plus simple est de considérer que (C.48) définit λ comme fonction de N, et d'écrire

$$\lambda = \sum_{m=1}^{\infty} b_m N^m . \tag{C.49}$$

On calcule les b_m en reportant (C.49) dans (C.48) et en identifiant terme à terme, ce qui donne[14]

$$b_1 = \frac{1}{a_1} , \tag{C.50.a}$$

$$b_2 = -2 \frac{a_2}{a_1} b_1^2 = -2 \frac{a_2}{a_1^3} , \tag{C.50.b}$$

$$b_3 = -\frac{3a_3 b_1^3 + 4a_2 b_1 b_2}{a_1} = \frac{1}{a_1^5} (8a_2^2 - 3a_3 a_1) , \tag{C.50.c}$$

etc.

En reportant (C.49) dans (C.44), on trouve

$$\frac{pV}{kT} = a_1 b_1 N + (a_1 b_2 + a_2 b_1^2) N^2 + (a_1 b_3 + 2a_2 b_1 b_2 + a_3 b_1^3) N^3 + \dots . \tag{C.51}$$

Les premiers coefficients du viriel s'en déduisent aussitôt en fonction des a_i grâce aux relations (C.50) :

$$B_1 = 1 , \tag{C.52.a}$$

14. Si on remplace λ par son expression (C.49), l'égalité (C.48) devient

$$N = a_1 [b_1 N + b_2 N^2 + b_3 N^3 + \dots] + 2a_2 [b_1 N + b_2 N^2 + \dots]^2 + 3a_3 [b_1 N + \dots]^3 + \dots .$$

Cette relation doit être vérifiée quel que soit N. Le coefficient de N au second membre doit donc être égal à 1, ceux de toutes les autres puissances doivent être nuls; d'où les formules (C.50).

$$B_2 = -V\frac{a_2}{a_1^2}\,, \tag{C.52.b}$$

$$B_3 = \frac{2V^2}{a_1^4}(2a_2^2 - a_3 a_1)\,, \tag{C.52.c}$$

etc.

On obtient finalement l'expression des coefficients du viriel directement en terme des intégrales de configuration Y_i en utilisant (C.45) et en remarquant que

$$a_1 = Y_1 = V : \tag{C.53}$$

$$B_1 = 1\,, \tag{C.54.a}$$

$$B_2 = \frac{1}{2V}(Y_1^2 - Y_2)\,, \tag{C.54.b}$$

$$B_3 = \frac{1}{3V^2}(Y_1^4 - 3Y_1^2 Y_2 - Y_1 Y_3 + 3Y_2^2)\,, \tag{C.54.c}$$

etc.

La méthode permet de calculer les coefficients suivants, mais les calculs deviennent vite extrêmement lourds.

En remplaçant Y_1 et Y_2 par leur expression (formule (C.40)), on peut écrire le *second coefficient du viriel* (C.54.b) sous la forme[15]

$$B_2(T) = 2\pi \int_0^\infty \mathrm{d}r\, r^2 [1 - \mathrm{e}^{-u(r)/kT}]\,; \tag{C.55}$$

on retrouve exactement le résultat du paragraphe II (formules (C.26) et (C.27)).

---

15. La relation (C.40) s'écrit pour deux particules

$$Y_2 = \int_V \mathrm{d}^3 r_1 \int_V \mathrm{d}^3 r_2\, \mathrm{e}^{-u(\vec{r}_2 - \vec{r}_1)/kT}\,. \tag{1}$$

Le changement de variables

$$\vec{r}\,' \widehat{=} \vec{r}_1\,,\quad \vec{r} \widehat{=} \vec{r}_2 - \vec{r}_1 \tag{2}$$

donne

$$Y_2 = \int_V \mathrm{d}^3 r' \int_V \mathrm{d}^3 r\, \mathrm{e}^{-u(\vec{r})/kT} = V \int_V \mathrm{d}^3 r\, \mathrm{e}^{-u(\vec{r})/kT}\,. \tag{3}$$

D'autre part, en reportant (3) et

$$Y_1 = \int_V \mathrm{d}^3 r = V \tag{4}$$

dans (C.54.b), on obtient

$$B_2 = \frac{1}{2V}\left[V \int_V \mathrm{d}^3 r - V \int_V \mathrm{d}^3 r\, \mathrm{e}^{-u(\vec{r})/kT}\right]. \tag{5}$$

Dans le cas où l'énergie d'interaction u ne dépend que de la distance r entre les deux particules, l'expression (5) s'écrit sous la forme (C.55), l'intégrale pouvant être étendue jusqu'à l'infini puisque la portée du potentiel u est très faible.

Dans le cas où on néglige les interactions à plus de deux corps, le coefficient B_3 prend lui aussi une forme relativement simple[16],[17] :

$$B_3(T) = -\frac{1}{3}\int d^3r \int d^3r' [e^{-u(r)/kT} - 1][e^{-u(r')/kT} - 1] \times [e^{-u(|\vec{r}-\vec{r}'|)/kT} - 1] . \qquad (C.56)$$

4. CONCLUSION

La méthode que nous venons d'exposer permet de calculer les termes successifs du développement du viriel jusqu'à n'importe quel ordre. Il est cependant impossible d'en déduire vraiment la somme de la série. En pratique, la méthode n'a donc d'intérêt que si les termes décroissent suffisamment vite pour qu'on puisse tronquer la somme à une valeur de q pas trop élevée. Elle n'est pas bien adaptée à l'étude des liquides : la densité y est trop forte, et chaque molécule est en constante interaction avec un grand nombre de ses voisines. Pour les gaz, par contre, elle donne d'excellents résultats (à condition évidemment que l'on dispose d'expressions satisfaisantes pour le potentiel $u(r)$). Pour l'argon à 25 °C, par exemple, les propriétés observées sont très bien décrites à partir des trois premiers termes du viriel, c'est-à-dire des coefficients B_2 et B_3; jusqu'à 100 atmosphères, ils rendent compte de l'équation d'état à mieux que 99,96 %. C'est seulement à des pressions considérables (c'est-à-dire pour des densités très élevées) que les déviations peuvent devenir importantes (37 % à 1 000 atmosphères pour l'argon).

---

16. En utilisant autant de fois que nécessaire le changement de variables (2) de la note 15, on peut écrire

$$Y_1^4 = V^2 \int d^3r \int d^3r' , \qquad (1)$$

$$3Y_1^2 Y_2 = 3V^2 \int d^3r \int d^3r' \, e^{-u(\vec{r}-\vec{r}')/kT} = 3V^2 \int d^3r \int d^3r' \, e^{-u(\vec{r})/kT}$$

$$= 3V^2 \int d^3r \int d^3r' \, e^{-u(\vec{r}')/kT}$$

$$= V^2 \int d^3r \int d^3r' [e^{-u(\vec{r})/kT} + e^{-u(\vec{r}')/kT} + e^{-u(\vec{r}-\vec{r}')/kT}] , \qquad (2)$$

$$Y_1 Y_3 = V^2 \int d^3r \int d^3r' \, e^{-u(\vec{r})/kT} . e^{-u(\vec{r}')/kT} . e^{-u(\vec{r}-\vec{r}')/kT} . \qquad (3)$$

En outre,

$$Y_2 = V \int d^3r \, e^{-u(\vec{r})/kT} = V \int d^3r' \, e^{-u(\vec{r}')/kT} = V \int d^3r \, e^{-u(\vec{r}-\vec{r}')/kT} , \qquad (4)$$

de sorte que

$$3Y_2^2 = V^2 \int d^3r \int d^3r' [e^{-u(\vec{r})/kT} . e^{-u(\vec{r}')/kT} + e^{-u(\vec{r})/kT} . e^{-u(\vec{r}-\vec{r}')/kT} + e^{-u(\vec{r}')/kT} . e^{-u(\vec{r}-\vec{r}')/kT}] . \qquad (5)$$

En prenant la combinaison *ad hoc* de (1), (2), (3) et (5) qui figure dans (C.54.c), on obtient (C.56).

17. Il existe une méthode graphique systématique pour déterminer le coefficient B_q lorsqu'on ne tient compte que des interactions à deux corps. Voir, par exemple, Mayer J. E. et Mayer M. G., *Statistical Mechanics,* John Wiley, New York (1963), chapitre 13.

COMPLÉMENT V.D

Le postulat fondamental sous sa forme « généralisée »

I. Énoncé « généralisé » du postulat fondamental

II. Applications du postulat « généralisé »
1. La distribution canonique déduite du postulat « généralisé »
2. Application la plus générale
a. Détermination de la distribution d'équilibre
b. Fonction de partition généralisée
c. Entropie d'équilibre

III. Discussion : les deux formes du postulat fondamental sont-elles équivalentes?

Sous la forme où nous l'avons énoncé au chapitre II puis appliqué dans les chapitres III et V, le postulat fondamental de la mécanique statistique privilégie les *systèmes isolés,* car les lois de conservation (énergie, nombre de particules...) s'expriment de façon simple lorsque les échanges entre le système considéré et l'extérieur sont impossibles. Nous avons cependant signalé dès le chapitre II (§ I.B.3.b) qu'il existe une formulation apparemment plus générale du postulat fondamental, qui peut en principe être utilisée quelle que soit la situation du système étudié. C'est cette *formulation « généralisée »* que nous allons présenter, illustrer et discuter dans ce complément.

I. ÉNONCÉ « GÉNÉRALISÉ » DU POSTULAT FONDAMENTAL

Reprenons d'abord le cas d'un *système isolé,* dont le nombre de particules N, l'énergie E, le volume V, ... sont fixés. Ces contraintes macroscopiques déterminent les *états microscopiques accessibles* au système. Rien ne privilégie *a priori* certains états accessibles par rapport aux autres : ils vérifient tous de la même façon les contraintes extérieures. Le postulat fondamental consiste donc à admettre que *tous les états accessibles sont également probables.* Nous avons déjà souligné (II, § I.B.3.a) que c'est le seul postulat possible dans ces conditions : *privilégier,* dans la distribution statistique choisie, *certains états accessibles par rapport aux autres serait erroné,* car on ne possède *aucune information permettant de le faire à bon escient.* Le contenu physique du postulat est donc en fin de compte le suivant : la donnée des paramètres définissant l'état macroscopique du système suffit à fonder la description statistique de ses états microscopiques. A partir de là, la formulation du postulat découle d'un principe très général : il faut *utiliser toute l'information disponible, et seulement celle-là.* La distribution qui vérifie cette condition est celle qui rend l'entropie statistique maximale.

C'est en appliquant ce même principe à d'autres situations que l'on aboutit à l'énoncé « généralisé » (réf. 1 et 2). Lorsque le système n'est pas isolé, les divers états microscopiques dans lesquels il peut se trouver *a priori* ne sont pas tous équivalents : si par exemple le système étudié échange de l'énergie avec un thermostat, un état microscopique dans lequel le système accapare presque toute l'énergie disponible n'est pas vraiment impossible, mais il est certainement moins probable que ceux où l'énergie est répartie de façon plus « équitable ». Il s'agit alors, à partir du *même contenu physique* que ci-dessus, de trouver la distribution statistique qui répond au *même principe :* utiliser toute l'information disponible, et seulement celle-là. La réponse est fournie par

la maximisation de l'entropie statistique, mesure de l'information manquante (I, § II.A) : choisir une distribution différente de celle qui donne le maximum serait erroné, car cela reviendrait à introduire une information supplémentaire, qu'on ne possède pas en réalité.

Nous remplacerons donc ici le postulat du chapitre II (§ I.B.1) par le suivant : *« A l'équilibre macroscopique*[1]*, la distribution statistique des états microscopiques est, parmi toutes celles qui vérifient les contraintes extérieures imposées au système, celle qui rend l'entropie statistique maximum ».*

Lorsqu'on l'applique à un système isolé, le présent énoncé redonne celui que nous avons utilisé jusqu'ici. Il est apparemment plus général, car il semble pouvoir s'adapter à toute autre situation. En pratique cependant, la nécessité d'expliciter les contraintes extérieures sous une forme exploitable réduit considérablement la portée de cette généralisation, comme nous allons le constater dans les applications. Nous analyserons cette question plus en détail au paragraphe III.

II. APPLICATIONS DU POSTULAT « GÉNÉRALISÉ »

1. LA DISTRIBUTION CANONIQUE DÉDUITE DU POSTULAT « GÉNÉRALISÉ »

Considérons un système dont le nombre de particules N, le volume V, ... sont fixés mais dont l'*énergie peut fluctuer* par échange de chaleur avec l'extérieur. Les contraintes liées à N, V, ... sont faciles à prendre en compte : il suffit, comme dans le cas du système isolé, d'exclure les états microscopiques qui ne les satisfont pas. Quant à l'énergie, elle n'est plus fixée, et la contrainte correspondante disparaît. Il faut pourtant la remplacer, car la supprimer simplement aboutirait à considérer comme également probables tous les états microscopiques ayant le nombre de particules N, le volume V, ..., mais d'énergie quelconque, ce qui est absurde physiquement : l'énergie fluctue, mais pas n'importe comment. Nous ne pouvons pas, ici, faire intervenir directement et explicitement un thermostat, car le postulat « généralisé » s'applique au système lui-même. Il faut donc trouver une condition qui ne fasse intervenir que les grandeurs liées au système. La solution consiste à *se donner la valeur moyenne de l'énergie du système.*

Cherchons en effet le maximum de l'entropie statistique du système,

$$S = -k \sum_{(\ell)} P_\ell \ln P_\ell \,, \tag{D.1}$$

en imposant aux probabilités P_ℓ non seulement d'être normalisées :

$$\sum_{(\ell)} P_\ell = 1 \,, \tag{D.2}$$

mais également de vérifier la condition supplémentaire[2]

$$\sum_{(\ell)} E_\ell P_\ell = \overline{E} \quad \text{donnée.} \tag{D.3}$$

1. Comme celui du chapitre II, ce postulat concerne les *systèmes de taille macroscopique.* Toutefois, il s'applique également aux *systèmes de taille microscopique* lorsque ceux-ci sont en contact avec un ou plusieurs systèmes macroscopiques, de sorte que la notion d'*équilibre macroscopique* ait un sens (c'est le cas dans la situation canonique, par exemple : voir chapitre III, paragraphe I.B.1 et paragraphe II.1 ci-dessous).

2. Dans les formules (D.1)-(D.3), la sommation porte seulement sur les états (ℓ) permis par les contraintes strictes concernant le nombre de particules, le volume, ...

Nous utilisons pour cela la méthode des multiplicateurs de Lagrange (appendice IV) : à partir de (D.1) et des conditions (D.2) et (D.3), nous construisons la fonction

$$Y \widehat{=} -k \sum_{(\ell)} P_\ell \ln P_\ell - a \left[\sum_{(\ell)} P_\ell - 1 \right] - b \left[\sum_{(\ell)} E_\ell P_\ell - \overline{E} \right], \tag{D.4}$$

où a et b ne dépendent pas des P_ℓ, et annulons ses dérivées par rapport à chacune des variables P_ℓ, comme si elles étaient indépendantes. Ceci donne

$$\frac{\partial Y}{\partial P_\ell} = -k (\ln P_\ell + 1) - a - b E_\ell = 0 . \tag{D.5}$$

Ces équations se résolvent aussitôt; leur solution $\{P_\ell^M\}$ est unique :

$$P_\ell^M = \mathrm{e}^{-1-\frac{a}{k}} \, \mathrm{e}^{-\frac{b}{k} E_\ell} . \tag{D.6}$$

On détermine ensuite les constantes a et b en reportant le résultat (D.6) dans les conditions (D.2) et (D.3) :

$$\mathrm{e}^{-1-\frac{a}{k}} \sum_{(\ell)} \mathrm{e}^{-\frac{b}{k} E_\ell} = 1 , \tag{D.7}$$

$$\mathrm{e}^{-1-\frac{a}{k}} \sum_{(\ell)} E_\ell \, \mathrm{e}^{-\frac{b}{k} E_\ell} = \overline{E} . \tag{D.8}$$

Il est clair d'après leur expression (D.6) que les probabilités P_ℓ^M trouvées sont bien positives; on peut d'autre part montrer[3] qu'elles rendent bien S maximum, et pas

3. Il suffit pour cela d'utiliser l'inégalité (complément I.G, formule (G.28))

$$S(P_\ell) \leqslant -k \sum_{(\ell)} P_\ell \ln P_\ell' , \tag{1}$$

valable pour deux distributions $\{P_\ell\}$ et $\{P_\ell'\}$ normalisées quelconques, et dans laquelle l'égalité est réalisée si et seulement si les deux distributions coïncident.

Considérons en effet une distribution $\{P_\ell\}$ normalisée et vérifiant la contrainte (D.3). On peut écrire, d'après (1),

$$S(P_\ell) \leqslant -k \sum_{(\ell)} P_\ell \ln P_\ell^M , \tag{2}$$

c'est-à-dire, en utilisant la formule (D.6),

$$S(P_\ell) \leqslant k \sum_{(\ell)} P_\ell \left[\frac{b}{k} E_\ell + 1 + \frac{a}{k} \right], \tag{3}$$

soit finalement, puisque $\{P_\ell\}$ vérifie les contraintes (D.2) et (D.3),

$$S(P_\ell) \leqslant b\overline{E} + (k + a) . \tag{4}$$

Mais le second membre n'est autre que l'entropie $S(P_\ell^M)$ associée à la solution $\{P_\ell^M\}$ des équations (D.5). Donc

$$S(P_\ell) \leqslant S(P_\ell^M) , \tag{5}$$

l'égalité n'ayant lieu que si $P_\ell = P_\ell^M$ pour tout (ℓ). $S(P_\ell^M)$ est donc bien le maximum de S lorsqu'on impose la contrainte (D.3).

seulement stationnaire. La solution (D.6)-(D.8) qui, si elle existe, est unique[4], est donc bien acceptable physiquement.

Il suffit alors de poser

$$\beta \widehat{=} \frac{b}{k} \tag{D.9}$$

et $$Z \widehat{=} e^{1+\frac{a}{k}} \tag{D.10}$$

pour identifier les formules (D.6) et (D.7) avec celles qui caractérisent la distribution canonique :

$$P_\ell^M \equiv P_\ell^c = \frac{1}{Z} e^{-\beta E_\ell} , \tag{D.11}$$

avec $$Z \widehat{=} \sum_{(\ell)} e^{-\beta E_\ell} . \tag{D.12}$$

Donc, dans le point de vue « généralisé » que nous adoptons dans ce complément, *le paramètre de température canonique β apparaît comme un multiplicateur de Lagrange* (de même d'ailleurs que $\ln Z$). Ce paramètre est *déterminé à partir de la valeur donnée de $\overline{E}$,* par la condition (D.8) :

$$\frac{1}{Z(\beta)} \sum_{(\ell)} E_\ell e^{-\beta E_\ell} = \overline{E} . \tag{D.13}$$

On peut considérer, comme au chapitre III (§ II.A.1), que la formule (D.12) définit Z comme une fonction du paramètre β : c'est la *fonction de partition* $Z(\beta)$. La relation (D.13) peut alors s'écrire sous la forme (cf. III, § II.B.1.a) :

$$-\frac{\partial}{\partial \beta} \ln Z(\beta) = \overline{E} . \tag{D.14}$$

Il n'existe pas toujours une valeur de β vérifiant (D.14); mais, s'il en existe une, cette valeur est unique (voir note 4).

Nous constatons par conséquent que l'étude d'un système en équilibre avec un thermostat peut être abordée de deux manières distinctes mais équivalentes : ou bien on applique le postulat fondamental tel qu'il a été énoncé au chapitre II et l'on considère que le thermostat impose sa température T au système; ou bien on utilise le postulat

---

4. La formule (D.6) découle de façon unique des équations (D.5). L'existence et l'unicité de la solution au problème de maximisation sont donc décidées par les équations (D.7)-(D.8) qui déterminent les paramètres a et b à partir de la donnée $\overline{E}$.

Ces équations peuvent ne pas avoir de solution. C'est le cas si la valeur de $\overline{E}$ que l'on veut imposer au système est en réalité impossible à atteindre, par exemple si l'on choisit pour $\overline{E}$ une énergie inférieure au niveau fondamental E_0 du système.

Mais *si les équations (D.7)-(D.8) ont une solution, celle-ci est unique.* Imaginons en effet qu'il existe deux couples $\{a, b\}$ et $\{a', b'\}$ de paramètres vérifiant ces équations. En reprenant le raisonnement de la note précédente, on démontre que les distributions $\{P_\ell^M\}$ et $\{P_\ell'^M\}$ correspondantes doivent être telles que

$$S(P_\ell'^M) \leqslant S(P_\ell^M) , \tag{1}$$

mais aussi

$$S(P_\ell^M) \leqslant S(P_\ell'^M) . \tag{2}$$

Chacune de ces inégalités doit donc être en fait une égalité, ce qui implique que les deux distributions sont identiques (c'est en effet seulement pour deux distributions identiques que l'inégalité (5) de la note 3 devient une égalité).

fondamental sous la forme «générale» du paragraphe I ci-dessus et l'on suppose que le contact avec le thermostat fixe la valeur moyenne $\overline{E}$ de l'énergie du système[5]. La relation entre ces deux points de vue est établie par la formule (D.13).

2. APPLICATION LA PLUS GÉNÉRALE

Le raisonnement du paragraphe précédent se généralise aisément. Imaginons un système à l'équilibre statistique dans une situation telle que les contraintes extérieures qui s'exercent sur lui consistent, soit à fixer un paramètre extérieur à une valeur certaine (éventuellement affectée d'une incertitude), soit à imposer la valeur moyenne d'une autre grandeur. De façon plus précise, *partant de la situation où le système est isolé, on remplace certaines des contraintes strictes* fixant les paramètres extérieurs *par des contraintes imposant seulement la valeur moyenne* des grandeurs correspondantes[6].

a. Détermination de la distribution d'équilibre

Les contraintes du premier type (contraintes strictes) sont entièrement prises en compte en excluant les états microscopiques qui ne les vérifient pas. Quant aux contraintes en valeur moyenne, supposons qu'elles portent sur n grandeurs[7] A^i :

$$\sum_{(\ell)} A^i_\ell P_\ell = \overline{A^i} \text{ fixée} ; \quad i = 1, 2, \ldots, n . \tag{D.15}$$

Conformément au postulat «généralisé», nous cherchons la distribution de probabilités $\{P^M_\ell\}$ (sur les états (ℓ) permis par les contraintes du premier type) qui rend l'entropie statistique maximale compte tenu des conditions (D.15). Appliquant à nouveau la méthode des multiplicateurs de Lagrange, nous formons la fonction

$$Y \mathrel{\widehat{=}} -k \sum_{(\ell)} P_\ell \ln P_\ell - a \left[\sum_{(\ell)} P_\ell - 1 \right] - \sum_{i=1}^{n} b_i \left[\sum_{(\ell)} A^i_\ell P_\ell - \overline{A^i} \right] , \tag{D.16}$$

où a et b_i $(i = 1, 2, \ldots, n)$ sont des constantes, et annulons ses dérivées par rapport à chacune des variables P_ℓ. On obtient ainsi les équations

$$\frac{\partial Y}{\partial P_\ell} = -k(\ln P_\ell + 1) - a - \sum_{i=1}^{n} b_i A^i_\ell = 0 , \tag{D.17}$$

---

5. Noter que cette deuxième méthode, plus élégante et plus commode formellement (voir par exemple le complément V.E, paragraphe I.3) présente l'inconvénient de ne pas intégrer d'emblée toutes les propriétés de la grandeur température : si par exemple deux systèmes isolés l'un de l'autre sont en équilibre avec le même thermostat, le premier point de vue permet d'affirmer aussitôt que ces deux systèmes ont même température; c'est moins évident dans le cadre du second point de vue.

6. Nous supposons bien sûr que ces diverses contraintes sont compatibles, ce qui exclut par exemple que l'on cherche à fixer simultanément, à des valeurs quelconques, les énergies moyennes de deux sous-systèmes en contact thermique entre eux (cf. note 5).

7. En pratique, les seules grandeurs physiques que l'on puisse fixer, de façon stricte ou en valeur moyenne, pour un système à l'équilibre macroscopique, sont des constantes du mouvement, c'est-à-dire des observables indépendantes du temps qui commutent avec l'hamiltonien H. Si les n observables A^i sont compatibles entre elles, c'est-à-dire si elles commutent deux à deux, on peut choisir les états (ℓ) comme états propres communs à H et aux A^i, de sorte que A^i_ℓ est alors la valeur propre de A^i correspondant à l'état (ℓ). Si les n grandeurs A^i ne sont pas compatibles entre elles, A^i_ℓ est de façon générale la valeur moyenne $\langle \ell | A^i | \ell \rangle$ de l'observable A^i dans l'état (ℓ), état propre de H.

qui donnent les probabilités cherchées sous la forme

$$P_\ell^M = \mathrm{e}^{-1-\frac{a}{k}}\, \mathrm{e}^{-\frac{1}{k}\sum_{i=1}^{n} b_i A_\ell^i} . \tag{D.18}$$

Les paramètres de Lagrange a et b_i sont déterminés en reportant cette expression dans la condition de normalisation et les contraintes (D.15) :

$$\mathrm{e}^{-1-\frac{a}{k}} \sum_{(\ell)} \exp\left[-\frac{1}{k} \sum_{i=1}^{n} b_i A_\ell^i \right] = 1 , \tag{D.19}$$

$$\mathrm{e}^{-1-\frac{a}{k}} \sum_{(\ell)} A_\ell^i \exp\left[-\frac{1}{k} \sum_{j=1}^{n} b_j A_\ell^j \right] = \overline{A^i} . \tag{D.20}$$

Les probabilités (D.18) sont visiblement positives; on montre aisément (cf. note 3) qu'elles donnent bien un maximum de l'entropie statistique. La solution trouvée est donc acceptable physiquement si elle existe[8].

La distribution canonique (paragraphe 1 ci-dessus) apparaît donc comme un cas particulier de la *« distribution canonique généralisée »* que nous venons d'obtenir. On peut mettre celle-ci sous forme plus compacte en posant

$$\lambda_i \mathrel{\widehat{=}} \frac{b_i}{k} ; \quad i = 1, 2, \ldots, n \tag{D.21}$$

et

$$\widetilde{Z} \mathrel{\widehat{=}} \sum_{(\ell)} \exp\left[-\sum_{i=1}^{n} \lambda_i A_\ell^i \right] . \tag{D.22}$$

Elle s'écrit alors

$$P_\ell^M = \frac{1}{\widetilde{Z}} \exp\left[-\sum_{i=1}^{n} \lambda_i A_\ell^i \right] . \tag{D.23}$$

Nous avons vu au chapitre V un exemple d'une telle distribution : l'énergie et le nombre de particules y fluctuent par échanges avec un réservoir. Ici, ces deux grandeurs sont fixées en valeur moyenne; on vérifie facilement que les multiplicateurs de Lagrange λ_i correspondants sont alors égaux à β et $-\beta\mu$ (μ étant le potentiel chimique)[9].

---

8. Ce sont les équations (D.19)-(D.20) qui peuvent ne pas avoir de solution $\{a ; b_i, i = 1, 2, \ldots, n\}$ pour certains ensembles de données $\{\overline{A^i}\}$ (cf. notes 4 et 6). Mais, lorsqu'elles ont une solution, celle-ci est unique (le raisonnement est analogue à celui de la note 4).

9. Le volume V du système, qui prend des valeurs continues et non pas discrètes, peut néanmoins être traité de façon analogue aux grandeurs A^i.

Supposons que V ne soit pas fixé de façon stricte. A chaque valeur de V est associé le même ensemble d'états propres (ℓ) de l'hamiltonien, dont cependant les caractéristiques, et en particulier les énergies $E_\ell(V)$, dépendent du volume V. Nous noterons

$$\mathrm{d}P_\ell(V) = w_\ell(V)\, \mathrm{d}V \tag{1}$$

la probabilité pour que le volume soit V à $\mathrm{d}V$ près et qu'en outre le système se trouve dans l'état microscopique (ℓ). Cette distribution de probabilités est normalisée :

$$\int_0^{+\infty} \mathrm{d}V \sum_{(\ell)} w_\ell(V) = 1 . \tag{2}$$

Nous cherchons, par la méthode développée dans ce complément, la distribution d'équilibre $w_\ell^M(V)$ qui maximise l'entropie statistique

$$S = -k \int_0^{+\infty} \mathrm{d}V \sum_{(\ell)} w_\ell(V) \ln w_\ell(V) , \tag{3}$$

.../...

b. Fonction de partition généralisée

Les A_ℓ^i étant des caractéristiques du système (valeurs des grandeurs A^i dans les divers états (ℓ)), la formule (D.22) peut être considérée comme définissant une «fonction de partition généralisée»

$$\widetilde{Z}(x;\lambda_1,\lambda_2,\ldots,\lambda_n)\widehat{=}\sum_{(\ell)}\exp\left[-\sum_{i=1}^{n}\lambda_i A_\ell^i\right], \tag{D.24}$$

où x représente les paramètres extérieurs qui continuent à être fixés de façon stricte. Les relations (D.20), destinées en principe à déterminer les paramètres λ_i à partir des données $\overline{A^i}$, peuvent être écrites sous la forme

$$-\frac{\partial}{\partial\lambda_i}\ln\widetilde{Z}(x;\lambda_1,\lambda_2,\ldots,\lambda_n)=\overline{A^i}, \tag{D.25}$$

qui généralise (D.14).

Il est donc particulièrement commode de prendre pour variables (outre les paramètres x fixés de façon stricte), non pas les n valeurs moyennes données $\overline{A^i}$, mais les n multiplicateurs de Lagrange λ_i qu'elles ont permis d'introduire. Dans ce point de vue, la fonction de partition apparaît comme la fonction fondamentale dont

---

compte tenu de la contrainte (2) et des conditions

$$\int_0^{+\infty}\mathrm{d}V\sum_{(\ell)}w_\ell(V)E_\ell(V)=\overline{E}, \tag{4.a}$$

$$\int_0^{+\infty}\mathrm{d}V\sum_{(\ell)}w_\ell(V)V=\overline{V}, \tag{4.b}$$

où $\overline{E}$ et $\overline{V}$ sont données (la définition de l'entropie statistique pour une distribution continue pose quelques problèmes (voir complément I.G, § III), mais ceux-ci n'influent pas sur le raisonnement présenté ici). Nous construisons la fonction

$$Y\widehat{=}-k\int_0^{+\infty}\mathrm{d}V\sum_{(\ell)}w_\ell(V)\ln w_\ell(V)-a\left[\int_0^{+\infty}\mathrm{d}V\sum_{(\ell)}w_\ell(V)-1\right]$$
$$-b\left[\int_0^{+\infty}\mathrm{d}V\sum_{(\ell)}w_\ell(V)E_\ell(V)-\overline{E}\right]-c\left[\int_0^{+\infty}\mathrm{d}V\sum_{(\ell)}w_\ell(V)V-\overline{V}\right]. \tag{5}$$

Le processus de maximisation doit se faire par rapport à l'ensemble des fonctions $w_\ell(V)$ dépendant des indices discrets (ℓ) et de la variable continue V. Il faut donc, en toute rigueur, faire appel à la technique de *dérivation fonctionnelle* (référence 3, p. 934) : il faut dériver la fonction Y par rapport à la fonction $w_\ell(V)$. En fait, cela ne présente ici aucune difficulté, et donne le même résultat que dans le cas purement discret :

$$-k\left[\ln w_\ell(V)+1\right]-a-bE_\ell(V)-cV=0. \tag{6}$$

Nous écrirons la solution sous la forme

$$w_\ell^M(V)=\frac{1}{\widetilde{Z}}\mathrm{e}^{-\beta[E_\ell(V)+pV]}, \tag{7}$$

avec $$\widetilde{Z}(\beta,p)=\int_0^{+\infty}\mathrm{d}V\sum_{(\ell)}\mathrm{e}^{-\beta[E_\ell(V)+pV]}. \tag{8}$$

On retrouve ainsi les formules du complément III.D (§ I.2).

on peut déduire toutes les propriétés macroscopiques du système. On introduit par commodité son logarithme, en définissant par exemple[10]

$$L(x; \lambda_1, \lambda_2, \ldots, \lambda_n) \widehat{=} -\ln \widetilde{Z}(x; \lambda_1, \lambda_2, \ldots, \lambda_n) . \tag{D.26}$$

La différentielle de la fonction L, qui s'écrit par définition

$$dL = \frac{\partial L}{\partial x} dx + \sum_{i=1}^{n} \frac{\partial L}{\partial \lambda_i} d\lambda_i , \tag{D.27}$$

prend d'après (D.25) la forme

$$dL = \frac{\partial L}{\partial x} dx + \sum_{i=1}^{n} \overline{A^i} \, d\lambda_i . \tag{D.28}$$

c. Entropie d'équilibre

Calculons, à partir de sa définition (D.1) et de l'expression (D.23) des probabilités d'équilibre, la valeur de l'entropie statistique dans la situation canonique généralisée :

$$\begin{aligned} S(P_\ell^M) &= k \sum_{(\ell)} P_\ell^M \left[\sum_{i=1}^{n} \lambda_i A_\ell^i + \ln \widetilde{Z} \right] \\ &= k \sum_{i=1}^{n} \lambda_i \overline{A^i} + k \ln \widetilde{Z} , \end{aligned} \tag{D.29}$$

puisque la distribution $\{P_\ell^M\}$ vérifie les contraintes (D.2) et (D.15). Cette formule peut être écrite sous la forme

$$\frac{1}{k} S = \sum_{i=1}^{n} \lambda_i \overline{A^i} - L , \tag{D.30}$$

la fonction L ayant été définie en (D.26)[11].

La relation (D.30) entre S/k et L est appelée une *«transformation de Legendre»*. Elle implique que *les variables les mieux adaptées à S sont (outre x) les* $\overline{A^i}$, et non plus les λ_i. Différentions en effet la relation (D.30) :

$$\frac{1}{k} dS = \sum_{i=1}^{n} \left(\lambda_i \, d\overline{A^i} + \overline{A^i} \, d\lambda_i \right) - dL . \tag{D.31}$$

---

10. Dans les cas particuliers canonique (chapitre III) et grand-canonique (chapitre V), la «fonction de partition généralisée» $\widetilde{Z}$ devient respectivement la fonction de partition ordinaire $Z(V, N, \ldots; \beta)$ et la grande fonction de partition $\Xi(V, \ldots; \beta, \mu)$. On préfère dans ces deux cas multiplier la fonction L par kT, ce qui conduit à l'énergie libre $F(V, N, \ldots; T)$ ou au grand potentiel $J(V, \ldots; T, \mu)$; on procède de même dans le cas de l'«ensemble T-p» (note 9 et complément III.D), la fonction kTL étant alors l'enthalpie libre $G(N, \ldots; T, p)$.

11. Le résultat (D.30) généralise la relation (cf. III, § II.B.3)

$$S^c = \frac{1}{T} \overline{E}^c - \frac{1}{T} F \tag{1}$$

liant, dans la situation canonique, l'entropie S^c à la valeur moyenne $\overline{E}^c$ de l'énergie et à l'énergie libre $F = -kT \ln Z$. Une relation analogue est valable dans la situation grand-canonique (cf. V, § II.B.3) :

$$S^G = \frac{1}{T} (\overline{E}^G - \mu \overline{N}^G) - \frac{1}{T} J , \tag{2}$$

où $J = -kT \ln \Xi$ est le grand potentiel.

Si l'on tient compte de l'expression (D.28) de dL, on trouve aussitôt

$$\frac{1}{k}\,\mathrm{d}S = -\frac{\partial L}{\partial x}\,\mathrm{d}x + \sum_{i=1}^{n} \lambda_i\,\mathrm{d}\overline{A^i}\,. \tag{D.32}$$

La différentielle de S s'exprime donc naturellement comme une forme linéaire en dx et d$\overline{A^i}$ ($i = 1, 2, \ldots, n$). En identifiant avec la formule de définition

$$\mathrm{d}S \widehat{=} \frac{\partial S}{\partial x}\,\mathrm{d}x + \sum_{i=1}^{n} \frac{\partial S}{\partial \overline{A^i}}\,\mathrm{d}\overline{A^i}\,, \tag{D.33}$$

on voit que

$$\frac{1}{k}\,\frac{\partial S}{\partial \overline{A^i}} = \lambda_i\,: \tag{D.34}$$

les multiplicateurs de Lagrange λ_i, variables canoniques généralisées lorsqu'on utilise $\widetilde{Z}$ ou L, sont ici donnés très simplement à partir des variations de l'entropie d'équilibre S provoquées par une modification des données $\overline{A^i}$.

III. DISCUSSION : LES DEUX FORMES DU POSTULAT FONDAMENTAL SONT-ELLES ÉQUIVALENTES?

Nous voulons comparer ici les deux énoncés du postulat fondamental de la mécanique statistique, celui qui a été introduit au chapitre II et la formulation «généralisée» qui fait l'objet de ce complément.

Notons en premier lieu que nous n'avons utilisé l'énoncé «généralisé», au paragraphe II, qu'en restreignant considérablement sa portée : pour les grandeurs devenues fluctuantes alors qu'elles étaient fixées lorsque le système était isolé, *nous avons supposé connues leurs valeurs moyennes.* Il est effectivement difficile de tenir compte de contraintes qui s'exprimeraient de façon moins précise. Surtout, on peut remarquer aussitôt qu'on n'obtiendrait pas la distribution canonique si l'on se donnait par exemple la valeur moyenne du carré de l'énergie, au lieu de celle de l'énergie elle-même[12]. Nous appellerons donc énoncé «généralisé» la forme plus restrictive que l'on applique véritablement, comme au paragraphe II.

Même réduit à cette forme, l'énoncé «généralisé» paraît effectivement, au premier abord, plus général que l'énoncé du chapitre II : ce dernier peut être considéré comme l'application du postulat «généralisé» au cas particulier où toutes les contraintes imposées au système le sont sous forme stricte (cas où le système est isolé); en outre, nous avons obtenu la distribution canonique en utilisant directement la méthode «généralisée», sans avoir à reconstituer un système global isolé pour pouvoir mettre en œuvre le postulat fondamental.

Mais c'est précisément cette facilité déconcertante d'utilisation qui pose problème. Prenons en effet le cas le plus simple, où seule l'énergie fluctue. Dans le cadre de l'énoncé «restreint», il apparaît clairement (III, § I.B.1) que la distribution canonique est valable seulement si le système considéré n'échange de la chaleur qu'avec un *thermostat* de taille suffisante pour que sa température reste invariable. Ici au contraire, il semble qu'on obtienne la distribution canonique dans tous les cas où l'énergie du système considéré fluctue, sans qu'il soit besoin de préciser avec quel

12. Le raisonnement conduirait alors à une probabilité de la forme

$$P'_\ell \propto \exp\left[-bE_\ell^2\right]\,.$$

système extérieur se font les échanges de chaleur : tout système, pourvu qu'il se trouve à l'équilibre statistique, possède une énergie moyenne $\bar{E}$ bien déterminée; les calculs du paragraphe II.1 montrent alors que ses états microscopiques (ℓ) devraient être distribués selon la formule canonique (D.11).

Or ceci est *manifestement erroné*. Pas à la limite thermodynamique, évidemment : si la taille du système est suffisante pour que les fluctuations y soient négligeables, il est équivalent de fixer son énergie de façon stricte ou en valeur moyenne, et on ne peut pas distinguer entre la distribution canonique et la distribution microcanonique. Mais considérons par exemple deux systèmes $\mathcal{S}$ et $\mathcal{S}'$ de taille comparable (macroscopique, mais pas thermodynamique), en contact thermique l'un avec l'autre, l'ensemble $\mathcal{S} \cup \mathcal{S}'$ étant isolé et à l'équilibre. La probabilité P_ℓ pour que le système $\mathcal{S}$ se trouve dans l'état microscopique (ℓ), d'énergie E_ℓ, est donnée par l'application du postulat fondamental au système global isolé : elle vaut (cf. III, § I.B.1)

$$P_\ell = \frac{\Omega'(E_t - E_\ell)}{\Omega_t(E_t)} , \tag{D.35}$$

où E_t et Ω_t sont l'énergie et le nombre d'états accessibles du système global $\mathcal{S} \cup \mathcal{S}'$, et Ω' le nombre d'états accessibles du système $\mathcal{S}'$. Les deux formulations du postulat s'accordent sur le résultat (D.35) puisqu'elles donnent toutes deux comme équiprobables tous les états accessibles du système global isolé. Comme il fallait s'y attendre, la probabilité P_ℓ dépend donc en général de l'autre système $\mathcal{S}'$, avec lequel le système étudié $\mathcal{S}$ interagit. Cette dépendance devient particulièrement simple si $\mathcal{S}'$ est beaucoup plus gros que $\mathcal{S}$; elle est alors caractérisée par un seul paramètre, la température canonique T :

$$\ln \Omega'(E_t - E_\ell) = \ln \Omega'(E_t) - \frac{1}{T} E_\ell \quad \text{si} \quad \mathcal{S}' \gg \mathcal{S} . \tag{D.36}$$

C'est seulement dans ce cas que l'application directe du postulat « généralisé » au système $\mathcal{S}$ donne le résultat correct[13]. Dans tous les autres cas, l'énoncé « généralisé » est en réalité incohérent : suivant qu'on l'applique au système global $\mathcal{S} \cup \mathcal{S}'$ ou au système $\mathcal{S}$ seul, on obtient le résultat correct (D.35) ou un résultat erroné (distribution canonique dans une situation où elle n'est pas valable).

---

13. On peut concrétiser ce raisonnement en prenant l'exemple simple de gaz parfaits. La densité d'états d'un gaz parfait de N particules ayant l'énergie E est de la forme (voir I, § III.B.2.b)

$$\rho(E) \propto E^{3N/2} . \tag{1}$$

La dépendance par rapport à E_ℓ de la probabilité (D.35) est donc ici connue explicitement :

$$P_\ell \propto \exp\left[\frac{3N'}{2} \ln \frac{E_t - E_\ell}{E_t}\right] \tag{2}$$

(nous avons divisé $E_t - E_\ell$ par E_t pour que l'argument du logarithme soit sans dimension). C'est *seulement dans le cas où $\mathcal{S}$ est très petit devant $\mathcal{S}'$* que l'on trouve la distribution canonique

$$P_\ell \propto \exp[-\beta E_\ell] . \tag{3}$$

Il suffit pour cela d'utiliser le fait que $E_\ell \ll E_t$ pour développer le logarithme :

$$\ln \frac{E_t - E_\ell}{E_t} \simeq -\frac{E_\ell}{E_t} , \tag{4}$$

et de remplacer E_t par

$$E_t = \frac{3}{2}(N + N')kT \simeq \frac{3}{2} N'kT . \tag{5}$$

Pour comprendre cette situation ambiguë, il faut distinguer entre probabilités « objectives » et probabilités « subjectives ». Prenons l'exemple du jeu de pile ou face, et supposons que la pièce utilisée, ou son mode de lancement, soit truqué : *objectivement,* les probabilités des deux résultats possibles sont différentes. Pour mesurer ces *probabilités « objectives »,* il faudrait lancer la pièce un grand nombre $\mathcal{N}$ de fois (appendice I, § I.3). Mais ce n'est souvent pas possible. Il faut alors se contenter de *probabilités « subjectives »,* que l'on évalue à partir des renseignements incomplets que l'on possède. Le principe est toujours le même : utiliser toute l'information disponible, mais seulement celle-là. Comme on est incapable de savoir dans quel sens opère le trucage, on ne peut que choisir des probabilités égales pour pile et face : si l'on attribuait une probabilité « subjective » plus grande à l'un des deux résultats qu'à l'autre, on s'exposerait à avoir ainsi favorisé celui qui est défavorisé objectivement. Il est clair que ces probabilités « subjectives » égales ne décrivent pas correctement la réalité, puisqu'elles diffèrent des véritables probabilités « objectives », inconnues mais différentes l'une de l'autre.

Revenons maintenant au postulat fondamental, et reprenons l'analyse de son premier énoncé (cf. II, § I.B.3). En premier lieu, c'est le seul postulat possible pour un système isolé : comme on ne possède aucune information permettant de distinguer les uns des autres les divers états microscopiques accessibles, on ne peut que les considérer comme équiprobables. Ensuite, comme pour tout postulat physique, il faut confronter celui-ci à l'expérience : il s'agit de savoir si les probabilités objectives des états microscopiques accessibles sont bien égales. Ce n'est pas évident *a priori;* les lois de la mécanique, classique ou quantique, suggèrent le contraire : si la préparation du système l'a placé, à l'instant t_0, dans un état microscopique particulier (forcément un état accessible), son état à un instant t ultérieur est parfaitement déterminé par la loi d'évolution dans le temps (relation fondamentale de la dynamique en mécanique classique, équation de Schrödinger en mécanique quantique); on s'attend donc à ce que seuls certains états microscopiques accessibles soient atteints. C'est le *comportement ergodique* des systèmes macroscopiques qui assure la validité du postulat fondamental : contrairement à ce que nous venons d'indiquer, un système macroscopique donné, isolé et à l'équilibre, passe successivement, au cours de son évolution temporelle (« agitation thermique »), par tous les états microscopiques accessibles, et ceci à des échelles de temps très rapides (cf. complément IV.B); il reste autant de temps, en moyenne, dans chacun d'eux.

L'énoncé « généralisé » du postulat fondamental fournit lui aussi des probabilités subjectives. Mais celles-ci, comme nous venons de le montrer, ne coïncident avec les probabilités objectives que dans certains cas. Plus précisément, le postulat « généralisé » ne donne des résultats conformes à la réalité que dans les situations « canoniques » : les grandeurs, telles que l'énergie, le nombre de particules, ... qui fluctuent et dont on se donne la valeur moyenne doivent être échangées, par le système considéré, avec un *réservoir* beaucoup plus gros que lui. C'est seulement dans ce cas que l'influence de l'extérieur sur le système étudié peut être caractérisée par un seul paramètre pour chaque grandeur fluctuante. Dans le cadre de l'énoncé initial du postulat, ce paramètre est la variable conjuguée de la grandeur (température pour l'énergie, potentiel chimique pour le nombre de particules, ...) dont le réservoir impose la valeur au système; pour l'énoncé « généralisé », ce paramètre s'introduit comme le multiplicateur de Lagrange associé à la contrainte imposant la valeur moyenne de la grandeur considérée. Les deux points de vue sont reliés par des formules du type de (D.13)-(D.14), ou plus généralement (D.25).

Lorsqu'on délimite ainsi le domaine d'applicabilité de l'énoncé « généralisé », *les deux formulations du postulat fondamental sont équivalentes.* Chacune d'elles présente pourtant des avantages. Celle du chapitre II est claire et rigoureuse dans son interprétation physique; c'est à elle qu'il faut se référer, comme nous venons de le faire, dans les cas ambigus. La formulation que nous avons décrite dans ce complément

permet de retrouver, par un calcul analytique simple et élégant, les distributions canoniques; cette simplicité est fort utile lorsqu'on cherche à écrire directement l'opérateur-densité d'un système placé dans une situation canonique (complément V.E, § I.3).

Références

1. E. T. Jaynes, *Physical Review 106,* p. 620 (1957) et *108,* p. 171 (1957).

2. A. Katz, *Principles of Statistical Mechanics,* Freeman, San Francisco (1967).

3. J. Bass, *Cours de Mathématiques,* tome II, Masson, Paris (1961).

COMPLÉMENT V.E

Utilisation du formalisme des opérateurs-densités

I. Exemples d'opérateurs-densités
   1. Relation entre les deux formalismes
   2. Les opérateurs-densités canoniques
      a. L'opérateur-densité microcanonique
      b. L'opérateur-densité canonique
      c. L'opérateur-densité grand-canonique
   3. Généralisation

II. Exemples d'utilisation de l'opérateur-densité : méthodes d'approximation
   1. Développement de perturbation
      a. Position du problème
      b. Développement de $\exp(-\beta H)$
      c. Développement en perturbations de la fonction de partition
   2. Méthode variationnelle

Le formalisme des opérateurs-densités permet de caractériser de façon compacte, par un objet théorique unique, l'état macroscopique d'un système quelconque, décrit par la mécanique quantique au niveau microscopique. Nous avons adopté dans cet ouvrage un formalisme plus accessible conceptuellement, bien adapté à la compréhension intuitive de la mécanique statistique et à l'étude de cas simples, mais moins efficace dans les développements théoriques et les applications de la mécanique statistique à des systèmes réels (notamment dans le traitement des interactions entre particules).

La définition et les principales propriétés des opérateurs-densités ont été introduites dans le complément I.H. Nous nous proposons ici d'appliquer ces considérations générales aux *situations canoniques,* et de montrer comment l'opérateur-densité fournit un cadre commode aux *méthodes d'approximation* le plus souvent nécessaires lorsque les interactions entre constituants microscopiques ne sont pas négligeables.

I. EXEMPLES D'OPÉRATEURS-DENSITÉS

1. RELATION ENTRE LES DEUX FORMALISMES

Le formalisme simplifié que nous utilisons principalement dans ce livre suppose connus les états microscopiques stationnaires du système étudié, c'est-à-dire les vecteurs propres $|\ell\rangle$ de son opérateur hamiltonien (appendice II, § II.2) :

$$H|\ell\rangle = E_\ell|\ell\rangle\ , \tag{E.1}$$

où les valeurs propres E_ℓ sont les énergies possibles pour le système considéré. Nous supposons que chacun des kets $|\ell\rangle$ est normé, et qu'ils sont orthogonaux entre eux :

$$\langle\ell|\ell'\rangle = \delta_{\ell\ell'}\ . \tag{E.2}$$

Leur ensemble constitue une base de l'espace des états du système, ce que l'on peut exprimer par la relation de fermeture

$$\sum_{(\ell)} |\ell\rangle\langle\ell| = \mathbb{1}\ . \tag{E.3}$$

L'état macroscopique du système est alors caractérisé par un ensemble $\{P_\ell\}$ de

probabilités, P_ℓ étant la probabilité pour que le système se trouve dans l'état stationnaire $|\ell\rangle$; cet ensemble de probabilités est bien sûr normalisé :

$$\sum_{(\ell)} P_\ell = 1 \ . \tag{E.4}$$

Dans le formalisme des opérateurs-densités, l'état macroscopique caractérisé par la distribution de probabilités $\{P_\ell\}$ est décrit par l'opérateur-densité

$$D = \sum_{(\ell)} P_\ell |\ell\rangle\langle\ell| \ . \tag{E.5}$$

La condition de normalisation (E.4) s'écrit sous la forme

$$\text{Tr}\, D = 1 \ , \tag{E.6}$$

et l'opérateur-densité (E.5) peut être rapporté en principe à n'importe quelle autre base de l'espace des états, dans laquelle il est représenté par une matrice-densité (voir complément I.H).

2. LES OPÉRATEURS-DENSITÉS CANONIQUES

Connaissant les distributions de probabilités $\{P_\ell\}$ associées aux situations microcanonique (II, § I.B.1.b), canonique (III, § I.B.1) et grand-canonique (V, § I.B.1), on déduit facilement, à partir de la formule générale (E.5), les opérateurs-densités correspondants.

a. L'opérateur-densité microcanonique

Lorsque le système considéré est isolé, son énergie E est fixée (à δE près). Ses états accessibles, que nous noterons ici $\{|\ell_a\rangle\}$, sont les états stationnaires $|\ell\rangle$ dont l'énergie E_ℓ est compatible avec cette contrainte, c'est-à-dire tels que

$$E \leqslant E_{\ell_a} \leqslant E + \delta E \ . \tag{E.7}$$

Ces états accessibles sont au nombre de Ω, et le postulat fondamental de la mécanique statistique les suppose tous équiprobables à l'équilibre (II, § I.B.1). La distribution microcanonique $\{P^*_\ell\}$ s'écrit donc

$$\begin{cases} P^*_{\ell_a} = \dfrac{1}{\Omega} & \text{pour tout } |\ell_a\rangle \text{ vérifiant (E.7),} \\ P^*_\ell = 0 & \text{pour tout } |\ell\rangle \text{ non accessible.} \end{cases} \tag{E.8}$$

L'opérateur-densité microcanonique D^*, qui décrit l'état macroscopique d'un système isolé à l'équilibre, a donc d'après (E.5) la forme

$$D^* = \sum_{(\ell_a)} \frac{1}{\Omega} |\ell_a\rangle\langle\ell_a| \ . \tag{E.9}$$

Le nombre $1/\Omega$, qui est le même pour tous les termes, peut être sorti du signe somme; l'opérateur restant,

$$\mathcal{P}_a \mathrel{\widehat{=}} \sum_{(\ell_a)} |\ell_a\rangle\langle\ell_a| \ , \tag{E.10}$$

est le *projecteur sur le sous-espace des états accessibles :* appliqué à une combinaison linéaire quelconque d'états accessibles $|\ell_a\rangle$, il la restitue sans la modifier (c'est-à-dire qu'il se comporte dans ce cas comme l'opérateur-identité); il donne au contraire le ket nul quand on l'applique à une combinaison linéaire quelconque d'états $|\ell\rangle$ non accessibles. En définitive,

$$\boxed{D^* = \frac{1}{\Omega} \mathcal{P}_a} \ . \tag{E.11}$$

Il est facile de s'assurer[1] que cette expression vérifie la condition de normalisation (E.6).

b. L'opérateur-densité canonique

Pour un système en équilibre avec un thermostat de température T, la distribution de probabilité de ses états stationnaires est la distribution canonique (III, § I.B.1)

$$P_\ell^c = \frac{1}{Z} e^{-\beta E_\ell} , \tag{E.12}$$

où $$Z \widehat{=} \sum_{(\ell)} e^{-\beta E_\ell} . \tag{E.13}$$

La formule (E.5) donne alors pour l'opérateur-densité canonique D^c l'expression

$$D^c = \sum_{(\ell)} \frac{1}{Z} e^{-\beta E_\ell} |\ell\rangle\langle\ell| . \tag{E.14}$$

La constante $1/Z$ sort de la somme. Celle-ci peut être mise sous une forme plus compacte si l'on remarque que l'équation aux valeurs propres (E.1) a pour conséquence[2]

$$e^{-\beta H} |\ell\rangle = e^{-\beta E_\ell} |\ell\rangle . \tag{E.15}$$

En effet, si l'on écrit

$$D^c = \frac{1}{Z} \sum_{(\ell)} e^{-\beta H} |\ell\rangle\langle\ell| , \tag{E.16}$$

---

1. Calculons la trace de D^* dans la base $\{|\ell\rangle\}$:

$$\text{Tr } D^* = \sum_{(\ell)} \langle\ell|D^*|\ell\rangle = \sum_{(\ell)} \frac{1}{\Omega} \langle\ell|\mathcal{I}_a|\ell\rangle . \tag{1}$$

Or $$\langle\ell|\mathcal{I}_a|\ell\rangle = 1 \quad \text{si } |\ell\rangle \text{ est accessible,}$$
$$= 0 \quad \text{si } |\ell\rangle \text{ n'est pas accessible.} \tag{2}$$

Donc $$\text{Tr } D^* = \frac{1}{\Omega} \sum_{(\ell_a)} 1 = \frac{1}{\Omega} \times \Omega = 1 \tag{3}$$

car la somme comporte Ω termes.

2. L'exponentielle d'opérateur $\exp(-\beta H)$ peut être définie soit à partir du développement de la fonction exponentielle en série entière :

$$e^{-\beta H} \widehat{=} \mathbb{1} - \beta H + \frac{1}{2!}(\beta H)^2 + \ldots + \frac{(-1)^n}{n!}(\beta H)^n + \ldots , \tag{1}$$

soit en se plaçant dans la base des vecteurs propres de H :

$$e^{-\beta H} |\ell\rangle \widehat{=} e^{-\beta E_\ell} |\ell\rangle . \tag{2}$$

Ces deux définitions possibles sont équivalentes. En effet, par application répétée (n fois) de H sur son état propre $|\ell\rangle$, on constate que

$$H^n |\ell\rangle = (E_\ell)^n |\ell\rangle . \tag{3}$$

La définition (1) donne donc

$$e^{-\beta H} |\ell\rangle = \left[1 - \beta E_\ell + \frac{1}{2!}(\beta E_\ell)^2 + \ldots + \frac{(-1)^n}{n!}(\beta E_\ell)^n + \ldots\right] |\ell\rangle = e^{-\beta E_\ell} |\ell\rangle . \tag{4}$$

on peut sortir $\exp(-\beta H)$, qui ne dépend plus de (ℓ), de la somme :

$$D^c = \frac{1}{Z} e^{-\beta H} \sum_{(\ell)} |\ell\rangle\langle\ell| \,. \tag{E.17}$$

Enfin, comme cette somme porte ici sur *tous les états stationnaires* (et pas seulement sur certains d'entre eux comme en (E.10)), elle s'effectue immédiatement grâce à la relation de fermeture (E.3) :

$$\boxed{D^c = \frac{1}{Z} e^{-\beta H}} \,. \tag{E.18}$$

La définition (E.13) de la fonction de partition Z, qui découle simplement de la normalisation de la distribution $\{P_\ell^c\}$, s'exprime en termes d'opérateur-densité par

$$\boxed{Z \hat{=} \mathrm{Tr}\, e^{-\beta H}} \,, \tag{E.19}$$

relation qui assure effectivement la normalisation de D^c sous la forme (E.6).

Les formules (E.11) et (E.18) sont analogues sur le plan fondamental : toutes deux permettent en principe d'utiliser une base quelconque de l'espace des états. Elles présentent pourtant une différence qui s'avère importante en pratique : l'opérateur-densité canonique (E.18) s'exprime directement à partir de l'hamiltonien H, alors que l'opérateur-densité microcanonique (E.11) fait intervenir le sous-espace des états accessibles, ce qui nécessite la connaissance des états propres de H. Or, lorsqu'on étudie des systèmes dans lesquels les interactions entre particules ne sont pas négligeables, il est fréquent qu'on connaisse avec une bonne approximation l'hamiltonien du système sans pouvoir cependant résoudre complètement son équation aux valeurs propres (E.1). Dans ces conditions, la formule (E.11) est pratiquement inutilisable, alors que (E.18) peut fournir des renseignements pertinents sur le comportement du système (voir paragraphe II ci-dessous).

c. L'opérateur-densité grand-canonique

Dans le cas où le système étudié échange de l'énergie mais aussi des particules avec un réservoir-thermostat de température T et potentiel chimique μ, il faut considérer ses états stationnaires pour tous les nombres de particules possibles, de 0 à $+\infty$. Les kets $|\ell\rangle$ décrivant ces états stationnaires sont alors vecteurs propres communs à l'hamiltonien H et à l'*opérateur nombre de particules*[3], que nous noterons $\hat{N}$:

$$H|\ell\rangle = E_\ell|\ell\rangle \,, \tag{E.20.a}$$

$$\hat{N}|\ell\rangle = N_\ell|\ell\rangle \,. \tag{E.20.b}$$

3. L'espace des états que l'on utilise ici est la somme directe (au sens des espaces vectoriels) des divers espaces d'états correspondant chacun à un nombre de particules fixé.

Les fluctuations du nombre de particules du système sont dues au couplage avec le réservoir-thermostat. Mais comme ce couplage est faible, on peut définir un hamiltonien H pour le système lui-même, et cet hamiltonien conserve le nombre de particules. On peut alors comprendre les égalités (E.20) de la manière suivante. On commence par choisir le nombre de particules du système, ce qui délimite un sous-espace de l'espace des états global; on cherche ensuite les vecteurs propres de H à l'intérieur de ce sous-espace, comme dans les situations microcanonique et canonique; on répète le procédé pour toutes les valeurs possibles du nombre de particules (entre 0 et $+\infty$). Finalement, les vecteurs $|\ell\rangle$ ainsi obtenus correspondent à un nombre de particules déterminé (noté ici N_ℓ), et à une énergie déterminée E_ℓ de ces N_ℓ particules.

H et $\widehat{N}$ étant des observables, l'ensemble des kets $|\ell\rangle$ constitue une base dans l'espace des états correspondant au problème, c'est-à-dire qu'il vérifie la relation de fermeture (E.3).

Nous savons que, à l'équilibre avec le réservoir-thermostat, les probabilités P_ℓ^G des divers états $|\ell\rangle$ du système sont données par (V, § I.B.1)

$$P_\ell^G = \frac{1}{\Xi} e^{-\beta(E_\ell - \mu N_\ell)}, \tag{E.21}$$

où $$\Xi \widehat{=} \sum_{(\ell)} e^{-\beta(E_\ell - \mu N_\ell)} \tag{E.22}$$

est la grande fonction de partition du système.

A partir de ces résultats, la formule (E.5) conduit, par un calcul analogue à celui du paragraphe b ci-dessus[4], aux expressions suivantes pour l'opérateur-densité grand-canonique D^G :

$$\boxed{D^G = \frac{1}{\Xi} e^{-\beta(H - \mu\widehat{N})}} \tag{E.23}$$

et pour la grande fonction de partition :

$$\boxed{\Xi \widehat{=} \mathrm{Tr}\, e^{-\beta(H - \mu\widehat{N})}}. \tag{E.24}$$

Ces formules sont du même type que leurs analogues canoniques : l'opérateur-densité grand-canonique et la grande fonction de partition s'expriment directement à partir des opérateurs H et $\widehat{N}$; il n'est donc pas vraiment indispensable de connaître les états stationnaires $|\ell\rangle$ pour utiliser ces formules. Le choix entre le formalisme canonique et le formalisme grand-canonique (équivalents à la limite thermodynamique; cf. V, § III.B) dépend du problème considéré (voir par exemple le chapitre VI).

3. GÉNÉRALISATION

Au paragraphe précédent, nous avons dans chaque cas déduit l'expression de l'opérateur-densité D de la distribution $\{P_\ell\}$ des états stationnaires. Celle-ci découle

---

4. D'après (E.5) et (E.21), l'opérateur-densité D^G s'écrit

$$D^G = \sum_{(\ell)} \frac{1}{\Xi} e^{-\beta(E_\ell - \mu N_\ell)} |\ell\rangle\langle\ell| . \tag{1}$$

Comme dans la note 2 ci-dessus, les équations (E.20) impliquent

$$e^{-\beta(H - \mu\widehat{N})}|\ell\rangle = e^{-\beta(E_\ell - \mu N_\ell)}|\ell\rangle . \tag{2}$$

On peut alors transformer l'expression (1) en

$$D^G = \frac{1}{\Xi} e^{-\beta(H - \mu\widehat{N})} \sum_{(\ell)} |\ell\rangle\langle\ell| , \tag{3}$$

où l'on reconnaît la relation de fermeture (E.3), de sorte que finalement

$$D^G = \frac{1}{\Xi} e^{-\beta(H - \mu\widehat{N})} . \tag{4}$$

La formule (E.24) donnant la grande fonction de partition Ξ découle ensuite de la condition de normalisation

$$\mathrm{Tr}\, D^G = 1. \tag{5}$$

elle-même de l'application, à chacune des situations considérées, du postulat fondamental de la mécanique statistique. Il est possible d'*obtenir directement l'opérateur-densité à partir du postulat fondamental;* il suffit pour cela d'exprimer celui-ci sous sa forme « généralisée » (complément V.D) et d'utiliser l'expression

$$S(D) = -k \operatorname{Tr}(D \ln D) \tag{E.25}$$

de l'entropie statistique en fonction de l'opérateur-densité D (complément I.H, § IV.1).

Les contraintes qui s'exercent sur le système, et dont nous devons tenir compte dans la recherche du maximum de l'entropie statistique, pourraient *a priori* se présenter sous des formes très diverses. Celles que l'on peut traiter simplement sont de deux types distincts (cf. complément V.D, § II).

Les contraintes du premier type concernent les paramètres extérieurs, fixés à une valeur certaine (éventuellement affectée d'une incertitude, mais non sujette à fluctuations; cf. II, § I.A.1). Ces contraintes sont prises en compte en restreignant l'espace sur lequel portera la recherche du maximum aux états qui les vérifient : l'opérateur densité D cherché ne peut avoir d'éléments de matrice non nuls qu'entre de tels états. Par exemple, pour un système isolé, le volume V, le nombre N de particules, ... et l'énergie E sont fixés (cette dernière à δE près); on n'admet donc que les états correspondant à ces valeurs (états accessibles).

Pour un système en équilibre avec un thermostat, V, N, ... sont encore fixés, mais pas l'énergie; nous avons vu dans le complément V.D (§ II.1) que le postulat « généralisé » doit être appliqué en imposant la valeur moyenne de l'énergie. De façon plus générale, nous allons supposer que le système considéré est en contact avec un certain nombre de réservoirs qui lui imposent la *valeur moyenne de n observables* A^i; en termes d'opérateur-densité, nous écrivons

$$\operatorname{Tr}(A^i D) = \overline{A^i} \text{ fixée}; \quad i = 1, 2, \ldots, n . \tag{E.26}$$

Il faut en outre tenir compte de la condition de normalisation

$$\operatorname{Tr} D = 1 . \tag{E.27}$$

Le problème est donc de trouver l'opérateur-densité D^M, défini dans l'espace des états que délimitent les contraintes du premier type, qui vérifie les contraintes (E.26) et (E.27) et pour lequel l'entropie statistique (E.25) est maximale dans ces conditions[5], [6].

La méthode est encore celle des multiplicateurs de Lagrange (appendice IV) : nous construisons la fonction

$$-k \operatorname{Tr}(D \ln D) - a [\operatorname{Tr} D - 1] - \sum_{i=1}^{n} b_i \left[\operatorname{Tr}(D A^i) - \overline{A^i} \right] \tag{E.28}$$

et écrivons que son accroissement, lorsque D devient $D + \mathrm{d}D$, est nul au premier ordre[7] :

$$\operatorname{Tr} \left\{ \mathrm{d}D \left[-k (\ln D + 1) - a - \sum_{i=1}^{n} b_i A^i \right] \right\} = 0 \quad \text{quel que soit } \mathrm{d}D. \tag{E.29}$$

---

5. Ce problème n'aura évidemment de solution que si chacune des contraintes est possible à réaliser physiquement, et si elles sont compatibles entre elles (cf. complément V.D, notes 4 et 6).

6. Nous avons traité un problème analogue dans le complément V.D (§ II.2). Cependant, nous travaillons ici dans un contexte plus général, puisque c'est l'opérateur-densité que nous cherchons, et non plus les probabilités d'états de base choisis à l'avance : les inconnues sont ici non seulement les valeurs propres de l'opérateur-densité, mais également la base de l'espace des états dans laquelle il est diagonal.

7. L'accroissement $\mathrm{d}D$ est ici un *opérateur*. De façon générale, l'accroissement d'une fonction $f(D)$ quelconque n'est pas donné par la règle ordinaire de dérivation, car D et $\mathrm{d}D$ ne commutent

.../...

Explicitons la trace dans une base orthonormée $\{|q\rangle\}$ quelconque de l'espace des états[8] :

$$\sum_{(q)}\sum_{(q')}\langle q|\mathrm{d}D|q'\rangle\langle q'|\left[-k(\ln D+1)-a-\sum_{i=1}^{n}b_iA^i\right]|q\rangle=0$$

quels que soient $\langle q|\mathrm{d}D|q'\rangle$. (E.30)

On en déduit[9]

$$\langle q'|\left[-k(\ln D+1)-a-\sum_{i=1}^{n}b_iA^i\right]|q\rangle=0 \quad \text{pour tous } |q\rangle \text{ et } |q'\rangle, \tag{E.31}$$

c'est-à-dire l'égalité entre opérateurs

$$-k(\ln D+1)-a-\sum_{i=1}^{n}b_iA^i=0\,. \tag{E.32}$$

---

pas nécessairement : $f(D)$ est définie par son développement en série entière, et l'on a, pour chaque terme de ce développement,

$$\mathrm{d}(D^p)=\sum_{j=0}^{p-1}D^j.\mathrm{d}D.D^{p-j-1}\,. \tag{1}$$

Dans le cas d'une *trace,* cependant, la différenciation redevient simple. Notons tout d'abord que

$$\mathrm{d}[\mathrm{Tr}\,f(D)]=\mathrm{Tr}[\mathrm{d}f(D)]\,, \tag{2}$$

car la base choisie pour expliciter la trace reste inchangée dans l'opération de différenciation. Calculons alors

$$\mathrm{Tr}[\mathrm{d}(D^p)]=\sum_{j=0}^{p-1}\mathrm{Tr}[D^j.\mathrm{d}D.D^{p-j-1}]\,. \tag{3}$$

Par suite de l'invariance de la trace par permutation cyclique, il vient finalement

$$\begin{aligned}\mathrm{Tr}[\mathrm{d}(D^p)]&=\sum_{j=0}^{p-1}\mathrm{Tr}[D^{p-j-1}.D^j.\mathrm{d}D]\\&=p\ \mathrm{Tr}(D^{p-1}.\mathrm{d}D)=p\ \mathrm{Tr}(\mathrm{d}D.D^{p-1})\,.\end{aligned} \tag{4}$$

L'accroissement de la fonction $\mathrm{Tr}(D\ln D)$ s'obtient donc par les règles habituelles de la dérivation, même si D et $\mathrm{d}D$ ne commutent pas.

8. La trace d'un opérateur est la somme de ses éléments de matrice diagonaux :

$$\begin{aligned}&\mathrm{Tr}\left\{\mathrm{d}D\left[-k(\ln D+1)-a-\sum_{i=1}^{n}b_iA^i\right]\right\}\\&\qquad=\sum_{(q)}\langle q|\left\{\mathrm{d}D\left[-k(\ln D+1)-a-\sum_{i=1}^{n}b_iA^i\right]\right\}|q\rangle\,.\end{aligned} \tag{1}$$

Pour faire apparaître explicitement les éléments de matrice $\langle q|\mathrm{d}D|q'\rangle$ de l'opérateur-accroissement $\mathrm{d}D$, il suffit d'intercaler entre $\mathrm{d}D$ et le crochet l'opérateur

$$\sum_{(q')}|q'\rangle\langle q'|\,, \tag{2}$$

qui est égal à l'identité puisque $\{|q\rangle\}$ constitue une base de l'espace des états (relation de fermeture). C'est ainsi que l'on obtient la formule (E.30).

9. Les éléments de matrice de $\mathrm{d}D$ ne sont pas tout à fait quelconques : cet opérateur doit être *hermitique* pour que $D+\mathrm{d}D$ le soit comme D, ce qui implique

$$\langle q|\,\mathrm{d}D|q'\rangle=[\langle q'|\,\mathrm{d}D|q\rangle]^* \tag{1}$$

.../...

L'opérateur-densité D^M cherché est donc de la forme

$$D^M = \mathrm{e}^{-1-a/k} \exp\left[-\sum_{i=1}^{n} \frac{b_i}{k} A^i \right] , \tag{E.33}$$

les paramètres a et b_i ($i = 1, 2, \ldots, n$) étant déterminés par les conditions

$$\mathrm{Tr}\, D^M = 1 , \tag{E.34.a}$$

$$\mathrm{Tr}(A^i D^M) = \overline{A^i} . \tag{E.34.b}$$

Si l'on pose

$$\lambda_i \mathrel{\widehat{=}} \frac{b_i}{k} , \tag{E.35}$$

la formule (E.33) et la condition (E.34.a) donnent

$$D^M = \frac{1}{\widetilde{Z}} \exp\left[-\sum_{i=1}^{n} \lambda_i A^i \right] , \tag{E.36}$$

où la *fonction de partition généralisée* $\widetilde{Z}(\lambda_1, \lambda_2, \ldots, \lambda_n)$ est définie par

$$\widetilde{Z}(\lambda_1, \lambda_2, \ldots, \lambda_n) \mathrel{\widehat{=}} \mathrm{Tr}\left\{ \exp\left[-\sum_{i=1}^{n} \lambda_i A^i \right] \right\} . \tag{E.37}$$

Les paramètres λ_i peuvent, si nécessaire, être calculés à partir des données $\overline{A^i}$ en résolvant les équations

$$\frac{1}{\widetilde{Z}} \mathrm{Tr}\left\{ A^i \exp\left[-\sum_{j=1}^{n} \lambda_j A^j \right] \right\} = \overline{A^i}; \quad i = 1, 2, \ldots, n , \tag{E.38}$$

équations que l'on peut aussi écrire d'après (E.37)

$$-\frac{\partial}{\partial \lambda_i} \ln \widetilde{Z}(\lambda_1, \lambda_2, \ldots, \lambda_n) = \overline{A^i}; \quad i = 1, 2, \ldots, n . \tag{E.39}$$

On montre facilement que l'opérateur-densité D^M trouvé est défini-positif[10], et que l'entropie statistique qui lui correspond est effectivement maximale, et pas

---

(où l'étoile note la conjugaison complexe). Malgré cette relation, *on peut raisonner comme si les éléments de matrice de* dD *étaient tous indépendants.* En effet, les éléments diagonaux sont réels, mais à part cela quelconques (la condition de normalisation de la trace a été incorporée dans la fonction (E.28) et ne doit donc plus être imposée par ailleurs). Quant aux éléments non diagonaux, ils sont complexes; pour deux éléments symétriques liés par la relation (1), on peut choisir indépendamment leur partie réelle et leur partie imaginaire, c'est-à-dire les deux paramètres réels

$$\frac{1}{2}[\langle q | \mathrm{d}D | q' \rangle + \langle q' | \mathrm{d}D | q \rangle] \tag{2.a}$$

et

$$\frac{1}{2i}[\langle q | \mathrm{d}D | q' \rangle - \langle q' | \mathrm{d}D | q \rangle] . \tag{2.b}$$

10. Supposons connus les vecteurs propres de l'opérateur

$$\sum_{i=1}^{n} \lambda_i A^i \tag{1}$$

(où les λ_i ont les valeurs déterminées par les équations (E.38)), c'est-à-dire les solutions de l'équation

$$\left[\sum_{i=1}^{n} \lambda_i A^i \right] | m \rangle = \alpha_m | m \rangle . \tag{2}$$

.../...

seulement stationnaire[11]. De plus, la solution du problème, si elle existe, est unique[12].

Les opérateurs-densités canonique (E.18) et grand-canonique (E.23) sont des cas particuliers de l'*opérateur-densité canonique généralisé* trouvé en (E.36) : le paramètre de température β et le potentiel chimique μ y sont déterminés à partir de la valeur moyenne de l'énergie et de celle du nombre de particules (voir complément V.D, § II). Dans le cas microcanonique, toutes les contraintes sont de type strict et seule subsiste, dans le calcul ci-dessus, la condition de normalisation (E.27) ($n=0$)[13]; c'est le nombre Ω d'états accessibles qui joue alors le rôle de la fonction de partition généralisée $\widetilde{Z}$.

II. EXEMPLES D'UTILISATION DE L'OPÉRATEUR-DENSITÉ : MÉTHODES D'APPROXIMATION

Par rapport au formalisme simplifié utilisant les probabilités P_ℓ des états stationnaires (ℓ), la description de l'état macroscopique d'un système par un opérateur-densité présente l'avantage d'être formellement plus compacte et surtout plus générale. Cet avantage se traduit dans la pratique, nous l'avons déjà souligné, par la *possibilité de choisir dans l'espace des états une base différente* de celle que constituent

---

Les kets $|m\rangle$, qui constituent une base orthonormée dans l'espace des états, sont également les vecteurs propres de D^M :

$$D^M|m\rangle=\frac{1}{\widetilde{Z}}\,e^{-\alpha_m}|m\rangle\,. \tag{3}$$

Dans la base $\{|m\rangle\}$, l'opérateur D^M est représenté par une *matrice diagonale dont les éléments sont réels et positifs,* car les valeurs propres α_m de l'opérateur hermitique (1) sont réelles.

Noter que si les n observables A^i commutent entre elles (et seulement dans ce cas), les kets $|m\rangle$ peuvent être pris comme vecteurs propres communs à ces n observables.

11. Étant donnés deux opérateurs-densités D et D' normalisés quelconques, on a toujours (complément I.H, formule (H.42))

$$S(D)\widehat{=}-k\ \mathrm{Tr}(D\ln D)\leqslant -k\ \mathrm{Tr}(D\ln D')\,, \tag{1}$$

cette inégalité devenant une égalité si et seulement si $D=D'$.

Choisissons alors un opérateur-densité quelconque D vérifiant les contraintes (E.26)-(E.27). L'entropie statistique qui lui correspond est telle que

$$S(D)\leqslant -k\ \mathrm{Tr}(D\ln D^M)\,. \tag{2}$$

Or, d'après (E.36) et les conditions (E.26)-(E.27), le second membre s'écrit

$$-k\ \mathrm{Tr}(D\ln D^M)=k\ \mathrm{Tr}\left\{D\left[\sum_{i=1}^{n}\lambda_i A^i+\ln\widetilde{Z}\right]\right\}=k\left[\sum_{i=1}^{n}\lambda_i\overline{A^i}+\ln\widetilde{Z}\right]$$
$$=S(D^M)\,. \tag{3}$$

Donc $$S(D)\leqslant S(D^M)\,, \tag{4}$$

l'égalité n'ayant lieu que si $D=D^M$.

12. Ce sont les équations (E.38), déterminant les λ_i à partir des données $\overline{A^i}$, qui décident si le problème admet des solutions. Les conclusions sont les mêmes que dans la note 4 du complément V.D, le raisonnement s'appuyant ici sur la relation (1) de la note précédente.

13. Nous avons montré au complément I.H (§ IV.3.b) que c'est bien l'opérateur-densité microcanonique (E.11) qui rend maximum dans ce cas l'entropie statistique.

les états stationnaires (ℓ). Ceci sera particulièrement utile dans les cas où ces derniers ne sont pas exactement connus; nous allons indiquer ici le principe de *méthodes d'approximation* qui peuvent être appliquées dans de tels cas.

1. DÉVELOPPEMENT DE PERTURBATION

a. Position du problème

Lorsqu'on cherche à appliquer la mécanique statistique à des systèmes réels, dans lesquels les interactions entre constituants élémentaires ne sont pas négligeables, on se trouve fréquemment confronté à la situation suivante : l'hamiltonien H du système est connu, au moins de façon approchée, mais son équation aux valeurs propres est trop compliquée pour être soluble. Pour pouvoir mettre en œuvre une méthode d'approximation dite « méthode des perturbations », on cherche alors à mettre l'hamiltonien H sous la forme

$$H = H_0 + H_i \,, \tag{E.40}$$

où l'on sait résoudre l'équation aux valeurs propres de H_0 :

$$H_0 |\ell_0\rangle = E^0_\ell |\ell_0\rangle \,, \tag{E.41}$$

et où H_i est petit devant H_0 :

$$H_i \ll H_0 \,. \tag{E.42}$$

H_0 est couramment appelé *« hamiltonien non perturbé »*, H_i étant alors la *« perturbation »*. Notons que, sauf rares exceptions, H_0 *et* H_i *ne commutent pas.*

La situation (E.40)-(E.42) est somme toute assez courante. L'opérateur H_0 peut représenter l'hamiltonien du système lorsqu'on néglige les interactions entre ses constituants microscopiques (gaz parfait, par exemple), et H_i l'ensemble des termes décrivant ces interactions, à condition que celles-ci soient faibles (c'est le cas pour un gaz réel ordinaire). Si cette condition n'est pas réalisée (gaz d'électrons dans un solide cristallin, par exemple), on essaie d'incorporer dans H_0, sans trop le compliquer pour que son équation aux valeurs propres reste soluble, une partie suffisamment importante des interactions pour que la perturbation résiduelle H_i vérifie l'inégalité (E.42). C'est ce que font notamment les méthodes dites du « champ moyen », qui d'ailleurs négligent ensuite la perturbation résiduelle (cf. compléments III.G à III.K).

Dans le formalisme simple fondé sur l'utilisation des états stationnaires, nous sommes pratiquement désarmés pour dépasser l'« approximation d'ordre zéro » où l'on néglige totalement H_i : on pourrait envisager d'utiliser les valeurs et vecteurs propres approchés de H fournis par la méthode des perturbations de la mécanique quantique; on obtiendrait ainsi le même résultat final qu'au paragraphe c ci-dessous, mais après des calculs beaucoup plus compliqués. On comprend aisément que, pour le problème qui vient d'être posé, la base la plus commode de l'espace des états est l'ensemble $\{|\ell_0\rangle\}$ des vecteurs propres de H_0. Par conséquent le formalisme de l'opérateur-densité, qui permet le choix de cette base, est certainement mieux adapté.

Nous allons supposer que le système étudié est en équilibre avec un thermostat de température T. Nous chercherons donc à évaluer de façon approchée la fonction de partition (cf. § I.2.b)

$$Z \mathrel{\widehat{=}} \mathrm{Tr}\,(e^{-\beta H}) = \sum_{(\ell_0)} \langle \ell_0 | e^{-\beta H} | \ell_0 \rangle \,. \tag{E.43}$$

La principale difficulté provient du fait que H_0 et H_i ne commutent pas, de sorte que

$$e^{-\beta H} \neq e^{-\beta H_0} . e^{-\beta H_i} \,. \tag{E.44}$$

Le développement de $\exp(-\beta H)$ au voisinage de $\exp(-\beta H_0)$ n'est donc pas immédiat.

b. Développement de $\exp(-\beta H)$

La méthode la plus commode pour obtenir ce développement est d'écrire en premier lieu une *équation intégrale* dont l'opérateur

$$u(\beta) \widehat{=} e^{-\beta H} \tag{E.45}$$

est la solution (en tant que fonction de β)[14].

Notons en premier lieu que $u(\beta)$ satisfait à l'équation différentielle[15]

$$\frac{d}{d\beta} u(\beta) = -H u(\beta) . \tag{E.46}$$

L'opérateur (E.45) est la solution de cette équation déterminée par la condition aux limites

$$u(0) = \mathbb{1} . \tag{E.47}$$

Regroupons alors l'équation différentielle (E.46) et la condition (E.47) en une équation intégrale. Plus précisément, nous savons que $u(\beta)$ va être « proche » de $\exp(-\beta H_0)$, c'est-à-dire que

$$\tilde{u}(\beta) \widehat{=} e^{+\beta H_0} u(\beta) \tag{E.48}$$

va être proche de l'identité. C'est donc à l'opérateur $\tilde{u}(\beta)$ que nous allons appliquer le raisonnement, à partir de l'équation différentielle[16]

$$\frac{d}{d\beta} \tilde{u}(\beta) = -e^{\beta H_0} H_i u(\beta) \tag{E.49}$$

complétée par la condition

$$\tilde{u}(0) = \mathbb{1} . \tag{E.50}$$

---

14. Cette équation intégrale est très semblable à celle que vérifie l'« opérateur d'évolution » en mécanique quantique. En effet, l'équation de Schrödinger

$$i\hbar \frac{d}{dt} |\psi(t)\rangle = H |\psi(t)\rangle , \tag{1}$$

qui régit l'évolution dans le temps du ket $|\psi(t)\rangle$ décrivant l'état (pur) d'un système quelconque, se résout formellement de façon très simple si l'hamiltonien H ne dépend pas du temps :

$$|\psi(t)\rangle = e^{-iHt/\hbar} |\psi(0)\rangle . \tag{2}$$

L'opérateur

$$U(t) \widehat{=} e^{-iHt/\hbar} , \tag{3}$$

dit « opérateur d'évolution », a une forme analogue à $u(\beta)$ défini par (E.45) : on passe de $U(t)$ à $u(\beta)$ en remplaçant $it/\hbar$ par le paramètre de température β, qui apparaît ainsi formellement comme un « temps imaginaire ».

15. La dérivation de $\exp(-\beta H)$ par rapport à β ne présente aucune difficulté, comme le montre immédiatement le développement en série entière. Bien entendu, H commute avec $\exp(-\beta H)$, de sorte que l'équation différentielle pourrait aussi bien s'écrire

$$\frac{d}{d\beta} u(\beta) = -u(\beta) . H .$$

16. On a en effet, d'après (E.46) et (E.40),

$$\frac{d}{d\beta} \left[e^{\beta H_0} u(\beta) \right] = e^{\beta H_0} \left[H_0 u(\beta) + \frac{d}{d\beta} u(\beta) \right] = e^{\beta H_0} \left[H_0 u(\beta) - H u(\beta) \right] = e^{\beta H_0} \left[-H_i u(\beta) \right] .$$

L'équation intégrale annoncée s'écrit alors[17]

$$\tilde{u}(\beta) = \mathbb{1} - \int_0^\beta d\beta' \, e^{\beta' H_0} H_i \, u(\beta') \,, \tag{E.51}$$

c'est-à-dire, pour $u(\beta)$ lui-même,

$$u(\beta) = e^{-\beta H_0} - \int_0^\beta d\beta' \, e^{-(\beta-\beta')H_0} H_i \, u(\beta') \,. \tag{E.52}$$

La *solution par itération* de l'équation intégrale (E.52) fournit le développement en perturbations de $u(\beta)$. En changeant simplement de notation, on obtient en effet à partir de (E.52) elle-même

$$u(\beta') = e^{-\beta' H_0} - \int_0^{\beta'} d\beta'' \, e^{-(\beta'-\beta'')H_0} H_i \, u(\beta'') \,. \tag{E.53}$$

Si l'on reporte cette expression dans l'équation intégrale, on la transforme en

$$u(\beta) = e^{-\beta H_0} - \int_0^\beta d\beta' \, e^{-(\beta-\beta')H_0} H_i \, e^{-\beta' H_0} + \int_0^\beta d\beta' \, e^{-(\beta-\beta')H_0} H_i \int_0^{\beta'} d\beta'' \, e^{-(\beta'-\beta'')H_0} H_i \, u(\beta'') \,. \tag{E.54}$$

On peut recommencer le procédé, ce qui donne

$$\begin{aligned} u(\beta) = {} & e^{-\beta H_0} - \int_0^\beta d\beta' \, e^{-(\beta-\beta')H_0} H_i \, e^{-\beta' H_0} \\ & + \int_0^\beta d\beta' \int_0^{\beta'} d\beta'' \, e^{-(\beta-\beta')H_0} H_i \, e^{-(\beta'-\beta'')H_0} H_i \, e^{-\beta'' H_0} \\ & - \int_0^\beta d\beta' \int_0^{\beta'} d\beta'' \int_0^{\beta''} d\beta''' \, e^{-(\beta-\beta')H_0} H_i \, e^{-(\beta'-\beta'')H_0} H_i \, e^{-(\beta''-\beta''')H_0} H_i u(\beta''') \,. \end{aligned} \tag{E.55}$$

Et ainsi de suite. Les trois premiers termes du second membre de (E.55) ne font intervenir que H_0 et H_i, connus. Le premier donne l'approximation d'ordre zéro à $u(\beta)$; le deuxième est du premier ordre en H_i, le troisième du second ordre; l'inconnue $u(\beta''')$ est repoussée au quatrième terme, qui est déjà d'ordre trois en H_i. Pour obtenir davantage de termes sous forme explicite, il suffit de poursuivre l'itération; on voit d'ailleurs déjà sur (E.55) comment se construit le terme général, d'ordre n en H_i :

$$\begin{aligned} (-1)^n \int_0^\beta d\beta_1 \int_0^{\beta_1} d\beta_2 \ldots \int_0^{\beta_{n-1}} d\beta_n \, & e^{-(\beta-\beta_1)H_0} H_i \\ & \times e^{-(\beta_1-\beta_2)H_0} H_i \ldots e^{-(\beta_{n-1}-\beta_n)H_0} H_i \, e^{-\beta_n H_0} \,. \end{aligned} \tag{E.56}$$

c. Développement en perturbations de la fonction de partition

La fonction de partition Z se calcule en prenant la trace de l'opérateur $u(\beta)$. Le développement que nous venons d'obtenir pour $u(\beta)$ se prête particulièrement bien à l'utilisation de la base $\{|\ell_0\rangle\}$.

---

17. Dériver le second membre de (E.51) par rapport à β, borne supérieure de l'intégrale, revient à prendre simplement l'expression à intégrer pour $\beta' = \beta$; d'où la relation (E.49) (le fait qu'il s'agisse d'opérateurs, et non de fonctions ordinaires, ne crée pas de difficultés; il impose seulement de respecter soigneusement l'ordre des facteurs). De plus, l'intégrale s'annulant pour $\beta = 0$, (E.51) vérifie également (E.50).

Prenons en effet, par exemple, le développement tronqué au second ordre, c'est-à-dire l'expression (E.55) dans laquelle nous négligeons le dernier terme. A cette approximation, la fonction de partition s'écrit[18]

$$Z \simeq Z_0 - \sum_{(\ell_0)} \int_0^\beta d\beta' \, e^{-(\beta-\beta')E^0_\ell} \langle \ell_0 | H_i | \ell_0 \rangle e^{-\beta' E^0_\ell} + \sum_{(\ell_0)} \sum_{(\ell'_0)} \int_0^\beta d\beta' \int_0^{\beta'} d\beta'' \, e^{-(\beta-\beta')E^0_\ell} \langle \ell_0 | H_i | \ell'_0 \rangle \, e^{-(\beta'-\beta'')E^0_{\ell'}} \times \langle \ell'_0 | H_i | \ell_0 \rangle \, e^{-\beta'' E^0_\ell} , \qquad \text{(E.57)}$$

où $$Z_0 \mathrel{\widehat{=}} \mathrm{Tr}\,(e^{-\beta H_0}) = \sum_{(\ell_0)} e^{-\beta E^0_\ell} \qquad \text{(E.58)}$$

est la fonction de partition à l'ordre zéro. Dans la formule (E.57), tous les termes sont en principe connus : ils ne font intervenir que les énergies non perturbées E^0_ℓ et les éléments de matrice de la perturbation H_i entre états non perturbés $|\ell_0\rangle$. Comme tous ces termes sont maintenant des nombres et non plus des opérateurs, et que les éléments de matrice de H_i sont indépendants de β, on peut effectuer explicitement les intégrales, ce qui donne[19]

$$Z \simeq Z_0 - \beta \sum_{(\ell_0)} e^{-\beta E^0_\ell} \langle \ell_0 | H_i | \ell_0 \rangle + \sum_{(\ell_0)} e^{-\beta E^0_\ell} \left\{ \frac{1}{2}\beta^2 \sum_{(\ell'_0) \,:\, E^0_{\ell'} = E^0_\ell} |\langle \ell_0 | H_i | \ell'_0 \rangle|^2 + \sum_{(\ell'_0) \,:\, E^0_{\ell'} \neq E^0_\ell} |\langle \ell_0 | H_i | \ell'_0 \rangle|^2 \, \frac{\beta}{E^0_{\ell'} - E^0_\ell} \right\} \qquad \text{(E.59)}$$

---

18. La trace de $u(\beta)$, somme de ses éléments de matrice diagonaux, est ici calculée dans la base $\{|\ell_0\rangle\}$:

$$Z = \mathrm{Tr}\; u(\beta) = \sum_{(\ell_0)} \langle \ell_0 | u(\beta) | \ell_0 \rangle . \qquad (1)$$

Les deux premiers termes de (E.57) s'en déduisent aussitôt, compte tenu du fait que

$$e^{-\beta H_0} | \ell_0 \rangle = e^{-\beta E^0_\ell} | \ell_0 \rangle \qquad (2.a)$$

et $$\langle \ell_0 | e^{-\beta H_0} = e^{-\beta E^0_\ell} \langle \ell_0 | . \qquad (2.b)$$

Dans le troisième terme, on intercale entre les deux opérateurs H_i la relation de fermeture de la base $\{|\ell_0\rangle\}$:

$$\langle \ell_0 | H_i \, e^{-(\beta'-\beta'')H_0} H_i | \ell_0 \rangle = \langle \ell_0 | H_i \left\{ \sum_{(\ell'_0)} |\ell'_0\rangle\langle\ell'_0| \right\} e^{-(\beta'-\beta'')H_0} H_i | \ell_0 \rangle = \sum_{(\ell'_0)} \langle \ell_0 | H_i | \ell'_0 \rangle \, e^{-(\beta'-\beta'')E^0_{\ell'}} \langle \ell'_0 | H_i | \ell_0 \rangle . \qquad (3)$$

19. Dans le deuxième terme de (E.57), les facteurs $\exp(\pm\beta' E^0_\ell)$ se compensent, faisant disparaître toute dépendance en β'. Dans le troisième terme, la fonction à intégrer s'écrit

$$e^{-\beta E^0_\ell} |\langle \ell_0 | H_i | \ell'_0 \rangle|^2 \, e^{\beta'(E^0_\ell - E^0_{\ell'})} \, e^{\beta''(E^0_{\ell'} - E^0_\ell)} . \qquad (1)$$

Deux cas sont à considérer. L'état (ℓ_0) étant fixé (on sommera ensuite sur (ℓ_0)), (ℓ'_0) peut être tel que

$$E^0_{\ell'} = E^0_\ell \qquad (2)$$

.../...

Nous avons seulement esquissé ici le principe de la méthode des perturbations. Reste, dans chaque problème particulier, à évaluer les éléments de matrice de la perturbation H_i, puis les sommes sur les états non perturbés ($\mathscr{E}_0$). C'est en général une tâche ardue, dont la difficulté croît très rapidement avec l'ordre du terme considéré. Des techniques, notamment diagrammatiques, ont été développées pour systématiser et faciliter les calculs (voir par exemple réf. 1 et 2). Malgré cela, on doit le plus souvent, dans les applications concrètes, se contenter des tout premiers termes du développement, car l'évaluation des termes d'ordre supérieur s'avère rapidement inextricable. Mais, si H_i est suffisamment petit[20], les corrections d'ordre croissant sont numériquement de plus en plus faibles, et les premiers termes donnent souvent une bonne approximation de la fonction de partition du système. Reste le problème, difficile voire impossible à analyser rigoureusement, de la convergence de la série de perturbations.

2. MÉTHODE VARIATIONNELLE

Outre le développement de perturbation, dont nous avons esquissé au paragraphe 1 l'application à la mécanique statistique, il existe en physique une autre

---

si le niveau E^0_ℓ est dégénéré; dans ce cas, la fonction (1) devient indépendante de β' et β'', de sorte que l'intégration donne simplement $\beta^2/2$. Pour les états ($\mathscr{E}'_0$) d'énergie différente de E^0_ℓ, l'intégrable double

$$\int_0^\beta d\beta'\, e^{\beta'(E^0_{\ell'}-E^0_\ell)}\int_0^{\beta'} d\beta''\, e^{\beta''(E^0_{\ell''}-E^0_\ell)} \tag{3}$$

se calcule facilement elle aussi :

$$\int_0^\beta d\beta'\, e^{\beta'(E^0_{\ell'}-E^0_\ell)}\frac{1}{E^0_{\ell'}-E^0_\ell}\left[e^{\beta'(E^0_{\ell''}-E^0_\ell)}-1\right] = \frac{1}{E^0_{\ell'}-E^0_\ell}\left\{\beta-\frac{1}{E^0_\ell-E^0_{\ell''}}\left[e^{\beta(E^0_{\ell''}-E^0_{\ell'})}-1\right]\right\}. \tag{4}$$

Le terme $\frac{1}{(E^0_{\ell'}-E^0_\ell)^2}\cdot[e^{\beta(E^0_{\ell''}-E^0_{\ell'})}-1]$ multiplié par $e^{-\beta E^0_{\ell''}}$ et sommé sur ($\mathscr{E}_0$) et ($\mathscr{E}'_0$) donne un résultat nul, d'où la formule (E-59).

20. Comme H_i et H_0 sont des opérateurs, la condition (E.42) demande à être précisée. En examinant la formule finale (E.59), on constate facilement que les éléments de matrice de H_i doivent au moins vérifier les *deux conditions* suivantes :

$$|\langle\mathscr{E}_0|H_i|\mathscr{E}'_0\rangle| \ll kT\,, \tag{1}$$

$$|\langle\mathscr{E}_0|H_i|\mathscr{E}'_0\rangle| \ll |E^0_{\ell'}-E^0_\ell|\,. \tag{2}$$

La première de ces conditions assure en effet que les nombres sans dimension $\beta|\langle\mathscr{E}_0|H_i|\mathscr{E}'_0\rangle|$, qui déterminent la correction du premier ordre et apparaissent également dans celle du deuxième ordre (ainsi que dans les termes d'ordre supérieur), sont petits devant 1. Notons que, la perturbation H_i et les états non perturbés ($\mathscr{E}_0$) étant donnés, la condition (1) ne peut être satisfaite qu'à *température suffisamment haute* (l'ordre de grandeur des températures acceptables dépendant évidemment du problème).

Quant à la condition (2), qui est l'expression correcte de (E.42), la dernière somme de la formule (E.59) en montre clairement la nécessité, puisque y apparaissent les termes

$$\frac{|\langle\mathscr{E}_0|H_i|\mathscr{E}'_0\rangle|^2}{(E^0_{\ell'}-E^0_\ell)^2} \tag{3}$$

et $$\beta|\langle\mathscr{E}_0|H_i|\mathscr{E}'_0\rangle|\times\frac{|\langle\mathscr{E}_0|H_i|\mathscr{E}'_0\rangle|}{E^0_{\ell'}-E^0_\ell}\,. \tag{4}$$

Remarquons que ce sont les *différences d'énergies non perturbées*, et non pas ces énergies elles-mêmes, qui fixent les limites imposées aux éléments de matrice de la perturbation.

classe de techniques d'approximation, dites « méthodes variationnelles ». Une telle méthode peut également être utilisée en mécanique statistique.

Reprenons en effet l'inégalité (H.42) du complément I.H : quels que soient les opérateurs-densités D et D' (normalisés), ils vérifient

$$S(D) \leqslant -k \text{ Tr } (D \ln D') , \tag{E.60}$$

l'égalité se produisant si et seulement si D et D' coïncident. Si l'on prend pour D' l'opérateur-densité canonique D^c du système considéré (formule (E.18)), cette inégalité s'écrit[21]

$$S(D) \leqslant \frac{1}{T} \overline{E}(D) + k \ln Z , \tag{E.61}$$

où $$\overline{E}(D) \widehat{=} \text{Tr } (DH) \tag{E.62}$$

est la valeur moyenne de l'énergie du système lorsque son état est décrit par l'opérateur-densité D. Nous écrirons la relation (E.61) sous la forme

$$\overline{E}(D) - TS(D) \geqslant F , \tag{E.63}$$

où $$F \widehat{=} -kT \ln Z \tag{E.64}$$

est l'énergie libre du système étudié, à l'équilibre canonique avec un thermostat de température T.

Explicitons alors la signification de l'inégalité (E.63). Le nombre de particules N, le volume V, ... du système sont fixés, ainsi que la température T imposée par le thermostat. Dans ces conditions, l'énergie libre F du système a une valeur bien déterminée. Mais, si nous ne savons pas résoudre exactement le problème, cette valeur nous est inconnue. On sait cependant qu'elle peut en principe être obtenue comme valeur minimale[22] du premier membre de (E.63).

Ces considérations servent de base à la méthode variationnelle. On choisit une famille d'opérateurs-densités D dont la forme paraît plausible du point de vue physique tout en étant suffisamment simple pour permettre le calcul du premier membre de (E.63). On prend alors pour valeur approchée de l'énergie libre la valeur minimale obtenue pour cette expression dans la famille d'opérateurs-densités choisie. Dans la pratique, cette « famille d'essai » est constituée par un ensemble d'opérateurs-densités dépendant d'un ou plusieurs paramètres ajustables, par rapport auxquels on minimise le premier membre de l'inégalité (E.63).

La méthode variationnelle est donc moins systématique et plus souple que la méthode des perturbations. Elle fait davantage appel à l'intuition physique, dans le choix de la famille d'essai (ce choix, en principe totalement libre, est limité en fait par les difficultés de calcul). Elle se prête encore plus difficilement à une évaluation de la validité de l'approximation qu'elle fournit.

---

21. L'expression (E.18) de D^c donne en effet

$$\ln D^c = -\beta H - \ln Z , \tag{1}$$

de sorte que

$$\begin{aligned} \text{Tr } (D \ln D^c) &= -\beta \text{ Tr } (DH) - (\text{Tr } D) \ln Z \\ &= -\beta \overline{E}(D) - \ln Z \end{aligned} \tag{2}$$

car D est normalisé selon (E.27).

22. Ce minimum, unique, se produit lorsque D est égal au véritable opérateur-densité canonique D^c (cf. formule III (II.27)).

Références

1. « Landau et Lifshitz, course of theoretical Physics », *Statistical Physics,* deuxième partie, par E.M. Lifshitz et L.P. Pitaevskii, Pergamon Press, Oxford (1980).

2. A. Abrikosov, L. Gor'kov et I. Dzyaloshinskii, *Quantum Field Theoretical Methods in Statistical Physics,* Pergamon Press, Oxford (1965).

COMPLÉMENT V.F

Fluctuations dans un système à l'équilibre statistique

I. Notions générales sur les fluctuations
1. Introduction : les variables internes et leurs fluctuations
2. Fluctuations des variables d'origine « mécanique »
3. Le problème des variables d'origine statistique

II. Fluctuations dans un fluide simple
1. Fluctuations locales du nombre de particules par unité de volume
a. Fluctuations du nombre de particules d'un sous-système
b. Fluctuations du volume d'un sous-système
c. Fluctuations locales de la densité
2. Fluctuations locales de l'énergie et de la température
a. Fluctuations de l'énergie d'un sous-système
b. « Fluctuations » locales de la température

I. NOTIONS GÉNÉRALES SUR LES FLUCTUATIONS

1. INTRODUCTION : LES VARIABLES INTERNES ET LEURS FLUCTUATIONS

Dans les différentes descriptions que nous avons données d'un système physique à l'équilibre, nous avons toujours distingué soigneusement entre les *paramètres extérieurs,* fixés, imposés en quelque sorte au système par l'opérateur, et les *variables internes* susceptibles de s'ajuster pour assurer l'équilibre statistique. Cette distinction dépend évidemment de la situation concrète dans laquelle est placé le système : une grandeur donnée peut être paramètre extérieur dans telle situation, variable interne dans telle autre. L'énergie, par exemple, est un paramètre extérieur lorsque le système est isolé; elle devient variable interne si le même système est mis en contact thermique avec un thermostat.

Bien que fixés, les paramètres extérieurs n'en sont pas pour autant connus avec une précision parfaite; leur valeur est toujours entachée d'une *incertitude* dont l'importance dépend du soin avec lequel le système a été préparé. Ces incertitudes se répercutent évidemment sur les variables internes, en particulier sur leur « valeur d'équilibre » (valeur la plus probable). Mais les variables internes sont en outre sujettes à des *fluctuations intrinsèques :* ce sont des variables aléatoires décrites par une *distribution de probabilité* même si l'on suppose que les paramètres extérieurs sont fixés exactement.

La connaissance de la distribution de probabilité d'une variable interne permet de calculer sa valeur la plus probable et sa valeur moyenne, mais elle permet aussi d'analyser ses fluctuations, c'est-à-dire ses écarts par rapport à l'une ou l'autre de ces valeurs. Pour des systèmes contenant un grand nombre de particules ($N \gg 1$) que nous considérerons ici, la valeur la plus probable y_m d'une variable interne y coïncide pratiquement avec sa valeur moyenne $\bar{y}$ (II, § II.A.2; III, § II.B.2; V, § II.B.2). Il est alors commode de caractériser les fluctuations de cette variable par *l'écart quadratique moyen* de sa distribution :

$$\Delta y \widehat{=} \left[\overline{(y-\bar{y})^2}\right]^{1/2} = \left[\overline{y^2} - \bar{y}^2\right]^{1/2} . \qquad \text{(F.1)}$$

Nous avons montré (II, note 16), et nous le vérifierons sur les exemples du paragraphe II, que le rapport entre l'écart quadratique moyen Δy et la valeur moyenne

$\bar{y}$ d'une variable interne quelconque se comporte (pour $N \gg 1$) comme $N^{-1/2}$; plus précisément,

$$\frac{\Delta y}{\bar{y}} = \frac{C}{\sqrt{N}}, \tag{F.2}$$

où C est en première approximation indépendant de N mais dépend du système *et* de la variable y considérée. Les fluctuations sont donc faibles en valeur relative, et s'annulent à la limite thermodynamique ($N \to \infty$); mais, les systèmes réels étant finis, elles existent. Il est donc intéressant de les étudier, éventuellement de les mesurer, et cela pour répondre à deux types de questions :

(i) A partir de quelles valeurs de N un objet devient-il un « système thermodynamique » dans lequel on peut négliger les fluctuations? Quelle est alors la précision de la description thermodynamique? La réponse est-elle indépendante de la grandeur physique à laquelle on s'intéresse?

(ii) Existe-t-il des situations où la « constante » C de la formule (F.2) peut devenir très grande, c'est-à-dire des cas où les fluctuations deviennent importantes même pour un système macroscopique?

Nous allons rappeler dans ce complément comment on peut calculer les « constantes » caractéristiques C et donner quelques exemples concrets de tels calculs au paragraphe II. La réponse aux questions (i) en découle immédiatement. Nous verrons également sur un exemple que les fluctuations peuvent devenir anormalement grandes (question (ii)) au voisinage de ce que l'on appelle un « point critique ».

2. FLUCTUATIONS DES VARIABLES D'ORIGINE « MÉCANIQUE »

Telles que nous les avons définies au paragraphe précédent, les variables internes susceptibles de fluctuer dans un système à l'équilibre sont les grandeurs, que nous qualifierons de « mécaniques », qui prennent une valeur déterminée dans chaque état microscopique du système. Par exemple, l'énergie, le volume, le nombre de particules, mais aussi le moment magnétique ou la quantité de mouvement d'un système sont des variables d'origine « mécanique », c'est-à-dire qu'elles peuvent être définies pour un système quelconque, décrit par la mécanique quantique ou la mécanique classique, sans faire appel à aucune notion de mécanique statistique ou de thermodynamique.

Soit y_ℓ la valeur de la variable interne y lorsque le système est dans l'état microscopique (ℓ). Si P_ℓ est la probabilité de cet état, la distribution de probabilité pour la variable y est donnée par (II, § II.A, III, § I.C, V, § I.C)

$$P(y) = \sum_{(\ell) \in \mathcal{L}(y)} P_\ell , \tag{F.3}$$

où $\mathcal{L}(y)$ est le sous-ensemble des états microscopiques (ℓ), défini par

$$(\ell) \in \mathcal{L}(y) \quad \text{ssi} \quad y_\ell = y . \tag{F.4}$$

Connaissant $P(y)$, on en déduit la valeur moyenne $\bar{y}$ et l'écart quadratique moyen Δy par les formules usuelles. La distribution de probabilité $P(y)$ dépend évidemment de la situation physique où se trouve le système : il faut prendre pour P_ℓ la distribution microcanonique (chapitre II) si le système est isolé (E, V, N, ... fixés), la distribution canonique (chapitre III) si le système est en équilibre thermique avec un thermostat (T, V, N, ... fixés), la distribution T-p (complément III.D) si le système est en équilibre thermique et mécanique avec un réservoir d'énergie et de volume (T, p, N, ... fixés), la distribution grand-canonique (chapitre V) si le système est en équilibre thermique et « chimique » avec un réservoir d'énergie et de particules (T, V, μ, ... fixés), etc.

Si on néglige les fluctuations (limite thermodynamique), les diverses distributions P_ℓ donnent toutes les mêmes résultats, c'est-à-dire les mêmes relations entre grandeurs physiques, qu'elles soient paramètres extérieurs ou valeurs moyennes de variables

internes. Mais les fluctuations autour des valeurs moyennes diffèrent radicalement suivant la situation dans laquelle est placé le système. Pour de grands systèmes ($N \gg 1$), nous avons montré que $P(y)$ peut être approximé par une gaussienne; à cette approximation, la distribution de la variable interne y est entièrement caractérisée par sa valeur moyenne $\bar{y}$ et son écart quadratique moyen Δy.

En pratique, on étudie le plus souvent les *fluctuations locales* de diverses grandeurs physiques : on délimite par la pensée, dans un système $\mathcal{S}$ globalement à l'équilibre, un sous-système Σ de petite taille et on analyse ses échanges fluctuants avec le reste du système $\mathcal{S}$, qui joue pour lui le rôle de réservoir. Par exemple, si l'on définit Σ en découpant dans $\mathcal{S}$ un petit volume V fixe, l'énergie E et le nombre de particules N de Σ seront des variables internes fluctuantes dont la distribution de probabilité sera déterminée par la température T et le potentiel chimique μ de $\mathcal{S}$ (distribution grand-canonique). Ou bien on caractérisera un sous-système Σ' en fixant son nombre de particules N; son volume V sera alors une variable interne comme son énergie E, et $\mathcal{S}$ lui imposera sa pression p et sa température T (distribution T-p).

3. LE PROBLÈME DES VARIABLES D'ORIGINE STATISTIQUE

Tout au long de ce livre, nous avons étudié certaines grandeurs physiques qui sont directement définies à partir des caractéristiques macroscopiques d'un système à l'équilibre. Ces grandeurs, telles que l'entropie, la température, le potentiel chimique, l'énergie libre, le grand potentiel, ..., n'ont pas de sens au niveau microscopique, c'est-à-dire pour un système qui se trouverait dans un état microscopique (ℓ) déterminé; elles sont d'origine fondamentalement statistique : elles n'existent que dans la mesure où un système macroscopique parcourt très rapidement tout un ensemble d'états microscopiques (ℓ). Leur définition même exclut qu'on leur attribue une distribution de probabilité; en toute rigueur cela n'a donc aucun sens de parler de leurs fluctuations.

Pourtant on sait bien (ou tout au moins on conçoit intuitivement) que, si on place un thermomètre de faibles dimensions dans un système à l'équilibre et si ce thermomètre est suffisamment précis, ses indications vont fluctuer au cours du temps autour d'une valeur moyenne égale à la température du système; autrement dit ce thermomètre met en évidence et mesure des « fluctuations locales de température ». De façon analogue, un manomètre de petite taille indiquera des « fluctuations locales de pression » dans un fluide à l'équilibre. Dans ce dernier cas, cependant, il n'y a pas nécessairement de contradiction : si la pression peut être définie directement de façon statistique (II, § III.B.1), elle peut aussi être considérée comme la valeur moyenne de la force fluctuante qui s'exerce, par unité de surface, sur la paroi sensible du manomètre (complément II.D, § I.3).

En réalité, les fluctuations d'une variable d'origine statistique enregistrées par un appareil de mesure ne font que traduire les fluctuations des variables mécaniques qui définissent l'état macroscopique de cet appareil. Pour préciser, considérons par exemple un sous-système Σ' de petite taille caractérisé par le nombre N, fixé, de particules qu'il contient. Son énergie E et son volume V, qui sont des grandeurs « mécaniques », fluctuent par suite des échanges avec le reste du système. A chaque couple de valeurs (E, V) on peut associer une température T et une pression p déterminées : ce sont, par définition, *la température et la pression microcanoniques* qu'aurait le système Σ' s'il était isolé et à l'équilibre, avec une énergie et un volume fixés aux valeurs E et V; cette température T et cette pression p sont données par les dérivées de l'entropie microcanonique S^* du système Σ' :

$$\frac{1}{T} \widehat{=} \frac{\partial S^*(E, V, N)}{\partial E}, \tag{F.5.a}$$

$$\frac{p}{T} \widehat{=} \frac{\partial S^*(E, V, N)}{\partial V}. \tag{F.5.b}$$

Lorsque E et V varient au cours du temps, par suite de leurs fluctuations, T et p varient également selon les relations (F.5). Évidemment ces valeurs « instantanées » de p et T n'ont de sens que si le système Σ' peut être considéré comme étant, à chaque instant, à l'équilibre microcanonique caractérisé par les valeurs E et V, c'est-à-dire si le temps d'établissement de l'équilibre local à l'intérieur de Σ' est beaucoup plus court que le temps caractéristique des échanges d'énergie et de volume entre Σ' et le reste du système. Si l'on parvient à construire des appareils de mesure (thermomètre ou manomètre) sensibles à ces valeurs « instantanées » de la température et de la pression, ils mettront en évidence des « fluctuations locales » de ces grandeurs statistiques; mais ces « fluctuations » seront un simple reflet des fluctuations (réelles, celles-là) des grandeurs « mécaniques » associées, énergie et volume.

Le problème des fluctuations des variables d'origine statistique a donné lieu à de longues controverses. Cette incompréhension entre différents auteurs provient en fait d'une différence de points de vue.

— Dans le premier, on s'en tient aux définitions de la mécanique statistique, basées sur la notion d'ensemble de Gibbs (I, § IV.B.2). Les variables d'origine statistique ne peuvent pas fluctuer. Par exemple, lorsque le système ou le sous-système considéré est en contact avec un réservoir d'énergie et de particules, sa température et son potentiel chimique sont fixés; ce sont des paramètres extérieurs (ou des multiplicateurs de Lagrange; cf. complément V.D, § II) caractérisant la distribution de probabilité des états microscopiques.

— Dans le second point de vue, on se réfère aux indications données par un appareil de mesure. Ces indications varient à cause des fluctuations des variables mécaniques dans le voisinage immédiat de l'appareil. Ce sont ces fluctuations réelles des variables mécaniques que l'on mesure sous forme de « fluctuations » des variables statistiques.

C'est ce second point de vue (à notre sens plus physique puisqu'on peut le confronter, au moins en principe, à l'expérience) que nous adopterons au paragraphe II. Il n'est en fait pas contradictoire avec le premier si l'on comprend bien sa signification et que l'on se garde d'étendre de façon abusive son interprétation.

II. FLUCTUATIONS DANS UN FLUIDE SIMPLE

Pour illustrer les notions précédentes, prenons l'exemple d'un fluide homogène constitué de molécules d'une seule espèce. Nous supposons qu'il contient suffisamment de particules pour qu'on puisse le considérer comme un système thermodynamique $\mathcal{S}$; nous allons calculer les *fluctuations locales* de quelques grandeurs caractéristiques de ce système.

1. FLUCTUATIONS LOCALES DU NOMBRE DE PARTICULES PAR UNITÉ DE VOLUME

a. Fluctuations du nombre de particules d'un sous-système

Considérons, à l'intérieur du système $\mathcal{S}$, un *petit volume V fixe*. Les particules qu'il contient constituent un sous-système Σ; celui-ci est en contact thermique et chimique avec le reste du système $\mathcal{S}$, qui joue pour Σ le rôle de réservoir d'énergie et de particules et lui impose sa température T et son potentiel chimique μ.

La probabilité pour que, à l'équilibre, Σ soit dans l'état microscopique (ℓ), d'énergie E_ℓ et de nombre de particules N_ℓ, est donnée par (V, formule (I.15))

$$P_\ell = \frac{1}{\Xi} e^{-(E_\ell - \mu N_\ell)/kT}, \tag{F.6}$$

où Ξ est la grande fonction de partition du système Σ :

$$\Xi \widehat{=} \sum_{(\ell)} \mathrm{e}^{-(E_\ell - \mu N_\ell)/kT} . \tag{F.7}$$

La valeur moyenne du nombre de particules contenues dans le volume V s'en déduit immédiatement :

$$\overline{N} \widehat{=} \frac{1}{\Xi} \sum_{(\ell)} N_\ell \, \mathrm{e}^{-(E_\ell - \mu N_\ell)/kT} = \frac{kT}{\Xi} \frac{\partial \Xi}{\partial \mu} ; \tag{F.8}$$

de même,

$$\overline{N^2} = \frac{1}{\Xi} \sum_{(\ell)} N_\ell^2 \, \mathrm{e}^{-(E_\ell - \mu N_\ell)/kT} = \frac{(kT)^2}{\Xi} \frac{\partial^2 \Xi}{\partial \mu^2} . \tag{F.9}$$

En dérivant les deux membres de l'expression (F.8) par rapport à μ, on obtient

$$\frac{\partial \overline{N}}{\partial \mu} = \frac{kT}{\Xi} \frac{\partial^2 \Xi}{\partial \mu^2} - \frac{kT}{\Xi^2} \left(\frac{\partial \Xi}{\partial \mu} \right)^2 , \tag{F.10}$$

d'où l'on tire, en utilisant (F.8) et (F.9),

$$\frac{\partial \overline{N}}{\partial \mu} = \frac{1}{kT} \left[\overline{N^2} - \overline{N}^2 \right] . \tag{F.11}$$

La fluctuation du nombre de particules dans Σ est donc caractérisée par l'écart quadratique moyen

$$\Delta N = \left[kT \left(\frac{\partial \overline{N}}{\partial \mu} \right)_{T, V} \right]^{1/2} . \tag{F.12}$$

$(\Delta N)^2$ étant par définition positif, on retrouve (III, § II.D.3.b) la condition de stabilité de l'équilibre qui impose à $\partial N / \partial \mu$ d'être positif ou, ce qui revient au même,

$$\left(\frac{\partial \mu}{\partial N} \right)_{T, V} > 0 . \tag{F.13}$$

Dans l'exemple simple d'un gaz parfait, N et μ sont reliés par (III, formule (IV.47) et complément III.B, formule (B.12))

$$\mu_{GP} = kT \left[\ln \frac{N}{V} + f(T) \right] , \tag{F.14}$$

où $f(T)$ est une fonction de la température seule (et de la nature du gaz). L'écart quadratique moyen vaut dans ce cas

$$(\Delta N)_{GP} = \sqrt{N} , \tag{F.15}$$

et la fluctuation relative

$$\left(\frac{\Delta N}{N} \right)_{GP} = \frac{1}{\sqrt{N}} . \tag{F.16}$$

b. Fluctuations du volume d'un sous-système

Parler de fluctuations de volume dans un sous-système Σ' de $\mathcal{S}$ n'a de sens que si nous définissons maintenant Σ' *en fixant le nombre de particules* N qui le constituent. Σ' est alors en contact *thermique et mécanique* avec le reste du système, qui lui impose *sa température* T *et sa pression* p.

La probabilité $\mathrm{d}P_\ell(V)$ pour que, à l'équilibre, Σ' ait un volume compris entre V et $V + \mathrm{d}V$ et soit dans l'état (ℓ), d'énergie E_ℓ, est donnée par la distribution T-p (complément III.D, formule (D.11)) :

$$\mathrm{d}P_\ell(V) = \frac{1}{\mathcal{Z}} \mathrm{e}^{-(E_\ell + pV)/kT} \, \mathrm{d}V , \tag{F.17}$$

où $$\widetilde{Z} \widehat{=} \int_0^\infty \mathrm{d}V \sum_{(\ell)} \mathrm{e}^{-(E_\ell + pV)/kT} . \tag{F.18}$$

La technique est la même qu'au paragraphe précédent. On écrit

$$\overline{V} = \frac{1}{\widetilde{Z}} \int_0^\infty \mathrm{d}V\, V \sum_{(\ell)} \mathrm{e}^{-(E_\ell + pV)/kT} = -\frac{kT}{\widetilde{Z}} \frac{\partial \widetilde{Z}}{\partial p} , \tag{F.19}$$

$$\overline{V^2} = \frac{1}{\widetilde{Z}} \int_0^\infty \mathrm{d}V\, V^2 \sum_{(\ell)} \mathrm{e}^{-(E_\ell + pV)/kT} = \frac{(kT)^2}{\widetilde{Z}} \frac{\partial^2 \widetilde{Z}}{\partial p^2} . \tag{F.20}$$

En dérivant les deux membres de (F.19) par rapport à p, on obtient facilement

$$\Delta V = \left[\overline{V^2} - \overline{V}^2\right]^{1/2} = \left[-kT \left(\frac{\partial \overline{V}}{\partial p}\right)_{T,N}\right]^{1/2} \tag{F.21}$$

Puisque $(\Delta V)^2$ est positif, $(\partial \overline{V}/\partial p)$ doit être négatif, ou encore (comme en III, § II.D.3.b)

$$\left(\frac{\partial p}{\partial V}\right)_{T,N} < 0 . \tag{F.22}$$

Pour un gaz parfait, le résultat est encore très simple. L'équation d'état $pV = NkT$ donne en effet

$$\left(\frac{\Delta V}{V}\right)_{GP} = \frac{1}{\sqrt{N}} . \tag{F.23}$$

c. Fluctuations locales de la densité

En pratique on s'intéresse rarement aux fluctuations du nombre de particules dans un volume V fixe, ou aux fluctuations du volume occupé par un nombre déterminé N de particules, mais plutôt aux fluctuations de la densité locale

$$\rho \widehat{=} \frac{N}{V} . \tag{F.24}$$

On peut déduire les fluctuations de ρ de l'un ou l'autre des deux calculs précédents. Dans celui du paragraphe a, V est fixé, ce qui donne

$$\frac{\Delta \rho}{\rho} = \frac{\Delta N}{N} = \left[\frac{kT}{N^2} \left(\frac{\partial N}{\partial \mu}\right)_{T,V}\right]^{1/2} \tag{F.25}$$

Dans le calcul du paragraphe b, c'est N qui est fixé et on obtient

$$\frac{\Delta \rho}{\rho} = \frac{\Delta V}{V} = \left[-\frac{kT}{V^2} \left(\frac{\partial V}{\partial p}\right)_{T,N}\right]^{1/2} \tag{F.26}$$

(dans (F.25) et (F.26) nous avons noté simplement N et V les valeurs moyennes du nombre de particules et du volume).

Le résultat ne doit évidemment pas dépendre de la méthode utilisée pour l'établir. Les deux expressions (F.25) et (F.26) sont effectivement équivalentes puisque, pour un fluide simple[1],

$$\left(\frac{\partial N}{\partial \mu}\right)_{T,V} = -\frac{N^2}{V^2} \left(\frac{\partial V}{\partial p}\right)_{T,N} . \tag{F.27}$$

---

1. Nous cherchons à relier

$$\left(\frac{\partial \mu}{\partial N}\right)_{T,V} = \frac{\partial^2 F}{\partial N^2} \tag{1}$$

.../...

Enfin il est facile de voir sur la formule (F.26) que les fluctuations de densité deviennent anormalement grandes au voisinage du point critique de la transition liquide-gaz (cf. complément III.G, § III et figure 3). La dérivée $(\partial p/\partial V)_{T,N}$ s'annule en effet en ce point :

$$\left(\frac{\partial p}{\partial V}\right)_{T,N} = 0 \quad \text{pour} \quad T = T_c, \quad p = p_c \,, \tag{F.28}$$

de sorte que $\Delta\rho/\rho \to \infty$ (évidemment, dans ce régime de « fluctuations critiques », l'écart quadratique moyen $\Delta\rho$ ne suffit pas à caractériser la distribution de probabilité de la densité ρ, qui n'est plus gaussienne). Ces fluctuations importantes de densité expliquent le phénomène de l'« opalescence critique »[2].

2. FLUCTUATIONS LOCALES DE L'ÉNERGIE ET DE LA TEMPÉRATURE

a. Fluctuations de l'énergie d'un sous-système

Pour calculer les fluctuations locales de l'énergie, on a encore le choix entre deux méthodes : on peut considérer soit le système Σ du paragraphe 1.a (volume fixé), soit le système Σ' du paragraphe 1.b (nombre de particules déterminé). Les expressions fournies par ces deux calculs peuvent être formellement différentes mais elles doivent, comme dans le cas de la densité, être physiquement équivalentes. Nous n'envisagerons

---

et $$\left(\frac{\partial p}{\partial V}\right)_{T,N} = -\frac{\partial^2 F}{\partial V^2} \tag{2}$$

(cf. III, formules (II.30)). Pour un *fluide simple*, l'énergie libre F est fonction de 3 variables seulement, T (intensive), V et N (extensives). Comme F est elle-même extensive, elle est nécessairement, à la limite thermodynamique, de la forme

$$F(T, V, N) = N f\left(T, \frac{V}{N}\right) , \tag{3}$$

où $f(T, u)$ est une fonction de deux variables intensives (le raisonnement est le même qu'au complément III.D, notes 9 et 10). On obtient alors

$$\frac{\partial F}{\partial N} = f\left(T, \frac{V}{N}\right) - \frac{V}{N}\frac{\partial f}{\partial u}\left(T, \frac{V}{N}\right) \tag{4}$$

puis $$\frac{\partial^2 F}{\partial N^2} = \frac{V^2}{N^3}\frac{\partial^2 f}{\partial u^2} , \tag{5}$$

et de même

$$\frac{\partial^2 F}{\partial V^2} = \frac{1}{N}\frac{\partial^2 f}{\partial u^2} . \tag{6}$$

La relation cherchée s'écrit donc

$$\frac{\partial^2 F}{\partial N^2} = \frac{V^2}{N^2}\frac{\partial^2 F}{\partial V^2} , \tag{7}$$

d'où l'on déduit (F.27), compte tenu de (1) et (2).

2. On peut atteindre le point critique en chauffant un tube scellé contenant du liquide et de la vapeur d'un corps pur, si la proportion des deux phases est convenablement choisie. Le ménisque qui les sépare, nettement visible au départ, disparaît « sur place » lorsqu'on parvient au point critique où liquide et vapeur ne se distinguent plus. Mais cette disparition est masquée par le phénomène d'opalescence, dû à la diffusion intense de la lumière provoquée par les hétérogénéités optiques importantes qui résultent des fluctuations de densité.

ici que la deuxième méthode, dans laquelle la distribution statistique des états microscopiques est donnée par l'expression (F.17).

On déduit facilement de cette expression, en posant $\beta \widehat{=} 1/kT$,

$$\overline{E}+p\overline{V}=\frac{1}{\widetilde{Z}}\int_0^\infty \mathrm{d}V \sum_{(\ell)} (E_\ell + pV)\, \mathrm{e}^{-\beta(E_\ell+pV)} = -\frac{1}{\widetilde{Z}}\frac{\partial \widetilde{Z}}{\partial \beta}, \tag{F.29}$$

et

$$\overline{(E+pV)^2}=\frac{1}{\widetilde{Z}}\int_0^\infty \mathrm{d}V \sum_{(\ell)} (E_\ell + pV)^2\, \mathrm{e}^{-\beta(E_\ell+pV)} = \frac{1}{\widetilde{Z}}\frac{\partial^2 \widetilde{Z}}{\partial \beta^2}. \tag{F.30}$$

L'écart quadratique moyen de $(E+pV)$ s'obtient encore en dérivant (F.29), ici par rapport à β :

$$\Delta(E+pV)=\left[-\frac{\partial(\overline{E}+p\overline{V})}{\partial\beta}\right]^{1/2}=\left[kT^2\left(\frac{\partial(\overline{E}+p\overline{V})}{\partial T}\right)_{p,N}\right]^{1/2} \tag{F.31}$$

Par définition, la dérivée de l'enthalpie

$$H\widehat{=}\overline{E}+p\overline{V} \tag{F.32}$$

par rapport à la température, à pression et nombre de particules fixés, est la *capacité calorifique à pression constante* C_p[3].

On a donc simplement

$$\Delta(E+pV)=[kT^2C_p]^{1/2}. \tag{F.33}$$

Pour calculer ΔE, il nous faut maintenant relier les fluctuations de $(E+pV)$ à celles de E. Or

$$\begin{aligned}\left[\Delta(E+pV)\right]^2 &= \overline{(E+pV)^2}-(\overline{E}+p\overline{V})^2 \\ &= (\Delta E)^2+(p\Delta V)^2+2p\,(\overline{E.V}-\overline{E}.\overline{V}).\end{aligned} \tag{F.34}$$

Nous connaissons déjà ΔV (formule (F.21)); il nous reste donc à calculer la quantité

$$\overline{E.V}-\overline{E}.\overline{V}=\overline{(E-\overline{E})(V-\overline{V})}, \tag{F.35}$$

qui mesure les *« corrélations »* entre E et V dans le système Σ'. Pour cela, exprimons la valeur moyenne du produit $V.(E+pV)$ à l'aide de $\widetilde{Z}$:

$$\overline{V.(E+pV)}=\frac{1}{\widetilde{Z}}\int_0^\infty \mathrm{d}V \sum_{(\ell)} V.(E_\ell+pV)\mathrm{e}^{-\beta(E_\ell+pV)}=\frac{1}{\widetilde{Z}}\frac{\partial}{\partial\beta}\left[\frac{1}{\beta}\frac{\partial\widetilde{Z}}{\partial p}\right]. \tag{F.36}$$

L'expression (F.19) de $\overline{V}$ fournit en outre la relation

$$\frac{\partial\overline{V}}{\partial\beta}=-\frac{1}{\widetilde{Z}}\frac{\partial}{\partial\beta}\left[\frac{1}{\beta}\frac{\partial\widetilde{Z}}{\partial p}\right]+\frac{1}{\beta\widetilde{Z}^2}\frac{\partial\widetilde{Z}}{\partial p}\frac{\partial\widetilde{Z}}{\partial\beta}. \tag{F.37}$$

---

3. $C_p\,\mathrm{d}T$ est la quantité de chaleur $đQ_p$ qu'il faut fournir au fluide pour augmenter sa température de $\mathrm{d}T$ à pression extérieure p constante. En effet (cf. complément II.D, § II), dans une transformation infinitésimale, l'énergie du système augmente de

$$\mathrm{d}\overline{E}=đQ+đW. \tag{1}$$

où $đQ$ et $đW$ sont la chaleur et le travail fournis au système par le milieu extérieur. Or

$$đW=-p\,\mathrm{d}\overline{V} \tag{2}$$

et donc

$$đQ=\mathrm{d}\overline{E}+p\,\mathrm{d}\overline{V} \tag{3}$$

Si la transformation a lieu à pression constante,

$$đQ_p=\mathrm{d}(\overline{E}+p\overline{V})_{p,N}\widehat{=}C_p\,\mathrm{d}T. \tag{4}$$

Regroupons alors les expressions (F.19), (F.29), (F.36) et (F.37); nous obtenons ainsi

$$\overline{V.(E+pV)}-\overline{V}.(\overline{E}+p\overline{V})=-\frac{\partial\overline{V}}{\partial\beta}=kT^2\left(\frac{\partial\overline{V}}{\partial T}\right)_{p,N}, \tag{F.38}$$

d'où nous tirons, en utilisant (F.21),

$$\overline{(E-\overline{E})(V-\overline{V})}=\overline{E.V}-\overline{E}.\overline{V}=kT\left[T\left(\frac{\partial\overline{V}}{\partial T}\right)_{p,N}+p\left(\frac{\partial\overline{V}}{\partial p}\right)_{T,N}\right]. \tag{F.39}$$

De (F.34) on déduit ensuite (à l'aide de (F.33), (F.21) et (F.39))

$$(\Delta E)^2=kT^2\left[C_p-\frac{p^2}{T}\left(\frac{\partial V}{\partial p}\right)_{T,N}-2p\left(\frac{\partial V}{\partial T}\right)_{p,N}\right]. \tag{F.40}$$

Si, à la place de C_p, on introduit la capacité calorifique à *volume constant* C_V, on transforme le résultat précédent en[4]

$$(\Delta E)^2=kT^2\left\{C_V-T\left(\frac{\partial V}{\partial p}\right)_{T,N}\left[\left(\frac{\partial p}{\partial T}\right)_{V,N}-\frac{p}{T}\right]^2\right\}. \tag{F.41}$$

Dans le cas d'un gaz parfait, cette formule devient simplement

$$(\Delta E)_{GP}=[kT^2C_V]^{1/2}. \tag{F.42}$$

Cette simplification est liée au fait que les fluctuations de E et V ne sont pas corrélées : le résultat (F.39) s'écrit ici

$$[\overline{(E-\overline{E})(V-\overline{V})}]_{GP}=0. \tag{F.43}$$

---

4. Des différentielles de l'énergie et de l'enthalpie

$$\mathrm{d}E=T\,\mathrm{d}S-p\,\mathrm{d}V+\mu\,\mathrm{d}N \tag{1}$$

et $$\mathrm{d}H=T\,\mathrm{d}S+V\,\mathrm{d}p+\mu\,\mathrm{d}N \tag{2}$$

on déduit

$$C_V\widehat{=}\left(\frac{\partial E}{\partial T}\right)_{V,N}=T\left(\frac{\partial S}{\partial T}\right)_{V,N} \tag{3}$$

et $$C_p\widehat{=}\left(\frac{\partial H}{\partial T}\right)_{p,N}=T\left(\frac{\partial S}{\partial T}\right)_{p,N}. \tag{4}$$

Pour relier C_p à C_V, il suffit donc de changer de variables dans la fonction S :

$$S(T,p,N)=S(T,V(T,p,N),N) \tag{5}$$

(V étant fonction de T, p et N à travers l'équation d'état du fluide). Dérivons les deux membres de (5) par rapport à T (à p et N fixés) :

$$\left(\frac{\partial S}{\partial T}\right)_{p,N}=\left(\frac{\partial S}{\partial T}\right)_{V,N}+\left(\frac{\partial S}{\partial V}\right)_{T,N}\left(\frac{\partial V}{\partial T}\right)_{p,N}. \tag{6}$$

On en déduit la « relation de Mayer »

$$C_p=C_V+T\left(\frac{\partial V}{\partial T}\right)_{p,N}\left(\frac{\partial p}{\partial T}\right)_{V,N}, \tag{7}$$

puisque l'unicité de la dérivée seconde $\partial^2F/\partial T\,\partial V$ entraîne

$$\left(\frac{\partial S}{\partial V}\right)_{T,N}=-\frac{\partial}{\partial V}\left(\frac{\partial F}{\partial T}\right)=\frac{\partial}{\partial T}\left(-\frac{\partial F}{\partial V}\right)=\left(\frac{\partial p}{\partial T}\right)_{V,N} \tag{8}$$

(cf. III, formules (II.28) et (II.30.a)).

D'autre part, si trois grandeurs x, y et z sont liées par une relation de la forme

$$f(x,y,z)=0, \tag{9}$$

.../...

b. « Fluctuations » locales de la température

Comme nous l'avons remarqué au paragraphe I, on ne mesure pas en général directement l'énergie contenue dans le système Σ'; on y place plutôt un petit thermomètre et on repère ses indications, car leur variation temporelle traduit les fluctuations de l'énergie contenue dans Σ'.

L'énergie moyenne $\overline{E}$ et le volume moyen $\overline{V}$ de Σ', qui coïncident pratiquement avec son énergie et son volume les plus probables, sont reliés à la température T du système global $\mathcal{S}$ par la « condition d'équilibre » entre Σ' et $\mathcal{S}$:

$$\frac{1}{T} = \frac{\partial S^*_{\Sigma'}}{\partial E}(\overline{E},\ \overline{V}) , \tag{F.44}$$

où $S^*_{\Sigma'}$ est l'entropie microcanonique du système Σ'. Si l'on admet que Σ' reste constamment en équilibre local (cf. § I.3) même lorsque son énergie fluctue de

$$\delta E \mathrel{\widehat{=}} E - \overline{E} \tag{F.45}$$

et son volume de

$$\delta V \mathrel{\widehat{=}} V - \overline{V} , \tag{F.46}$$

on peut définir à chaque instant sa température $T_{\Sigma'}$: elle est reliée à son énergie E et son volume V instantanés par

$$\frac{1}{T_{\Sigma'}} = \frac{\partial S^*_{\Sigma'}}{\partial E}(E,\ V) , \tag{F.47}$$

qui donne $T_{\Sigma'}$ comme fonction des paramètres E et V. La variation

$$\delta T \mathrel{\widehat{=}} T_{\Sigma'} - T \tag{F.48}$$

de la température de Σ', associée aux fluctuations δE et δV, est donc de la forme

$$\delta T = \left(\frac{\partial T}{\partial E}\right)_V \delta E + \left(\frac{\partial T}{\partial V}\right)_E \delta V , \tag{F.49}$$

où les dérivées de T sont à prendre pour les valeurs $\overline{E}$ et $\overline{V}$ de l'énergie et du volume.

---

l'une quelconque d'entre elles peut être considérée comme fonction des deux autres et les dérivées partielles correspondantes vérifient

$$\left(\frac{\partial x}{\partial y}\right)_z = -\left(\frac{\partial x}{\partial z}\right)_y \left(\frac{\partial z}{\partial y}\right)_x . \tag{10}$$

Il suffit pour le montrer de différentier la relation (9) :

$$\frac{\partial f}{\partial x}\,\mathrm{d}x + \frac{\partial f}{\partial y}\,\mathrm{d}y + \frac{\partial f}{\partial z}\,\mathrm{d}z = 0 , \tag{11}$$

de sorte que, par exemple,

$$\left(\frac{\partial x}{\partial y}\right)_z = -\frac{\partial f/\partial y}{\partial f/\partial x} . \tag{12}$$

L'application de la relation (10) à V, T, p, liées par l'équation d'état, donne

$$\left(\frac{\partial V}{\partial T}\right)_{p,\,N} = -\left(\frac{\partial V}{\partial p}\right)_{T,\,N}\left(\frac{\partial p}{\partial T}\right)_{V,\,N} , \tag{13}$$

et permet de transformer la relation de Mayer (7) en

$$C_p = C_V - T\left(\frac{\partial V}{\partial p}\right)_{T,\,N}\left(\frac{\partial p}{\partial T}\right)^2_{V,\,N} . \tag{14}$$

La relation (F.41) découle de (F.40) si l'on y reporte (13) et (14).

Or (cf. III, formule (II.11))

$$\left(\frac{\partial T}{\partial E}\right)_V = \frac{1}{(\partial E/\partial T)_V} = \frac{1}{C_V} \tag{F.50}$$

et[5] $$\left(\frac{\partial T}{\partial V}\right)_E = \frac{1}{C_V}\left[p - T\left(\frac{\partial p}{\partial T}\right)_{V,N}\right], \tag{F.51}$$

où p est la pression du système global S[6]. On a donc

$$\delta T = \frac{1}{C_V}\left[\delta E + p\delta V - T\left(\frac{\partial p}{\partial T}\right)_{V,N}\delta V\right]. \tag{F.52}$$

On en déduit l'écart quadratique moyen associé aux « fluctuations de température » :

$$(\Delta T)^2 = \overline{(\delta T)^2} = \frac{1}{C_V^2}\left[\overline{(\delta E + p\delta V)^2} + T^2\left(\frac{\partial p}{\partial T}\right)^2_{V,N}\overline{(\delta V)^2} - 2T\left(\frac{\partial p}{\partial T}\right)_{V,N}\overline{\delta V.(\delta E + p\delta V)}\right] \tag{F.53}$$

et la corrélation entre les variations de température et de volume :

$$\overline{\delta T.\delta V} = \frac{1}{C_V}\left[\overline{(\delta E + p\delta V).\delta V} - T\left(\frac{\partial p}{\partial T}\right)_{V,N}\overline{(\delta V)^2}\right]. \tag{F.54}$$

Les valeurs moyennes qui figurent dans ces expressions ont été calculées aux paragraphes précédents. Le résultat final est très simple : d'une part, les fluctuations de température et de volume sont indépendantes, puisque[7]

$$\overline{\delta T.\delta V} = \overline{(T_{\Sigma'} - T).(V - \overline{V})} = 0 ; \tag{F.55}$$

---

5. Si l'on applique l'identité (10) de la note 4 à T, E, V (à N constant), on obtient

$$\left(\frac{\partial T}{\partial V}\right)_{E,N} = -\left(\frac{\partial T}{\partial E}\right)_{V,N}\left(\frac{\partial E}{\partial V}\right)_{T,N} = -\frac{1}{C_V}\left(\frac{\partial E}{\partial V}\right)_{T,N}. \tag{1}$$

De la différentielle de E (formule (1) de la note 4) on déduit ensuite

$$\left(\frac{\partial E}{\partial V}\right)_{T,N} = T\left(\frac{\partial S}{\partial V}\right)_{T,N} - p = T\left(\frac{\partial p}{\partial T}\right)_{V,N} - p \tag{2}$$

(d'après l'égalité (8) de la note 4).

6. Dans les formules (F.50) et (F.51), l'énergie et le volume du sous-système Σ' ont les valeurs $\overline{E}$ et $\overline{V}$; c'est pourquoi la température et la pression qui y figurent sont celles du système global S (cf. (F.44)).

Comme la température $T_{\Sigma'}$, la pression $p_{\Sigma'}$ « fluctue » autour de la valeur imposée par S. Ces « fluctuations » se calculent très simplement à partir de celles de la température $T_{\Sigma'}$ et du volume V si l'on connaît l'équation d'état du fluide constituant Σ'.

7. Reportons dans (F.54) les expressions trouvées en (F.21) et (F.38) :

$$\overline{\delta T.\delta V} = \frac{kT^2}{C_V}\left[\left(\frac{\partial V}{\partial T}\right)_{p,N} + \left(\frac{\partial p}{\partial T}\right)_{V,N}\left(\frac{\partial V}{\partial p}\right)_{T,N}\right].$$

Le second membre est nul d'après la relation (13) de la note 4.

d'autre part, l'écart quadratique moyen ΔT s'écrit[8]

$$\Delta T = \left[\overline{T_{\Sigma'}^2} - T^2\right]^{1/2} = \left[\frac{kT^2}{C_V}\right]^{1/2}. \tag{F.56}$$

Ici encore le rapport $\Delta T/T$ est proportionnel à $N^{-1/2}$ puisque C_V est proportionnel à N pour $N \gg 1$.

8. Compte tenu des résultats (F.21), (F.33) et (F.38), la formule (F.53) s'écrit

$$(\Delta T)^2 = \frac{kT^2}{C_V^2}\left\{C_p - T\left(\frac{\partial p}{\partial T}\right)_{V,N}\left[\left(\frac{\partial p}{\partial T}\right)_{V,N}\left(\frac{\partial V}{\partial p}\right)_{T,N} + 2\left(\frac{\partial V}{\partial T}\right)_{p,N}\right]\right\}.$$

On aboutit à (F.56) en utilisant les relations (13) et (14) de la note 4.

Exercices V

EXERCICES SIMPLES SUR LES FERMIONS ET LES BOSONS

V.1. On veut préciser sur un exemple très simple la différence dans le décompte des états d'un système suivant qu'il est constitué de bosons ou de fermions. On imagine pour cela un ensemble de *A sites indépendants et discernables.* Chaque site peut accueillir 0, 1 ou 2 particules; il offre à chaque particule 2 états individuels possibles, (1) et (2), d'énergies respectives ε_1 et ε_2. On néglige les interactions entre les particules, même lorsque deux d'entre elles se trouvent sur le même site.

1) On suppose d'abord que les particules étudiées sont des *bosons.*

a) Énumérer les états possibles pour un système de 2 particules localisées sur le même site.

b) Écrire la grande fonction de partition d'un site, et en déduire celle de l'ensemble du système.

2) Mêmes questions dans le cas de *fermions* (pour que la comparaison soit plus claire, on suppose que les états individuels (1) et (2) tiennent compte du spin).

3) Calculer dans chacun des deux cas le nombre d'occupation moyen d'un site. Obtient-on la somme des nombres d'occupation moyens des deux états? Pourquoi?

V.2. On considère un système constitué de A sites *indépendants* sur lesquels peuvent se fixer des particules identiques provenant d'un réservoir-thermostat. Chaque site (i) $(i = 1, 2, \ldots, A)$ peut accueillir 0 ou 1 particule, à laquelle il offre une série d'états possibles (λ).

1) Écrire la grande fonction de partition du système.

2) On pourrait penser que ce système est constitué de particules indépendantes, leurs états individuels possibles étant caractérisés par (i, λ), où (i) repère le site et (λ) l'état sur ce site. Montrer que l'application des formules du chapitre V, paragraphe IV.A ne donne pas le résultat correct trouvé à la question 1, même si les particules sont des fermions. Pourquoi les particules ne sont-elles pas vraiment indépendantes dans ce système?

3) Étudier le cas particulier où il existe un seul état possible par site.

V.3. On considère un système de *fermions* indépendants, de spin 1/2, dans lequel, en l'absence de champ magnétique, tous les états à une particule ont la même énergie $\varepsilon_0 = 0$. Soit $2N'$ la dégénérescence du niveau d'énergie, et N le nombre de fermions contenus dans le système.

Sous l'action d'un champ magnétique extérieur, le niveau d'énergie se dédouble : chacun des deux nouveaux niveaux d'énergie $\varepsilon_- = -10^{-4}$ eV et $\varepsilon_+ = 10^4$ eV est N' fois dégénérés. Une mesure du moment magnétique total montre qu'un tiers des particules se trouve dans l'état d'énergie ε_+.

Écrire les équations permettant de trouver la température et le potentiel chimique en fonction de N et N'. Montrer que, pour une valeur de N' donnée, ces équations n'ont de solution que lorsque N est inférieur à une valeur N_{max} que l'on exprimera en fonction de N'.

Déterminer la température et le potentiel chimique dans le cas où $N' = 2N$.

Pour quelles valeurs de N'/N pourrait-on utiliser la statistique de Maxwell-Boltzmann ?

V.4. On considère un gaz parfait de N molécules en équilibre à la température T dans un récipient de volume V. Soit un petit domaine de volume $v \ll V$ à l'intérieur du récipient.

1) Exprimer la grande fonction de partition des molécules contenues dans ce petit domaine, à l'aide de la fonction de partition de l'une d'entre elles et du potentiel chimique du gaz (qui joue le rôle de réservoir).

2) En déduire le nombre moyen $\bar{n}$ de molécules contenues dans le volume v.

3) Calculer la probabilité $P(n)$ pour que le volume v contienne n molécules et montrer qu'elle est donnée par la distribution de Poisson (cf. exercice I.10)

$$P(n) = \bar{n}^n \frac{e^{-\bar{n}}}{n!} .$$

Comment le fait que $v \ll V$ intervient-il dans ce résultat?

V.5. Gaz parfaits faiblement dégénérés. On considère un gaz constitué de N particules identiques, de masse m et de spin s, *indépendantes* et *libres*. Ce gaz est enfermé dans un récipient de volume V et maintenu à la température T. On suppose valable la *limite thermodynamique*. On veut étudier ce gaz dans le domaine où *les nombres d'occupation moyens de tous les états individuels* (λ) *sont très faibles :*

$$\overline{N}_\lambda \ll 1 \quad \text{pour tout } (\lambda). \tag{1}$$

1) Montrer que l'hypothèse (1) se traduit par une condition sur la fugacité

$$\varphi \widehat{=} e^{\beta\mu} . \tag{2}$$

2) On cherche d'abord les propriétés du gaz en se limitant à l'ordre le plus bas en φ.

a) Calculer dans cette limite, en fonction de N, V et T, le potentiel chimique μ du gaz, puis son énergie E et sa pression p.

b) Écrire la condition de validité de ces résultats en termes de la longueur d'onde thermique

$$\Lambda \widehat{=} \left(\frac{2\pi\hbar^2}{mkT}\right)^{1/2} \tag{3}$$

et de la densité de particules N/V.

3) On veut maintenant évaluer les termes correctifs aux résultats précédents, en poussant le développement jusqu'à l'ordre suivant en

$$u \widehat{=} \frac{1}{2s+1} \frac{N}{V} \Lambda^3 . \tag{4}$$

a) Écrire l'expression de $\overline{N}_\lambda$ lorsqu'on conserve le terme du deuxième ordre en φ (on distinguera le cas des bosons et celui des fermions).

b) Calculer la fugacité φ au second ordre en u.

c) En déduire le terme correctif à l'énergie E et à la pression p trouvées en 2.*a*.

d) Montrer que ces termes correctifs indiquent une attraction effective entre les particules si ce sont des bosons, et une répulsion effective si ce sont des fermions.

4) Le développement ainsi amorcé est l'analogue, pour un gaz parfait quantique, du développement du viriel pour un gaz réel classique (complément V.C et exercice V.18). Esquisser le calcul permettant de déterminer le terme suivant.

ÉQUILIBRES

V.6. Adsorption. On veut étudier, dans un modèle très simple, l'*adsorption d'un gaz sur une paroi solide* dans le cas *(«physisorption»)* où les atomes adsorbés ne sont pas fixés en des sites localisés (comme au complément V.A) mais peuvent se déplacer librement sur la paroi : on suppose qu'ils constituent, sur la paroi d'aire s, un gaz parfait monoatomique à 2 dimensions; l'énergie de chacun d'eux est de la forme

$$\varepsilon(\vec{k}) = \frac{\hbar^2 \vec{k}^2}{2m} - \varepsilon_0 ,$$

où $\varepsilon_0(>0)$ est l'énergie de liaison de l'atome sur la paroi et $\vec{k}$ un vecteur d'onde permis parallèle à la paroi.

1) On se place dans le cadre *grand-canonique :* l'ensemble des atomes adsorbés est en équilibre avec un réservoir-thermostat de potentiel chimique μ et de température T.

a) Calculer, dans l'approximation de Maxwell-Boltzmann, la grande fonction de partition Ξ et le grand potentiel J du système des atomes adsorbés. En déduire leur nombre (moyen) N_a en fonction de T, μ et s.

b) Le réservoir est un gaz parfait classique à 3 dimensions, constitué des mêmes particules que l'adsorbat. Écrire sa fonction de partition canonique et son énergie libre; en déduire son potentiel chimique μ, que l'on exprimera en fonction de la pression p et de la température T.

c) Calculer le nombre d'atomes adsorbés par unité de surface, en fonction de la température T et de la pression p du gaz-réservoir.

2) *a*) Montrer que l'énergie (moyenne) de l'ensemble des atomes adsorbés peut s'écrire

$$E_a = N_a(kT - \varepsilon_0) .$$

b) En déduire l'énergie E_T du *système global* constitué par le gaz-réservoir et l'adsorbat. Calculer la capacité calorifique à volume constant C_V de ce système global (comme au paragraphe I.3.b du complément V.A, on interprétera physiquement les divers termes contribuant à C_V, et on les exprimera en fonction de p et T).

3) On considère maintenant le cas où le gaz-réservoir à 3 dimensions est constitué de *molécules* H_2 et l'adsorbat d'*atomes* H. On reprend le problème à partir du résultat de la question 1.*a*. En utilisant les formules (A.51) et (A.53) du complément V.A, calculer à nouveau, dans ce cas, le nombre moyen d'atomes adsorbés par unité de surface en fonction de T et p.

4) Retrouver les résultats des questions précédentes dans le cadre *canonique,* en considérant l'adsorbat et le gaz à 3 dimensions comme deux sous-systèmes d'un système global dont le nombre total d'atomes est fixé, et en calculant le nombre le plus probable d'atomes adsorbés.

V.7. Équilibre de sublimation. On considère un récipient de volume V contenant un solide en équilibre à la température T avec sa vapeur (équilibre de sublimation). On suppose que la vapeur est un gaz parfait constitué de particules monoatomiques sans structure interne (gaz rare).

1) Le solide joue le rôle de réservoir de particules pour le gaz; il lui impose donc son potentiel chimique μ_s. Calculer la grande fonction de partition du gaz. Calculer sa pression $p(T, \mu_s)$ et son entropie par atome $s_g(T, \mu_s)$.

2) Pour préciser les propriétés du système, il faut déterminer le potentiel chimique du solide. Dans les modèles les plus simples (cf. complément III.E), les propriétés d'un solide à l'équilibre ne dépendent que de la température T et du nombre de particules $N_s(\gg 1)$ qui le composent, pas de son volume. A une bonne approximation, on peut donc considérer que la fonction de partition Z_s du solide ne dépend que de T et N_s (voir cependant l'exercice III.41). Dans l'état fondamental, l'énergie du solide vaut $E_0 = -N_s\varepsilon_0$, où $\varepsilon_0(>0)$ est la chaleur latente de sublimation par atome à $T=0$ (on a choisi pour origine l'énergie d'un atome au repos dans la vapeur).

a) Montrer que, pour T tendant vers zéro,

$$\ln Z_s(T, N_s) \simeq \frac{N_s\varepsilon_0}{kT} ,$$

et estimer la dépendance en N_s et T des premiers termes correctifs (on admettra que la dégénérescence de l'état fondamental est de l'ordre de l'unité).

b) Soit $C_s(T)$ la capacité calorifique par atome de solide (on sait que $C_s(T)/T < a$, où a est une constante, lorsque $T \to 0$). Montrer que l'énergie moyenne du solide à la température T est liée à C_s par

$$E(T, N_s) = -N_s \left[\varepsilon_0 - \int_0^T C_s(T')\, \mathrm{d}T' \right].$$

En déduire l'expression de $\ln Z_s$ et de μ_s.

c) En reportant cette expression de μ_s dans $p(T, \mu_s)$, montrer que la pression de vapeur ne dépend que de la température et l'exprimer en fonction des grandeurs mesurables que sont ε_0 et $C_s(T)$. Montrer que $p(T)$ est une fonction croissante de la température (on notera que $E(T, N_s)$ doit être négative pour que le solide soit stable).

d) Calculer l'entropie $s_s(T)$ par atome de solide. En déduire la chaleur latente de sublimation par atome

$$\ell(T) = T\left[s_g(T) - s_s(T)\right].$$

Montrer qu'elle vérifie l'équation de Clapeyron

$$\ell(T) = T v_g \frac{\mathrm{d}p}{\mathrm{d}T},$$

où v_g est le volume par atome dans le gaz (le volume par atome dans le solide est négligeable devant v_g).

3) On peut obtenir des résultats plus précis si on se donne un modèle microscopique pour le solide. Calculer explicitement $p(T)$ et $\ell(T)$ à l'aide du modèle d'Einstein (complément III.E, § I).

V.8. On considère un système comportant deux types de particules, (1) et (2).

1) Montrer que sa grande fonction de partition peut s'écrire

$$\Xi(T, V, \mu_1, \mu_2) = \sum_{N_1=0}^{\infty} \sum_{N_2=0}^{\infty} Z(T, V, N_1, N_2)\, \mathrm{e}^{(\mu_1 N_1 + \mu_2 N_2)/kT},$$

où $Z(T, V, N_1, N_2)$ est la fonction de partition du système lorsque le nombre de particules des deux types est fixé à N_1 et N_2.

2) On suppose le système en équilibre à la température T et avec un réservoir de particules de chaque type.

a) Exprimer la valeur moyenne de N_1, de N_2 et de l'énergie, ainsi que la pression et l'entropie.

b) Montrer que

$$\overline{N_1 N_2} - \overline{N}_1 . \overline{N}_2 = kT \frac{\partial \overline{N}_1}{\partial \mu_2} = kT \frac{\partial \overline{N}_2}{\partial \mu_1}.$$

3) Que devient ce dernier résultat dans le cas d'un mélange idéal (système dans lequel les interactions entre les molécules (1) et (2) sont négligeables)?

V.9. On étudie une solution très diluée de N' molécules de soluté (masse moléculaire m') dans N molécules de solvant (masse moléculaire m), en équilibre avec un thermostat à la température T.

En l'absence de champ de force extérieur, les potentiels chimiques dépendent de la pression p, de la température T et de la concentration

$$c \widehat{=} \frac{N'}{N + N'} \simeq \frac{N'}{N}.$$

Ils s'écrivent (complément III.F, § II.3)

$\mu = \mu_0(p, T) - kTc$ pour le solvant,

$\mu' = \psi(p, T) + kT \ln c$ pour le soluté.

La solution se trouve dans le champ de pesanteur terrestre, de sorte que la pression et la concentration dépendent de l'altitude z. Déterminer $p(z)$ et $c(z)$ à partir des hypothèses suivantes :

(i) $\partial\mu_0/\partial p$ indépendant de p (cette quantité est le volume v occupé par une molécule de solvant; l'hypothèse est donc celle de l'incompressibilité du solvant);

(ii) $\partial\psi/\partial p$ indépendant de p (c'est le volume v' occupé par une molécule de soluté dans la solution; l'hypothèse est donc celle de l'incompressibilité de la solution).

Interpréter les différentes contributions à $p(z)$ et $c(z)$ (pression hydrostatique, pression osmotique, poussée d'Archimède).

V.10. Mesure des masses des macromolécules

1) On veut mesurer la masse moléculaire d'une substance constituée de macromolécules. Pour cela on les disperse en solution dans un liquide de masse volumique ρ_0, contenu dans un tube vertical. Lorsqu'on ajoute à la solution une masse M de macromolécules, son volume augmente de V. On appelle $\rho = M/V$ la masse volumique des macromolécules dans la solution.

a) Quelle est l'énergie potentielle d'une macromolécule située à l'altitude z (on considérera que l'effet du solvant se réduit à la poussée d'Archimède sur le volume occupé par la macromolécule)?

b) En déduire la dépendance en z du nombre $n(z)$ de macromolécules par unité de volume dans le tube, à la température T.

c) On mesure, par des méthodes optiques, le rapport $n(z)/n(z_0)$ à différentes altitudes. Sachant que l'on peut détecter des variations $[n(z) - n(z_0)]/n(z_0)$ d'environ 10 %, estimer, pour $\rho/\rho_0 = 0{,}75$, l'ordre de grandeur des masses que l'on peut déterminer par cette méthode.

2) Pour pouvoir mesurer des masses moléculaires plus petites, on utilise des ultracentrifugeuses. Tout se passe comme si l'accélération de la pesanteur était remplacée par l'accélération centrifuge.

a) Comment varie le nombre $n(r)$ de macromolécules par unité de volume, dans une centrifugeuse tournant à la vitesse angulaire ω, en fonction de la distance r à l'axe de rotation?

b) Estimer les masses mesurables par cette méthode (on prendra $\rho/\rho_0 = 0{,}75$, $\omega = 10\,000$ tours/min et $5\ \text{cm} < r < 10\ \text{cm}$).

V.11. On considère un gaz (quantique) constitué de particules identiques faiblement couplées. On note (λ) leurs états individuels possibles, d'énergie ε_λ. Ces particules interagissent seulement lors de collisions de durée négligeable. On suppose pour l'instant que seules sont à prendre en compte les collisions binaires : on note $a_{\lambda', \nu'; \lambda, \nu}$ la probabilité par unité de temps pour que deux particules, initialement dans les états individuels (λ) et (ν), passent, de façon pratiquement instantanée, dans les états (λ') et (ν'). Les probabilités de transition $a_{\lambda', \nu'; \lambda, \nu}$ sont symétriques dans l'échange de (λ') et (ν') ou de (λ) et (ν) (particules identiques), symétriques dans l'échange des paires (λ, ν) et (λ', ν') (renversement du temps), et elles conservent l'énergie totale :

$$\varepsilon_{\lambda'} + \varepsilon_{\nu'} = \varepsilon_\lambda + \varepsilon_\nu \,. \qquad (1)$$

1) Le système, qui est en contact avec un réservoir-thermostat de température T et potentiel chimique μ, est initialement hors d'équilibre. Soit $\overline{N}_\lambda(t)$ le nombre

d'occupation moyen de l'état (λ) à l'instant t. On veut établir, en adaptant le raisonnement heuristique du chapitre IV, § I.B.1, l'équation d'évolution des $\overline{N}_\lambda$ au cours du temps.

a) Supposons d'abord que les particules étudiées sont des *fermions*. Le principe de Pauli interdit la transition $(\lambda, \nu) \longrightarrow (\lambda', \nu')$ si l'un des états (λ') et (ν') est déjà occupé ($N_{\lambda'} = 1$ ou $N_{\nu'} = 1$). Montrer que l'équation cherchée s'écrit dans ce cas

$$\frac{d\overline{N}_\lambda}{dt} = \sum_{(\nu)} \sum_{(\lambda'),\,(\nu')} \left[a_{\lambda,\,\nu;\,\lambda',\,\nu'} \overline{N}_\lambda \overline{N}_{\nu'} (1 - \overline{N}_\lambda)(1 - \overline{N}_\nu) - a_{\lambda',\,\nu';\,\lambda,\,\nu} \overline{N}_\lambda \overline{N}_\nu (1 - \overline{N}_{\lambda'})(1 - \overline{N}_{\nu'}) \right] . \qquad (2)$$

b) Si les particules considérées sont des *bosons*, on peut montrer que l'équation a une forme analogue :

$$\frac{d\overline{N}_\lambda}{dt} = \sum_{(\nu)} \sum_{(\lambda'),\,(\nu')} \left[a_{\lambda,\,\nu;\,\lambda',\,\nu'} \overline{N}_\lambda \overline{N}_{\nu'} (1 + \overline{N}_\lambda)(1 + \overline{N}_\nu) - a_{\lambda',\,\nu';\,\lambda,\,\nu} \overline{N}_\lambda \overline{N}_\nu (1 + \overline{N}_{\lambda'})(1 + \overline{N}_{\nu'}) \right] . \qquad (3)$$

Ici, $1 + N_\lambda$ est l'élément de matrice, élevé au carré, de l'opérateur de création qui ajoute une particule à l'état (λ), lorsque celui-ci en comporte déjà N_λ (C.D.L., p. 497).

2) On se propose maintenant de vérifier que la distribution de Fermi-Dirac est bien une solution d'équilibre de l'équation (2), et celle de Bose-Einstein de l'équation (3).

a) Montrer que, par suite de la conservation de l'énergie (1), ces distributions vérifient l'égalité

$$\overline{N}_\lambda \overline{N}_{\nu'} (1 \pm \overline{N}_\lambda)(1 \pm \overline{N}_\nu) = \overline{N}_\lambda \overline{N}_\nu (1 \pm \overline{N}_{\lambda'})(1 \pm \overline{N}_{\nu'}) \qquad (4)$$

(le signe supérieur se réfère au cas des bosons, le signe inférieur à celui des fermions).

b) En déduire, grâce à la propriété d'invariance par renversement du temps, que la distribution de Fermi-Dirac annule bien le second membre de l'équation (2) pour tout (λ), et que celle de Bose-Einstein fait de même pour l'équation (3).

3) L'hypothèse que les collisions à plusieurs particules sont négligeables est-elle nécessaire au raisonnement?

V.12. On veut discuter la distribution des diverses espèces de particules a_α ($\alpha = 1, 2, \ldots, A$) présentes dans l'Univers à une étape de son évolution où elles étaient toutes en équilibre entre elles. On suppose que ces particules constituent A gaz parfaits indépendants, et on note

$$n_\alpha(\vec{p}, m_s) \frac{d^3p}{(2\pi\hbar)^3} \qquad (1)$$

le nombre moyen, par unité de volume, de particules a_α qui ont pour impulsion $\vec{p}$ à d^3p près et $m_s\hbar$ pour composante du spin le long d'un axe déterminé; le nombre total N_α, par unité de volume, de particules a_α s'écrit

$$N_\alpha = \sum_{m_s} \int \frac{d^3p}{(2\pi\hbar)^3} n_\alpha(\vec{p}, m_s) . \qquad (2)$$

Les particules de différentes espèces participent à des réactions réversibles :

$$a_{1i} + a_{2i} + \ldots + a_{pi} \rightleftarrows a_{1f} + a_{2f} + \ldots + a_{qf} \; ; \qquad (3)$$

on notera (i) l'état microscopique initial et (f) l'état final d'une telle réaction (le même type de particules peut intervenir plusieurs fois dans chacun de ces états). Ces processus sont susceptibles de faire varier les nombres N_α au cours du temps; on veut s'assurer que l'Univers est bien à l'équilibre statistique lorsque la distribution de chaque type de particules est celle de Fermi-Dirac ou de Bose-Einstein, suivant les cas.

1) Dans un premier temps, on suppose pour simplifier que les A gaz parfaits sont suffisamment dilués pour que l'*approximation de Maxwell-Boltzmann* leur soit applicable. Dans ce cas, l'évolution dans le temps du nombre N_α est donnée par une équation que nous écrirons sous la forme simplifiée

$$\frac{dN_\alpha}{dt} = \sum_{(i):\,\alpha\in(i)} \sum_{(f)} \left[w_{if} \prod_{\gamma\in(f)} n_\gamma - w_{fi} \prod_{\delta\in(i)} n_\delta \right] : \tag{4}$$

on somme sur les états (i) comportant la particule a_α choisie, et sur tous les états (f) (la somme comprend une intégration sur les impulsions et une sommation sur les états de spin); w_{fi} est la probabilité de transition par unité de temps de l'état (i) à l'état (f).

a) Retrouver dans l'équation (4) la forme générale d'une équation maîtresse (chapitre IV, § I.B.1) : identifier les termes de « peuplement » et de « dépeuplement » du sous-système des particules a_α et analyser la structure de chacun de ces termes.

b) Les réactions (3) conservent l'énergie totale. Montrer que, à l'équilibre, les potentiels chimiques des particules qui y participent vérifient également

$$\sum_{\gamma\in(f)} \mu_\gamma = \sum_{\delta\in(i)} \mu_\delta \,. \tag{5}$$

En déduire que, si les particules sont distribuées selon la loi

$$n_\alpha(\vec{p}, m_s) = e^{-\beta[\varepsilon_\alpha(\vec{p})-\mu_\alpha]}, \tag{6}$$

où $\varepsilon_\alpha(\vec{p})$ est l'énergie (indépendante du spin) d'une particule a_α d'impulsion $\vec{p}$, on a, pour chaque réaction, l'égalité

$$\prod_{\gamma\in(f)} n_\gamma = \prod_{\delta\in(i)} n_\delta \,. \tag{7}$$

c) Si toutes les interactions sont invariantes par renversement du temps, les probabilités de transition par unité de temps sont symétriques :

$$w_{if} = w_{fi} \qquad \forall (i) \text{ et } (f) \,. \tag{8}$$

Montrer que, dans ce cas, la distribution de Maxwell-Boltzmann (6) annule les dN_α/dt donnés par (4) et correspond donc bien à l'équilibre statistique.

d) Si l'égalité (8) n'est pas satisfaite, quelle condition doivent vérifier les w_{fi} pour que la distribution (6) donne quand même dans (4) l'équilibre statistique? Comme la normalisation des probabilités implique (chapitre IV, note 3)

$$\sum_{(f)\neq(i)} w_{fi} = -w_{ii} \,, \tag{9}$$

montrer que la condition cherchée s'écrit

$$\sum_{(f)\neq(i)} w_{if} = -w_{ii} \tag{10}$$

(en théorie des collisions, cette relation découle de l'« unitarité de la matrice S », valable même si certaines interactions ne sont pas invariantes par renversement du temps).

2) Les résultats de la question 1 se généralisent au cas où les *effets quantiques* ne sont pas négligeables. L'équation d'évolution de N_αs'écrit alors (voir exercice V.11)

$$\frac{dN_\alpha}{dt} = \sum_{(i):\,\alpha\in(i)} \sum_{(f)} \left[w_{if} \prod_{\delta\in(i)} (1 \pm n_\delta) \prod_{\gamma\in(f)} n_\gamma - w_{fi} \prod_{\gamma\in(f)} (1 \pm n_\gamma) \prod_{\delta\in(i)} n_\delta \right], \tag{11}$$

où l'on prend le signe supérieur pour les bosons et le signe inférieur pour les fermions.

a) Établir, sur les mêmes bases qu'en 1.*b*, que l'égalité

$$\prod_{\delta\in(i)}\prod_{\gamma\in(f)}(1\pm n_\delta)n_\gamma=\prod_{\gamma\in(f)}\prod_{\delta\in(i)}(1\pm n_\gamma)n_\delta \tag{12}$$

est vérifiée par les distributions quantiques

$$n_\alpha(\vec{p},m_s)=\frac{1}{e^{\beta[\varepsilon_\alpha(\vec{p})-\mu_\alpha]}\mp 1}\,. \tag{13}$$

b) Montrer que, si l'invariance par renversement du temps est valable (relation (8)), les distributions de Fermi-Dirac et Bose-Einstein (13) correspondent bien à l'équilibre statistique.

c) Comment s'écrit ici, pour chaque état (i), la condition que doivent vérifier les w_{fi} pour que les distributions (13) donnent encore l'équilibre même si la relation (8) n'est pas satisfaite?

FLUCTUATIONS

V.13. Un récipient de volume V maintenu à la température T est en communication avec un réservoir de particules identiques sans interaction, de potentiel chimique μ. Calculer l'écart quadratique moyen ΔN du nombre de particules contenues dans le récipient lorsqu'il s'agit de fermions, puis de bosons, enfin de particules « classiques ».

V.14. On considère une solution en équilibre à la température T, contenant au total N_2^0 molécules de soluté pour N_1^0 molécules de solvant. On se propose de calculer les fluctuations locales de la proportion de soluté $x\widehat{=}N_2/N_1$. Pour cela on considère un sous-système Σ contenant un nombre fixé N_1 de molécules de solvant et on étudie comment fluctue le nombre N_2 de molécules de soluté dans Σ.

1) Montrer que les paramètres extérieurs qui caractérisent l'état macroscopique de Σ sont N_1, T, la pression p et le potentiel chimique μ_2 du soluté dans la solution.

2) En déduire que la probabilité P_ℓ pour que Σ soit dans un état microscopique (ℓ) d'énergie E_ℓ, de volume v et contenant N_2 molécules de solvant est proportionnelle à $\exp-(E_\ell-\mu_2N_2+pv)/kT$.

3) Calculer le nombre moyen $\overline{N}_2$ de particules de soluté dans Σ, ainsi que son écart quadratique moyen ΔN_2.

4) En déduire Δx et montrer qu'on peut l'écrire sous la forme

$$(\Delta x)^2=kT/N_1(\partial\mu_2/\partial x)_{p,T}\,.$$

Calculer Δx dans le cas d'une solution diluée ($x\ll 1$) (voir complément III.F, § II.3).

V.15. Fluctuations de densité dans l'atmosphère et bleu du ciel. On considère un gaz en équilibre à la température T, et un petit domaine de volume v à l'intérieur du gaz.

1) Soit N le nombre de molécules du gaz contenues dans v. Calculer $\overline{N}$ et $\overline{N^2}$ et en déduire que l'écart quadratique moyen ΔN vérifie

$$(\Delta N)^2=kT\left(\frac{\partial\overline{N}}{\partial\mu}\right)_{v,T},$$

où μ est le potentiel chimique du gaz.

2) On peut montrer (complément V.F, note 1), que

$$\left(\frac{\partial\mu}{\partial\overline{N}}\right)_{v,T}=-\frac{v^2}{N^2}\left(\frac{\partial p}{\partial v}\right)_{N,T}.$$

En déduire que les fluctuations de la densité $\rho = N/v$ de particules dans v sont telles que

$$\frac{\Delta\rho}{\rho} = \left(\frac{kT\varkappa}{v}\right)^{1/2},$$

où $\varkappa \widehat{=} -\frac{1}{v}\left(\frac{\partial v}{\partial p}\right)_{N,T}$ est la compressibilité isotherme du gaz.

3) *Application à la diffusion de la lumière par l'atmosphère*

Lorsqu'une onde lumineuse (de longueur d'onde λ) traverse une région de l'espace de volume v dont la permittivité diélectrique relative ε diffère de la valeur moyenne $\bar{\varepsilon}$ qu'elle a dans l'environnement, une partie de la lumière est diffusée. On peut montrer que l'intensité lumineuse diffusée à la distance r, dans la direction θ par rapport à l'onde incidente (d'intensité I_0), est donnée par

$$I(\theta) = \frac{\pi^2 v^2}{2\lambda^4}\,\frac{1+\cos^2\theta}{r^2}\left(\overline{\varepsilon^2} - \bar{\varepsilon}^2\right) I_0 .$$

D'autre part, ε et ρ sont reliés par la formule de Clausius-Mossotti :

$$\frac{\varepsilon - 1}{\varepsilon + 2} = \alpha\rho ,$$

où α est une constante qui ne dépend que de la nature du gaz.

a) Calculer l'écart quadratique moyen de la permittivité du gaz contenu dans v.

b) En déduire l'intensité lumineuse totale diffusée à travers une sphère de rayon R centrée sur le petit volume v considéré.

c) Montrer que le résultat obtenu permet d'expliquer le bleu du ciel, ainsi que la couleur rouge du soleil à son lever et à son coucher.

SYSTÈMES EN INTERACTION

V.16. Un solide comporte N sites, donc chacun est susceptible d'accueillir zéro, un ou deux électrons. Lorsqu'un seul électron est piégé sur un site, il a pour énergie ε_0 (<0), la même pour les deux orientations possibles du spin. Lorsque deux électrons se trouvent sur le même site, leurs spins sont antiparallèles et l'énergie de la paire vaut $2\varepsilon_0 + g$, où l'énergie d'interaction g est positive. On néglige les interactions entre électrons situés sur des sites différents.

1) L'ensemble est en équilibre avec un thermostat à la température T et un réservoir d'électrons de potentiel chimique μ. Calculer le nombre moyen d'électrons piégés et leur énergie moyenne. Comparer le nombre moyen d'électrons seuls dans un piège et le nombre moyen d'électrons piégés par paires.

2) En fait, le nombre d'électrons est fixé, égal au nombre N de sites. Calculer le potentiel chimique des électrons et le nombre moyen de sites vides, de sites contenant un seul électron et de sites totalement occupés. Que valent ces nombres moyens à très basse et très haute température? Pouvait-on prévoir ces résultats?

3) Le solide est placé dans un champ magnétique $\vec{B}$. Calculer son moment magnétique moyen lorsqu'il est en contact avec un réservoir d'électrons, puis lorsque le nombre d'électrons est fixé, égal à N.

V.17. La capacité calorifique à volume constant C_V d'un gaz parfait est, comme son énergie, indépendante du volume (complément III.B, § I.3). On veut étudier ici la façon dont elle est modifiée lorsque le gaz s'écarte de l'état parfait.

Pour un gaz réel, l'équation d'état peut être écrite sous la forme du « développement du viriel » (complément V.C) :

$$p = \frac{NkT}{V}\left[1 + B_2(T)\frac{N}{V} + B_3(T)\left(\frac{N}{V}\right)^2 + \ldots\right]. \tag{1}$$

Indépendamment de leur signification théorique, les « coefficients du viriel » $B_q(T)$ peuvent être mesurés expérimentalement, pour chaque gaz, en fonction de la température T.

1) Montrer que la capacité calorifique C_V d'un système quelconque peut s'écrire

$$C_V = -T\left(\frac{\partial^2 F}{\partial T^2}\right)_{V,N}, \tag{2}$$

où F est l'énergie libre du système. En déduire la relation

$$\left(\frac{\partial C_V}{\partial V}\right)_{T,N} = T\left(\frac{\partial^2 p}{\partial T^2}\right)_{V,N}. \tag{3}$$

2) Si l'équation d'état de van der Waals :

$$\left(p + a\frac{N^2}{V^2}\right)(V - Nb) = NkT \tag{4}$$

décrit bien les propriétés du gaz considéré, comment sa capacité calorifique C_V dépend-elle du volume (cf. complément III.G, § II.2.c)?

3) On suppose maintenant que l'équation d'état du gaz étudié a été mise sous la forme (1).

a) Calculer dans ce cas la dérivée $(\partial C_V/\partial V)_{T,N}$.

b) Intégrer l'expression trouvée par rapport à V, à T constante. Montrer que la « constante d'intégration », qui est en réalité une fonction *a priori* inconnue de T, n'est autre que la capacité calorifique $C_V^{(0)}(T)$ du gaz lorsque sa densité est suffisamment faible pour qu'il soit parfait.

c) En déduire l'expression de la capacité calorifique du gaz réel :

$$C_V(T, V) = C_V^{(0)}(T) - Nk\left[\frac{N}{V}\frac{\mathrm{d}}{\mathrm{d}T}\left(T^2\frac{\mathrm{d}B_2}{\mathrm{d}T}\right) + \frac{1}{2}\left(\frac{N}{V}\right)^2\frac{\mathrm{d}}{\mathrm{d}T}\left(T^2\frac{\mathrm{d}B_3}{\mathrm{d}T}\right) + \ldots\right]. \tag{5}$$

d) Que vaut le premier terme correctif si l'on y reporte l'expression approchée de B_2 obtenue au complément V.C, formule (C.32)? Comparer ce résultat à celui de la question 2; discuter à partir de l'expression exacte de B_2 (formules (C.26)-(C.27)).

V.18. Un fluide est constitué de N molécules monoatomiques enfermées dans un récipient de volume V à la température T. Ces particules interagissent deux à deux avec une énergie potentielle d'interaction à courte portée $u(r_{ij})$ qui ne dépend que de la distance r_{ij} entre les atomes i et j considérés (cf. figure 1 du complément III.G).

1) Montrer que la fonction de partition de ce système peut se développer sous la forme

$$Z = \frac{1}{N!}\frac{1}{\lambda^{3N}}\int \mathrm{d}^3 r_1\, \mathrm{d}^3 r_2 \ldots \mathrm{d}^3 r_N \left[1 + \sum_{i<j} f_{ij} + \sum_{i<j}\sum_{\substack{i'<j' \\ (i,j)\neq(i',j')}} f_{ij} f_{i'j'} + \ldots\right],$$

où λ est la longueur d'onde thermique de de Broglie et

$$f_{ij} \widehat{=} \mathrm{e}^{-u(r_{ij})/kT} - 1 .$$

2) En ne gardant que les deux premiers termes du développement, déterminer la première correction à l'équation d'état du gaz parfait. Comparer le résultat obtenu avec le développement du viriel (complément V.C) et avec l'équation de van der Waals (complément III.G); on pourra supposer

$$u(r) \simeq \infty \qquad \text{pour} \quad r < \rho_1 ,$$
$$u(r)/kT \ll 1 \qquad \text{pour} \quad r > \rho_1 .$$

V.19. Équilibre isotherme d'un gaz réel dans le champ de pesanteur. On veut évaluer les corrections à la « formule du nivellement barométrique » (complément V.B, § II.1.b) dans le cas où le gaz n'est pas parfait. On prend l'équation d'état du gaz, comme dans l'exercice V.17, sous la forme

$$p = nkT\left[1 + B_2(T)\, n + B_3(T)\, n^2 + \ldots\right], \tag{1}$$

où $n \widehat{=} N/V$ est la densité du gaz.

1) Exprimer la condition d'équilibre du gaz dans le champ de pesanteur sous forme d'une équation différentielle pour $n(z)$, densité à l'altitude z (cf. note 7 du complément V.B).

2) On cherche $n(z)$ sous la forme d'un développement :

$$n(z) = n_0\, \mathrm{e}^{-mgz/kT}\left[1 + b_2(z)\, n_0 + b_3(z)\, n_0^2 + \ldots\right], \tag{2}$$

où n_0 est la densité à l'altitude $z = 0$.

a) Trouver l'équation différentielle donnant $b_2(z)$, et l'intégrer.

b) Indiquer comment on peut calculer de proche en proche les corrections d'ordre supérieur.

V.20. Modèle du gaz sur réseau. Pour étudier (de façon très simplifiée) le comportement des gaz réels, on utilise parfois le modèle du « gaz sur réseau ». Dans ce modèle, le volume V du récipient est découpé en petits cubes élémentaires de volume v_0 (de l'ordre de grandeur d'un volume atomique); les centres des cubes forment donc les sites d'un réseau cubique simple, les atomes du « gaz » ne pouvant occuper que l'un des $N_0 = V/v_0$ sites, à raison d'un atome au plus par site. L'interaction entre les atomes est supposée à courte portée et limitée aux plus proches voisins; elle est prise en compte de façon approchée par l'hamiltonien

$$H = -\varepsilon \sum_{(i,\, j)_v} n_i n_j ,$$

avec $n_i = 1$ si le site (i) est occupé par un atome, $n_i = 0$ s'il est vide (la somme porte sur toutes les paires de plus proches voisins).

On se propose d'étudier ce système dans le cadre du formalisme grand-canonique(*).

1) On considère tout d'abord le cas d'un « gaz » parfait, c'est-à-dire le cas $\varepsilon = 0$.

a) Calculer la grande fonction de partition d'un site; en déduire la grande fonction de partition $\Xi(T, \mu)$ du système global.

b) Calculer la pression $p(T, \mu)$ du système ainsi que le nombre moyen $N(T, \mu)$ d'atomes dans le récipient.

c) En éliminant μ entre ces deux grandeurs, déterminer l'équation d'état du « gaz ». Que devient cette équation dans la limite des faibles densités ($N \ll N_0$)?

---

* Ce modèle est complètement équivalent à celui que nous avons utilisé pour étudier les mélanges binaires; on peut donc aussi lui appliquer les méthodes du complément III.K (voir aussi les exercices III.39 et III.40).

2) On considère maintenant le cas d'un «gaz» réel, pour lequel $\varepsilon > 0$. On traite le problème dans l'approximation de champ moyen : on remplace dans H le produit $n_i n_j$ par $(n_i n + n n_j - n^2)$, où n est la valeur moyenne (supposée indépendante du site considéré) du nombre d'occupation d'un site.

a) Montrer que, dans cette approximation,

$$H \simeq H_{cm} = -6\varepsilon n \sum_i n_i + 3\varepsilon N_0 n^2 ,$$

et calculer la grande fonction de partition $\Xi(T, \mu, n)$ *pour n fixé,* ainsi que le grand potentiel $J(T, \mu, n)$.

b) Montrer que n doit vérifier une équation d'auto-cohérence et déterminer cette équation (pour cela on peut, soit calculer le nombre moyen $N = N_0 n$ de particules dans le système à partir de Ξ, soit considérer n comme une variable interne qui, à l'équilibre, rend J minimum).

c) Calculer la pression du système. En éliminant μ à l'aide de l'équation d'auto-cohérence, montrer que l'équation d'état du système s'écrit

$$p = \frac{kT}{v_0} \ln \frac{1}{1 - N v_0 / V} - 3\varepsilon v_0 \frac{N^2}{V^2} .$$

d) Donner l'allure générale des isothermes $p = f(V)$ à N et T fixés. Montrer que, si T est inférieure à une température critique T_c, le système peut devenir instable $\left(\text{compressibilité isotherme } \chi = -\frac{1}{V}\left(\frac{\partial V}{\partial p}\right)_T \text{ négative}\right)$. Déterminer les coordonnées T_c, p_c, V_c du point critique (c'est le point où s'annulent à la fois $(\partial p / \partial V)_T$ et $(\partial^2 p / \partial V^2)_T$) et montrer qu'il correspond à un nombre d'occupation moyen $n_c = 1/2$.

3) En pratique, c'est le nombre N d'atomes dans le gaz qui est fixé et non pas le potentiel chimique. Celui-ci est alors déterminé *a posteriori.* Pour $T > T_c$, n est fixé $(= N/N_0)$ et l'équation d'auto-cohérence détermine μ sans difficulté; toutes les grandeurs physiques s'en déduisent immédiatement. Le domaine $T < T_c$ est plus délicat puisque, pour certaines valeurs de N, le système ne peut rester homogène et se sépare en deux phases ayant des n différents (mais le même potentiel chimique).

a) Montrer qu'au point critique le potentiel chimique vaut $\mu_0 = -3\varepsilon$.

b) Supposons μ fixé à la valeur μ_0 et posons $n = \frac{1}{2}(1 + m)$ (avec $-1 \leqslant m \leqslant 1$). Montrer que l'équation d'auto-cohérence s'écrit alors

$$m = \text{th} \frac{3\varepsilon m}{2kT} ,$$

et que, pour $T < T_c$, elle admet trois solutions : 0, $m_0(T)$ et $-m_0(T)$. Montrer que la solution $m = 0$ correspond à un état instable du système et que les deux solutions $\pm m_0(T)$ correspondent toutes deux à un minimum de J, avec $J(m_0) = J(-m_0)$. Pour $\mu = \mu_0$ on a donc deux solutions possibles, équivalentes.

c) Supposons maintenant que μ soit fixé à une valeur différente de μ_0. Montrer que, même si l'équation d'auto-cohérence admet plusieurs solutions, il n'existe qu'un minimum absolu de J, et donc une seule valeur de m correspondant à un état d'équilibre stable du système (on pourra s'inspirer de la note 26 du complément III.J, car le calcul du grand potentiel J est ici analogue à celui de l'énergie libre d'un ferromagnétique en présence d'un champ extérieur).

d) En déduire que, pour $T < T_c$ et pour toutes les valeurs de N telles que

$$\frac{1}{2}[1 - m_0(T)] < \frac{N}{N_0} < \frac{1}{2}[1 + m_0(T)] ,$$

le potentiel chimique reste fixé à la valeur μ_0 et que le système se sépare en deux phases : une phase « liquide » de « densité » élevée $n_\ell = [1 + m_0(T)]/2$ et une phase « gazeuse » de faible « densité » $n_g = [1 - m_0(T)]/2$. Quel est alors, pour N donné, le volume occupé par chacune des deux phases?

e) Montrer que, lorsqu'on fait varier N/N_0 entre les deux limites précédentes à température constante, la pression reste constante et vaut

$$p(T) = \frac{kT}{2v_0} \ln \frac{4}{1 - m_0^2(T)} - \frac{kT_c}{2v_0}\left[1 + m_0^2(T)\right] ;$$

$p(T)$ est la « pression de vapeur saturante » à la température T; il lui correspond un « palier de liquéfaction » sur l'isotherme dans le diagramme (p, V). Dessiner l'allure de la courbe donnant p en fonction de T.

f) Calculer la « chaleur latente de vaporisation » par atome à la température T (c'est-à-dire la chaleur qu'il faut fournir pour faire passer, à température et pression constantes, un atome de l'état « liquide » à l'état « gazeux ») :

$$\ell(T) = T\left[\frac{S(n_g)}{N_0 n_g} - \frac{S(n_\ell)}{N_0 n_\ell}\right] ,$$

où $S(n)$ est l'entropie du système lorsque sa densité est n. Montrer qu'elle vérifie « l'équation de Clapeyron »

$$\ell = T\left[\frac{v_0}{n_g} - \frac{v_0}{n_\ell}\right]\frac{\mathrm{d}p}{\mathrm{d}T} .$$

UTILISATION DE L'OPÉRATEUR-DENSITÉ

V.21. On considère un système de N spins 1/2 indépendants placés dans un champ magnétique uniforme $\vec{B}_0$. L'hamiltonien de ce système s'écrit

$$H_0 = -g\mu_B B_0 \sum_{i=1}^{N} S_i^z ,$$

où S_i^z est la composante, dans la direction du champ, de l'opérateur décrivant le spin (i).

1) Donner l'expression générale de l'opérateur-densité canonique D_0 du système.

2) Montrer que, pour ce système de particules indépendantes, l'opérateur D_0 se factorise :

$$D_0 = \prod_{i=1}^{N} d_i .$$

3) Calculer explicitement la fonction de partition

$$Z_0 = \mathrm{Tr}\ \mathrm{e}^{-\beta H_0}$$

du système, ainsi que la valeur moyenne de la composante sur Oz de l'un des spins,

$$\overline{S_i^z} = \mathrm{Tr}\ D_0 S_i^z .$$

4) En déduire l'aimantation $\mathcal{M}$ du système, puis sa susceptibilité magnétique $\chi = \lim_{B_0 \to 0} \mathcal{M}/B_0$.

V.22. On considère un système de N spins 1/2 placés dans un champ magnétique uniforme $\vec{B}_0$. Les spins interagissent entre eux suivant le modèle d'Ising; l'hamiltonien du système s'écrit donc

$$H = -g\mu_B B_0 \sum_{i=1}^{N} S_i^z - J \sum_{(i,\,j)_v} S_i^z S_j^z , \qquad J > 0 , \tag{1}$$

où la deuxième somme porte sur les paires de plus proches voisins. L'ensemble est en équilibre avec un thermostat à la température T.

On se propose d'étudier ce problème en utilisant le formalisme des opérateurs-densités, et plus précisément d'en donner une solution approchée par la méthode variationnelle (complément V.E, § II.2).

Pour cela on choisit pour famille d'essai les opérateurs-densités canoniques $D_0(X)$ associés à l'hamiltonien

$$H_0 = -g\mu_B X \sum_{i=1}^{N} S_i^z , \tag{2}$$

X étant un paramètre. On sait que l'énergie libre F du système vérifie

$$F \leqslant \overline{E}(D_0) - TS(D_0) \quad \text{quel que soit } X, \tag{3}$$

où $\overline{E}(D_0)$ et $S(D_0)$ sont l'énergie moyenne et l'entropie du système, d'hamiltonien H, lorsqu'il est décrit par l'opérateur-densité $D_0(X)$.

1) Montrer que l'inégalité (3) peut encore s'écrire

$$F \leqslant \mathrm{Tr}\left[D_0(H - H_0)\right] - kT \ln Z_0 , \tag{4}$$

où Z_0 est la fonction de partition qu'aurait le système si son hamiltonien était H_0.

2) Montrer que, pour $i \neq j$,

$$\mathrm{Tr}\ D_0 S_i^z S_j^z = (\mathrm{Tr}\ D_0 S_i^z)(\mathrm{Tr}\ D_0 S_j^z) . \tag{5}$$

3) En utilisant les résultats de l'exercice V.21, en déduire l'expression du second membre de l'inégalité (4) en fonction de X.

4) Pour quelle valeur X_0 de X le second membre de (4) est-il minimum? Exprimer l'énergie libre approchée ainsi obtenue en fonction de X_0. Comparer les résultats avec l'équation d'auto-cohérence et l'énergie libre calculées dans l'approximation de champ moyen (complément III.J, § II.1); en déduire une interprétation physique de X_0.

VI

I. GAZ DE FERMIONS INDÉPENDANTS

A. Propriétés générales des gaz de fermions indépendants
   1. Le facteur de Fermi : variations et courbe représentative
   2. Expression des diverses grandeurs physiques
      a. Formules grand-canoniques rigoureuses
      b. Cas des systèmes thermodynamiques. Interprétation canonique des formules
      c. Exemple : gaz de fermions libres

B. Comportement à basse température d'un système thermodynamique de fermions libres
   1. Propriétés du système à température nulle
      a. Le potentiel chimique à $T = 0$
      b. L'énergie à $T = 0$
      c. Pression quantique
   2. Développement des grandeurs physiques à basse température
      a. Méthode de développement
      b. Potentiel chimique
      c. Énergie et chaleur spécifique
      d. Grand potentiel, entropie, énergie libre
      e. Pression et équation d'état
   3. Domaine de validité de l'approximation

II. GAZ DE BOSONS INDÉPENDANTS

A. Propriétés générales des gaz de bosons indépendants
   1. Le facteur de Bose : variations et courbe représentative
   2. Expression des diverses grandeurs physiques
      a. Formules grand-canoniques rigoureuses
      b. Cas des systèmes thermodynamiques. Interprétation canonique des formules
      c. Exemple : gaz de bosons libres

B. Comportement à basse température d'un système thermodynamique de bosons libres
   1. Le phénomène de condensation de Bose
      a. Étude critique de la relation entre le nombre de particules et le potentiel chimique
      b. Répartition des particules entre les états individuels
      c. Nombre de particules condensées
      d. Importance physique du phénomène
   2. Expression des grandeurs physiques dans le cas de condensation de Bose
      a. Énergie et chaleur spécifique
      b. Grand potentiel, entropie, énergie libre
      c. Pression et équation d'état

C. La transition de phase et ses caractéristiques
   1. Comportement des grandeurs physiques à nombre de particules et volume constants
   2. Étude de la transition de phase
      a. Changement de variables et de point de vue
      b. Courbes caractérisant la transition de phase
      c. Caractérisation des deux phases en présence

III. GAZ DE PHOTONS. THERMODYNAMIQUE DU RAYONNEMENT

A. Ondes électromagnétiques et photons
   1. Modes propres électromagnétiques d'une enceinte
      a. Caractérisation des modes propres
      b. Densité spectrale de modes propres
   2. Quantification des modes propres : photons
      a. Impulsion, énergie et état de spin d'un photon
      b. Densité d'états individuels dans un gaz de photons
   3. Propriétés statistiques des photons
      a. Caractéristiques particulières d'un gaz de photons
      b. Fonction de distribution des photons

B. Lois du rayonnement thermique
   1. Loi de Planck
      a. Démonstration de la loi de Planck
      b. Discussion : validité de la loi
      c. Courbe de répartition spectrale de l'énergie
      d. Limite de basse fréquence et formule classique
   2. Propriétés du rayonnement thermique
      a. Énergie totale : loi de Stéfan-Boltzmann
      b. Grand potentiel
      c. Pression, entropie, énergie libre

CHAPITRE VI

Gaz parfaits quantiques

Nous nous proposons d'étudier dans ce chapitre les *systèmes de particules identiques,* en prenant correctement en compte les *effets quantiques* liés au postulat de symétrisation (appendice II, § IV.2). Cette étude est considérablement facilitée si l'on suppose, ce que nous ferons ici, que *les interactions entre particules du système sont négligeables,* de sorte que ces particules sont *indépendantes.* Il suffit alors de connaître la densité d'états individuels pour exprimer les grandeurs physiques à partir des fonctions de distribution quantiques (de Fermi-Dirac ou de Bose-Einstein, suivant les cas) que nous avons écrites à la fin du chapitre précédent (V, § IV.B).

L'exemple le plus simple est celui où les particules du système sont sans structure interne et *libres,* c'est-à-dire sans *aucune interaction,* ni entre elles, ni avec des objets situés en-dehors du système considéré. C'est principalement sur cet exemple que nous illustrerons les considérations générales. La densité d'états individuels s'écrit alors (I, formule (III.26))

$$\rho(\varepsilon) = A V \varepsilon^{1/2} , \tag{0.1}$$

$$\text{avec} \quad A = \frac{2s+1}{4\pi^2} \left(\frac{2m}{\hbar^2}\right)^{3/2} \tag{0.2}$$

(V est le volume du système, m la masse d'une particule et s son spin); l'énergie ε_0 du niveau fondamental individuel est dans ce cas nulle[1] :

$$\varepsilon_0 = 0 . \tag{0.3}$$

Soulignons d'emblée que l'hypothèse d'indépendance des particules, essentielle dans ce qui suit, limite assez sévèrement la validité des résultats que nous allons obtenir : les effets quantiques se manifestent principalement à basse température, et les systèmes réels où les interactions peuvent vraiment être négligées dans ces conditions sont extrêmement rares. Ajoutons que la prise en compte, même approchée, des interactions entre particules soulève de redoutables difficultés, qu'il n'est pas question d'aborder ici (c'est le sujet de ce que l'on appelle le « problème à N corps »). Il existe cependant des cas (électrons dans un solide cristallin, par exemple; voir complément VI.A) où l'on obtient une bonne approximation en traitant les interactions entre particules du système de façon globale et moyenne : les diverses particules sont considérées comme indépendantes, chacune d'elles étant soumise au potentiel moyen créé par l'ensemble des autres (c'est l'approximation du « champ moyen »). Disons également, pour justifier l'étude que nous allons entreprendre malgré ses limites, qu'elle fait apparaître des phénomènes et des comportement cruciaux pour la compréhension qualitative et semi-quantitative de nombreux systèmes réels.

---

1. A strictement parler, ceci n'est vrai que si l'on a choisi des conditions aux limites périodiques (appendice II, § II.4.c). Pour des conditions aux limites strictes (par exemple dans une enceinte cubique d'arête $L = V^{1/3}$), on aurait

$$\varepsilon_0 = \frac{3\pi^2\hbar^2}{2mV^{2/3}}$$

(cf. appendice II, formule (34)). La différence est inappréciable, car nous considérerons toujours dans la suite des gaz enfermés dans une enceinte de *volume macroscopique* ($\varepsilon_0 \simeq 10^{-12}$ eV pour un électron enfermé dans une boîte de 1 mm^3).

Nous allons considérer séparément les gaz de *fermions* indépendants (première partie), puis ceux de *bosons* (deuxième partie). La troisième partie sera consacrée à une espèce particulière de bosons, les *photons* : les photons sont des particules réellement indépendantes, dont le nombre n'est pas conservé; on peut calculer exactement les propriétés des gaz de photons, qui sont celles du *rayonnement du corps noir.*

I. GAZ DE FERMIONS INDÉPENDANTS

Nous considérons ici les *systèmes de fermions identiques et indépendants.* Nous indiquons d'abord (§ A) comment on peut de façon générale déterminer leurs propriétés et les grandeurs macroscopiques qui leur sont associées. Les expressions obtenues ne permettant pas un calcul analytique explicite pour des valeurs quelconques des deux paramètres T et μ, nous développons au paragraphe B une *approximation valable à basse température,* domaine où les effets quantiques se font pleinement sentir.

A. Propriétés générales des gaz de fermions indépendants

1. LE FACTEUR DE FERMI : VARIATIONS ET COURBE REPRÉSENTATIVE

Les gaz de fermions suivent la statistique de Fermi-Dirac : le nombre moyen $\overline{N}_\lambda^F$ de particules occupant l'état individuel (λ), d'énergie ε_λ, est donné par (V, formule (IV.13))

$$\overline{N}_\lambda^F = \frac{1}{e^{\beta(\varepsilon_\lambda - \mu)} + 1} \,. \tag{I.1}$$

$\beta = 1/kT$ est positif, et le potentiel chimique μ n'est ici soumis *a priori* à aucune contrainte.

Notons tout d'abord que, comme le nombre d'occupation N_λ ne peut prendre, dans le cas des fermions, que les valeurs 0 et 1, on a nécessairement

$$0 \leqslant \overline{N}_\lambda^F \leqslant 1 \,, \tag{I.2}$$

ce que l'on vérifie aisément sur l'expression (I.1)

De plus $\overline{N}_\lambda^F$ est une *fonction décroissante de l'énergie* ε_λ, à β et μ fixés. Ce comportement est qualitativement analogue à celui auquel nous sommes habitués depuis le chapitre III : les états individuels sont d'autant moins peuplés que leur énergie est plus élevée.

Si l'on oublie provisoirement le fait que les valeurs possibles de ε_λ sont discrètes et qu'elles sont bornées inférieurement par le niveau fondamental individuel ε_0, on peut introduire la fonction

$$N^F(\varepsilon;\, T,\, \mu) \widehat{=} \frac{1}{e^{\beta(\varepsilon - \mu)} + 1} \quad (\beta = 1/kT) \tag{I.3}$$

qui interpole les valeurs de $\overline{N}_\lambda^F$, et tracer la courbe représentant cette fonction de l'énergie ε, à T et μ fixés. On voit facilement que cette courbe est *symétrique*[2] par

2. En effet, prenons une énergie ε' telle que

$$\varepsilon' - \mu = -(\varepsilon - \mu) \,. \tag{1}$$

On a alors

$$N^F(\varepsilon';\, T,\, \mu) = \frac{1}{e^{-\beta(\varepsilon - \mu)} + 1} = 1 - N^F(\varepsilon;\, T,\, \mu) \,, \tag{2}$$

rapport au point d'abscisse $\varepsilon = \mu$ et d'ordonnée 1/2, et qu'elle a l'allure représentée sur la figure 1.

Précisons le rôle des paramètres T et μ. Le potentiel chimique μ fixe le centre de la courbe, pour lequel N^F vaut 1/2; $\overline{N}^F_\lambda$ est donc supérieur ou inférieur à 1/2 suivant que ε_λ se place au-dessous ou au-dessus de μ. Modifier le potentiel chimique μ, à température constante, revient à faire glisser la courbe vers la droite ou vers la gauche, sans déformation, par rapport à l'échelle des énergies.

Quant à la température T, son rôle peut être caractérisé en remarquant que N^F passe de valeurs proches de 1 à des valeurs proches de 0 sur un domaine d'énergie dont la largeur est de l'ordre de kT (fig. 1.a)[3]. Lorsque T tend vers zéro, la courbe se rapproche (fig. 1.b) d'une «marche d'escalier» discontinue : à la limite en effet[4],

$$N^F(\varepsilon; T, \mu) \xrightarrow[T \to 0]{} \begin{cases} 1 \text{ si } \varepsilon < \mu , \\ 0 \text{ si } \varepsilon > \mu . \end{cases} \tag{I.4}$$

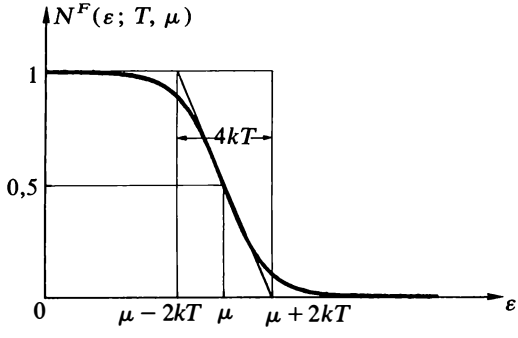


FIGURE 1. a

Courbe représentative de la fonction de Fermi (I.3) en fonction de l'énergie, à β et μ fixés; cette courbe est symétrique par rapport au point (μ, 1/2), où sa tangente a pour pente $-\beta/4$[3].

3. La dérivée de $N^F(\varepsilon)$,

$$\frac{dN^F}{d\varepsilon} = \frac{-\beta\, e^{\beta(\varepsilon-\mu)}}{[e^{\beta(\varepsilon-\mu)}+1]^2} , \tag{1}$$

donne pour pente de la tangente à la courbe en son centre

$$\frac{dN^F}{d\varepsilon}(\varepsilon = \mu) = -\frac{\beta}{4} = -\frac{1}{4kT} . \tag{2}$$

On en déduit facilement que cette tangente coupe les horizontales d'ordonnées 0 et 1 aux points d'abscisses $\mu \pm 2kT$ (voir fig. 1.a). Ceci fixe semi-quantitativement la largeur du domaine en ε sur lequel N^F varie de façon importante.

4. Quand $T \to 0$, $\beta \to +\infty$ et

$$e^{\beta(\varepsilon-\mu)} \to +\infty \text{ si } \varepsilon - \mu > 0 ,$$
$$\to 0 \quad \text{si } \varepsilon - \mu < 0 .$$

D'où le comportement (I.4) de N^F.

Au contraire, la transition entre 1 et 0 est d'autant plus progressive que la température est plus haute (fig. 1.b).

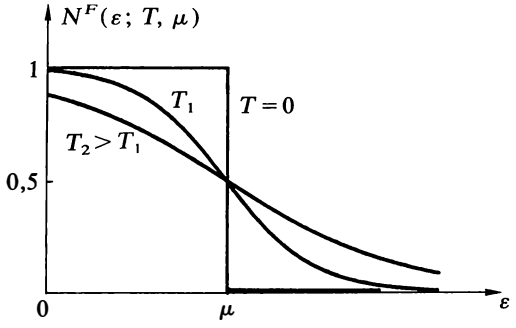


FIGURE 1. b

Influence de la température sur la fonction de Fermi : pour $T=0$, la courbe a la forme d'une « marche d'escalier » abrupte (formule (I.4)); lorsque la température croît, cette « marche » s'arrondit progressivement, la variation de N^F devenant plus lente.

2. EXPRESSION DES DIVERSES GRANDEURS PHYSIQUES

a. Formules grand-canoniques rigoureuses

Les formules que nous avons obtenues au chapitre V (§ IV.A.2) peuvent être facilement transformées pour donner les propriétés d'un système de fermions indépendants en fonction du facteur de Fermi (I.1).

Le *nombre moyen de particules* $\overline{N}$ et l'*énergie moyenne* $\overline{E}$ sont déjà sous cette forme dans les expressions V (IV.7) et V (IV.9) :

$$\overline{N} = \sum_{(\lambda)} \overline{N}_\lambda^F , \tag{I.5}$$

$$\overline{E} = \sum_{(\lambda)} \overline{N}_\lambda^F \varepsilon_\lambda . \tag{I.6}$$

Pour les autres grandeurs, il suffit d'utiliser la relation V (IV.14) liant la grande fonction de partition ξ_λ^F de l'état (λ) et son nombre d'occupation moyen $\overline{N}_\lambda^F$. On obtient ainsi, à partir des résultats du chapitre V (§ IV.A.2), les expressions suivantes[5] pour le *grand potentiel :*

$$J = kT \sum_{(\lambda)} \ln\left(1 - \overline{N}_\lambda^F\right) , \tag{I.7}$$

5. Les expressions de J et p découlent directement des formules V (IV.6) et V (IV.11), dans lesquelles on reporte V (IV.14).

Quant à l'entropie S, la formule V (IV.10) donne

$$S = k \sum_{(\lambda)} \left[-\ln\left(1 - \overline{N}_\lambda^F\right) + \beta(\varepsilon_\lambda - \mu)\, \overline{N}_\lambda^F \right] . \tag{1}$$

.../...

pour l'*entropie* grand-canonique :

$$S = -k \sum_{(\lambda)} \left[(1 - \overline{N}^F_\lambda) \ln (1 - \overline{N}^F_\lambda) + \overline{N}^F_\lambda \ln \overline{N}^F_\lambda \right], \tag{I.8}$$

et pour la *pression* grand-canonique[6] :

$$p = -kT \frac{\partial}{\partial V} \sum_{(\lambda)} \ln (1 - \overline{N}^F_\lambda). \tag{I.9}$$

Ces formules permettent de calculer les propriétés grand-canoniques d'un système quelconque de fermions identiques *indépendants* si l'on connaît l'ensemble des états (λ) possibles pour chacun de ces fermions et le spectre des énergies individuelles ε_λ correspondantes.

b. Cas des systèmes thermodynamiques. Interprétation canonique des formules

Pour des systèmes suffisamment grands (limite thermodynamique), on peut simplifier les formules du paragraphe précédent : d'une part, *le spectre des énergies individuelles ε_λ est pratiquement continu*[7], ce qui permet de remplacer la somme sur les états individuels (λ) par une intégrale sur l'énergie ε, à condition de connaître la *densité d'états individuels* $\rho(\varepsilon)$; d'autre part, *les fluctuations sont négligeables,* de sorte que les variables internes (comme le nombre de particules et l'énergie) ne s'écartent pas de leur valeur moyenne.

Cette seconde simplification permet de réinterpréter les résultats obtenus ci-dessus dans le langage canonique, c'est-à-dire de les appliquer à des *systèmes dont le nombre N de particules est fixé.* Précisons ce point. Reprenons pour cela la formule (I.5); à la limite thermodynamique, on peut remplacer au premier membre la valeur moyenne $\overline{N}$ par le nombre N de particules, et exprimer la somme sur les états individuels comme une intégrale sur l'énergie :

$$N = \int_{\varepsilon_0}^{+\infty} d\varepsilon \rho(\varepsilon) N^F(\varepsilon;\, T,\, \mu) = \int_{\varepsilon_0}^{+\infty} d\varepsilon \rho(\varepsilon) \frac{1}{e^{\beta(\varepsilon - \mu)} + 1} \quad \text{(lim. thermo.)}. \tag{I.10}$$

La relation (I.10) a été déduite du formalisme grand-canonique. Elle donne le nombre N de particules d'un *système thermodynamique* en équilibre avec un réservoir qui fixe le potentiel chimique μ (et la température T). Mais, pour de tels systèmes, les descriptions canonique et grand-canonique sont équivalentes (V, § III.B), de sorte que

---

Mais on voit d'après (I.1) que

$$e^{\beta(\varepsilon_\lambda - \mu)} = \frac{1 - \overline{N}^F_\lambda}{\overline{N}^F_\lambda}, \tag{2}$$

de sorte que

$$\beta(\varepsilon_\lambda - \mu) = \ln (1 - \overline{N}^F_\lambda) - \ln \overline{N}^F_\lambda. \tag{3}$$

Il suffit de reporter cette expression dans la formule (1) ci-dessus pour obtenir le résultat (I.8).

6. Ce sont bien sûr les niveaux d'énergie individuels ε_λ qui, dans l'expression (I.1), dépendent du volume V du système.

7. Dans certains cas, par exemple celui des électrons dans un cristal (complément I.F, § II), le spectre des énergies individuelles présente une *structure en bandes :* les énergies des divers états individuels possibles se regroupent dans des domaines d'extension finie (bandes permises), séparés par des intervalles, eux aussi finis, auxquels ne correspond aucun état (bandes interdites). Les formules qui suivent sont applicables à de tels cas : dans les bandes permises, les valeurs de ε_λ sont suffisamment serrées pour qu'on puisse définir une densité d'états; dans les bandes interdites, la densité d'états est nulle.

si le nombre de particules du système est fixé à N, son potentiel chimique μ est déterminé par la relation (I.10) (à condition bien sûr que la *limite thermodynamique* soit atteinte).

Arrêtons-nous un instant pour analyser la démarche précédente. Nous avions constaté, au chapitre III (§ IV.A.3), que les systèmes de particules identiques, indépendantes et indiscernables ne peuvent être traités exactement dans le cadre canonique : c'est la contrainte sur le nombre total de particules qui empêche le calcul explicite de la fonction de partition. Le formalisme grand-canonique, dans lequel cette contrainte n'existe plus, permet un traitement exact de ces systèmes. Pour les systèmes thermodynamiques, dans lesquels les fluctuations sont négligeables, les formules grand-canoniques se traduisent immédiatement dans le langage canonique (N fixé au lieu de μ)[8].

Considérons un système thermodynamique de fermions indépendants, dont on a fixé la température T et le nombre N de particules. Sa densité d'états individuels $\rho(\varepsilon)$ dépend du volume V (et éventuellement d'autres paramètres extérieurs), mais pas de N et T; ces deux dernières variables figurent explicitement dans la condition (I.10) qui détermine le potentiel chimique μ; celui-ci apparaît donc comme une *fonction implicite de T, N, V, ... :*

$$\mu = \mu(T, N, V, \ldots) , \quad \text{solution de (I.10).} \tag{I.11}$$

Il est facile de montrer[9] que μ est une *fonction croissante de N* (à T, V, ... fixés). Les autres grandeurs associées à ce système peuvent être obtenues à partir des formules du paragraphe a, transformées en intégrales sur l'énergie individuelle ε, et dans lesquelles μ *est à remplacer par sa valeur tirée de la condition* (I.10). Par exemple, l'énergie E du système est ici donnée par

$$\begin{cases} E = \displaystyle\int_{\varepsilon_0}^{+\infty} \mathrm{d}\varepsilon\, \varepsilon \rho(\varepsilon) N^F(\varepsilon; T, \mu) = \int_{\varepsilon_0}^{+\infty} \mathrm{d}\varepsilon\, \varepsilon \rho(\varepsilon) \dfrac{1}{\mathrm{e}^{\beta(\varepsilon-\mu)}+1} , \\ \mu \text{ solution de (I.10)} . \end{cases} \tag{I.12}$$

Bien entendu, le fait que μ n'est connu qu'implicitement complique considérablement la situation (voir l'exemple du paragraphe c ci-dessous).

c. Exemple : gaz de fermions libres

Illustrons les considérations précédentes sur le cas de fermions identiques *libres et sans structure interne.* La densité d'états individuels est alors donnée par (0.1)-(0.2) (rappelons que, pour des fermions, s est demi-entier : $s = 1/2$, ou $3/2$, ...).

---

8. Nous avions su calculer la fonction de partition canonique dans l'approximation de Maxwell-Boltzmann (III, § IV.A.3.c) qui, pour une densité N/V fixée, est valable à suffisamment haute température. Les résultats que nous obtenons ici ne sont pas limités ainsi à un domaine particulier de température ou de potentiel chimique; la seule condition concerne la validité de la limite thermodynamique, qui est pratiquement assurée pour tout système macroscopique (cf. III, § III.A).

9. Dans la relation (I.10), augmentons N à T, V, ... constants. L'intégrale du second membre doit alors croître, ce qu'elle ne peut faire que si μ croît (la croissance de μ entraîne celle du facteur de Fermi pour tout ε; cf. fig. 1). Ce résultat n'est d'ailleurs que l'illustration, dans ce cas particulier, d'une propriété tout à fait générale (cf. formule III (II.52)).

On peut aussi montrer que, si la densité d'états individuels $\rho(\varepsilon)$ est une fonction non décroissante de ε, alors μ *est une fonction décroissante de* T à N, V... fixés. En effet, une augmentation de T se traduirait, si μ restait invariable, par une augmentation de l'intégrale (le facteur de Fermi croîtrait pour $\varepsilon > \mu$ et décroîtrait de façon symétrique pour $\varepsilon < \mu$, mais le domaine d'intégration favorise la région $\varepsilon > \mu$, puisqu'il ne commence qu'en $\varepsilon = 0$ et s'étend par contre jusqu'à $+\infty$); pour empêcher la croissance de l'intégrale (N est maintenu constant), il faut que μ décroisse.

Le potentiel chimique est déterminé à partir de T, N et V par la condition (I.10), qui s'écrit ici

$$N = AV \int_0^{+\infty} d\varepsilon \frac{\varepsilon^{1/2}}{e^{\beta(\varepsilon-\mu)}+1} . \tag{I.13}$$

Notons que μ ne dépend pas séparément de N et V, mais seulement de la *densité de particules* N/V[10]. Mais l'intégrale du second membre n'est pas calculable analytiquement comme fonction des deux paramètres β et μ. On peut cependant l'évaluer numériquement sans difficulté et donc déterminer μ par *ajustement numérique :* pour travailler avec des variables sans dimension, introduisons la fugacité

$$\varphi \widehat{=} e^{\beta\mu} \tag{I.14}$$

et posons

$$x \widehat{=} \beta\varepsilon \ ; \tag{I.15}$$

l'équation (I.13) à résoudre prend alors la forme

$$\frac{4\pi^2}{2s+1}\left(\frac{\hbar^2\beta}{2m}\right)^{3/2}\frac{N}{V} = \int_0^{+\infty} dx \frac{x^{1/2}}{1+e^x/\varphi} , \tag{I.16}$$

où l'inconnue est la fugacité φ. La figure 2 montre que, *quelles que soient la température T et la densité* N/V (positives), l'équation (I.16) admet *une solution unique,* à partir de laquelle, connaissant T, on calcule aussitôt μ.

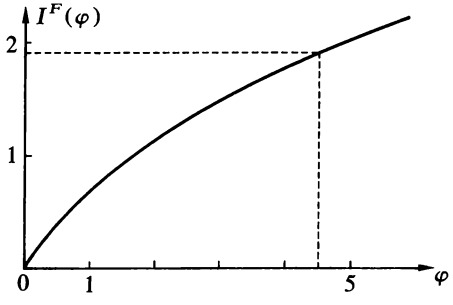


FIGURE 2

Solution graphique de l'équation (I.13), prise sous la forme (I.16). La courbe représente

$$I^F(\varphi) \widehat{=} \int_0^{+\infty} dx \frac{x^{1/2}}{1+e^x/\varphi}$$

en fonction de $\varphi \widehat{=} \exp(\beta\mu)$. c'est une fonction monotone croissante de 0 à $+\infty$, qu'il faut couper par l'horizontale d'ordonnée positive

$$\frac{N}{V}\frac{\beta^{3/2}}{A} = \frac{4\pi^2}{2s+1}\left(\frac{\hbar^2\beta}{2m}\right)^{3/2}\frac{N}{V}$$

pour obtenir φ (et donc μ) en fonction de T, N et V. Il existe une solution et une seule pour chaque ensemble $\{T, N, V\}$.

10. μ est une grandeur intensive (complément II.A, § II). N et V sont les seules grandeurs extensives du problème. Par conséquent μ ne peut être fonction que de N/V.

On peut ensuite reporter la valeur ainsi obtenue pour μ (ou directement celle de φ) dans les formules donnant les diverses grandeurs physiques. C'est ainsi que l'*énergie* du système s'obtient par

$$\begin{cases} E = AV \displaystyle\int_0^{+\infty} \mathrm{d}\varepsilon \, \frac{\varepsilon^{3/2}}{\mathrm{e}^{\beta(\varepsilon-\mu)}+1} , \\ \mu \text{ solution de (I.13).} \end{cases} \tag{I.17}$$

Pour calculer les autres grandeurs, le plus simple est sans doute de commencer par le *grand potentiel J :*

$$J = AVkT \int_0^{+\infty} \mathrm{d}\varepsilon \, \varepsilon^{1/2} \ln\left[1 - N^F(\varepsilon ; T, \mu)\right] , \tag{I.18}$$

que l'on peut réécrire, à l'aide d'une intégration par parties[11],

$$J = -\frac{2}{3} AV \int_0^{+\infty} \mathrm{d}\varepsilon \, \frac{\varepsilon^{3/2}}{\mathrm{e}^{\beta(\varepsilon-\mu)}+1} . \tag{I.19}$$

On reconnaît, à un facteur près, l'énergie E du système (formule (I.17)) :

$$J = -\frac{2}{3} E . \tag{I.20}$$

Le grand potentiel est ici proportionnel à l'énergie.

---

11. $$1 - N^F(\varepsilon ; T, \mu) = \frac{1}{1+\mathrm{e}^{-\beta(\varepsilon-\mu)}} . \tag{1}$$

On a donc

$$\frac{\partial}{\partial \varepsilon} \ln\left[1 - N^F(\varepsilon ; T, \mu)\right] = \beta \frac{\mathrm{e}^{-\beta(\varepsilon-\mu)}}{1+\mathrm{e}^{-\beta(\varepsilon-\mu)}} = \beta N^F(\varepsilon ; T, \mu) . \tag{2}$$

L'intégration par parties s'effectue alors simplement :

$$J = \left[\frac{2}{3} AVkT\varepsilon^{3/2} \ln\left[1 - N^F(\varepsilon ; T, \mu)\right]\right]_{\varepsilon=0}^{\varepsilon=+\infty} - \frac{2}{3} AVkT\beta \int_0^{+\infty} \varepsilon^{3/2} N^F(\varepsilon ; T, \mu) \mathrm{d}\varepsilon . \tag{3}$$

Le terme tout intégré peut ensuite être déterminé à partir de la formule (1) :

$$\varepsilon^{3/2} \ln\left[1 - N^F(\varepsilon ; T, \mu)\right] = -\varepsilon^{3/2} \ln\left[1 + \mathrm{e}^{-\beta(\varepsilon-\mu)}\right] . \tag{4}$$

Cette expression est visiblement nulle pour $\varepsilon = 0$; pour ε tendant vers $+\infty$, on peut la développer sous la forme

$$-\varepsilon^{3/2} \mathrm{e}^{-\beta(\varepsilon-\mu)} \xrightarrow[\varepsilon \to +\infty]{} 0 . \tag{5}$$

Le terme tout intégré est donc nul, d'où le résultat (I.19).

Noter que le facteur $-2/3$ qui figure dans la formule (I.20) est ainsi directement relié au fait que la densité d'états $\rho(\varepsilon)$ varie en $\varepsilon^{1/2}$.

Si l'on change d'origine pour l'énergie potentielle individuelle, c'est-à-dire si l'énergie ε d'une particule d'impulsion $\hbar\vec{K}$ s'écrit

$$\varepsilon = \varepsilon_0 + \frac{\hbar^2 \vec{K}^2}{2m} , \tag{6}$$

la formule (I.20) devient

$$J = -\frac{2}{3}(E - N\varepsilon_0) , \tag{7}$$

où E et N sont donnés, en fonction des variables grand-canoniques T, μ, V, par (I.10) et (I.12), avec $\rho(\varepsilon) = AV(\varepsilon - \varepsilon_0)^{1/2}$.

Du grand potentiel on déduit l'*entropie S* par la relation

$$S = -\left(\frac{\partial J}{\partial T}\right)_{\mu,\, V} . \tag{I.21}$$

Le calcul se fait sans difficulté à partir de la forme (I.19) de J, qui fait intervenir explicitement μ :

$$\begin{cases} S = k\beta^2 \times \dfrac{2}{3} AV \displaystyle\int_0^{+\infty} \mathrm{d}\varepsilon\, \varepsilon^{3/2} \frac{(\varepsilon - \mu)\, \mathrm{e}^{\beta(\varepsilon-\mu)}}{[\mathrm{e}^{\beta(\varepsilon-\mu)}+1]^2} , \\ \mu \text{ solution de (I.13).} \end{cases} \tag{I.22}$$

L'expression de l'*énergie libre* thermodynamique

$$F = E - TS \tag{I.23}$$

découle aussitôt des formules (I.17) et (I.22).

Quant à la *pression p* du gaz de fermions, on l'obtient à l'aide d'une des deux relations

$$p = -\left(\frac{\partial J}{\partial V}\right)_{T,\, \mu} \tag{I.24}$$

ou

$$p = -J/V , \tag{I.25}$$

qui sont ici équivalentes (voir V, § II.C.2). Il vient ainsi

$$p = \frac{2}{3}\frac{E}{V} = \frac{2}{3} A \int_0^{+\infty} \mathrm{d}\varepsilon\, \frac{\varepsilon^{3/2}}{\mathrm{e}^{\beta(\varepsilon-\mu)}+1} . \tag{I.26}$$

On vérifie bien que, si l'on garde les variables grand-canoniques T, μ, V, la pression p dépend seulement de μ et T, pas du volume V (cf. V, formule (II.36)). Cependant, pour aboutir à l'*équation d'état* du gaz (relation entre la pression, la température, le volume et le nombre de particules), il faut éliminer μ au profit de N (ce qui introduit également le volume V). Ceci n'étant possible qu'implicitement, l'équation d'état du gaz est donnée sous forme paramétrique :

$$\begin{cases} p = \dfrac{2s+1}{6\pi^2} \left(\dfrac{2m}{\hbar^2}\right)^{3/2} \displaystyle\int_0^{+\infty} \mathrm{d}\varepsilon\, \frac{\varepsilon^{3/2}}{\mathrm{e}^{\beta(\varepsilon-\mu)}+1} , \\ \mu \text{ solution de (I.13).} \end{cases} \tag{I.27}$$

On est apparemment loin de l'équation d'état simple des gaz parfaits classiques!

La prise en compte du postulat de symétrisation quantique modifie donc profondément l'expression des grandeurs physiques associées à un gaz parfait. Il est pourtant facile de montrer[12] que, comme on s'y attend, *on retrouve, dans le domaine des hautes températures, les propriétés des gaz parfaits classiques.* C'est surtout à basse température, nous allons le voir au paragraphe B, que les effets quantiques sont importants.

---

12. Si $\mathrm{e}^{\beta(\varepsilon-\mu)} \gg 1$ pour tout $\varepsilon \geqslant 0$, (1)

la condition (I.13) se simplifie en

$$N = AV \mathrm{e}^{\beta\mu} \int_0^{+\infty} \mathrm{d}\varepsilon\, \varepsilon^{1/2}\, \mathrm{e}^{-\beta\varepsilon} . \tag{2}$$

L'intégrale se calcule facilement : si l'on pose

$$y \mathrel{\widehat{=}} (\beta\varepsilon)^{1/2} , \tag{3}$$

elle s'écrit

$$\int_0^{+\infty} \mathrm{d}\varepsilon\, \varepsilon^{1/2}\, \mathrm{e}^{-\beta\varepsilon} = \frac{2}{\beta^{3/2}} \int_0^{+\infty} y^2\, \mathrm{d}y\, \mathrm{e}^{-y^2} = \frac{2}{\beta^{3/2}} \frac{\sqrt{\pi}}{4} \tag{4}$$

.../...

B. Comportement à basse température d'un système thermodynamique de fermions libres

Les formules que nous venons d'écrire font intervenir des intégrales qui ne sont pas calculables analytiquement; il n'est donc pas possible d'obtenir des expressions explicites pour les diverses grandeurs physiques. L'exemple simple du gaz de *fermions libres* (§ A.2.c) illustre bien cette difficulté. C'est encore à cet exemple que nous nous limiterons dans tout ce paragraphe B.

A haute température, nous le savons (V, § IV.C), la statistique de Fermi-Dirac se réduit à celle de Maxwell-Boltzmann; le gaz de fermions libres devient alors un *gaz*

(voir formulaire). L'équation (2) donne alors

$$e^{\beta\mu} = \frac{N}{V}\frac{2\beta^{3/2}}{A\sqrt{\pi}}, \tag{5}$$

c'est-à-dire, si l'on remplace A par sa valeur (0.2),

$$\mu = -kT \ln\left[\frac{V}{N}(2s+1)\left(\frac{m}{2\pi\hbar^2\beta}\right)^{3/2}\right]. \tag{6}$$

On reconnaît là l'expression III (IV.47) du potentiel chimique d'un gaz parfait monoatomique.

La formule (I.17) devient de même, dans la limite (1),

$$E = AV\, e^{\beta\mu} \int_0^{+\infty} d\varepsilon\, \varepsilon^{3/2}\, e^{-\beta\varepsilon}. \tag{7}$$

Or
$$\int_0^{+\infty} d\varepsilon\, \varepsilon^{3/2}\, e^{-\beta\varepsilon} = \frac{2}{\beta^{5/2}} \int_0^{+\infty} dy\, y^4\, e^{-y^2} = \frac{2}{\beta^{5/2}}\frac{3}{8}\sqrt{\pi} \tag{8}$$

(voir formulaire). Compte tenu de (5), on retrouve bien le résultat III (IV.37) attendu :

$$E = \frac{3}{2} NkT. \tag{9}$$

On en déduit aussitôt, grâce à (I.26), la pression p et l'équation d'état :

$$p = \frac{NkT}{V}. \tag{10}$$

Enfin, l'expression (I.22) de l'entropie s'écrit ici

$$S = \frac{2}{3} k\beta^2 AV\, e^{\beta\mu} \int_0^{+\infty} d\varepsilon\, \varepsilon^{3/2}(\varepsilon - \mu)\, e^{-\beta\varepsilon}. \tag{11}$$

En utilisant (5), (6), (8) et

$$\int_0^{+\infty} d\varepsilon\, \varepsilon^{5/2}\, e^{-\beta\varepsilon} = \frac{2}{\beta^{7/2}} \int_0^{+\infty} dy\, y^6\, e^{-y^2} = \frac{2}{\beta^{7/2}}\frac{15}{16}\sqrt{\pi}, \tag{12}$$

on obtient

$$S = Nk\left\{\frac{5}{2} + \ln\left[\frac{V}{N}(2s+1)\left(\frac{m}{2\pi\hbar^2\beta}\right)^{3/2}\right]\right\}, \tag{13}$$

c'est-à-dire la formule de Sackur-Tétrode III (IV.43).

Si le gaz étudié se trouvait dans la situation grand-canonique (μ fixé), la condition (1) exigerait que μ soit négatif et β suffisamment grand, c'est-à-dire T suffisamment faible. Mais c'est à *nombre de particules fixé* (situation canonique) que nous appliquons les résultats précédents; ils sont alors valables si (cf. formules (1) et (5))

$$T \gg \frac{2\pi\hbar^2}{km}\left[\frac{1}{2s+1}\frac{N}{V}\right]^{2/3} \tag{14}$$

parfait classique (cf. note 12). Pour mettre en évidence les *effets quantiques* engendrés par le facteur de Fermi, nous allons étudier les propriétés du système à *basse température.*

1. PROPRIÉTÉS DU SYSTÈME A TEMPÉRATURE NULLE

Commençons par choisir une température suffisamment basse pour que le facteur de Fermi coïncide pratiquement avec sa limite de température nulle (I.4) (« marche d'escalier » abrupte de la figure 1.b). Bien sûr, la température n'est jamais à proprement parler nulle. Mais il existe toujours un domaine de températures, dont la borne supérieure varie considérablement d'un cas à l'autre, où l'*approximation de température nulle* donne une idée correcte des principales propriétés du système; nous préciserons cette affirmation au cours des calculs qui vont suivre, et la discuterons de façon plus détaillée au paragraphe 3.

a. Le potentiel chimique à $T=0$

Si la température est suffisamment basse pour que la limite (I.4) du facteur de Fermi soit pratiquement atteinte, la formule (I.10) s'écrit

$$N=\int_0^{\mu_0} \mathrm{d}\varepsilon\,\rho(\varepsilon)=AV\int_0^{\mu_0}\varepsilon^{1/2}\,\mathrm{d}\varepsilon\,, \tag{I.28}$$

où μ_0 est le potentiel chimique à température nulle, que l'on appelle souvent *« niveau de Fermi »;* on trouve immédiatement, compte tenu de l'expression (0.2) de A,

$$\mu_0=\frac{\hbar^2}{2m}\left(\frac{6\pi^2}{2s+1}\frac{N}{V}\right)^{2/3}. \tag{I.29}$$

On écrit souvent le résultat sous la forme

$$\mu_0=\frac{\hbar^2 K_F^2}{2m}\,, \tag{I.30.a}$$

ce qui définit le *« vecteur d'onde de Fermi »* K_F et donne sa valeur :

$$K_F=\left(\frac{6\pi^2}{2s+1}\frac{N}{V}\right)^{1/3}. \tag{I.30.b}$$

Les vecteurs d'onde $\vec{K}$ de module inférieur à K_F correspondent tous à des états individuels[13] occupés à $T=0$, ceux dont le module est supérieur à K_F à des états individuels vides. Si l'on trace, dans l'espace des $\vec{K}$, la *« sphère de Fermi »*[14] centrée à l'origine et de rayon K_F (fig. 3), elle sépare les états occupés (vecteurs d'onde intérieurs à la sphère) des états non occupés (vecteurs $\vec{K}$ d'extrémité extérieure à la sphère) pour T tendant vers zéro.

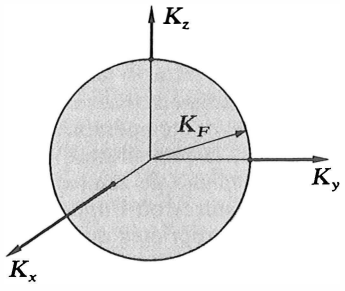
A partir de μ_0 (dont les dimensions sont celles d'une énergie), on peut aussi définir la *« température de Fermi »* T_F, ou *« température de dégénérescence »*, par la relation

$$kT_F \widehat{=} \mu_0\,. \tag{I.31}$$

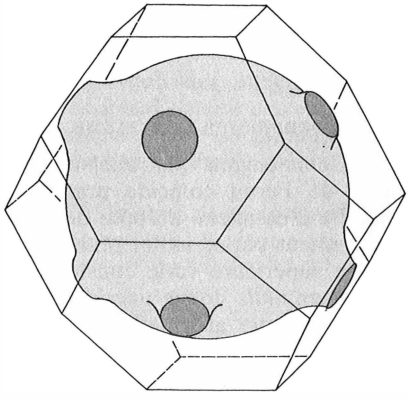
---

13. Rappelons qu'un état individuel est caractérisé par un vecteur d'onde $\vec{K}$ et la valeur m_s de la projection du spin suivant Oz; pour un spin $s=1/2$, par exemple, il y a donc deux états distincts pour chaque vecteur $\vec{K}$.

14. Dans des cas plus compliqués comme celui des électrons d'un solide cristallin (complément I.F, § II), on peut encore définir un *« espace réciproque »* (espace des $\vec{K}$) et, dans celui-ci, une *« surface de Fermi »* séparant comme ici les vecteurs d'onde associés à des états occupés à $T=0$ de ceux qui correspondent à des états vides. Cette surface n'est pas toujours simple (cf. fig. 3.b).



(a)



(b)

FIGURE 3
La surface de Fermi est une sphère pour un gaz de fermions libres (fig. a). Elle est plus compliquée pour les électrons de conduction d'un métal : celle de l'argent est représentée sur la figure b; pour d'autres exemples, voir N.W. Ashcroft et N.D. Mermin, *Solid State Physics*, Holt-Saunders, (1981), chapitre 15.

On conçoit — nous le vérifierons au paragraphe 2 — que l'*approximation de température nulle peut être utilisée si*

$$T \ll T_F \,. \tag{I.32}$$

b. L'énergie à $T=0$

Lorsque la température tend vers zéro, l'énergie totale du système tend vers sa valeur minimale

$$E_0 = \int_0^{\mu_0} \mathrm{d}\varepsilon\, \varepsilon \rho(\varepsilon) = AV \int_0^{\mu_0} \varepsilon^{3/2}\, \mathrm{d}\varepsilon = \frac{2}{5} AV \mu_0^{5/2} \,, \tag{I.33}$$

qui s'écrit aussi[15]

$$E_0 = \frac{3}{5} NkT_F \,. \tag{I.34}$$

Cette énergie caractérise le *niveau fondamental du système.* Sa valeur est une conséquence directe du *principe d'exclusion de Pauli :* comme les électrons dans un atome, les N fermions indépendants qui constituent le système sont placés dans les N états individuels d'énergie la plus basse possible (la relation (I.29) entre μ_0 et N revient alors à « compter » ces N états individuels, pour un volume V fixé).

c. Pression quantique

Le grand potentiel J du système est ici proportionnel à E (formule (I.20)); sa valeur à $T=0$, qui s'obtient à partir de (I.33), n'est donc pas nulle. Par conséquent,

15. Le résultat doit être proportionnel à $\mu_0 = kT_F$ pour préserver les dimensions, et l'on fait facilement apparaître N à partir de (I.28) :

$$N = AV \frac{2}{3} \mu_0^{3/2} \,. \tag{1}$$

D'où la formule (I.34).

la pression d'un gaz de fermions libres ne s'annule pas lorsque la température tend vers zéro : la pression résiduelle p_0, appelée « pression quantique », est donnée par

$$p_0 = \frac{2}{3}\frac{E_0}{V} = \frac{4}{15} A \mu_0^{5/2} = \frac{2}{15}\frac{\hbar^2}{m}\left(\frac{4\pi^2}{2s+1}\right)^{2/3}\left(\frac{3}{2}\frac{N}{V}\right)^{5/3}. \tag{I.35}$$

Elle est due elle aussi au principe de Pauli : même à $T=0$, les particules ont une impulsion non nulle (comprise entre zéro et l'impulsion de Fermi $\hbar K_F$), de sorte que leurs chocs contre les parois ne cessent pas.

2. DÉVELOPPEMENT DES GRANDEURS PHYSIQUES A BASSE TEMPÉRATURE

Pour améliorer l'approximation de température nulle et préciser ses conditions de validité, nous allons calculer les premiers *termes correctifs* à cette approximation. Ceux-ci vont résulter, pour chacune des grandeurs physiques, d'un *développement en puissances de la température* ; ce développement, dû à Sommerfeld, nous permettra également d'évaluer les grandeurs qui, comme la chaleur spécifique ou l'entropie, découlent d'une dérivation par rapport à la température.

a. Méthode de développement

Comme la limite du facteur de Fermi pour T tendant vers zéro est discontinue (formule (I.4)), la recherche des corrections de température non nulle doit être menée avec prudence. Le raisonnement s'appuie sur la constatation suivante (fig. 1) : à basse température, *seuls les états d'énergie ε voisine de μ voient leur nombre d'occupation modifié* de façon appréciable par rapport à la situation de température nulle (comme nous le préciserons à propos de l'énergie et de la chaleur spécifique (§ c.β), ceci est une conséquence directe du principe de Pauli).

Le système que nous considérons étant de taille macroscopique, les grandeurs physiques associées sont données par des *intégrales sur l'énergie individuelle ε*. Après transformation éventuelle par intégration par parties, elles sont toutes de la forme (§ A.2.c)[16]

$$g(T, \mu) \widehat{=} \int_0^{+\infty} \mathrm{d}\varepsilon\, N^F(\varepsilon; T, \mu) f(\varepsilon)\,, \tag{I.36}$$

où $f(\varepsilon)$ est une fonction régulière, continue et dérivable pour $\varepsilon \geqslant 0$ (la dépendance par rapport au volume V est sous-entendue). En général, ces intégrales ne sont pas calculables analytiquement sans approximation.

La valeur de g pour $T=0$ se déduit immédiatement de la limite (I.4) de N^F[17] :

$$g(0, \mu) = \int_0^{\mu} \mathrm{d}\varepsilon\, f(\varepsilon)\,. \tag{I.37}$$

Pour évaluer les termes correctifs, nous avons donc à calculer

$$\delta g(T, \mu) \widehat{=} g(T, \mu) - g(0, \mu)\,, \tag{I.38}$$

qui s'écrit, d'après (I.36) et (I.37),

$$\delta g(T, \mu) = \int_0^{\mu} \mathrm{d}\varepsilon \left[N^F(\varepsilon; T, \mu) - 1\right] f(\varepsilon) + \int_{\mu}^{+\infty} \mathrm{d}\varepsilon\, N^F(\varepsilon; T, \mu) f(\varepsilon)\,, \tag{I.39}$$

16. C'est le potentiel chimique μ qui figure dans l'expression des diverses grandeurs physiques. Ce n'est qu'après avoir calculé les intégrales correspondantes que l'on pourra éliminer μ et faire apparaître le nombre N de particules.

17. Si μ était négatif, $g(0, \mu)$ serait simplement nul. Dans tous les cas que nous aurons à considérer, μ sera positif au voisinage de $T=0$ (cf. formule (I.29)).

c'est-à-dire

$$\delta g(T, \mu) = \int_o^{+\infty} d\varepsilon \, \delta N^F(\varepsilon; T, \mu) f(\varepsilon) , \tag{I.40}$$

avec $\delta N^F(\varepsilon; T, \mu) \widehat{=} N^F(\varepsilon; T, \mu) - 1$ si $\varepsilon < \mu$

$$\widehat{=} N^F(\varepsilon; T, \mu) \quad \text{si} \quad \varepsilon > \mu . \tag{I.41}$$

Dans le domaine des basses températures qui nous intéresse ici :

$$kT \ll \mu , \tag{I.42}$$

la fonction $\delta N^F(\varepsilon)$ définie par (I.41) a l'allure représentée sur la figure 4. Elle n'est appréciablement différente de zéro que dans un domaine de largeur quelques kT centré en $\varepsilon = \mu$, et elle est symétrique par rapport au point $(\mu, 0)$.

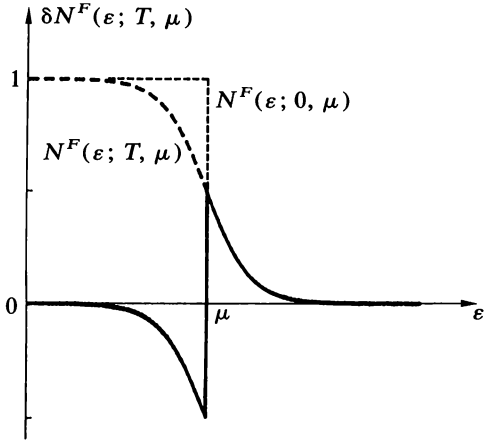


FIGURE 4

Courbe représentative de la fonction

$\delta N^F(\varepsilon; T, \mu) \widehat{=} N^F(\varepsilon; T, \mu) - N^F(\varepsilon; 0, \mu)$.

Dans la formule (I.40), la fonction à intégrer est donc nulle partout ailleurs que dans un voisinage de $\varepsilon = \mu$, d'extension quelques kT. Si la condition (I.42) est satisfaite, on peut repousser la borne inférieure de 0 à $-\infty$ sans changer pratiquement la valeur de l'intégrale. De plus la fonction $f(\varepsilon)$, elle, varie peu dans l'intervalle $\sim kT$ qui contribue seul à l'intégrale; on obtiendra donc une bonne approximation en développant $f(\varepsilon)$ autour de $\varepsilon = \mu$. Ces deux manipulations conduisent, nous allons le voir, à des intégrales calculables analytiquement.

Posons en effet

$$x \widehat{=} \beta(\varepsilon - \mu) . \tag{I.43}$$

En utilisant la symétrie de la fonction N^F, ainsi que son expression explicite (I.3), on trouve facilement [18]

18. Puisque la borne inférieure de l'intégrale peut être sans inconvénient prise à $-\infty$, on a d'après (I.3) et (I.43)

$$\delta g(T, \mu) = kT \int_{-\infty}^{0} dx \left[\frac{1}{e^x + 1} - 1 \right] f(xkT + \mu) + kT \int_0^{+\infty} dx \, \frac{1}{e^x + 1} f(xkT + \mu) . \tag{1}$$

.../...

$$\delta g(T,\mu)=kT\int_0^{+\infty}\mathrm{d}x\,\frac{1}{\mathrm{e}^x+1}\left[f(\mu+xkT)-f(\mu-xkT)\right]. \tag{I.44}$$

Il suffit alors de développer le crochet en puissances de xkT pour en déduire le développement de δg en puissances de kT, donc celui de $g(T,\mu)$. Le résultat s'écrit[19]

$$g(T,\mu)=g(0,\mu)+\frac{\pi^2}{6}(kT)^2 f'(\mu)+\mathcal{O}(T^4), \tag{I.45}$$

où $f'(\varepsilon)$ est la dérivée de la fonction $f(\varepsilon)$[20].

Mais (cf. note 2)

$$\frac{1}{\mathrm{e}^x+1}-1=-\frac{1}{\mathrm{e}^{-x}+1}. \tag{2}$$

Si, dans la première intégrale de la formule (1), on change la variable d'intégration en son opposé, on obtient

$$\delta g(T,\mu)=-kT\int_0^{+\infty}\mathrm{d}x\,\frac{1}{\mathrm{e}^x+1}f(\mu-xkT)+kT\int_0^{+\infty}\mathrm{d}x\,\frac{1}{\mathrm{e}^x+1}f(\mu+xkT), \tag{3}$$

c'est-à-dire l'expression (I.44).

19. La figure 4 montre que, pour T faible, les valeurs de $x=(\varepsilon-\mu)/kT$ qui contribuent à l'intégrale (I.44) donnant δg ne dépassent pas quelques unités. Dans ces conditions, xkT reste petit devant l'intervalle d'énergie sur lequel $f(\varepsilon)$ varie de façon appréciable, et l'on peut utiliser le développement

$$f(\mu\pm xkT)=f(\mu)\pm xkTf'(\mu)+\frac{1}{2!}(xkT)^2 f''(\mu)+\ldots \tag{1}$$

(où $f'(\varepsilon)$, $f''(\varepsilon)$, ... sont les dérivées successives de $f(\varepsilon)$), qui donne

$$f(\mu+xkT)-f(\mu-xkT)=2xkTf'(\mu)+\frac{2}{3!}(xkT)^3 f'''(\mu)+\ldots. \tag{2}$$

En reportant ce développement dans (I.44), on obtient

$$\delta g(T,\mu)=2(kT)^2 f'(\mu)\int_0^{+\infty}x\,\mathrm{d}x\,\frac{1}{\mathrm{e}^x+1}+\frac{1}{3}(kT)^4 f'''(\mu)\int_0^{+\infty}x^3\,\mathrm{d}x\,\frac{1}{\mathrm{e}^x+1}$$
$$+\frac{1}{60}(kT)^6 f^{(5)}(\mu)\int_0^{+\infty}x^5\,\mathrm{d}x\,\frac{1}{\mathrm{e}^x+1}+\ldots. \tag{3}$$

Les intégrales définies qui restent à calculer figurent dans les tables :

$$\int_0^{+\infty}\frac{x\,\mathrm{d}x}{\mathrm{e}^x+1}=\frac{\pi^2}{12},\quad \int_0^{+\infty}\frac{x^3\,\mathrm{d}x}{\mathrm{e}^x+1}=\frac{7}{120}\pi^4,\quad \int_0^{+\infty}\frac{x^5\,\mathrm{d}x}{\mathrm{e}^x+1}=\frac{31}{252}\pi^6,\ \ldots. \tag{4}$$

Ceci détermine les coefficients du développement (3) :

$$\delta g(T,\mu)=\frac{\pi^2}{6}f'(\mu)(kT)^2+\frac{7\pi^4}{360}f'''(\mu)(kT)^4+\frac{31\pi^6}{15\,120}f^{(5)}(\mu)(kT)^6+\ldots, \tag{5}$$

avec $$\frac{\pi^2}{6}\simeq 1{,}645,\quad \frac{7\pi^4}{360}\simeq 1{,}894,\quad \frac{31\pi^6}{15\,120}\simeq 1{,}971. \tag{6}$$

20. On peut exprimer le résultat en disant qu'on a développé le facteur de Fermi $N^F(\varepsilon;T,\mu)$ «au sens des distributions» :

$$N^F(\varepsilon;T,\mu)=\theta(\mu-\varepsilon)-\frac{\pi^2}{6}(kT)^2\delta'(\varepsilon-\mu)+\mathcal{O}(T^4), \tag{1}$$

.../...

b. Potentiel chimique

Appliquons d'abord la formule (I.45) à la condition (I.13) qui détermine le potentiel chimique μ à partir du nombre N de particules du système; il vient aussitôt

$$N = \frac{2}{3} AV\mu^{3/2} + \frac{\pi^2}{6} (kT)^2 \frac{AV}{2} \mu^{-1/2} + \mathcal{O}(T^4) . \tag{I.46}$$

Il est facile d'inverser cette relation[21] pour obtenir le *développement de μ en puissances de T :*

$$\mu = \mu_0 \left[1 - \frac{\pi^2}{12} \frac{k^2 T^2}{\mu_0^2} + \mathcal{O}(T^4) \right] , \tag{I.47}$$

que l'on peut écrire, en utilisant la définition (I.31) de la température de Fermi T_F, sous la forme

$$\mu = kT_F \left\{ 1 - \frac{\pi^2}{12} \left(\frac{T}{T_F} \right)^2 + \mathcal{O}\left[\left(\frac{T}{T_F} \right)^4 \right] \right\} \tag{I.48}$$

(le paramètre sans dimension du développement à basse température est T/T_F, voir paragraphe 3).

Compte tenu de (I.29), l'expression (I.47) donne μ comme fonction de T, N et V. Elle explicite donc, dans le cas particulier envisagé ici (fermions libres, basse température), la relation (I.11) qui est seulement implicite dans le cas général (§ A.2.b). On vérifie sans peine que μ est bien une fonction décroissante de T à N et V fixés, et une fonction croissante de N à T et V fixés (cf. note 9).

c. Énergie et chaleur spécifique

α. *Énergie du système à basse température*

Nous connaissons l'énergie du système à température nulle (formule (I.34)). A température basse mais non nulle, les termes correctifs sont donnés par l'application de la formule générale (I.45) à l'expression (I.17) :

$$E = \frac{2}{5} AV\mu^{5/2} + \frac{\pi^2}{6} (kT)^2 \times \frac{3}{2} AV\mu^{1/2} + \mathcal{O}(T^4) . \tag{I.49}$$

---

où $\theta(x)$ est la « fonction de Heaviside » :

$$\theta(x) = 0 \quad \text{si} \quad x < 0 ,$$
$$= 1 \quad \text{si} \quad x > 0 , \tag{2}$$

et $\delta'(x)$ la dérivée de la « fonction de Dirac » $\delta(x)$, définie par

$$\int dx\, \delta'(x) F(x) \widehat{=} - \int dx\, \delta(x) F'(x) = -F'(0) . \tag{3}$$

21. Posons en effet

$$\mu = \mu_0 [1 + aT^2 + \mathcal{O}(T^4)] \tag{1}$$

et déterminons le coefficient a en annulant les termes en T^2 de la relation (I.46) :

$$\frac{2}{3} AV\mu_0^{3/2} \times \frac{3}{2} aT^2 + \frac{\pi^2}{12} (kT)^2 AV\mu_0^{-1/2} = 0 \tag{2}$$

donne aussitôt

$$a = -\frac{\pi^2}{12} \frac{k^2}{\mu_0^2} . \tag{3}$$

Mais *attention :* dans cette expression figure le potentiel chimique μ correspondant à la température T, pas μ_0. En y reportant le développement (I.48), on obtient finalement[22]

$$E = E_0 \left\{ 1 + \frac{5\pi^2}{12} \left(\frac{T}{T_F} \right)^2 + \mathcal{O}\left[\left(\frac{T}{T_F} \right)^4 \right] \right\}. \tag{I.50}$$

β. Chaleur spécifique à volume constant

La *capacité calorifique à volume constant* C_V et la *chaleur spécifique* c_V correspondantes ont été définies au chapitre III (§ II.B.1.b). En dérivant par rapport à T, à N et V constants, l'expression (I.50) de E, on obtient aussitôt

$$\boxed{c_V = \eta T + \mathcal{O}(T^3)} \ , \tag{I.51}$$

avec (cf. formules (I.34), (I.31) et (I.29))

$$\eta = \frac{\pi^2 k^2 m}{\hbar^2} N_A \left(\frac{2s+1}{6\pi^2} \frac{V}{N} \right)^{2/3} = \frac{\pi^2}{2} \frac{N_A k}{T_F} \ . \tag{I.52}$$

La chaleur spécifique c_V d'un gaz de fermions libres est donc, à basse température, *proportionnelle à T* (N et V étant fixés). Ce résultat est radicalement différent de celui que nous avons trouvé pour un gaz parfait classique monoatomique (III, formule (IV.38)) :

$$c_V^{\text{class}} = \frac{3}{2} R = \frac{3}{2} N_A k \ . \tag{I.53}$$

Comme l'approximation classique de Maxwell-Boltzmann est valable à haute température, pour N et V fixés (V, § IV.C et III, § IV.A.3.c), les variations de la chaleur spécifique c_V d'un gaz de fermions libres en fonction de la température ont l'allure représentée sur la figure 5.

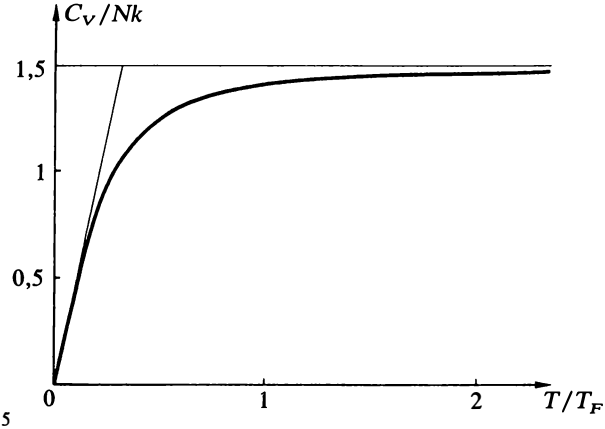


FIGURE 5

Variation avec la température de la chaleur spécifique à volume constant c_V d'un gaz de fermions libres sans structure interne. A haute température, l'approximation classique de Maxwell-Boltzmann est valable et c_V tend vers $3R/2$. A basse température, les effets quantiques (principe de Pauli) dominent et c_V part de zéro proportionnellement à T.

22. Il vient en effet

$$E = \frac{2}{5} AV\mu_0^{5/2} \left[1 - \frac{5}{2} \frac{\pi^2}{12} \left(\frac{T}{T_F} \right)^2 \right] + \frac{3\pi^2}{12} AV(kT)^2 \mu_0^{1/2} + \mathcal{O}(T^4) \ .$$

C'est essentiellement le *principe d'exclusion de Pauli* qui régit le comportement à basse température du gaz de fermions libres que nous étudions ici. On peut le comprendre semi-quantitativement de la façon suivante. A température nulle, tous les états individuels d'énergie ε inférieure à μ_0 sont occupés, tous ceux d'énergie ε supérieure à μ_0 sont vides. A une température T non nulle, le thermostat « offre » à chaque particule une énergie d'excitation de l'ordre de kT. Si T est suffisamment basse pour que la répartition de température nulle puisse être prise comme situation de départ, les particules se trouvant initialement dans un état dont l'énergie ε est telle que

$$0 \leqslant \varepsilon \leqslant \mu_0 - kT \tag{I.54}$$

ne pourront pas être excitées thermiquement (c'est-à-dire « profiter de l'offre » du thermostat) car ceci les amènerait à un état déjà occupé, qui leur est donc interdit par le principe de Pauli. Par conséquent, *seuls les fermions d'énergie voisine de* μ_0 *participent à l'augmentation d'énergie du système* lorsque la température devient non nulle.

On peut retrouver sur cette base l'*ordre de grandeur de la chaleur spécifique* à basse température. Le nombre N_a de particules « actives », c'est-à-dire susceptibles d'être excitées thermiquement, vaut approximativement[23]

$$N_a \sim \rho(\mu_0) \times kT \sim N \frac{kT}{\mu_0} . \tag{I.55}$$

L'énergie du système à basse température est donc sensiblement égale à[24]

$$E(T) \sim E_0 + N_a kT \sim E_0 + N \frac{k^2 T^2}{\mu_0} . \tag{I.56}$$

On en déduit par dérivation une chaleur spécifique de l'ordre de

$$c_V \sim 2 \frac{N_A k^2}{\mu_0} T . \tag{I.57}$$

On retrouve le fait que c_V est proportionnelle à T à basse température, avec un coefficient de proportionnalité du bon ordre de grandeur (voir formule (I.52)).

Il suffit de regrouper les termes pour aboutir à (I.50).

Noter que E diffère ainsi de E_0 pour deux raisons : correction explicite de température (donnée par (I.49)) et correction de potentiel chimique (donnée par (I.48)).

23. Le nombre d'occupation N^F de chaque état individuel étant de l'ordre de 1 (1/2 pour $\varepsilon = \mu$), on compte les états dont l'énergie est contenue dans une bande de largeur quelques kT autour de $\varepsilon = \mu$; à basse température, kT est petit et $\mu \simeq \mu_0$, de sorte que

$$N_a \simeq \rho(\mu_0) \times kT = AV\mu_0^{1/2} kT . \tag{1}$$

Mais on sait d'après (I.28) que

$$N = \frac{2}{3} AV\mu_0^{3/2} . \tag{2}$$

On a donc (en omettant toujours les facteurs de l'ordre de 1)

$$N_a \simeq N \frac{kT}{\mu_0} . \tag{3}$$

24. Le théorème d'équipartition de l'énergie n'est bien sûr pas valable dans le domaine de température considéré ici (III. § V.C); mais chaque particule « active » acquiert une énergie de l'ordre de kT.

Ce raisonnement semi-quantitatif explicite le mécanisme physique sous-jacent à la formule (I.51). Il montre clairement que c_V est proportionnelle à T parce que, non seulement l'énergie thermique de chaque particule, mais aussi le nombre de particules «actives», sont proportionnels à kT.

d. Grand potentiel, entropie, énergie libre

Le *grand potentiel* J du gaz de fermions libres s'obtient à partir de la relation (I.20), valable à toute température, et de l'expression (I.49) de l'énergie E, puisque nous voulons J en fonction des *variables grand-canoniques* T, μ et V :

$$J = -\frac{4}{15} AV\mu^{5/2} - \frac{\pi^2}{6} (kT)^2 AV\mu^{1/2} + \mathcal{O}(T^4) . \tag{I.58}$$

L'*entropie* S s'en déduit aussitôt par la formule[25]

$$S = -\left(\frac{\partial J}{\partial T}\right)_{\mu, V} , \tag{I.59}$$

qui donne

$$S = \frac{\pi^2}{3} k^2 AV\mu^{1/2} T + \mathcal{O}(T^3) . \tag{I.60}$$

Pour revenir aux variables canoniques T, N, V, il suffit de remplacer μ par μ_0[26], qui est fonction de N/V par la relation (I.29) :

$$S = \pi^2 \frac{mk^2}{\hbar^2} N \left(\frac{2s+1}{6\pi^2} \frac{V}{N}\right)^{2/3} T + \mathcal{O}(T^3) . \tag{I.61}$$

On note sur cette formule que, contrairement à E et J, *l'entropie S tend vers zéro avec la température.* Ce comportement est conforme au *«principe de Nernst»*, ou *«troisième principe de la thermodynamique»*, selon lequel l'entropie de tout système doit s'annuler quand la température tend vers zéro[27].

Connaissant l'énergie et l'entropie, on calcule aussitôt l'*énergie libre* thermodynamique :

$$F = E - TS . \tag{I.62}$$

On obtient ainsi[28]

$$F = E_0 \left\{ 1 - \frac{5\pi^2}{12} \left(\frac{T}{T_F}\right)^2 + \mathcal{O}\left[\left(\frac{T}{T_F}\right)^4\right] \right\} , \tag{I.63}$$

où E_0 et T_F sont données par (I.34) et (I.31).

---

25. On peut également retrouver la formule (I.46) à partir de

$$N = -\left(\frac{\partial J}{\partial \mu}\right)_{T, V} .$$

26. Le deuxième terme du développement (I.47) de μ doit être ici négligé, car il donne une contribution en T^3.

27. La formule de Sackur-Tétrode III (IV.43), qui donne l'entropie d'un gaz parfait classique, ne satisfait pas au principe de Nernst. Rien d'étonnant à cela : l'approximation de Maxwell-Boltzmann est valable à haute température, et l'on ne peut pas extrapoler ses résultats vers $T = 0$ sans tenir compte des effets quantiques.

28. Le terme en T^2 provient à la fois de E et de TS :

$$\frac{5\pi^2}{12} E_0 \left(\frac{T}{T_F}\right)^2 - \frac{\pi^2}{3} AVk^2 T^2 \mu_0^{1/2} . \tag{1}$$

.../...

e. Pression et équation d'état

Comme toujours, la *pression* p se déduit aussitôt de l'expression (I.58) du grand potentiel J :

$$p = -\left(\frac{\partial J}{\partial V}\right)_{T,\mu} = -\frac{J}{V}, \tag{I.64}$$

ce qui donne ici

$$p = A\mu^{5/2}\left\{\frac{4}{15} + \frac{\pi^2}{6}\left(\frac{kT}{\mu}\right)^2 + \mathcal{O}\left[\left(\frac{kT}{\mu}\right)^4\right]\right\}. \tag{I.65}$$

Il suffit ensuite de remplacer μ en fonction de T, N, V (formules (I.47) et (I.29)) pour obtenir l'*équation d'état* du gaz[29] :

$$p^3\left(\frac{V}{N}\right)^5 = \frac{36\pi^4}{125(2s+1)^2}\frac{\hbar^6}{m^3}\left[1 + \frac{5\pi^2 m^2 k^2 T^2}{\hbar^4}\left(\frac{2s+1}{6\pi^2}\frac{V}{N}\right)^{4/3} + \mathcal{O}(T^4)\right]. \tag{I.66}$$

Cette équation d'état est effectivement très différente de celle des gaz parfaits classiques. On remarque en particulier que *la pression ne tend pas vers zéro avec la température :* la formule (I.66) donne en effet pour $T = 0$ la valeur

$$p_0 = \left[\frac{36\pi^4}{125(2s+1)^2}\right]^{1/3}\frac{\hbar^2}{m}\left(\frac{N}{V}\right)^{5/3} \tag{I.67}$$

que nous avons trouvée au paragraphe 1.c (formule (I.35)).

3. DOMAINE DE VALIDITÉ DE L'APPROXIMATION

Les formules que nous avons écrites aux paragraphes 1 et 2 précédents donnent une bonne approximation des grandeurs physiques correspondantes si la température est suffisamment basse; mais cette affirmation demande à être précisée.

Or $$E_0 = \frac{2}{5}AV\mu_0^{5/2} \tag{2}$$

(formule (I.33)), ce qui permet de remplacer AV en fonction de E_0 :

$$\frac{5\pi^2}{12}E_0\left(\frac{T}{T_F}\right)^2 - \frac{\pi^2}{3}\frac{5}{2}E_0\frac{k^2T^2}{\mu_0^2} = -\frac{5\pi^2}{12}E_0\left(\frac{T}{T_F}\right)^2. \tag{3}$$

29. La cohérence du développement en T demande que μ soit simplement remplacé par $\mu_0 = kT_F$ dans le deuxième terme de (I.65). Dans le premier, on reporte la formule (I.47) :

$$p = A\mu_0^{5/2}\left[1 - \frac{5\pi^2}{24}\left(\frac{T}{T_F}\right)^2 + \mathcal{O}(T^4)\right]\left[\frac{4}{15} + \frac{\pi^2}{6}\left(\frac{T}{T_F}\right)^2 + \mathcal{O}(T^4)\right]$$

$$= \frac{4A}{15}\mu_0^{5/2}\left[1 + \frac{5\pi^2}{12}\left(\frac{T}{T_F}\right)^2 + \mathcal{O}(T^4)\right]. \tag{1}$$

On explicite ensuite $\mu_0 = kT_F$ selon la formule (I.29) et A selon (0.2) :

$$p = \left(\frac{2m}{\hbar^2}\right)^{3/2}\frac{2s+1}{15\pi^2}\left(\frac{\hbar^2}{2m}\right)^{5/2}\left(\frac{6\pi^2}{2s+1}\frac{N}{V}\right)^{5/3}$$

$$\times\left\{1 + \frac{5\pi^2}{12}(kT)^2\left(\frac{2m}{\hbar^2}\right)^2\left(\frac{2s+1}{6\pi^2}\frac{V}{N}\right)^{4/3} + \mathcal{O}(T^4)\right\}. \tag{2}$$

On élève enfin les deux membres au cube, pour simplifier la relation :

$$p^3 = \left(\frac{2s+1}{15\pi^2}\right)^3\left(\frac{\hbar^2}{2m}\right)^3\left(\frac{6\pi^2}{2s+1}\frac{N}{V}\right)^5\left[1 + \frac{5\pi^2}{4}(kT)^2\left(\frac{2m}{\hbar^2}\right)^2\left(\frac{2s+1}{6\pi^2}\frac{V}{N}\right)^{4/3} + \mathcal{O}(T^4)\right]. \tag{3}$$

Le résultat (I.66) s'en déduit aussitôt.

Nous avons constaté dès le paragraphe 2.b que, *pour N et V fixés*[30], *le paramètre sans dimension du développement est* T/T_F, où la *température de Fermi* T_F est donnée par (I.31) et (I.29) :

$$\boxed{T_F = \frac{\hbar^2}{2mk}\left(\frac{6\pi^2}{2s+1}\frac{N}{V}\right)^{2/3}} \quad . \tag{I.68}$$

Donc l'approximation de basse température est valable pour

$$T \ll T_F \, . \tag{I.69}$$

Notons que, dans toutes les formules du paragraphe 2, les termes négligés comportent un facteur T^2 de plus que ceux qui sont pris en compte; on s'attend donc à ce que l'approximation soit acceptable dès que T est sensiblement inférieure à T_F.

Or ce même système de fermions libres se comporte à *haute température* comme un *gaz parfait classique.* De façon plus précise (formule III (IV.34)), l'approximation de Maxwell-Boltzmann est valable pour

$$T \gg \frac{2\pi\hbar^2}{km}\left[\frac{N}{(2s+1)V}\right]^{2/3} \, . \tag{I.70}$$

On constate que, à un facteur près voisin de 1, l'expression figurant dans (I.70) est *égale à la température de Fermi*[31].

Pour un gaz de fermions libres identiques dont on fixe le nombre N, ainsi que le volume V et la température T, la situation d'ensemble est donc la suivante. Si T est nettement inférieure à la température de Fermi T_F, le système constitue ce que l'on appelle un *«gaz dégénéré»,* dans lequel les *effets quantiques* engendrés par la statistique de Fermi-Dirac se manifestent clairement; dans ce cas, les développements de basse température que nous avons écrits au paragraphe 2 ci-dessus décrivent correctement les propriétés du système. A des températures très supérieures à cette même température de Fermi T_F, le système devient un *gaz parfait classique,* dans lequel les effets du facteur de Fermi ne sont plus sensibles, et qui est décrit par les formules du chapitre III (§ IV.B.3). Bien entendu, le domaine des températures de l'ordre de T_F est plus compliqué que les deux précédents, aucune des deux approximations n'y étant valable[32].

---

30. Si l'on fixe au contraire μ et V, il apparaît sur la formule (I.58) que le paramètre sans dimension est kT/μ.

31. Ceci n'est bien sûr pas dû au hasard : à partir de N, V, des constantes fondamentales $\hbar$ et k et de la masse m des particules, on ne peut construire qu'une expression T_0 qui ait les dimensions d'une température et qui soit indépendante de la taille du système (c'est-à-dire fonction de N et V seulement à travers leur rapport N/V). En effet

$$kT_0 = \left(\frac{N}{V}\right)^a \hbar^b m^c \tag{1}$$

doit avoir les dimensions d'une énergie, ce qui impose

$$a = \frac{2}{3} \, , \quad b = 2 \, , \quad \text{et} \quad c = -1 \, . \tag{2}$$

Le reste est affaire de facteurs numériques, dont l'expérience montre qu'ils sont pratiquement toujours du même ordre de grandeur (ici, pour $s = 1/2$, on a $(3\pi^2)^{2/3}/2 \simeq 4{,}78$ dans (I.68) et $2\pi/2^{2/3} \simeq 3{,}96$ dans (I.70)).

32. On peut envisager d'améliorer ces approximations : pour le développement de basse température, on calculera les termes d'ordre supérieur en T; pour l'approximation de Maxwell-Boltzmann, on tiendra compte d'un ou plusieurs termes correctifs (V, note 27). Mais les améliorations ainsi obtenues n'élargissent que faiblement les deux domaines de validité.

Soulignons pour terminer que la température de Fermi T_F croît avec la densité N/V du gaz[33] et qu'elle est inversement proportionnelle à la masse des particules qui le constituent. Son ordre de grandeur peut varier considérablement d'un système à l'autre. Pour les électrons de conduction d'un métal, par exemple, la température de Fermi est de plusieurs dizaines de milliers de kelvins (complément VI.A, § II.1.b); c'est dire que le gaz formé par ces électrons est toujours fortement dégénéré, et que l'approximation de basse température y est toujours valable (le métal n'existe en tant que tel qu'à des températures très inférieures à la température de Fermi). Dans les étoiles que l'on nomme «naines blanches» existe également un gaz d'électrons totalement dégénéré, bien que la température y atteigne 10^7 K (complément VI.B). Pour les systèmes constitués d'atomes ou de molécules, la masse m est plusieurs milliers de fois supérieure à celle de l'électron et la température de Fermi est de l'ordre du kelvin (2,7 K pour l'isotope 3 de l'hélium); sauf cas très particulier (l'hélium 3 en est un; voir complément VI.D), les matériaux se solidifient à des températures bien supérieures, de sorte que l'étude précédente ne leur est pas applicable. Mais il existe des corps célestes («étoiles à neutrons») où les densités et les pressions sont si élevées qu'elles créent un système dégénéré de neutrons (complément VI.B, référence de la note 21).

II. GAZ DE BOSONS INDÉPENDANTS

Nous étudions maintenant les *systèmes de bosons identiques et indépendants.* Après avoir décrit leurs propriétés dans le cas général (§ A), nous examinerons l'exemple simple des bosons libres. Aux basses températures où les effets quantiques sont importants (§ B), nous constaterons que les systèmes de bosons se comportent de manière radicalement différente des systèmes de fermions.

A. Propriétés générales des gaz de bosons indépendants

1. LE FACTEUR DE BOSE : VARIATIONS ET COURBE REPÉSENTATIVE

Les propriétés des systèmes de bosons indépendants sont régies par la distribution de Bose-Einstein V (IV.22), c'est-à-dire par la fonction

$$N^B(\varepsilon;\, T,\, \mu) = \frac{1}{e^{\beta(\varepsilon-\mu)}-1}\,, \tag{II.1}$$

dépendant de la variable ε (considérée ici comme continue) et des deux paramètres T (ou $\beta = 1/kT$) et μ. Le nombre d'occupation N^B devant être positif pour tout $\varepsilon \geqslant \varepsilon_0$ (où ε_0 est le niveau fondamental individuel), le potentiel chimique μ est ici soumis à la condition

$$\mu < \varepsilon_0\,. \tag{II.2}$$

A température T et potentiel chimique μ fixés, $N^B(\varepsilon)$ est une *fonction constamment décroissante* de l'énergie individuelle ε : de deux états individuels (λ)

33. A chaque température T on peut faire correspondre une *longueur d'onde thermique de de Broglie* (formule III (IV.31))

$$\Lambda = \left(\frac{2\pi\hbar^2}{mkT}\right)^{1/2}.$$

Si on les exprime à l'aide de Λ, on constate que les inégalités (I.69) et (I.70) comparent cette longueur d'onde à $(V/N)^{1/3}$, qui représente la *distance moyenne entre les particules.* Ainsi, le système est classique si la longueur d'onde de de Broglie associée à une particule est très inférieure à la distance séparant en moyenne deux particules; les effets quantiques se font sentir lorsque ces deux longueurs deviennent du même ordre de grandeur, et *a fortiori* si Λ est supérieure à $(V/N)^{1/3}$.

ayant des énergies ε_λ différentes, le plus peuplé est ici aussi celui dont l'énergie est la plus basse. Les variations de N^B comme fonction de ε, et l'influence des paramètres T et μ, sont représentées sur la figure 6 : le rôle du paramètre μ apparaît clairement sur la figure 6.a, celui de T sur la figure 6.b. Lorsque μ est fixé, une diminution

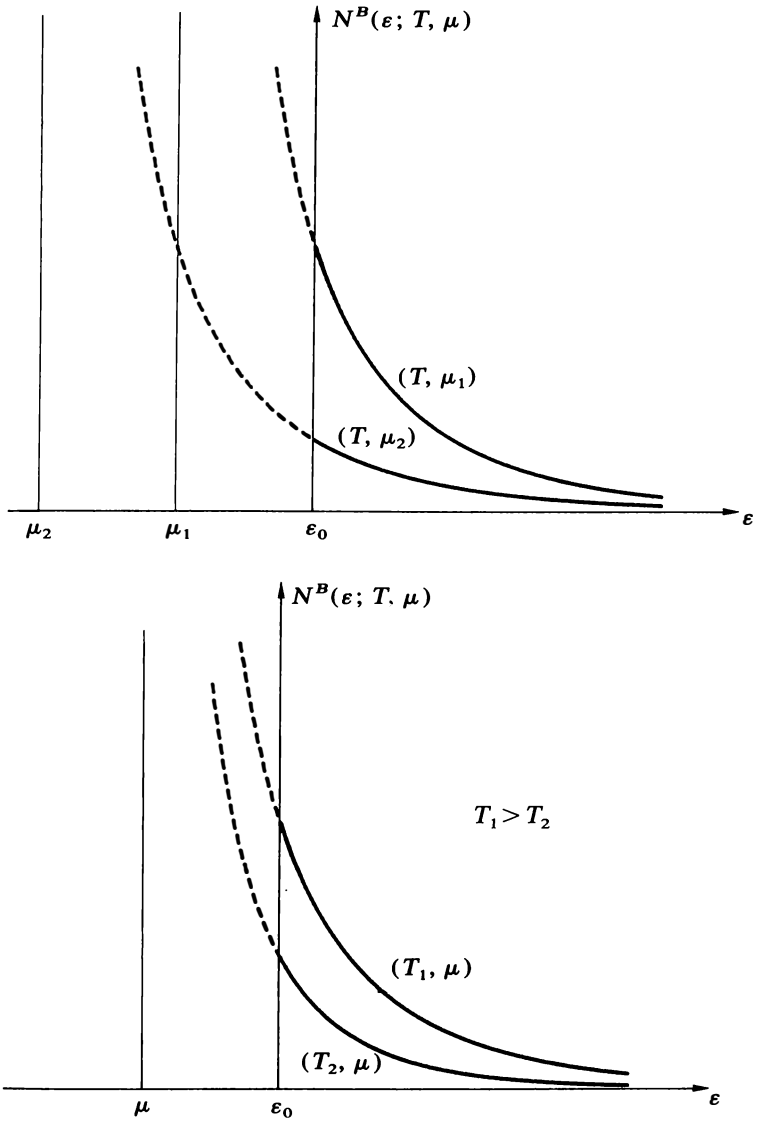


FIGURE 6

Variations de la fonction $N^B(\varepsilon; T, \mu)$ lorsque ε varie à T et μ fixés (ε_0 est le niveau fondamental individuel). La figure a montre l'influence du paramètre μ : changer le potentiel chimique μ, à température T fixée, revient à faire glisser la courbe sans déformation le long de l'axe des énergies. D'autre part (fig. b), la valeur de $N^B(\varepsilon; T, \mu)$ diminue, pour tout ε, si l'on diminue T à μ fixé : la courbe associée à une température T_2 est située entièrement au-dessous de celle qui correspond à une température T_1 supérieure à T_2.

progressive de la température rapproche la courbe de ses asymptotes, contre lesquelles elle se « plaque » de plus en plus à basse température. A la limite, lorsque T tend vers zéro à μ fixé, N^B tend vers zéro pour tout $\varepsilon \geqslant \varepsilon_0$: le système se vide de ses particules au profit du réservoir. Mais *attention :* nous raisonnons ici dans le cadre grand-canonique, c'est-à-dire à *potentiel chimique μ fixé;* le comportement du système à basse température est très différent dans le cas où c'est au contraire le nombre N de particules qui est fixé (voir paragraphe B ci-dessous).

2. EXPRESSION DES DIVERSES GRANDEURS PHYSIQUES

a. Formules grand-canoniques rigoureuses

Les grandeurs physiques associées à un système de bosons identiques indépendants s'expriment simplement à partir du facteur de Bose V (IV.22) :

$$\overline{N}_\lambda^B = \frac{1}{e^{\beta(\varepsilon_\lambda - \mu)} - 1} . \tag{II.3}$$

Il est facile en effet d'écrire toutes les formules du chapitre V (§ IV.A.2) en fonction des nombres d'occupation moyens $\overline{N}_\lambda^B$ des états individuels (λ).

Le *nombre moyen de particules* $\overline{N}$ et l'*énergie moyenne* $\overline{E}$ sont directement donnés par V (IV.7) et V (IV.9) :

$$\overline{N} = \sum_{(\lambda)} \overline{N}_\lambda^B , \tag{II.4}$$

$$\overline{E} = \sum_{(\lambda)} \overline{N}_\lambda^B \varepsilon_\lambda . \tag{II.5}$$

Pour ce qui est du *grand potentiel* J, de l'*entropie* S et de la *pression* p, on trouve facilement[34] les expressions suivantes :

$$J = -kT \sum_{(\lambda)} \ln\left(1 + \overline{N}_\lambda^B\right) , \tag{II.6}$$

---

34. Pour transformer l'expression V (IV.6) du grand potentiel J, il suffit de se rappeler (formule V (IV.23)) que

$$\xi_\lambda^B = 1 + \overline{N}_\lambda^B . \tag{1}$$

Quant à l'entropie S, la formule V (IV.10) donne ici, compte tenu encore de (1),

$$S = k \sum_{(\lambda)} \left[\ln\left(1 + \overline{N}_\lambda^B\right) + \beta \overline{N}_\lambda^B (\varepsilon_\lambda - \mu)\right] . \tag{2}$$

Mais comme

$$e^{\beta(\varepsilon_\lambda - \mu)} = \frac{1 + \overline{N}_\lambda^B}{\overline{N}_\lambda^B} , \tag{3}$$

le deuxième terme de (2) peut lui aussi s'écrire en fonction de $\overline{N}_\lambda^B$:

$$S = k \sum_{(\lambda)} \left[\ln\left(1 + \overline{N}_\lambda^B\right) + \overline{N}_\lambda^B \ln \frac{1 + \overline{N}_\lambda^B}{\overline{N}_\lambda^B}\right] . \tag{4}$$

L'expression (II.7) en découle aussitôt. Enfin, la pression p se déduit du grand potentiel J par

$$p = -\left(\frac{\partial J}{\partial V}\right)_{\mu, T} , \tag{5}$$

ce qui donne immédiatement (II.8).

$$S = k \sum_{(\lambda)} \left[(1 + \overline{N}_\lambda^B) \ln (1 + \overline{N}_\lambda^B) - \overline{N}_\lambda^B \ln \overline{N}_\lambda^B \right], \tag{II.7}$$

$$p = kT \frac{\partial}{\partial V} \sum_{(\lambda)} \ln (1 + \overline{N}_\lambda^B). \tag{II.8}$$

Les formules (II.6)-(II.8) ressemblent à leurs homologues pour les fermions; mais attention aux différences de signe.

Les formules que nous venons d'écrire donnent les propriétés grand-canoniques de tout système de bosons identiques *indépendants*. Elles ne font intervenir aucune autre hypothèse ou approximation, et serviront donc de référence dans la suite.

b. Cas des systèmes thermodynamiques. Interprétation canonique des formules

Lorsque la taille du système est suffisante pour que, d'une part, la séparation entre deux niveaux d'énergie individuels consécutifs soit inappréciable et, d'autre part, les fluctuations soient négligeables *(limite thermodynamique),* les formules du paragraphe précédent se simplifient, ainsi que leur interprétation.

L'argument est au départ analogue à celui du paragraphe I.A.2.b pour les fermions : la formule (II.4) devient ici

$$N = \int_{\varepsilon_0}^{+\infty} d\varepsilon \, \rho(\varepsilon) N^B(\varepsilon;\, T,\, \mu) = \int_{\varepsilon_0}^{+\infty} d\varepsilon \, \rho(\varepsilon) \frac{1}{e^{\beta(\varepsilon - \mu)} - 1}, \tag{II.9}$$

où ε_0 est le niveau fondamental individuel et $\rho(\varepsilon)$ la densité d'états individuels; comme pour les fermions, la relation (II.9) peut être considérée comme une *équation déterminant le potentiel chimique μ à partir du nombre de particules N* du système. A la limite thermodynamique, on peut donc utiliser les formules grand-canoniques du paragraphe a (dans lesquelles on remplace la somme sur les états individuels (λ) par une intégrale sur l'énergie individuelle ε) *en les interprétant dans le cadre canonique :* le potentiel chimique μ qui y figure n'est plus alors une variable indépendante, mais une fonction des variables canoniques T, N, V, ..., fonction donnée de façon implicite par la relation ci-dessus :

$$\mu = \mu(T,\, N,\, V,\, \ldots), \quad \text{solution de (II.9).} \tag{II.10}$$

Comme dans le cas des fermions et pour les mêmes raisons (note 9), μ est une *fonction croissante de N* (à T, V, ... fixés); il est facile de montrer de façon analogue que μ est une *fonction constamment décroissante de T* (à N, V, ... fixés).

Il faut cependant tenir compte d'une *différence fondamentale* avec les systèmes de fermions indépendants : *ici, les valeurs permises au potentiel chimique μ sont bornées supérieurement par ε_0*. La condition donnant μ s'écrit donc plus précisément

$$\begin{cases} N = \displaystyle\int_{\varepsilon_0}^{+\infty} d\varepsilon \, \rho(\varepsilon) \frac{1}{e^{\beta(\varepsilon - \mu)} - 1}, & \text{(II.11.a)} \\ \mu < \varepsilon_0 . & \text{(II.11.b)} \end{cases}$$

Dans certains cas (voir l'exemple du paragraphe c ci-dessous), les valeurs choisies pour T, N, V, ... sont telles que l'intégrale du second membre de (II.11.a) reste inférieure à N pour toutes les valeurs de μ vérifiant (II.11.b); l'équation n'admet pas alors de solution. Nous analyserons plus loin (§ B) le comportement du système dans une telle situation. Mais tant qu'il existe une valeur de μ inférieure à ε_0 qui satisfait à l'égalité (II.11.a), le comportement du système est « normal » : cette valeur (dont on montre facilement qu'elle est unique, voir paragraphe c) est celle du potentiel chimique du système. On la reporte ensuite dans l'expression des autres grandeurs physiques. Par exemple, l'énergie E du système est alors donnée par

$$\begin{cases} E = \displaystyle\int_{\varepsilon_0}^{+\infty} d\varepsilon \, \rho(\varepsilon) \varepsilon \frac{1}{e^{\beta(\varepsilon - \mu)} - 1}, \\ \mu \text{ solution de (II.11).} \end{cases} \tag{II.12}$$

c. Exemple : gaz de bosons libres

Prenons l'exemple d'un système thermodynamique de bosons identiques *libres et sans structure interne.* La densité d'états individuels est donnée dans ce cas par les formules (0.1)-(0.2), le spin s des particules étant maintenant entier : $s = 0, 1, \ldots$

La condition (II.11) qui détermine le potentiel chimique μ à partir des variables canoniques T, N, V s'écrit ici

$$\begin{cases} N = AV \displaystyle\int_0^{+\infty} \mathrm{d}\varepsilon \, \frac{\varepsilon^{1/2}}{\mathrm{e}^{\beta(\varepsilon-\mu)}-1} , & \text{(II.13.a)} \\ \mu < 0 . & \text{(II.13.b)} \end{cases}$$

Il faut résoudre cette condition par *ajustement numérique,* car l'intégrale qui y figure n'est pas calculable analytiquement pour l'ensemble des valeurs permises aux deux paramètres β et μ. Introduisons les variables sans dimension

$$\varphi \widehat{=} \mathrm{e}^{\beta\mu} , \qquad \text{(II.14)}$$

$$x \widehat{=} \beta\varepsilon . \qquad \text{(II.15)}$$

L'équation à résoudre (II.13.a) et la condition supplémentaire (II.13.b) prennent alors la forme

$$\begin{cases} \dfrac{4\pi^2}{2s+1} \left(\dfrac{\hbar^2\beta}{2m}\right)^{3/2} \dfrac{N}{V} = \displaystyle\int_0^{+\infty} \mathrm{d}x \, \frac{x^{1/2}}{-1+\mathrm{e}^x/\varphi} , & \text{(II.16.a)} \\ 0 \leqslant \varphi < 1 , & \text{(II.16.b)} \end{cases}$$

l'inconnue étant la fugacité φ. La figure 7 représente la fonction

$$I^B(\varphi) \widehat{=} \int_0^{+\infty} \mathrm{d}x \, \frac{x^{1/2}}{-1+\mathrm{e}^x/\varphi} \qquad \text{(II.17)}$$

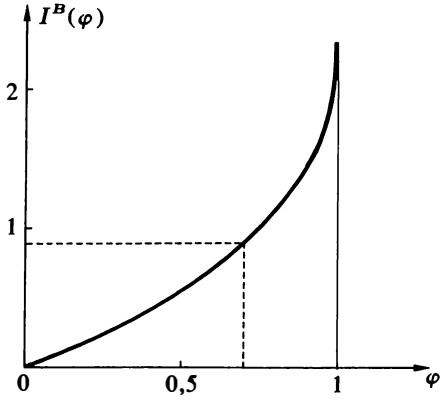


FIGURE 7

Solution graphique de l'équation (II.16). La courbe représente le second membre $I^B(\varphi)$ de (II.16.a) en fonction de la fugacité φ. Il faut couper cette courbe par une droite horizontale, d'ordonnée

$$\frac{4\pi^2}{2s+1} \left(\frac{\hbar^2\beta}{2m}\right)^{3/2} \frac{N}{V} .$$

On obtient une solution unique si cette valeur est inférieure à $I^B(1) \simeq 2{,}315$, pas de solution si elle lui est supérieure.

dans l'intervalle $0 \leqslant \varphi < 1$. Le point important est que *la valeur de cette fonction pour* $\varphi = 1$ *est finie*[35]. On trouve cette valeur dans les tables[36] :

$$I^B(1) \simeq \frac{\sqrt{\pi}}{2} \times 2{,}612\ldots \simeq 2{,}315\ldots \tag{II.18}$$

Comme, de façon évidente, $I^B(\varphi)$ est une fonction croissante de φ dans l'intervalle $[0, 1]$, la condition (II.16) ne peut être satisfaite que si

$$\frac{4\pi^2}{2s+1}\left(\frac{\hbar^2\beta}{2m}\right)^{3/2}\frac{N}{V} \leqslant 2{,}612\,\frac{\sqrt{\pi}}{2}\,. \tag{II.19}$$

Pour N *et* V *fixés,* cette restriction s'écrit

$$T \geqslant T_B\,, \tag{II.20}$$

où la *température de Bose* T_B est donnée par[37], [38]

$$\boxed{T_B \mathrel{\widehat{=}} \frac{1}{k}\,\frac{2\pi\hbar^2}{m}\left[\frac{N}{V}\,\frac{1}{2s+1}\,\frac{1}{2{,}612}\right]^{2/3}} \tag{II.21}$$

---

35. L'intégrale

$$\int_0^{+\infty} \mathrm{d}x\,\frac{x^{1/2}}{\mathrm{e}^x - 1} \tag{1}$$

est en effet convergente : seule la borne inférieure pourrait poser problème, et l'on a alors

$$\frac{x^{1/2}}{\mathrm{e}^x - 1} \underset{x \to 0}{\sim} \frac{x^{1/2}}{x} = \frac{1}{x^{1/2}}\,, \tag{2}$$

qui est bien intégrable.

36. L'intégrale $I^B(1)$ fait intervenir une *«fonction* ζ *de Riemann»,* dont la définition générale s'écrit

$$\int_0^{+\infty} \mathrm{d}x\,\frac{x^y}{\mathrm{e}^x - 1} \mathrel{\widehat{=}} \Gamma(y+1)\,\zeta(y+1)\,, \tag{1}$$

où Γ est la fonction bien connue :

$$\Gamma(y) \mathrel{\widehat{=}} \int_0^{+\infty} \mathrm{d}x\; x^{y-1}\,\mathrm{e}^{-x}\,. \tag{2}$$

On a ici

$$\int_0^{+\infty} \mathrm{d}x\,\frac{x^{1/2}}{\mathrm{e}^x - 1} = \Gamma\left(\frac{3}{2}\right)\zeta\left(\frac{3}{2}\right) \tag{3}$$

avec $\Gamma\left(\frac{3}{2}\right) = \frac{\sqrt{\pi}}{2}\,, \quad \zeta\left(\frac{3}{2}\right) = 2{,}612\ldots$ (4)

(voir H.B. Dwight, *Mathematical Tables of Elementary and Some Higher Mathematical Functions,* Dover, New York (1958)).

37. La température de Bose est donc celle qui correspond, dans la relation (II.13.a), à $\mu = 0$:

$$N = AV\int_0^{+\infty} \mathrm{d}\varepsilon\,\frac{\varepsilon^{1/2}}{\mathrm{e}^{\beta_B\varepsilon} - 1} = \frac{AV}{\beta_B^{3/2}}\,I^B(1) = 2{,}612\,\frac{\sqrt{\pi}}{2}\,\frac{AV}{\beta_B^{3/2}}\,,$$

avec $\beta_B \mathrel{\widehat{=}} \frac{1}{kT_B}\,.$

De façon plus générale, un raisonnement analogue à celui de la note 35 montre qu'il existe une température de Bose non nulle pour tout système de bosons identiques indépendants dont la densité d'états $\rho(\varepsilon)$ tend vers zéro pour ε tendant vers ε_0.

38. On retrouve bien sûr, à un facteur numérique près, la même expression qu'en (I.68) (température de Fermi). Comme nous l'avons déjà souligné dans la note 31, ceci était inévitable compte tenu des propriétés dimensionnelles et de l'intensivité de la grandeur température.

.../...

Nous examinerons au paragraphe B comment il convient de modifier les formules pour traiter le cas où T est inférieure à T_B. Mais si la condition (II.19) est satisfaite, il est évident d'après la figure 7 que l'équation (II.16) admet une solution φ et une seule, à partir de laquelle on détermine le potentiel chimique μ du système (relation (II.14)).

On peut ensuite calculer, toujours pour $T \geqslant T_B$, les autres grandeurs physiques. L'*énergie* E du système s'obtient par

$$\begin{cases} E = AV \displaystyle\int_0^{+\infty} \mathrm{d}\varepsilon \, \frac{\varepsilon^{3/2}}{\mathrm{e}^{\beta(\varepsilon-\mu)}-1}, \\ \mu \text{ solution de (II.13).} \end{cases} \tag{II.22}$$

Quant au *grand potentiel* J, il s'écrit

$$J = -kT.AV \int_0^{+\infty} \mathrm{d}\varepsilon \, \varepsilon^{1/2} \ln\left[1 + N^B(\varepsilon; T, \mu)\right]. \tag{II.23}$$

Une intégration par parties transforme cette expression en[39]

$$J = -\frac{2}{3} AV \int_0^{+\infty} \mathrm{d}\varepsilon \, \frac{\varepsilon^{3/2}}{\mathrm{e}^{\beta(\varepsilon-\mu)}-1}. \tag{II.24}$$

Comme la température de Fermi T_F pour les fermions (note 33), la température de Bose T_B correspond à une longueur d'onde thermique de de Broglie du même ordre de grandeur que la distance moyenne $(V/N)^{1/3}$ entre les particules :

si $\quad \Lambda_B \widehat{=} \left[\dfrac{2\pi\hbar^2}{mkT_B}\right]^{1/2}$,

la formule (II.21) peut s'écrire

$$\frac{1}{2s+1} \frac{N}{V} \Lambda_B^3 = 2{,}612 \, .$$

Mais il faut noter une *différence importante* avec le cas des fermions. T_B constitue ici une *limite bien tranchée entre deux domaines* de températures : pour $T \geqslant T_B$, le comportement du gaz est «normal», c'est-à-dire fondé sur la relation (II.13); pour $T < T_B$ se produit un phénomène nouveau (condensation de Bose) qui affecte profondément les propriétés du système (§ B ci-dessous). Dans le cas des fermions, T_F donnait un ordre de grandeur des températures pour lesquelles les effets quantiques étaient négligeables ($T \gg T_F$) ou au contraire importants ($T \ll T_F$); mais rien de marquant ne se produisait pour $T = T_F$ précisément.

39. On a effet

$$1 + N^B(\varepsilon; T, \mu) = 1 + \frac{1}{\mathrm{e}^{\beta(\varepsilon-\mu)}-1} = \frac{1}{1-\mathrm{e}^{-\beta(\varepsilon-\mu)}}, \tag{1}$$

de sorte que

$$\frac{\partial}{\partial\varepsilon} \ln\left[1 + N^B(\varepsilon; T; \mu)\right] = -\frac{\partial}{\partial\varepsilon} \ln\left[1 - \mathrm{e}^{-\beta(\varepsilon-\mu)}\right] = -\beta \frac{1}{\mathrm{e}^{\beta(\varepsilon-\mu)}-1}. \tag{2}$$

Effectuons alors une intégration par parties dans l'intégrale (II.23) :

$$J = -AVkT \left\{-\frac{2}{3} \varepsilon^{3/2} \ln\left[1 - \mathrm{e}^{-\beta(\varepsilon-\mu)}\right]\right\}_0^{+\infty} - AVkT \int_0^{+\infty} \mathrm{d}\varepsilon \, \frac{2}{3} \varepsilon^{3/2} \beta \frac{1}{\mathrm{e}^{\beta(\varepsilon-\mu)}-1}. \tag{3}$$

Le terme tout intégré est nul : il l'est visiblement à la borne inférieure $\varepsilon = 0$; à la borne supérieure, on trouve en développant le logarithme

$$\frac{2}{3} \varepsilon^{3/2} \mathrm{e}^{-\beta(\varepsilon-\mu)} \xrightarrow[\varepsilon \to +\infty]{} 0. \tag{4}$$

D'où le résultat (II.24).

On notera que la valeur du facteur de proportionnalité $(-2/3)$ entre J et E est directement reliée à celle de l'exposant $(1/2)$ de ε dans la densité d'états $\rho(\varepsilon)$.

On constate alors que *le grand potentiel est simplement proportionnel à l'énergie* (cf. formule (II.22)) :

$$J = -\frac{2}{3} E . \tag{II.25}$$

On en déduit l'*entropie* S par

$$S = -\left(\frac{\partial J}{\partial T}\right)_{\mu, V} . \tag{II.26}$$

En utilisant l'expression (II.24) de J en fonction de μ, on obtient

$$\begin{cases} S = k\beta^2 \frac{2}{3} AV \int_0^{+\infty} d\varepsilon\, \varepsilon^{3/2} \frac{(\varepsilon - \mu)\, e^{\beta(\varepsilon-\mu)}}{[e^{\beta(\varepsilon-\mu)} - 1]^2} , \\ \mu \text{ solution de (II.13).} \end{cases} \tag{II.27}$$

A partir de E et S, on calcule aussitôt l'*énergie libre*

$$F = E - TS . \tag{II.28}$$

Enfin, la *pression* p se déduit également du grand potentiel :

$$p = -\frac{J}{V} = \frac{2}{3}\frac{E}{V} \tag{II.29}$$

d'après (II.25). L'*équation d'état* du gaz de bosons libres est donc connue sous forme paramétrique à partir de (II.22) :

$$\begin{cases} p = \frac{2s+1}{6\pi^2} \left(\frac{2m}{\hbar^2}\right)^{3/2} \int_0^{+\infty} d\varepsilon \frac{\varepsilon^{3/2}}{e^{\beta(\varepsilon-\mu)} - 1} , \\ \mu \text{ solution de (II.13).} \end{cases} \tag{II.30}$$

Rappelons que les formules précédentes sont valables dans le cas où la condition (II.19) est satisfaite, c'est-à-dire, si N et V sont fixés, pour $T \geqslant T_B$. On montre sans peine que *les propriétés du gaz de bosons libres deviennent à haute température celles d'un gaz parfait classique*[40].

B. Comportement à basse température d'un système thermodynamique de bosons libres

Le comportement d'un gaz de bosons à basse température est *radicalement différent de celui d'un gaz de fermions*. Cette différence essentielle, due à un simple changement de signe dans la fonction de distribution, traduit en fait la distinction fondamentale entre les deux types de particules : un état individuel donné peut être occupé par une seule particule s'il s'agit de fermions, mais par un nombre quelconque dans le cas de bosons. A basse température où les particules tendent à s'accumuler dans les états de faible énergie, cette distinction devient cruciale.

1. LE PHÉNOMÈNE DE CONDENSATION DE BOSE

Pour illustrer les *effets quantiques* qui se manifestent à basse température, reprenons le gaz de bosons identiques libres considéré au paragraphe A.2.c. *Le nombre de particules N, le volume V et la température T sont fixés.* La taille du système (caractérisée par N ou V) est supposée suffisamment grande pour que la *limite*

40. Le raisonnement est le même que dans le cas des fermions (note 12).

thermodynamique soit atteinte. Il s'agit de compléter l'analyse en étudiant le cas où la température est inférieure à la température de Bose (II.21) :

$$T < T_B \,. \tag{II.31}$$

a. Étude critique de la relation entre le nombre de particules et le potentiel chimique

Nous avons vu que, si T est supérieure à la température de Bose T_B, le potentiel chimique μ du gaz est donné à partir de N, T et V par la relation (II.13). Lorsque la température décroît, le potentiel chimique croît et se rapproche, par valeurs inférieures, de sa valeur limite nulle; il atteint cette borne supérieure pour $T = T_B$. Que se passe-t-il lorsque T franchit cette température critique T_B?

Rappelons tout d'abord que *la relation* (II.13) *entre* μ *et* N *cesse d'être valable telle quelle pour* $T < T_B$: la condition (II.19) n'est plus satisfaite, c'est-à-dire que l'intégrale figurant au second membre de (II.13.a) reste inférieure à N pour toutes les valeurs permises de μ ($\mu < \varepsilon_0 = 0$).

Revenons alors à la relation exacte (II.4). La limite thermodynamique étant atteinte, il est licite de remplacer comme nous l'avons fait plus haut la valeur moyenne $\overline{N}$ par *la* valeur N, fixée, du nombre de particules. Dans l'autre membre de l'égalité, nous avons en toute rigueur à calculer une somme (discrète) sur les états individuels (λ). Considérant que l'écart entre deux niveaux individuels consécutifs est très faible lorsque le système est macroscopique, nous avons remplacé la somme discrète par une intégrale sur l'énergie ε. Ceci est parfaitement valable, et donne une excellente approximation, *sauf si l'expression à sommer prend, pour un terme particulier associé à un état* (λ_0) *déterminé, une valeur très différente* de celle des termes qui correspondent dans la somme aux états voisins (c'est-à-dire d'énergie voisine); si c'est le cas, l'interpolation qu'effectue l'intégrale par rapport à la véritable somme discrète pourra être grossièrement incorrecte.

Il nous faut donc rechercher si certains états individuels particuliers peuvent donner une contribution singulière à la somme (II.4). Comme nous examinons ici le domaine des basses températures, c'est évidemment vers les états individuels de plus faible énergie que nous devons nous tourner.

b. Répartition des particules entre les états individuels

α. *Population de l'état fondamental individuel*

Comme $\overline{N}_\lambda^B$ (formule (II.3)) est une fonction toujours décroissante de ε_λ, l'état individuel dont le nombre d'occupation moyen est le plus grand est toujours l'état de plus faible énergie (ici $\varepsilon_0 = 0$)[41]. Or la contribution des états d'énergie voisine de zéro à l'intégrale (II.13.a) est très petite, puisque la densité d'états $\rho(\varepsilon)$ tend vers zéro avec ε. Ceci est sans importance dans des conditions « normales », c'est-à-dire tant que le nombre d'occupation de l'état fondamental individuel est négligeable devant le nombre total N de particules, qui est d'ordre macroscopique (c'est-à-dire comparable au nombre d'Avogadro). Mais existe-t-il des situations dans lesquelles ce nombre d'occupation soit lui même d'ordre macroscopique?

Le nombre d'occupation moyen de l'état fondamental s'écrit

$$N^B(\varepsilon_0 = 0;\, T,\, \mu) = \frac{1}{e^{-\beta\mu} - 1} \,. \tag{II.32}$$

Est-il donc possible que l'on ait

$$\frac{1}{e^{-\beta\mu} - 1} = N_0 \,, \text{ avec } N_0 \text{ macroscopique?} \tag{II.33}$$

41. Cet état fondamental est unique dans le cas de bosons libres de spin s nul et sans structure interne. Mais le raisonnement serait analogue s'il existait plusieurs états distincts associés au niveau fondamental individuel $\varepsilon_0 = 0$ (par exemple si s est non nul).

Il faut pour cela que $\exp(-\beta\mu)$ soit extrêmement proche de 1 :

$$e^{-\beta\mu} = 1 + \frac{1}{N_0} . \tag{II.34}$$

Ceci implique que $-\beta\mu$ soit extrêmement voisin de zéro[42] :

$$-\beta\mu = \frac{1}{N_0} , \tag{II.35}$$

c'est-à-dire finalement

$$\mu = -\frac{kT}{N_0} , \text{ avec } N_0 \text{ macroscopique.} \tag{II.36}$$

Ainsi, *la population de l'état fondamental est effectivement d'ordre macroscopique si le potentiel chimique du système est suffisamment proche de sa valeur limite nulle.*

Il importe cependant de noter que l'ordre de grandeur du résultat (II.36) est *totalement inhabituel* : l'énergie kT, nous le savons (III, § I.B.2.b), est déjà d'ordre microscopique; lorsqu'on la divise par N_0 macroscopique, on obtient des *valeurs extraordinairement petites.* En toute autre occasion, on aurait à bon droit négligé un terme de cet ordre. Si nous le conservons provisoirement (bien qu'aucune mesure directe ne puisse évidemment permettre en pratique de le distinguer de $\mu = 0$), c'est qu'il est nécessaire à la compréhension du phénomène que nous étudions.

Ce phénomène est connu sous le nom *condensation de Bose :* lorsque le potentiel chimique du système vérifie (II.36), un nombre N_0 macroscopique de particules sont « condensées » dans l'état fondamental individuel.

β. Population des autres états individuels

Nous allons maintenant montrer que, dans la situation caractérisée par (II.36), *aucun autre état individuel* (pas même ceux qui correspondent au premier niveau excité) *n'a une population comparable à celle de l'état fondamental.*

Il est tout d'abord facile de montrer[43] que les états individuels dont l'énergie ε_λ est supérieure à kT ou du même ordre sont occupés par un nombre de particules très petit devant N_0. Ceci n'a bien sûr rien de surprenant.

Mais considérons maintenant un état (λ) dont l'énergie ε_λ est très inférieure à kT. Nous poserons

$$\varepsilon_\lambda = \eta_\lambda kT , \tag{II.37}$$

avec donc

$$\eta_\lambda \ll 1 . \tag{II.38}$$

De tels états existent, pour une température donnée (sauf pour des volumes V très petits, cf. I, note 18). On s'attendrait à ce que leur population soit certes inférieure à celle de l'état fondamental, mais du même ordre. Voyons ce qu'il en est en réalité.

---

42. Il faudrait en toute rigueur écrire

$$-\beta\mu = \ln\left(1 + \frac{1}{N_0}\right) = \frac{1}{N_0}\left(1 - \frac{1}{2N_0} + \ldots\right) .$$

Mais les termes correctifs sont totalement négligeables devant 1, pour N_0 de l'ordre de 10^{20} ou plus.

43. Si $\varepsilon_\lambda \geqslant kT$, on a aussi $(\mu < 0)$ $\varepsilon_\lambda - \mu \geqslant kT$, c'est-à-dire $\beta(\varepsilon_\lambda - \mu) \geqslant 1$.
Par conséquent, le nombre d'occupation moyen d'un tel état (λ) est inférieur à 1 :

$$\overline{N}^B_\lambda = \frac{1}{e^{\beta(\varepsilon_\lambda - \mu)} - 1} \leqslant \frac{1}{e - 1} < 1 ,$$

et donc incomparablement plus petit que N_0.

Comme $\beta(\varepsilon_\lambda - \mu)$ est petit devant 1, le nombre moyen d'occupation d'un tel état vaut

$$\overline{N}_\lambda^B = \frac{1}{e^{\beta(\varepsilon_\lambda - \mu)} - 1} \simeq \frac{1}{\beta(\varepsilon_\lambda - \mu)}, \tag{II.39}$$

c'est-à-dire, d'après (II.35) et (II.37),

$$\overline{N}_\lambda^B \simeq \frac{1}{\eta_\lambda + 1/N_0}. \tag{II.40}$$

Si donc η_λ est de l'ordre de 1/10 ou 1/100, le terme $1/N_0$ est totalement négligeable devant lui et $\overline{N}_\lambda^B$ est de l'ordre de 10 ou 100, c'est-à-dire environ 10^{20} fois plus petit que N_0. Pour que $\overline{N}_\lambda^B$ prenne une valeur macroscopique, comparable à la population N_0 de l'état fondamental, il faudrait que η_λ soit lui-même de l'ordre de $1/N_0$.

Or *ceci est impossible* pour un système thermodynamique dont le nombre de particules N et le volume V sont fixés à des valeurs macroscopiques. Le premier niveau excité individuel ε_1 est en effet tel que[44]

$$\frac{\varepsilon_1}{kT} \geqslant \frac{C}{N^{2/3}}, \tag{II.41}$$

où C est une constante de l'ordre de 1. Cette borne est certes très petite : pour $N \simeq 10^{23}$, elle est inférieure à 10^{-15}. Mais elle permet quand même d'exclure la possibilité que η_λ soit du même ordre que $1/N_0$:

$$\eta_\lambda = \frac{\varepsilon_\lambda}{kT} \geqslant \frac{\varepsilon_1}{kT} \geqslant \frac{1}{N^{2/3}} \gg \frac{1}{N_0} \tag{II.42}$$

(typiquement, il faut comparer $1/N^{2/3} \simeq 10^{-15}$ à $1/N_0 \simeq 10^{-23}$). *Donc le nombre d'occupation de tous les états individuels excités reste négligeable devant celui de l'état fondamental* (par des facteurs de l'ordre de 10^{-7} ou 10^{-8}).

Ainsi, *le phénomène de condensation de Bose concerne uniquement l'état fondamental individuel.*

c. Nombre de particules condensées

Les résultats du paragraphe b fournissent la solution à la question que nous nous posions au paragraphe a. De façon générale en effet, la relation exacte (II.4) peut être écrite de la manière suivante : la contribution à la somme sur (λ) de l'ensemble des états (λ) autres que l'état fondamental est « normale », c'est-à-dire qu'elle peut être correctement évaluée par une intégrale sur l'énergie individuelle ε, comme en (II.13.a); celle de l'état fondamental est dans certains cas « anormale », c'est-à-dire d'ordre macroscopique, et il faut alors la rajouter explicitement, « à la main » en quelque sorte.

---

44. Le premier niveau excité ε_1 d'une particule libre enfermée dans une boîte cubique de volume $V = L^3$ a pour expression (I, formules (III.21)-(III-22))

$$\varepsilon_1 = \frac{2\pi^2\hbar^2}{m} \frac{1}{V^{2/3}}. \tag{1}$$

On peut donc écrire, pour $T \leqslant T_B$,

$$\frac{\varepsilon_1}{kT} \geqslant \frac{\varepsilon_1}{kT_B} = \frac{C}{N^{2/3}}, \tag{2}$$

d'après (II.21); la constante C vaut (pour $s = 0$)

$$C = \pi \times (2{,}612)^{2/3} \simeq 5{,}96\ldots. \tag{3}$$

Précisons. Pour $T > T_B$, il existe toujours une valeur de μ négative vérifiant (II.13); la population de l'état fondamental est alors « normale », c'est-à-dire négligeable devant N, et il n'est pas nécessaire de l'écrire explicitement[45]. Pour $T < T_B$ au contraire, l'intégrale sur l'énergie individuelle ε reste nettement inférieure à N pour toutes les valeurs permises de μ; il faut donc que se produise le phénomène de condensation de Bose, car la relation (II.4) ne peut être vérifiée que si la contribution de l'état fondamental est d'ordre macroscopique. La condition fixant le potentiel chimique μ à partir du nombre N de particules (la température T et le volume V étant fixés) s'écrit donc, pour $T < T_B$,

$$N = N_0 + AV \int_0^{+\infty} d\varepsilon \, \frac{\varepsilon^{1/2}}{e^{\beta(\varepsilon-\mu)} - 1} , \tag{II.43.a}$$

N_0 et μ étant reliés par (II.36) :

$$\mu = -\frac{kT}{N_0} . \tag{II.43.b}$$

Mais cette valeur de μ est tellement proche de zéro, puisque N_0 est macroscopique, que l'intégrale ne la distingue pas de $\mu = 0$[46]. Les propriétés thermodynamiques du système sont en fait correctement données par

$$\begin{cases} \mu = 0 , & \text{(II.44.a)} \\ N = N_0 + AV \displaystyle\int_0^{+\infty} d\varepsilon \, \frac{\varepsilon^{1/2}}{e^{\beta\varepsilon} - 1} , \quad \text{pour } T \leqslant T_B . & \text{(II.44.b)} \end{cases}$$

Les égalités (II.44) remplacent donc la relation (II.13) dans le domaine où celle-ci ne pouvait plus être valable, c'est-à-dire pour toute température T inférieure à la température de Bose T_B.

Ainsi, *le potentiel chimique thermodynamique reste nul* dans tout le domaine des températures inférieures à T_B. C'est *le nombre N_0 de particules « condensées »* dans l'état fondamental qui *s'ajuste,* pour chaque température, *de manière à satisfaire* (II.44.b) : quand la température baisse, l'intégrale diminue; il faut donc, pour $T < T_B$, que N_0 augmente pour que la somme des deux termes reste constante.

Il doit être clair d'après ce qui précède que *la condensation de Bose se produit parce que le nombre total de particules est fixé.* Si l'on avait fixé le potentiel chimique μ (à une valeur négative), et non pas le nombre de particules, c'est ce dernier, ou plutôt sa valeur moyenne $\overline{N}$, qui s'ajusterait pour satisfaire la relation (II.13.a) à chaque

---

45. C'est ce que confirme indirectement le raisonnement du paragraphe b.α : lorsque μ est nettement différent de zéro, le nombre d'occupation moyen de l'état fondamental individuel est effectivement très petit devant N, de sorte que la relation (II.13.a) est valable sans modification. Ceci reste vrai si T s'approche de T_B par valeurs supérieures : μ tend alors vers zéro par valeurs négatives, mais si $\mu = -kT/10$ ou $-kT/100$, le nombre d'occupation de l'état fondamental n'est jamais que 10 ou 100, ce qui reste négligeable devant 10^{20} ou 10^{23}. Nous reviendrons plus loin (note 48) sur ce qui se passe lorsque T est égale à T_B ou extrêmement proche.

46. Comme $-\beta\mu = 1/N_0$, l'exponentielle figurant dans l'intégrale peut s'écrire

$$e^{\beta(\varepsilon-\mu)} \simeq e^{\beta\varepsilon}(1 - \beta\mu) = e^{\beta\varepsilon}\left(1 + \frac{1}{N_0}\right) .$$

N_0 étant ici d'ordre macroscopique, la parenthèse vaut 1 à environ 10^{-20} près, de sorte que l'exponentielle considérée ne se distingue pas de $\exp(\beta\varepsilon)$.

C'est seulement pour l'état fondamental, associé à $\varepsilon = 0$, que la distinction entre la valeur précise (II.43.b) et $\mu = 0$, a de l'importance : $\exp \beta\varepsilon$ étant alors strictement égale à 1, c'est le terme en $1/N_0$ qui détermine le nombre d'occupation. Mais la contribution de l'état fondamental a été mise à part, dans (II.43.a).

température T; $\overline{N}$ décroîtrait régulièrement lorsqu'on abaisserait la température à μ et V fixés; il n'y aurait pas de température limite comme T_B, et le système se viderait de ses particules pour T tendant vers zéro. Mais, dans la plupart des situations concrètes, c'est le nombre de particules N du système, et non son potentiel chimique, qui est fixé.

L'intégrale qui figure dans (II.44.b) est la même que celle que nous avons calculée, pour $T = T_B$, au paragraphe A.2.c. Le nombre N_0 de particules condensées dans l'état fondamental individuel a donc pour expression[47]

$$\boxed{N_0 = N\left[1 - \left(\frac{T}{T_B}\right)^{3/2}\right]} \quad \text{pour} \quad T \leqslant T_B \,. \tag{II.45}$$

On voit qu'il croît continûment de 0 à N quand T décroît de T_B à 0, N et V étant maintenus constants[48] (fig. 8).



FIGURE 8

Variation avec la température de la proportion N_0/N de particules condensées dans l'état fondamental individuel, pour un gaz de bosons identiques libres en nombre N fixé (la température de Bose T_B est donnée à partir de N et du volume V par la formule (II.21)).

47. En introduisant comme en (II.15) la variable sans dimension

$$x \widehat{=} \beta\varepsilon \,, \tag{1}$$

on transforme l'intégrale à calculer en

$$\frac{AV}{\beta^{3/2}} \int_0^{+\infty} \mathrm{d}x \, \frac{x^{1/2}}{\mathrm{e}^x - 1} = AV(kT)^{3/2} I^B(1) \,, \tag{2}$$

où $I^B(\varphi)$ est définie par (II.17). Or la température de Bose T_B est telle que (note 37)

$$N = AV(kT_B)^{3/2} I^B(1) \,. \tag{3}$$

On a donc

$$AV \int_0^{+\infty} \mathrm{d}\varepsilon \, \frac{\varepsilon^{1/2}}{\mathrm{e}^{\beta\varepsilon} - 1} = N \frac{T^{3/2}}{T_B^{3/2}} \,, \tag{4}$$

d'où la relation (II.45).

48. Le raisonnement qui nous a conduits à (II.45) suppose N_0 macroscopique. On pourrait lui objecter que, pour T voisine de T_B (mais inférieure), N_0 n'est plus macroscopique puisqu'il tend vers zéro quand $T \longrightarrow T_B{}^-$. De façon plus précise,

$$N_0 \underset{T \to T_B^-}{\sim} N \times \frac{3}{2} \frac{T_B - T}{T_B} \,;$$

d. Importance physique du phénomène

Ainsi, le phénomène de condensation de Bose fait jouer à l'état fondamental individuel un rôle tout à fait singulier : cet état est le seul à pouvoir être occupé par un nombre macroscopique N_0 de bosons, de l'ordre du nombre total N de particules du système, et tendant vers N lorsque T tend vers zéro. Les autres états individuels, même ceux dont l'énergie est très proche de $\varepsilon_0 = 0$, ont une population pratiquement « normale », c'est-à-dire qui ne diffère pas de façon radicale de celle qu'ils avaient pour $T > T_B$.

Il s'agit donc d'un phénomène « original », radicalement différent de ceux que nous avons étudiés jusqu'ici : c'est la première fois que nous rencontrons une situation dans laquelle *un ensemble macroscopique de particules se trouve dans un état quantique pur.* Habituellement en effet, l'état quantique d'un système macroscopique est un mélange statistique d'états purs (I, § III.A.2) : si par exemple, comme ici, les particules constituant le système sont indépendantes, leur répartition entre les divers états individuels possibles se fait de telle sorte que, à chaque instant, un état individuel donné n'est occupé que par un nombre microscopique (c'est-à-dire de l'ordre de 1) d'entre elles; ce nombre fluctue donc fortement (en valeur relative) au cours du temps et l'état quantique du système dans son ensemble ne peut être spécifié comme état pur. Dans le cas de la condensation de Bose au contraire, les N_0 bosons « condensés » se trouvent tous dans le même état quantique (pur) individuel; le système qu'ils constituent, macroscopique puisque N_0 l'est, se trouve dans l'état quantique pur construit par produit (tensoriel) des N_0 états individuels; même si N_0 fluctue au cours du temps, ses fluctuations sont négligeables comme celles de toutes les grandeurs macroscopiques (limite thermodynamique)[49]. On peut donc s'attendre à ce que, *dans les systèmes thermodynamiques de bosons, des effets quantiques se manifestent directement à l'échelle macroscopique,* à basse température.

C'est effectivement ce que montre l'expérience (voir complément VI.D). Les *propriétés de l'hélium à basse température,* plus précisément celles de l'isotope ^{4}He (le plus abondant dans l'hélium naturel), en sont un exemple frappant. Ce corps se liquéfie, sous la pression atmosphérique, vers 4 K. Si l'on abaisse encore la température, toujours sous la pression atmosphérique, on constate qu'il ne se solidifie pas mais que, à 2,2 K, ses propriétés changent de façon dramatique et deviennent extrêmement curieuses : au-dessous de cette température critique, on a affaire à un nouveau type de liquide, dit *« superfluide »*. Ce comportement très particulier est dû au fait que les particules constituant l'hélium 4 sont des bosons[50]. Un autre exemple est le phénomène de *supraconductivité,* que l'on observe à basse température pour

---

on peut donc toujours, en principe, choisir $T_B - T$ suffisamment petit pour que le second membre soit d'ordre microscopique (c'est-à-dire de l'ordre de 1) malgré le facteur N. En pratique cependant, la précision des mesures sur T, N et V est loin d'être suffisante pour qu'on puisse contrôler des différences $(T_B - T)$ de cet ordre.

Des arguments analogues sont valables pour le domaine $T > T_B$: la population de l'état fondamental y reste négligeable devant N, même lorsque $T \longrightarrow T_{B^+}$.

Ainsi, les deux descriptions coïncident en $T = T_B$ (égalité valable à 1 % ou 1 ‰ près au mieux) : on peut dire soit que la population de l'état fondamental y est « normale », soit qu'il y a condensation de Bose, mais avec un nombre de particules condensées très petit devant N.

49. *Attention :* les N_0 particules « condensées » dans l'état fondamental sont identiques aux $(N - N_0)$ qui occupent les autres états individuels, et n'en sont *pas discernables.*

50. Comme tous les « gaz rares », l'hélium est inerte chimiquement, de sorte que les particules qui le constituent au niveau microscopique sont des atomes isolés. Ces atomes contiennent chacun 2 électrons (le numéro atomique de l'hélium est $Z = 2$). Le noyau de l'isotope ^{4}He est composé de 2 protons et 2 neutrons (nombre de masse $A = 4$). L'atome d'hélium 4, formé d'un *nombre pair de particules de spin* 1/2, a forcément un spin entier, ce qui en fait obligatoirement un

.../...

certains métaux (mais aussi quelques autres corps) : leur résistance électrique, qui décroît d'abord de façon normale si l'on diminue la température, tombe brusquement à une valeur nulle lorsqu'on atteint une certaine température critique, et s'y maintient pour les températures inférieures. Ici aussi, l'explication de ce comportement inattendu fait intervenir des bosons : les électrons de conduction se groupent deux par deux pour former des entités de spin nul (« paires de Cooper »)[51].

Si l'origine de ces effets ne fait aucun doute, leur description quantitative est plus compliquée que celle de la condensation de Bose simple que nous étudions ici. Nous nous sommes en effet limités à des systèmes de bosons indépendants (et même libres). Comme il faut se placer à basse température pour provoquer la condensation de Bose, l'énergie caractéristique kT correspondante est faible. Dans un système réel, les *énergies d'interaction* entre les particules ne sont alors pas négligeables devant kT[52]; l'approximation des particules indépendantes est dans ce cas mauvaise, et les phénomènes observés sont quantitativement assez différents de ceux que l'on calculerait à partir de la théorie développée ici. Ils restent cependant toujours spectaculaires : lorsqu'on arrive à atteindre des températures inférieures à la température de Bose, les effets quantiques se manifestent à l'échelle macroscopique.

2. EXPRESSION DES GRANDEURS PHYSIQUES DANS LE CAS DE CONDENSATION DE BOSE

Dans le domaine où se produit la condensation de Bose (c'est-à-dire, si N et V sont fixés, pour $T < T_B$), *le potentiel chimique μ du système reste constamment nul* (cf. (II.44.a)). L'expression des autres grandeurs thermodynamiques se déduit des formules du paragraphe A.2.a par la même méthode que pour le nombre de particules (§ 1.c) : on calcule à part le terme associé à l'état fondamental individuel, et on évalue le reste de la somme sur les états (λ) de la façon habituelle pour un système macroscopique, c'est-à-dire en la remplaçant par une intégrale sur l'énergie individuelle ε; dans cette intégrale, la fonction de distribution $N^B(\varepsilon; T, \mu)$ est prise pour $\mu = 0$.

a. Énergie et chaleur spécifique

L'*énergie du système* s'évalue ainsi à partir de (II.5). Il est clair que le terme correspondant à l'état fondamental individuel est nul, puisque l'énergie ε_0 de cet état l'est : *les N_0 particules condensées ne contribuent pas à l'énergie* du système. Celle-ci provient donc seulement des $(N - N_0)$ particules qui ne sont pas condensées, et vaut

$$E = AV \int_0^{+\infty} d\varepsilon \frac{\varepsilon^{3/2}}{e^{\beta\varepsilon} - 1} \quad \text{pour } T \leqslant T_B \,. \tag{II.46}$$

---

boson. Au contraire, le noyau de l'isotope ^{3}He ne comporte qu'un neutron ($A = 3$), ce qui fait de l'atome correspondant un fermion (nombre impair de particules de spin 1/2). Effectivement, *l'hélium 3 ne devient pas superfluide* comme son isotope bosonique (sauf à température beaucoup plus basse, où se produit un appariement des atomes ^{3}He analogue à celui des électrons dans un supraconducteur).

51. On peut s'étonner que deux électrons, qui se repoussent électrostatiquement, puissent ainsi s'apparier. L'explication du phénomène fait intervenir le réseau d'ions positifs dans lequel se meuvent les électrons. On peut en donner une idée grossière en raisonnant comme suit.

En premier lieu, les ions positifs écrantent les charges des deux électrons, annulant pratiquement leur répulsion électrostatique. Mais le passage d'un électron déforme très légèrement le réseau dans son voisinage; les ions positifs, attirés par ce premier électron, attirent à leur tour le deuxième. C'est ainsi que, par l'intermédiaire du réseau, l'interaction résiduelle entre deux électrons est faiblement attractive.

52. Par exemple, la transition superfluide se produit dans un domaine où l'hélium 4 est *liquide,* et non plus gazeux. Or, dans un liquide, l'énergie potentielle d'interaction entre les particules est du même ordre de grandeur que leur énergie cinétique.

On trouve facilement[53]

$$E = 1{,}341 \times \frac{3\sqrt{\pi}}{4} AV(kT)^{5/2} \quad \text{pour } T \leqslant T_B \,, \tag{II.47.a}$$

que l'on peut écrire finalement, en faisant intervenir T_B,

$$E = 0{,}770\,1 NkT \left(\frac{T}{T_B}\right)^{3/2} \quad \text{pour } T \leqslant T_B \,. \tag{II.47.b}$$

La *capacité calorifique à volume constant* s'en déduit aussitôt :

$$C_V \widehat{=} \left(\frac{\partial E}{\partial T}\right)_{N,\,V} = \frac{5}{2} . 0{,}770\,1 Nk \left(\frac{T}{T_B}\right)^{3/2} \quad \text{pour } T \leqslant T_B \,, \tag{II.48}$$

ainsi que la *chaleur spécifique à volume constant :*

$$c_V = 1{,}925 R \left(\frac{T}{T_B}\right)^{3/2} \quad \text{pour } T \leqslant T_B \tag{II.49}$$

($R \widehat{=} N_A k$ est la constante des gaz parfaits).

Le résultat (II.47) est à comparer avec NkT (ordre de grandeur attendu pour l'énergie d'excitation thermique d'un ensemble de N particules), et le résultat (II.49) avec R (la chaleur spécifique c_V tend vers $3R/2$ à haute température, domaine de validité de la mécanique classique). On voit que les valeurs pour $T = T_B$ sont du bon ordre de grandeur : $0{,}77 NkT_B$ pour l'énergie et $1{,}9R$ pour la chaleur spécifique. Mais, lorsque la température T s'abaisse, l'énergie décroît plus rapidement qu'on ne s'y attendrait (comme $T^{5/2}$), et la chaleur spécifique décroît elle aussi au lieu de rester constante, pour tendre vers zéro avec T. Ce comportement reflète bien sûr le fait que les particules se condensent en nombre de plus en plus grand au fur et à mesure que s'abaisse la température, et qu'elles cessent alors de contribuer à l'énergie du système comme nous l'avons souligné ci-dessus : l'expression (II.47.b) est effectivement de l'ordre de $(N - N_0)kT$ (voir formule (II.45)).

---

53. En utilisant à nouveau le changement de variable

$$x \widehat{=} \beta\varepsilon \,, \tag{1}$$

on transforme l'expression (II.46) en

$$E = \frac{AV}{\beta^{5/2}} \int_0^{+\infty} \mathrm{d}x \, \frac{x^{3/2}}{\mathrm{e}^x - 1} \,, \tag{2}$$

produit de grandeurs dimensionnées par une intégrale définie sans dimension que l'on trouve dans les tables :

$$\int_0^{+\infty} \mathrm{d}x \, \frac{x^{3/2}}{\mathrm{e}^x - 1} = \Gamma\left(\frac{5}{2}\right) \zeta\left(\frac{5}{2}\right) \tag{3}$$

d'après la définition (1) de la note 36, avec

$$\Gamma\left(\frac{5}{2}\right) = \frac{3\sqrt{\pi}}{4} \,, \quad \zeta\left(\frac{5}{2}\right) = 1{,}341 \ldots \,. \tag{4}$$

On peut ensuite utiliser la relation (note 37)

$$\frac{N}{AV} \beta_B^{3/2} = 2{,}612 \, \frac{\sqrt{\pi}}{2} \tag{5}$$

pour remplacer AV en fonction de N et T_B :

$$E = 1{,}341 \, \frac{3\sqrt{\pi}}{4} AV(kT)^{5/2} = 0{,}770\,1 NkT (T/T_B)^{3/2} \,. \tag{6}$$

b. Grand potentiel, entropie, énergie libre

Le *grand potentiel* du gaz de bosons libres peut être calculé, pour $T \leqslant T_B$, en appliquant à la formule (II.6) la méthode indiquée ci-dessus :

$$J = -kT \ln(1+N_0) - kT.AV \int_0^{+\infty} d\varepsilon\, \varepsilon^{1/2} \ln\left[1+N^B(\varepsilon;\, T,\, 0)\right]$$

pour $T \leqslant T_B$. (II.50)

Le premier terme du second membre, associé à l'état fondamental individuel, est ici seulement de l'ordre de $kT \ln N$ (c'est-à-dire quelques dizaines de fois kT); il est donc totalement négligeable devant le terme suivant, qui est d'ordre NkT[54], de sorte que le grand potentiel thermodynamique du système s'écrit en fait

$$J = -AVkT \int_0^{+\infty} d\varepsilon\, \varepsilon^{1/2} \ln\left[1+N^B(\varepsilon;\, T,\, 0)\right] \quad \text{pour } T \leqslant T_B . \tag{II.51}$$

La relation

$$J = -\frac{2}{3} E , \tag{II.52}$$

que nous avons obtenue plus haut en effectuant une intégration par parties sur une expression de cette forme (formule (II.23)) reste donc valable : *le grand potentiel est simplement proportionnel à l'énergie*[55].

On peut en déduire l'*énergie libre thermodynamique* F par la relation

$$F = J + \mu N \tag{II.53}$$

(cf. V (III.20)). Comme le potentiel chimique μ est identiquement nul dans tout le domaine qui nous intéresse, on a simplement[56]

$$F = J = -\frac{2}{3} E \quad \text{pour } T \leqslant T_B . \tag{II.54}$$

L'énergie libre coïncide avec le grand potentiel et elle est donc elle aussi *proportionnelle à l'énergie.*

Quant à l'*entropie thermodynamique,* elle découle simplement du résultat précédent par la formule III (II.28) :

$$S = -\left(\frac{\partial F}{\partial T}\right)_{N,\, V} , \tag{II.55}$$

---

54. Ce second terme est en effet proportionnel à V, c'est-à-dire d'ordre N à densité N/V fixée. Autrement dit, il est bien *extensif* comme doit l'être le grand potentiel J à la limite thermodynamique, alors que le premier terme se situe au niveau de ceux que l'on néglige dans cette limite.

55. *Attention :* dans le domaine où il y a condensation de Bose, le potentiel chimique μ reste constamment nul; il n'est donc pas possible ici d'exprimer le grand potentiel J comme fonction de ses variables « naturelles » T, μ et V.

56. Le potentiel chimique étant ici multiplié par le nombre total de particules N, on pourrait se demander si le terme correspondant est bien nul. Nous savons qu'en fait (formule (II.36))

$$\mu = -\frac{kT}{N_0} ,$$

où N_0 est le nombre de particules condensées. Comme N_0 et N sont du même ordre de grandeur, le produit μN est de l'ordre de kT, c'est-à-dire absolument inappréciable devant J qui est d'ordre macroscopique.

qui donne[57]

$$S = \frac{5}{3}\frac{E}{T} \quad \text{pour } T \leqslant T_B \,, \tag{II.56}$$

c'est-à-dire

$$S = 1{,}283 Nk \left(\frac{T}{T_B}\right)^{3/2} \quad \text{pour } T \leqslant T_B \,. \tag{II.57}$$

Il est intéressant de souligner que *les N_0 particules condensées ne contribuent pas à l'entropie du système.* Bien sûr, la raison technique en est simplement que l'expression (II.57), qui est d'ailleurs proportionnelle à $(N - N_0)$ (cf. formule (II.45)), provient uniquement de termes « normaux » dans J et E, grandeurs auxquelles les N_0 particules condensées donnent déjà une contribution nulle. Mais, si l'on se réfère à l'interprétation microcanonique de l'entropie comme logarithme du nombre d'états accessibles (interprétation que l'on peut aussi bien utiliser ici puisqu'il s'agit d'un système thermodynamique), on constate qu'une fraction importante des particules du système ne sont ainsi pas prises en compte dans le dénombrement des états accessibles. L'explication est la suivante : le sous-système formé par les N_0 particules condensées se trouve dans un état quantique pur bien déterminé (cf. § 1.d), de sorte que la contribution qu'il apporte au nombre d'états accessibles est simplement un facteur 1, et son entropie est donc nulle.

c. Pression et équation d'état

La *pression* du gaz de bosons se déduit par exemple de son énergie libre :

$$p = -\left(\frac{\partial F}{\partial V}\right)_{N,\,T} \,. \tag{II.58}$$

On trouve ainsi[58]

$$p = \frac{2}{3}\frac{E}{V} = 1{,}341\,(2s+1)\left(\frac{m}{2\pi\hbar^2}\right)^{3/2} (kT)^{5/2} \quad \text{pour } T \leqslant T_B \,. \tag{II.59}$$

Ce résultat a une *forme très particulière : la pression est indépendante de N et V, et dépend seulement de la température T.* Dans le domaine où se manifeste la condensation de Bose, l'*équation d'état* d'un gaz de bosons libres se réduit donc à une *relation entre p et T,* dont N et V sont absents. Ceci va nous permettre, dans le paragraphe suivant, de donner une interprétation physique simple au phénomène de condensation de Bose.

---

57. En effet, d'après (II.54),

$$S = -\left(\frac{\partial F}{\partial T}\right)_{N,\,V} = +\frac{2}{3}\left(\frac{\partial E}{\partial T}\right)_{N,\,V} \quad \text{pour } T \leqslant T_B \,. \tag{1}$$

Il suffit alors d'utiliser l'expression (II.47) de E pour trouver

$$S = \frac{2}{3} \times \frac{5}{2}\frac{E}{T} = \frac{5}{3}\frac{E}{T} \quad \text{pour } T \leqslant T_B \,. \tag{2}$$

58. Compte tenu de (II.54), la relation (II.58) s'écrit dans le cas présent

$$p = \frac{2}{3}\left(\frac{\partial E}{\partial V}\right)_{N,\,T} \quad \text{pour } T \leqslant T_B \,. \tag{1}$$

Mais la formule (II.46) montre que l'énergie E est simplement proportionnelle au volume V du système, de sorte que

$$p = \frac{2}{3}\frac{E}{V} \tag{2}$$

comme pour $T > T_B$ (cf. relation (II.29)). L'expression (II.59) s'en déduit aussitôt.

C. La transition de phase et ses caractéristiques

1. COMPORTEMENT DES GRANDEURS PHYSIQUES A NOMBRE DE PARTICULES ET VOLUME CONSTANTS

Le phénomène de condensation de Bose affecte profondément les propriétés du système : les formules du paragraphe B.2, qui caractérisent ces propriétés dans le domaine où se produit la condensation de Bose, ont une structure nettement différente de celles qui sont valables lorsque la population de l'état fondamental reste d'ordre microscopique (situation que nous avons qualifiée de « normale » (§ A.2.c)). Lorsque, en modifiant les paramètres extérieurs, on lui fait franchir la frontière entre ces deux domaines, le système subit ce que l'on appelle une *« transition de phase »* : à nombre de particules N et volume V fixés, la transition de phase se produit lorsque la température T atteint la température de Bose T_B, qui dépend de N, V, et de la nature du gaz (formule (II.21)).

Les principales grandeurs thermodynamiques (potentiel chimique μ, énergie E, grand potentiel J, entropie S, énergie libre F et pression p) *sont continues en* $T = T_B$, à N et V fixés. En effet, le potentiel chimique μ, qui est déterminé pour $T \geqslant T_B$ par la relation (II.13), s'annule lorsque T tend vers T_B par valeurs supérieures, et l'on sait qu'il reste nul pour $T < T_B$. Ceci assure en même temps que la valeur de E déduite de la formule (II.22) coïncide pour $T = T_B$ avec celle qui provient de (II.46)-(II.47) :

$$E(T_B) = 0{,}770\,1 N k T_B \,. \tag{II.60}$$

La continuité de J et p en découle aussitôt, puisque les relations (II.25) et (II.29) sont valables dans les deux domaines (voir (II.54) et (II.59)). On montre facilement[59] que l'entropie S est également continue, et donc aussi F. Pour la plupart des grandeurs précédentes, la continuité était d'ailleurs nécessaire : on voit mal l'énergie, par exemple, « sauter » brusquement d'une valeur à une autre lorsque la température passe par T_B. En revanche, leurs dérivées peuvent être discontinues en $T = T_B$. On montre pourtant (complément VI.C) que la chaleur spécifique à volume constant c_V, dérivée de l'énergie par rapport à la température (à un facteur près), est continue; c'est au niveau de sa propre dérivée que se situe la discontinuité. La figure 9 représente les variations en

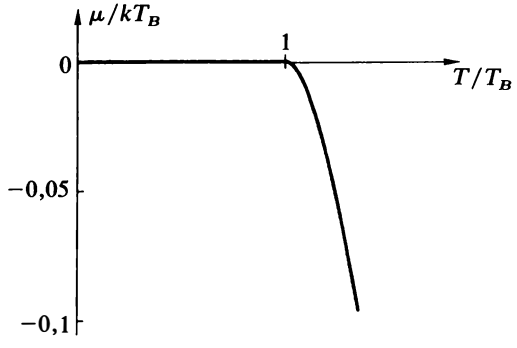


FIGURE 9

Comportement du potentiel chimique μ d'un gaz de bosons libres au voisinage de $T = T_B$, à nombre de particules et volume fixés. μ est continu au point de transition, ainsi que sa dérivée première (voir complément VI.C, formule (C.9)).

59. En effet, la formule (II.27) peut aussi s'écrire (cf. formule (2) de la note 34)

$$S = kAV \int_0^{+\infty} d\varepsilon\, \varepsilon^{1/2} \{\ln[1 + N^B(\varepsilon; T, \mu)] + \beta(\varepsilon - \mu) N^B(\varepsilon; T, \mu)\} \,, \tag{1}$$

.../...

fonction de la température T, à N et V fixés, du potentiel chimique μ, la figure 10 celles de l'énergie E et de la chaleur spécifique c_V du système. Les courbes montrent clairement un changement de comportement en $T = T_B$, même si ces grandeurs sont continues en ce point.

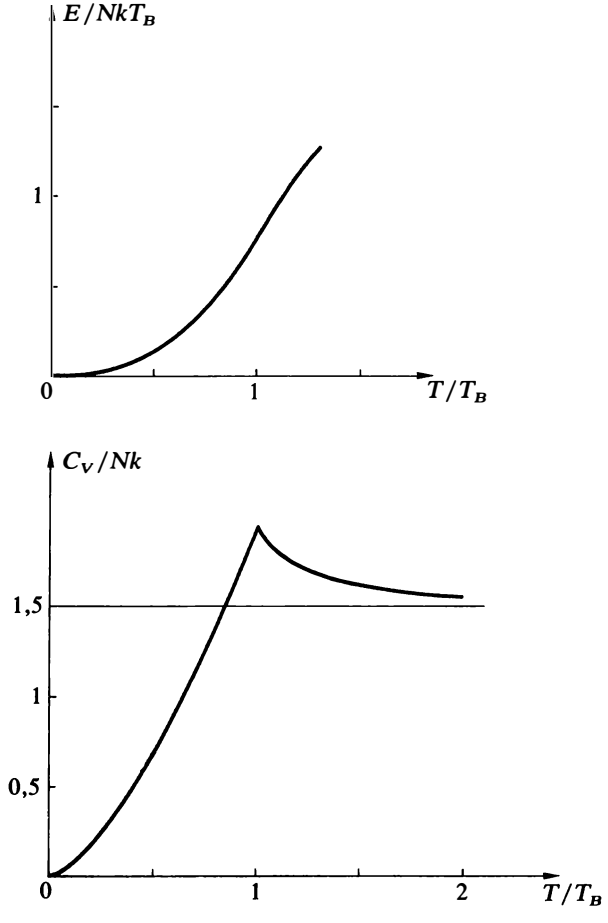


FIGURE 10

Variations, en fonction de la température T, de l'énergie E (fig. a) et de la chaleur spécifique c_V (fig. b) d'un gaz de bosons libres, le nombre de particules et le volume étant fixés. Ces deux grandeurs sont continues en $T = T_B$, mais la dérivée première de c_V par rapport à T, ainsi que la dérivée seconde de E qui lui est proportionnelle, sont discontinues (complément VI.C, formules (C.3) et (C.14)).

qui n'est autre qu'une explicitation de la relation

$$S = \frac{1}{T}(-J + E - \mu N) \tag{2}$$

(cf. formule V (II.23)). J, E et μ étant continus en $T = T_B$ (N est fixé), S l'est également. De même $F = E - TS$ est continue puisque E et S le sont.

2. ÉTUDE DE LA TRANSITION DE PHASE

a. Changement de variables et de point de vue

Nous avons jusqu'ici exprimé les grandeurs physiques en fonction des variables canoniques N, V et T, et nous avons analysé le comportement du système à N et V fixés; la condensation de Bose se produit alors lorsque T devient inférieure à la température de Bose T_B, fonction elle aussi de N et V.

Or l'étude des transitions de phase montre de façon générale que leurs propriétés apparaissent plus clairement si l'on ne fixe pas le volume total du système, et si l'on utilise comme variables, outre le *nombre de particules* N qui reste *fixé,* la *pression* p et la *température* T. Examinons en effet les transitions de phase les plus courantes, c'est-à-dire les *«changements d'état» d'un corps pur* tel que l'eau : vaporisation ou liquéfaction, fusion ou solidification, sublimation. On sait que, *à pression constante, la transition de phase se produit à une température déterminée :* sous la pression atmosphérique par exemple, l'eau liquide et sa vapeur sont en équilibre à 100° Celsius (la glace et l'eau liquide à 0 °C), quel que soit par ailleurs le volume du système, qui change lorsque varie la proportion relative de liquide et de vapeur (ou de solide et de liquide), et quel que soit le nombre total de molécules.

Pour analyser la transition de phase qui nous intéresse ici, nous allons donc changer de point de vue : *le nombre de particules N restera fixé, mais pas le volume V.* Nous pourrons ainsi comparer la condensation de Bose avec par exemple la transition liquide-vapeur, dont les caractéristiques sont bien connues.

b. Courbes caractérisant la transition de phase

α. Isothermes dans le plan (V, p)

Examinons en premier lieu le comportement du gaz de bosons libres à *température T et nombre de particules N fixés.* Supposons que le système se trouve initialement dans le domaine «normal», et qu'*on diminue progressivement son volume V.* La condition (II.13) fixant le potentiel chimique μ montre aussitôt que ce dernier croît et se rapproche de zéro. La pression p, donnée par la formule (II.30), croît également.

Lorsqu'on parvient à une certaine valeur V_B du volume, le potentiel chimique μ s'annule. Cette valeur V_B est d'ailleurs, d'après (II.13), telle que (cf. relation (II.19))

$$\frac{4\pi^2}{2s+1}\left(\frac{\hbar^2\beta}{2m}\right)^{3/2}\frac{N}{V_B}=2{,}612\,\frac{\sqrt{\pi}}{2}\,, \tag{II.61}$$

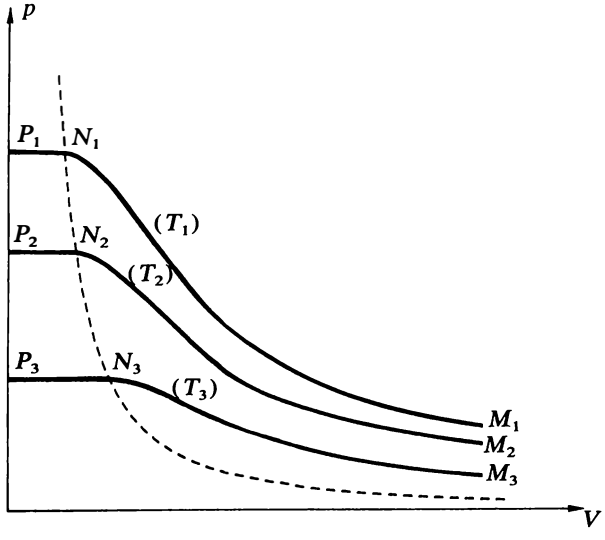
où cette fois β et N sont fixés. La pression vaut alors

$$p_B=\frac{2s+1}{6\pi^2}\left(\frac{2m}{\hbar^2}\right)^{3/2}\int_0^{+\infty}\mathrm{d}\varepsilon\,\frac{\varepsilon^{3/2}}{\mathrm{e}^{\beta\varepsilon}-1}=1{,}341\,(2s+1)\left(\frac{m}{2\pi\hbar^2}\right)^{3/2}(kT)^{5/2} \tag{II.62}$$

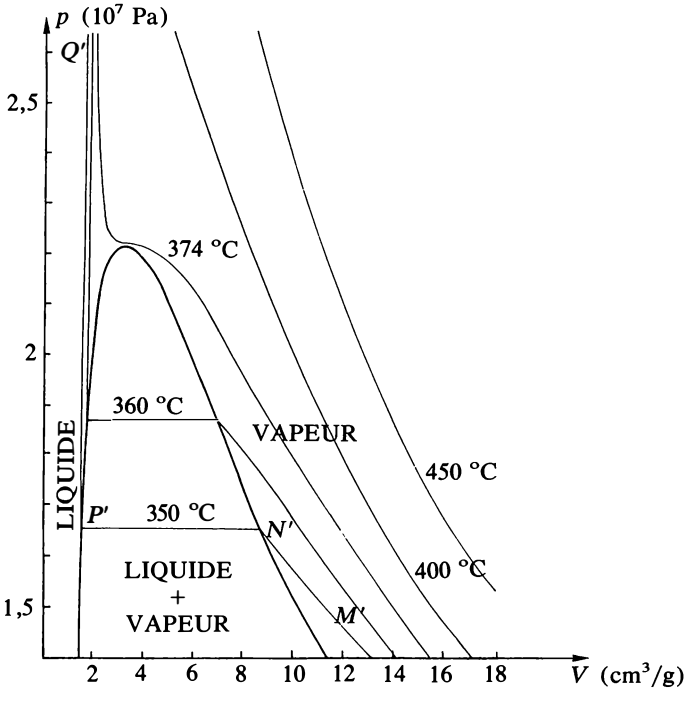
(voir formule (II.59)).

Si l'on continue ensuite à diminuer le volume V à β et N fixés, la relation (II.13.a) ne peut plus être satisfaite et doit être remplacée par (II.44.b) : pour $V<V_B$, il y a *condensation de Bose,* et le nombre N_0 de particules condensées croît au fur et à mesure que V décroît. Mais alors *la pression reste constante* et égale à p_B, puisque d'après (II.59) elle dépend seulement de T, pas de N ni de V.

La figure 11 résume le comportement que nous venons de décrire et le compare à celui d'un corps pur lors de la transition vapeur-liquide. On constate une analogie certaine, mais aussi des différences, que nous discuterons plus loin (§ c).



(a)



(b)

FIGURE 11

Courbes isothermes (pression en fonction du volume, à température constante) pour un système de bosons libres (a) et pour l'eau au voisinage de la transition liquide vapeur (b). Sur la figure a, les portions de courbes *MN* correspondent au comportement « normal » du système, les paliers *NP* à la condensation de Bose. Dans la figure b, l'eau est entièrement sous forme gazeuse le long de l'arc *M'N'*, entièrement liquide le long de *P'Q'*; le segment *N'P'* est le « palier de liquéfaction » sur lequel eau liquide et vapeur d'eau sont en équilibre.

β. Courbe de transition et d'équilibre dans le plan (T, p)

La figure 12 représente la pression de condensation p_B en fonction de la température T choisie (relation (II.62)), en même temps que le « diagramme d'état » de l'eau, constitué par la courbe de vaporisation (ensemble des points d'équilibre entre l'eau liquide et sa vapeur), la courbe de fusion (équilibre entre la glace et l'eau liquide) et la courbe de sublimation (équilibre entre glace et vapeur d'eau). Ici aussi apparaissent clairement des analogies, et des différences.

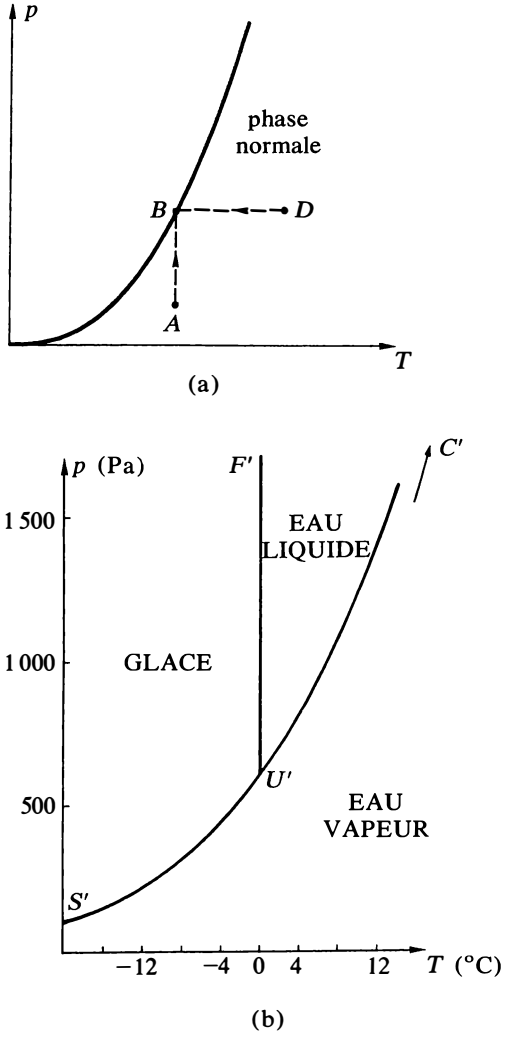


FIGURE 12

Courbes de transition de phase pour un gaz de bosons libres (a) et pour l'eau (b). On a représenté sur la figure a une évolution du système à température constante (trajet *AB*) et une évolution à pression constante (*DB*). Sur la figure b, *U'C'* est la courbe de vaporisation, limitée par le « point triple » *U'* et le « point critique » *C'*; *U'F'* est la courbe de fusion et *S'U'* la courbe de sublimation, limitées toutes deux en *U'* mais ne comportant pas de point critique.

L'évolution du système lorsqu'on diminue son volume à température constante (§ α ci-dessus) s'effectue dans le plan (T, p) sur une verticale d'abscisse T fixée (fig. 12.a) : la pression croît tant que le système se trouve dans le domaine « normal »; la condensation de Bose commence à se produire lorsque le point représentatif de l'état du système atteint la courbe de transition donnant p_B en fonction de T; il s'y maintient ensuite sans plus bouger pendant que le volume continue à décroître et que le nombre de particules condensées augmente.

On pourrait également envisager de faire parcourir au système une portion d'horizontale, sur la figure 12.a : dans le domaine « normal », on peut diminuer la température du système à pression constante; lorsque le point représentatif atteint la courbe de transition, la condensation de Bose s'amorce; la température est alors « bloquée », et le point représentatif reste sur la courbe de transition même si l'on continue à retirer de l'énergie au système, à pression constante (auquel cas son volume décroît[60] et le nombre de particules condensées croît).

De façon plus générale, *le domaine « normal » correspond à la portion du plan* (T, p) *située à droite de la courbe de transition; s'il y a condensation de Bose, le point représentant l'état du système se situe obligatoirement sur la courbe de transition.* Notons que la portion du plan située à gauche de la courbe n'est jamais atteinte (voir § c ci-dessous).

c. Caractérisation des deux phases en présence

L'interprétation que nous venons de dégager pour la relation (II.62) est tout à fait semblable à celle qui vaut pour les courbes composant le diagramme d'état de la figure 12.b : la courbe de vaporisation est, dans le plan (T, p), l'ensemble des points d'équilibre entre la phase liquide et la phase gazeuse du corps pur considéré, la courbe de fusion joue le même rôle pour les phases solide et liquide, et la courbe de sublimation pour les phases solide et gazeuse. Les plages comprises entre ces courbes de transition correspondent aux états homogènes, c'est-à-dire constitués d'une phase unique, liquide, solide ou gazeuse suivant le cas.

Ainsi se trouve expliquée la forme très particulière de l'« équation d'état » (II.59) du système de bosons libres dans le domaine où il y a condensation de Bose : cette relation est en fait l'équation de *la courbe du plan* (T, p) *sur laquelle coexistent deux phases en équilibre.* Celles-ci sont faciles à identifier par analogie : il s'agit d'une *« phase condensée »*, constituée par les N_0 particules occupant l'état fondamental individuel, et une *« phase normale »* comprenant les $(N - N_0)$ particules réparties dans les autres états individuels[61].

---

60. Dans le cas où la pression est fixée, le rapport E/V reste constant d'après (II.29) et (II.59). Dans le domaine « normal », β et μ varient tous deux de façon que l'expression (II.30) garde une valeur constante; mais l'intégrale figurant dans (II.13.a) croît, pour que son produit par V reste lui aussi constant. La condensation de Bose commence toujours lorsque le potentiel chimique μ s'annule.

61. La condensation de Bose accompagne donc une transition de phase *« du premier ordre »*, comme la transition liquide-vapeur. L'affirmation contraire, qui figure dans certains ouvrages, est erronée. Il faut dire que les transitions de phase observées à basse température dans les systèmes réels de bosons sont en général d'un type différent, dit *« du second ordre » :* pour l'hélium-4 par exemple, il existe bien une courbe de transition, dans le plan (T, p), entre la phase liquide normale et la phase superfluide, mais le système la traverse « sans s'y arrêter », c'est-à-dire qu'on n'observe pas de palier sur les isothermes du plan (V, p); il n'y a pas non plus de chaleur latente de transition. Pour plus de détails, voir complément VI.D.

Cette analogie fondamentale entre la condensation de Bose et les changements d'états ordinaires d'un corps pur s'accompagne cependant de différences somme toute assez importantes. En premier lieu, *la phase condensée et la phase normale* d'un système de bosons libres *ne sont pas, lorsqu'elles coexistent, séparées dans l'espace* comme un liquide et sa vapeur. La raison en est que les particules condensées se trouvent dans un état quantique individuel (l'état fondamental) dont la fonction d'onde est non nulle dans tout le volume V; cette *délocalisation quantique de la phase condensée* aboutit à une interpénétration totale entre les deux phases : la phase « normale » occupe le volume V comme le ferait un gaz ordinaire; quant aux N_0 particules de la phase condensée, chacune d'elles a une probabilité de présence (quantique) répartie uniformément dans ce même volume V.

D'autre part, la comparaison entre les figures 11.a et 11.b, ou le fait que la partie gauche du plan (T, p) est inaccessible sur la figure 12.a, montre qu'*on ne peut pas observer la phase condensée seule.* Autrement dit, on ne peut pas effectuer la transformation complète de N particules initialement toutes normales en un système de N particules condensées d'où la phase normale aurait disparu, sauf dans des conditions limites peu réalistes : volume nul si la température est fixée (voir figure 11.a), température nulle si le volume est maintenu constant (cf. formule (II.45)). Cette propriété est évidemment liée au *caractère trop simpliste du modèle* que nous avons étudié ici : particules strictement ponctuelles et sans aucune interaction[62]. Mais toute modification de ces hypothèses dans un sens plus réaliste complique considérablement le modèle (voir complément VI.D).

III. GAZ DE PHOTONS. THERMODYNAMIQUE DU RAYONNEMENT

Nous considérons dans cette troisième partie les *systèmes de photons.* Les photons, particules de masse nulle, sont les *quanta associés au champ électromagnétique;* le comportement d'un gaz de photons en équilibre peut être analysé expérimentalement à travers les propriétés thermodynamiques du rayonnement électromagnétique. Les photons sont des *bosons dont le nombre n'est pas conservé,* ce qui leur confère un comportement statistique très particulier : leur potentiel chimique est identiquement nul[63].

On appelle *« corps noir »* une cavité fermée contenant un *gaz de photons en équilibre thermique avec un thermostat.* Pour le réaliser, il suffit de maintenir à la température T une enceinte de volume V; il s'établit alors spontanément à l'intérieur de cette enceinte un système d'ondes électromagnétiques, c'est-à-dire du point de vue

62. On peut s'étonner qu'une transition de phase apparaisse dans un système dont les constituants n'ont pas d'interactions mutuelles : c'est en effet l'attraction entre les molécules d'eau qui est responsable de l'existence d'une phase liquide, dont les propriétés physiques sont nettement différentes de celles de la phase gazeuse. Mais on peut montrer (exercices V.5 et VI.1) que les statistiques quantiques induisent une sorte d'« interaction effective » entre particules identiques, répulsive s'il s'agit de fermions, attractive dans le cas de bosons. La transition de phase associée à la condensation de Bose apparaît ainsi comme résultant de cette « attraction effective » entre les particules du système.

63. Il existe d'autres bosons dont le nombre n'est pas conservé : les phonons, par exemple, quanta d'excitation des vibrations d'un réseau cristallin (cf. complément III.E, § V). Les phonons ne sont pas de véritables particules, puisqu'ils n'existent pas en-dehors du cristal auquel ils sont attachés (ce sont des « quasi-particules »), mais leur comportement statistique est analogue à celui des photons (leurs interactions ne sont toutefois pas totalement négligeables).

quantique un système de photons (§ A). Les grandeurs physiques associées à ce système (§ B) peuvent être mesurées en observant les caractéristiques du rayonnement émis vers l'extérieur par une ouverture de taille négligeable percée dans la paroi de l'enceinte (*« rayonnement du corps noir »*, ou *« rayonnement thermique »*, voir complément VI.F, § II.3). Ces caractéristiques sont bien connues depuis la fin du siècle dernier. Rappelons que c'est leur analyse qui amena Planck à formuler en 1900 l'« hypothèse des quanta », point de départ de la révolution scientifique qui allait conduire à la découverte de la mécanique quantique.

A. Ondes électromagnétiques et photons

Nous commençons par déterminer, en utilisant la théorie classique de l'électromagnétisme, les états quantiques individuels que peuvent occuper les photons du système. Nous précisons ensuite les propriétés statistiques des ensembles de photons, pour en déduire enfin (§ B) les lois qui régissent le rayonnement du corps noir.

1. MODES PROPRES ÉLECTROMAGNÉTIQUES D'UNE ENCEINTE

Considérons une enceinte vide, que nous supposons pour simplifier cubique, d'arête L (cette hypothèse ne restreint pas la généralité des résultats). Du point de vue classique, le champ électromagnétique le plus général qui peut exister dans cette cavité est donné par les équations de Maxwell. Ces équations étant linéaires, leur solution générale est une *superposition linéaire de solutions élémentaires monochromatiques vérifiant les conditions aux limites* imposées par les parois de la cavité; ces solutions monochromatiques sont les *« modes propres »* de la cavité.

a. Caractérisation des modes propres

Si les parois sont parfaitement réfléchissantes (ce qui est le cas lorsqu'elles sont constituées de conducteurs parfaits), elles imposent des conditions aux limites de type strict : annulation, sur les parois, de la composante tangentielle du champ électrique et de la composante normale du champ magnétique. Comme toujours (cf. appendice II, § II.4.c.α), la situation réelle est beaucoup plus compliquée si on l'analyse en détail, mais il n'est pas nécessaire de la connaître exactement pour déterminer les propriétés macroscopiques du système.

Comme au chapitre I (§ III.B.2), nous choisirons des *conditions aux limites périodiques*. Leur avantage est de permettre l'utilisation d'ondes progressives. On sait en particulier que, dans une onde progressive plane et monochromatique, la dépendance par rapport aux coordonnées d'espace $\vec{r}$ et au temps t des diverses composantes des champs électrique et magnétique est, en notation complexe, de la forme

$$\exp i\left[\vec{K}\,.\vec{r}-\omega t\right]. \tag{III.1}$$

Les équations de Maxwell (dans le vide) imposent au vecteur d'onde $\vec{K}$ d'être lié à la pulsation ω par la relation

$$|\vec{K}|=\frac{\omega}{c}, \tag{III.2}$$

où c est la vitesse de la lumière. En outre, les conditions aux limites restreignent les vecteurs d'onde possibles : leurs composantes doivent être telles que[64]

$$K_x=n_x\frac{2\pi}{L},\quad K_y=n_y\frac{2\pi}{L},\quad K_z=n_z\frac{2\pi}{L}, \tag{III.3}$$

64. Les conditions (III.3) assurent que l'expression (III.1) reprend la même valeur en $x=L$ qu'en $x=0$ (pour tous y et z), en $y=L$ qu'en $y=0$ (pour tous x et z), en $z=L$ qu'en $z=0$ (pour tous x et y), et vérifie donc les conditions aux limites périodiques.

où n_x, n_y, n_z sont des *entiers positifs, négatifs ou nuls*[65].

Les équations de Maxwell impliquent que, pour chaque vecteur d'onde $\vec{K}$, les champs électrique $\vec{\mathcal{E}}$ et magnétique $\vec{\mathcal{B}}$ de l'onde sont constamment perpendiculaires à $\vec{K}$ et perpendiculaires entre eux (transversalité des ondes électromagnétiques). Ceci permet *deux états de polarisation indépendants,* dans le plan orthogonal à $\vec{K}$[66].

Ainsi, *un mode propre électromagnétique* de l'enceinte cubique considérée *est caractérisé par un vecteur d'onde* $\vec{K}$ *vérifiant* (III.3) *et un état de polarisation* déterminé; sa pulsation est donnée par la relation (III.2). Le champ électromagnétique le plus général qui peut exister dans cette enceinte est une superposition linéaire des divers modes propres ainsi définis.

b. Densité spectrale de modes propres

Pour les applications macroscopiques, l'important est de connaître le *nombre de modes propres dont la pulsation est comprise entre* ω *et* $\omega + d\omega$. Ce nombre, de la forme $\tilde{\rho}(\omega)\,d\omega$, se calcule facilement à partir des résultats du paragraphe a : on trouve[67] que la densité spectrale $\tilde{\rho}(\omega)$ correspondante s'écrit

$$\tilde{\rho}(\omega) = \frac{V}{\pi^2 c^3}\,\omega^2\ , \tag{III.4}$$

où V est le volume de l'enceinte.

2. QUANTIFICATION DES MODES PROPRES : PHOTONS

a. Impulsion, énergie et état de spin d'un photon

Si l'on s'en tenait à la physique classique, l'amplitude de chaque mode propre (c'est-à-dire celle des champs $\vec{\mathcal{E}}$ et $\vec{\mathcal{B}}$ correspondants), et par conséquent son énergie, pourraient prendre *a priori* des valeurs quelconques, réparties continûment de 0 à $+\infty$. En réalité, l'énergie d'un champ électromagnétique, comme celle d'un système matériel, est quantifiée.

---

65. Le « mode » caractérisé par $n_x = n_y = n_z = 0$, c'est-à-dire $\vec{K} = \vec{0}$ et donc $\omega = 0$ d'après (III.2), doit d'emblée être mis à part : il existe quelles que soient les dimensions de l'enceinte, et correspond à des champs $\vec{\mathcal{E}}$ et $\vec{\mathcal{B}}$ uniformes et indépendants du temps, mais à part cela quelconques. Il serait d'ailleurs exclu si l'on choisissait des conditions aux limites strictes.

66. On peut choisir par exemple, pour les caractériser, deux vecteurs unitaires orthogonaux à $\vec{K}$ et orthogonaux entre eux. On prendra alors, soit les deux polarisations rectilignes indépendantes définies par ces deux vecteurs, soit les polarisations circulaires droite et gauche que l'on peut leur associer.

67. Le nombre de modes propres dont le vecteur d'onde $\vec{K}$ a un module compris entre K et $K + dK$ vaut d'après (III.3)

$$2 \times \left(\frac{L}{2\pi}\right)^3 \times 4\pi K^2\,dK = \frac{L^3}{\pi^2}\,K^2\,dK\ . \tag{1}$$

Le calcul est en tous points semblable à celui de la note 19 du chapitre I, le facteur 2 supplémentaire étant dû au fait qu'à chaque vecteur d'onde possible sont associés deux modes propres distincts, correspondant aux deux états de polarisation indépendants. Il suffit ensuite de changer de variable selon (III.2) pour obtenir

$$\frac{L^3}{\pi^2}\,K^2\,dK = \frac{L^3}{\pi^2}\,\frac{\omega^2}{c^2}\,\frac{d\omega}{c}\ , \tag{2}$$

c'est-à-dire le résultat (III.4).

Plus précisément, à une *onde* électromagnétique de *vecteur d'onde* $\vec{K}$ et de *pulsation* ω on associe des *particules,* appelées *photons,* dont l'*impulsion* $\vec{p}$ et l'*énergie* ε sont données par les *relations de Planck-Einstein*[68]

$$\vec{p} = \hbar \vec{K} , \tag{III.5.a}$$

$$\varepsilon = \hbar \omega . \tag{III.5.b}$$

La relation ondulatoire (III.2) devient alors pour les photons

$$\varepsilon = |\vec{p}\,| \, c . \tag{III.6}$$

Si l'on compare cette relation à celle qui lie, en relativité, l'énergie ε et l'impulsion $\vec{p}$ d'une particule libre de masse m :

$$\varepsilon = \sqrt{\vec{p}^{\,2} c^2 + m^2 c^4} , \tag{III.7}$$

on constate que *le photon est une particule de masse nulle :*

$$m = 0 . \tag{III.8}$$

Les deux polarisations possibles pour une onde électromagnétique de vecteur d'onde $\vec{K}$ se traduisent, pour les photons correspondants d'impulsion $\vec{p} = \hbar \vec{K}$, par *deux états de spin possibles.* Cependant, *attention :* une étude plus détaillée du champ électromagnétique et des photons associés montre que le spin de ces derniers est

$$s = 1 , \tag{III.9}$$

et non pas 1/2 comme pourrait le faire croire l'existence de seulement deux états indépendants; c'est que, pour une particule de masse nulle, on peut montrer de façon très générale qu'il existe seulement, quelle que soit la valeur du spin s, deux états de spin possibles et non pas $(2s+1)$[69].

b. Densité d'états individuels dans un gaz de photons

Un mode propre classique, caractérisé par son vecteur d'onde $\vec{K}$ et sa polarisation, *apparaît donc comme un état individuel possible pour les photons* enfermés dans l'enceinte considérée[70] : un photon se trouvant dans cet état a pour impulsion $\vec{p} = \hbar \vec{K}$ et pour énergie

$$\varepsilon = \hbar \omega = \hbar |\vec{K}\,| \, c , \tag{III.10}$$

son état de spin étant déterminé par la polarisation du mode propre (note 69).

---

68. L'énergie d'une onde de pulsation ω est alors $N_\omega \hbar \omega$, où N_ω est le nombre de photons correspondant.

69. Pour une particule de masse nulle, l'impulsion $\vec{p}$ ne s'annule jamais (puisque la vitesse d'une telle particule a pour module c dans tous les référentiels d'inertie). On peut donc toujours la choisir comme axe de quantification. Les deux états possibles sont alors caractérisés par les valeurs $+s\hbar$ et $-s\hbar$ pour la projection du spin le long de cet axe. Dans le cas du photon, ces deux états correspondent aux polarisations circulaires droite et gauche, respectivement, de l'onde électromagnétique associée.

70. Il faut toutefois exclure le « mode » classique caractérisé par $\vec{K} = \vec{0}$ (cf. note 65), car une particule de masse nulle telle que le photon ne peut pas avoir une impulsion nulle (c'est-à-dire se trouver au repos).

A la densité spectrale de modes propres (III.4) *est ainsi associée une densité d'états individuels pour les photons*[71]

$$\rho(\varepsilon)=\frac{V}{\pi^2\hbar^3c^3}\,\varepsilon^2\ , \tag{III.11}$$

telle que, comme toujours, le nombre d'états individuels dont l'énergie est comprise entre ε et $\varepsilon+\mathrm{d}\varepsilon$ soit égal à $\rho(\varepsilon)\,\mathrm{d}\varepsilon$.

3. PROPRIÉTÉS STATISTIQUES DES PHOTONS

a. Caractéristiques particulières d'un gaz de photons

Puisque leur spin est entier (cf. (III.9)), *les photons sont des bosons.*

De plus, les photons n'ont *rigoureusement aucune interaction mutuelle*[72], et constituent donc un gaz parfait. En effet, raisonnons d'abord dans le cadre ondulatoire classique. Les équations de Maxwell dans le vide étant parfaitement linéaires, le champ électromagnétique le plus général qui peut s'établir dans l'enceinte est une superposition des divers modes possibles (§ 1.a). L'énergie de ce champ électromagnétique est exactement égale à la somme des énergies des modes propres[73]. Lorsqu'on passe ensuite à la situation quantique, l'énergie d'un mode propre de pulsation ω est forcément un nombre entier de fois l'énergie $\hbar\omega$ d'un photon associé à ce mode propre. Finalement donc, l'énergie du gaz de photons qui constitue le champ électromagnétique régnant dans l'enceinte est, sans aucune approximation, la somme des énergies des divers photons; ceux-ci sont rigoureusement indépendants[74].

Les photons possèdent en outre une propriété très particulière, dont nous n'avons pas encore rencontré d'exemple, et qui affecte profondément leur comportement statistique : *leur nombre n'est pas conservé.*

Pour préciser ce qu'il faut entendre par là, analysons la manière dont le champ électromagnétique (c'est-à-dire le gaz de photons) s'établit à l'intérieur de l'enceinte d'un corps noir et parvient à l'équilibre thermique. Les parois étant maintenues à la température T, l'agitation thermique met en mouvement les particules chargées (électrons, essentiellement) qu'elles contiennent. Ces mouvements désordonnés créent (classiquement) des champs électromagnétiques aléatoires, qui se propagent librement à l'intérieur de l'enceinte et induisent à leur tour des mouvements de charges dans les

---

71. Il suffit d'effectuer, sur l'expression de $\tilde{\rho}(\omega)\mathrm{d}\omega$ donnée par (III.4), le changement de variable $\varepsilon=\hbar\omega$. On trouve ainsi

$$\tilde{\rho}(\omega)\,\mathrm{d}\omega=\frac{V}{\pi^2c^3}\,\frac{\varepsilon^2}{\hbar^2}\,\frac{\mathrm{d}\varepsilon}{\hbar}\ ,$$

d'où l'on tire aussitôt (III.11).

72. Les photons du système n'interagissent pas entre eux, mais ils interagissent avec les parois de l'enceinte. C'est grâce à cette interaction que le système « se thermalise », c'est-à-dire se met en équilibre thermique avec le thermostat qui maintient l'enceinte à une température T fixe.

73. Voir L. Landau et E. Lifchitz, *Théorie du Champ,* Mir, Moscou (1966), paragraphe 52.

74. Cette affirmation doit en réalité être légèrement nuancée. En effet, lorsqu'on quantifie le champ électromagnétique en tenant compte de son couplage avec les particules chargées (électrodynamique quantique), s'introduit la possibilité que deux photons interagissent par émission et réabsorption quasi-immédiate d'une paire virtuelle électron-positron, par exemple. Cet effet est cependant extrêmement faible : par rapport au couplage principal entre matière et rayonnement (responsable de la thermalisation des photons au contact des parois de l'enceinte), qui est déjà faible, il comporte un facteur α^2 supplémentaire, où $\alpha=1/137$ est la constante de structure fine. L'interaction entre photons est donc totalement négligeable dans l'étude du corps noir.

parois. Un couplage, relativement faible pour une enceinte de taille macroscopique, existe ainsi entre le gaz de photons constituant le champ électromagnétique et la matière constituant les parois[75]. C'est à travers ce couplage que les photons enfermés dans le corps noir finissent par atteindre une situation d'équilibre (cf. § B.1.b).

Le couplage qui vient d'être décrit est, du point de vue énergétique, du même type que ceux que nous avons rencontrés jusqu'ici : les parois du corps noir, maintenues à la température T, font partie du thermostat avec lequel le gaz de photons échange de l'énergie; comme toujours, *l'énergie totale est conservée* dans ces échanges. Mais l'interaction entre matière et rayonnement (quantifié) se fait par *absorption ou émission de photons,* de sorte que *leur nombre total ne reste pas constant. Attention :* ce n'est plus une situation où le nombre de particules constituant le gaz étudié peut fluctuer par échanges avec un réservoir, le système global gaz + réservoir gardant un nombre total de particules invariable, mais bien un *problème physique d'un caractère radicalement différent,* dans lequel le nombre total de particules n'est plus conservé.

b. Fonction de distribution des photons

Pour résoudre le problème ainsi posé, nous considérons comme toujours le *système isolé* constitué par la réunion du gaz de photons et du thermostat, et nous lui appliquons le *postulat fondamental* selon lequel tous ses états accessibles sont également probables. Ces états accessibles sont caractérisés par l'énergie totale E_{tot}, et d'autres paramètres extérieurs tels que le volume; mais *le nombre total de photons ne figure pas parmi les paramètres extérieurs,* car il n'est pas fixé bien que le système global soit isolé.

Il suffit alors de reprendre pas à pas le raisonnement du chapitre V, paragraphe I.B.1, en éliminant à chaque étape les conséquences de la conservation du nombre de particules[76]. La probabilité P_ℓ^P pour que le gaz de photons se trouve dans l'état microscopique (ℓ) (d'énergie E_ℓ et nombre de photons N_ℓ) est proportionnelle au nombre d'états accessibles du thermostat caractérisés par l'énergie $E_{\text{tot}} - E_\ell$, mais ayant par ailleurs un nombre quelconque de photons[77] :

$$P_\ell^P \propto \Omega_{\mathcal{R}} (E_{\mathcal{R}} = E_{\text{tot}} - E_\ell, \ldots) . \tag{III.12}$$

---

75. Les parois étant constituées d'atomes, on pourrait penser que les raies caractéristiques de ces atomes vont apparaître de façon privilégiée dans le rayonnement du corps noir, qui dépendrait ainsi de la nature des parois. Il n'en est rien en réalité. A l'équilibre, le nombre de photons d'une fréquence donnée est déterminé seulement par la température du thermostat (voir paragraphe b ci-dessous). En effet, la structure des parois est toujours suffisamment riche pour que toutes les fréquences soient couplées au thermostat (les différences dans l'intensité du couplage étant sans importance à l'équilibre) : déjà, la juxtaposition d'atomes dans un réseau cristallin remplace les niveaux discrets des électrons par des bandes d'énergie pratiquement continues (complément I.F, § II); mais il y a aussi les impuretés, les défauts, les vibrations cristallines, les effets de surface, ...

76. Le nombre de particules du système étudié (le gaz de photons) n'étant pas fixé, c'est à la situation grand-canonique que l'on peut le plus directement se référer pour une comparaison. Mais répétons qu'il s'agit ici d'un problème nouveau, dont les données ne s'identifient pas à celles de la situation grand-canonique puisqu'il n'y a pas à proprement parler de réservoir de photons, seulement un réservoir d'énergie. Le lien exact avec le formalisme grand-canonique sera précisé plus loin (formule (III.17) et note 78).

77. Le nombre d'états $\Omega_{\mathcal{R}}(E_{\mathcal{R}}, \ldots)$ est par conséquent obtenu de la façon suivante : l'énergie $E_{\mathcal{R}}$ étant fixée (à $\delta E_{\mathcal{R}}$ près), on calcule le nombre d'états du thermostat ayant cette énergie et un nombre de photons égal à $N_{\mathcal{R}}^P$, puis *on somme sur tous les nombres* $N_{\mathcal{R}}^P$ *possibles* (à $E_{\mathcal{R}}$ constante). Il n'y a plus ainsi, dans la fonction $\Omega_{\mathcal{R}}$, de variable correspondant au nombre de photons.

On introduit comme toujours l'entropie microcanonique du thermostat :

$$S^*_{\mathcal{R}}(E_{\mathcal{R}}, \ldots) \widehat{=} k \ln \Omega_{\mathcal{R}}(E_{\mathcal{R}}, \ldots) , \tag{III.13}$$

puis on développe $S^*_{\mathcal{R}}(E_{\mathcal{R}} = E_{\text{tot}} - E_\ell, \ldots)$ autour de $E_{\mathcal{R}} = E_{\text{tot}}$, en se limitant au premier ordre en E_ℓ (les termes d'ordre supérieur sont négligeables par définition du thermostat) :

$$S^*_{\mathcal{R}}(E_{\mathcal{R}} = E_{\text{tot}} - E_\ell, \ldots) = S^*_{\mathcal{R}}(E_{\mathcal{R}} = E_{\text{tot}}) - \frac{1}{T^*_{\mathcal{R}}} E_\ell . \tag{III.14}$$

Comme la température T imposée au gaz de photons est par définition

$$T \widehat{=} T^*_{\mathcal{R}} , \tag{II.15}$$

on trouve finalement

$$P^P_\ell \propto \mathrm{e}^{-E_\ell / kT} : \tag{III.16}$$

la probabilité P^P_ℓ d'un état microscopique (ℓ) du gaz de photons ne dépend pas du nombre N_ℓ de photons correspondant à cet état, mais seulement de son énergie E_ℓ.

On peut exprimer ce résultat de façon simple si l'on considère (III.16) comme un cas particulier de la distribution grand-canonique V(I.15) : on voit alors que *le potentiel chimique d'un gaz de photons est toujours nul* à l'équilibre[78], [79] :

$$\boxed{\mu^P \equiv 0} . \tag{III.17}$$

Cette interprétation fournit aussitôt le nombre d'occupation moyen $\overline{N}^P_\lambda$ des états individuels (λ) possibles pour chacun des photons du gaz : les photons étant des *bosons identiques, sans interaction mutuelle, de potentiel chimique constamment nul,* il suffit de prendre la distribution de Bose-Einstein V (IV.22) pour $\mu = 0$:

$$\boxed{\overline{N}^P_\lambda = \frac{1}{\mathrm{e}^{\beta \varepsilon_\lambda} - 1}} , \tag{III.18}$$

où ε_λ est l'énergie de l'état individuel (λ)[80].

Notons que *le phénomène de condensation de Bose* (§ II.B) *ne peut pas se produire pour des photons :* il n'est pas possible ici de fixer le nombre total de particules du gaz; c'est au contraire son potentiel chimique qui reste constant selon (III.17).

---

78. D'après la façon dont est défini ici le nombre d'états accessibles $\Omega_{\mathcal{R}}(E_{\mathcal{R}}, \ldots)$ du thermostat (note 77), son entropie microcanonique (III.13) ne fait pas intervenir de variable correspondant au nombre de photons. Le potentiel chimique microcanonique $\mu^*_{\mathcal{R}}$ du thermostat vis-à-vis des photons, défini comme toujours par

$$\frac{\mu^*_{\mathcal{R}}}{T^*_{\mathcal{R}}} \widehat{=} - \frac{\partial S^*_{\mathcal{R}}}{\partial N} ,$$

est par conséquent nul. Ainsi, dire que le potentiel chimique d'un système de photons est identiquement nul reste conforme à la définition générale V (I.5.b) du potentiel chimique grand-canonique. Finalement, la non conservation du nombre de photons s'exprime analytiquement sous la forme (III.17).

79. La situation microcanonique est en général impossible à réaliser pour un gaz de photons, car il faudrait supprimer ses interactions avec les parois du récipient qui l'enferme. Un tel système de photons isolé existe pourtant dans l'univers : c'est le « rayonnement fossile à $3K$ » (complément VI.G). Dans ce cas, comme il n'y a plus d'interaction avec la matière, le nombre de photons du gaz ne peut plus varier. Son potentiel chimique reste cependant nul.

80. Rappelons que, pour des photons, l'état individuel d'énergie nulle est exclu (cf. notes 65 et 70). La divergence de la formule (III.18) pour $\varepsilon_\lambda = 0$ ne crée donc pas de difficulté.

B. Lois du rayonnement thermique

Un corps noir est un système macroscopique : le nombre de photons qu'il contient est énorme. Les mesures que l'on effectue sur le rayonnement du corps noir sont donc essentiellement de nature classique. Les propriétés des gaz de photons établies au paragraphe A permettent de comprendre les résultats de ces mesures.

1. LOI DE PLANCK

La loi fondamentale, dont se déduisent toutes les autres, concerne la *densité spectrale d'énergie électromagnétique* $u(\omega; T)$ dans le corps noir : $u(\omega; T)\mathrm{d}\omega$ représente la densité d'énergie électromagnétique (par unité de volume), située, à la température T, dans la bande de pulsations $(\omega, \omega + \mathrm{d}\omega)$. Max Planck découvrit d'abord cette loi de manière empirique, avant de la démontrer à partir de la notion de *quantification*. Nous verrons effectivement (paragraphe d ci-dessous) que la densité spectrale d'énergie observée dans le corps noir est en contradiction flagrante avec les prédictions de la théorie ondulatoire classique de l'électromagnétisme.

a. Démonstration de la loi de Planck

Dans le volume macroscopique V du corps noir, les énergies individuelles possibles pour chacun des photons du gaz peuvent être considérées comme pratiquement continues (cf. formules (III.6) et (III.3)). Lorsque l'équilibre est établi à la température T du thermostat, le nombre d'occupation moyen d'un état individuel d'énergie ε est donné par (III.18); nous l'écrirons

$$N^P(\varepsilon; T) = \frac{1}{\mathrm{e}^{\varepsilon/kT} - 1} . \tag{III.19}$$

Nous supposons le système suffisamment grand pour que les fluctuations y soient négligeables *(limite thermodynamique)*. Dans ces conditions, le nombre $\mathrm{d}N(\varepsilon; T)$ de photons du gaz qui, à la température T, ont une énergie individuelle comprise entre ε et $\varepsilon + \mathrm{d}\varepsilon$ s'obtient en multipliant le nombre $N^P(\varepsilon; T)$ de photons qui occupent chacun des états individuels de cet intervalle d'énergie par le nombre $\rho(\varepsilon)\mathrm{d}\varepsilon$ de tels états $(\rho(\varepsilon)$ étant la densité d'états individuels$)$:

$$\mathrm{d}N(\varepsilon; T) = N^P(\varepsilon; T)\rho(\varepsilon)\mathrm{d}\varepsilon . \tag{III.20}$$

Comme chacun d'eux a l'énergie ε (à $\mathrm{d}\varepsilon$ près), l'énergie totale $\mathrm{d}E(\varepsilon; T)$ de cet ensemble de photons vaut

$$\mathrm{d}E(\varepsilon; T) = \varepsilon\ \mathrm{d}N(\varepsilon; T) . \tag{III.21}$$

Or l'énergie ε d'un photon est simplement proportionnelle à la pulsation ω de l'onde électromagnétique associée (relation de Planck (III.5.b)); l'expression (III.21) représente alors une fraction de l'énergie électromagnétique existant dans le corps noir à la température T, la fraction $\mathrm{d}E(\omega; T)$ portée par les ondes dont la pulsation est comprise entre ω et $\omega + \mathrm{d}\omega$. On calcule facilement $\mathrm{d}E(\omega; T)$ à partir des formules (III.21), (III.20), (III.11) et (III.5.b); on trouve ainsi que *cette énergie est proportionnelle au volume V de l'enceinte :*

$$dE(\omega; T) = Vu(\omega; T)\mathrm{d}\omega , \tag{III.22}$$

La *densité spectrale par unité de volume* $u(\omega; T)$ ayant pour expression

$$\boxed{u(\omega; T) = \frac{\hbar}{\pi^2 c^3} \frac{\omega^3}{\mathrm{e}^{\hbar\omega/kT} - 1}} \tag{III.23}$$

Ce résultat constitue la *loi de Planck pour le rayonnement du corps noir*[81].

b. Discussion : validité de la loi

Insistons sur le fait que la loi de Planck (III.23) n'est valable que dans les cas où le rayonnement est à l'*équilibre thermique* avec un thermostat de température T.

Cette restriction est importante dans la pratique, car les phénomènes d'émission ou d'absorption de rayonnement par la matière se produisent le plus souvent dans des conditions hors d'équilibre : dans le cas par exemple de l'éclairage par une lampe électrique ou du chauffage électrique par rayonnement infrarouge, il y a transformation irréversible (et donc hors d'équilibre) d'énergie électrique en énergie de rayonnement; de même, le rayonnement solaire est produit par les réactions nucléaires qui ont lieu à l'intérieur du soleil et qui consument peu à peu sa substance; au niveau microscopique également, l'émission d'un photon par un atome excité est très souvent un retour irréversible de l'atome à son état fondamental (émission spontanée hors d'équilibre). Dans le cas du corps noir, au contraire, le rayonnement est confiné à l'intérieur d'une enceinte fermée (on laisse éventuellement une fraction négligeable de ce rayonnement s'échapper à l'extérieur pour y être soumise aux mesures) et peut ainsi parvenir à l'équilibre thermique avec les parois.

La loi de Planck, que nous avons déduite au paragraphe a de considérations théoriques, est *parfaitement vérifiée par l'expérience,* dans tout le domaine de températures accessible.

c. Courbe de répartition spectrale de l'énergie

Traçons la courbe représentative de $u(\omega; T)$ comme fonction de la pulsation ω des ondes, à température T fixée. Aux faibles fréquences, c'est-à-dire pour $\hbar\omega \ll kT$, on peut développer l'exponentielle, et on trouve facilement la *«formule de Rayleigh-Jeans»* (voir paragraphe d ci-dessous) :

$$u(\omega; T) \simeq \frac{kT}{\pi^2 c^3}\,\omega^2 \quad \text{pour } \hbar\omega \ll kT\,. \tag{III.24}$$

La courbe part donc de $\omega = 0$ comme une parabole. A haute fréquence au contraire, le terme 1 du dénominateur devient négligeable devant l'exponentielle, de sorte que

$$u(\omega; T) \simeq \frac{\hbar}{\pi^2 c^3}\,\omega^3\, e^{-\hbar\omega/kT} \quad \text{pour } \hbar\omega \gg kT\,. \tag{III.25}$$

La fonction décroît donc exponentiellement dans la partie supérieure du spectre (*«loi de Wien»,* découverte empiriquement en 1893). On calcule enfin facilement la pulsation ω_m qui rend $u(\omega)$ maximale[82] :

$$\omega_m = 2{,}821\,\frac{kT}{\hbar}\,. \tag{III.26}$$

---

81. Ce résultat est ici exact, et non pas approché comme pour les fermions et les bosons, car les interactions entre photons sont nulles (cf. notes 73 et 74).

82. Pour T fixée, le maximum de $u(\omega; T)$ s'obtient en calculant par exemple sa dérivée logarithmique :

$$\frac{1}{u}\frac{\partial u}{\partial \omega} = \frac{3}{\omega} - \frac{\hbar}{kT}\frac{e^{\hbar\omega/kT}}{e^{\hbar\omega/kT}-1} = \frac{3}{\omega} - \frac{\hbar}{kT}\frac{1}{1-e^{-\hbar\omega/kT}}\,. \tag{1}$$

La valeur ω_m de la pulsation qui annule cette dérivée est donc donnée par l'équation

$$e^{-x_m} = 1 - \frac{x_m}{3}\,, \tag{2}$$

où l'on a posé

$$x_m = \frac{\hbar\omega_m}{kT}\,. \tag{3}$$

.../...

La courbe a alors l'allure représentée sur la figure 13.

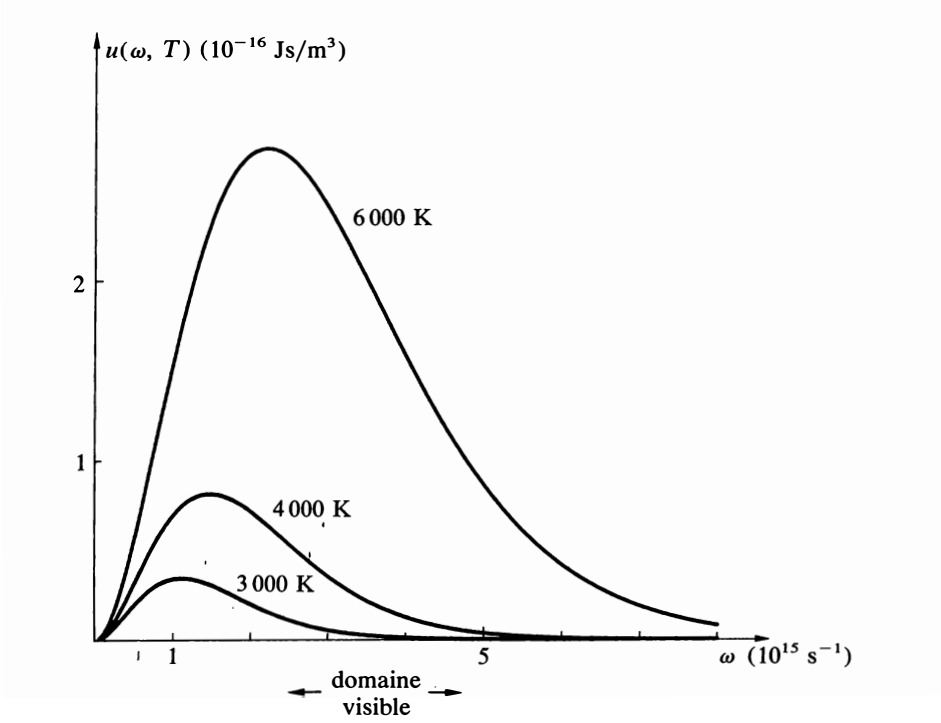


FIGURE 13

Densité spectrale d'énergie (par unité de volume) dans le corps noir, pour trois températures différentes. Outre le déplacement du maximum vers les hautes fréquences, on constate, lorsque T croît, une augmentation rapide de l'énergie totale, représentée par l'aire sous la courbe (l'énergie totale est proportionnelle à T^4. Voir formule (III.35)).

La forme générale de cette courbe est la même quelle que soit la température T, mais ses caractéristiques varient avec T. C'est ainsi que (formule (III.26)) *la position ω_m du maximum se déplace proportionnellement à T;* ce résultat, initialement déduit de l'étude empirique des courbes expérimentales, porte le nom de *« loi du déplacement de Wien »*. Il est facile de calculer par exemple[83] que le maximum de la

L'équation (2) admet une seule racine positive (outre $x_m = 0$), que l'on peut déterminer numériquement ou graphiquement :

$$x_m = 2{,}821\,44\ldots\ . \tag{4}$$

D'où le résultat (III.26).

83. Le domaine visible s'étend approximativement, en termes de longueurs d'onde λ, de $\lambda_V = 4\,000$ Å à $\lambda_R = 8\,000$ Å. Les pulsations correspondantes sont

$$\omega_V = \frac{2\pi c}{\lambda_V} \simeq 4{,}7 \times 10^{15}\ \text{s}^{-1} \tag{1}$$

.../...

courbe atteint la limite inférieure du domaine visible pour $T \simeq 6000$ K. Il apparaît immédiatement sur l'expression (III-23) que $u(\omega; T)$ croît rapidement avec T pour ω fixée quelconque; en d'autres termes, la courbe associée à une température T_2 est tout entière située au-dessus de celle qui correspond à T_1 si $T_2 > T_1$. De façon plus précise, nous montrerons plus loin que *l'aire située sous la courbe varie comme* T^4 (formule (III-35)).

d. Limite de basse fréquence et formule classique

Dans la *théorie ondulatoire classique* de l'électromagnétisme, chacun des modes propres de la cavité se comporte, du point de vue énergétique et statistique, comme un oscillateur harmonique à une dimension de même pulsation[84]. Or on sait, grâce au *théorème d'équipartition de l'énergie* (III, § V.C), lui aussi valable dans le cadre classique, que l'énergie moyenne d'un tel oscillateur harmonique est égale à kT. La physique classique prédit donc une densité spectrale d'énergie égale au produit de kT par la densité de modes propres :

$$u^{\mathrm{cl}}(\omega; T) = kT \frac{1}{V} \tilde{\rho}(\omega) . \tag{III.27}$$

---

et $$\omega_R = \frac{2\pi c}{\lambda_R} \simeq 2{,}4 \times 10^{15}\ \mathrm{s}^{-1} . \tag{2}$$

On a donc

$$\omega_m \simeq \omega_R \quad \text{pour } T \simeq \frac{\hbar \omega_R}{2{,}821 k} \simeq 6\,300\ \mathrm{K} . \tag{3}$$

De même, le maximum se situe au milieu du spectre visible ($\lambda \simeq 6000$ Å) pour $T \simeq 8500$ K.

84. A l'intérieur de la cavité, le champ électrique $\vec{\mathcal{E}}(\vec{r}, t)$ est solution de l'équation de propagation

$$\Delta \vec{\mathcal{E}}(\vec{r}, t) - \frac{1}{c^2} \frac{\partial^2}{\partial t^2} \vec{\mathcal{E}}(\vec{r}, t) = \vec{0} . \tag{1}$$

Il doit en outre vérifier les conditions aux limites imposées par les parois. Un mode propre $\{\vec{K}, \sigma\}$ est caractérisé par une fonction des coordonnées d'espace $\vec{\mathcal{E}}_{\vec{K}, \sigma}(\vec{r})$ qui vérifie ces mêmes conditions aux limites et satisfait à l'équation

$$\Delta \vec{\mathcal{E}}_{\vec{K}, \sigma}(\vec{r}) + \vec{K}^2 \vec{\mathcal{E}}_{\vec{K}, \sigma}(\vec{r}) = \vec{0} . \tag{2}$$

On peut montrer que le champ $\vec{\mathcal{E}}(\vec{r}, t)$ le plus général vérifiant les conditions aux limites peut être développé sur l'ensemble des fonctions $\vec{\mathcal{E}}_{\vec{K}, \sigma}(\vec{r})$:

$$\vec{\mathcal{E}}(\vec{r}, t) = \sum_{\vec{K}, \sigma} q_{\vec{K}, \sigma}(t) \vec{\mathcal{E}}_{\vec{K}, \sigma}(\vec{r}) . \tag{3}$$

Si l'on reporte cette expression dans l'équation (1), on voit (les modes propres étant indépendants les uns des autres) que les coefficients $q_{\vec{K}, \sigma}(t)$ doivent être solution de

$$\frac{\mathrm{d}^2}{\mathrm{d}t^2} q_{\vec{K}, \sigma}(t) + c^2 \vec{K}^2 q_{\vec{K}, \sigma}(t) = 0 . \tag{4}$$

On reconnaît l'équation caractéristique d'un *oscillateur harmonique à une dimension,* de pulsation $\omega = c|\vec{K}|$ (voir aussi la référence de la note 73).

A chaque mode propre est ainsi associé un oscillateur harmonique à une dimension de même pulsation que lui. Ces divers oscillateurs sont *découplés* les uns des autres. Le corps noir devient donc équivalent à une assemblée d'oscillateurs harmoniques à une dimension, non couplés. La quantification des modes propres maintient l'analogie : l'énergie d'un mode propre de pulsation ω ne peut valoir que $N\hbar\omega$, avec N entier, comme celle d'un oscillateur harmonique quantique à une dimension de même pulsation (à condition toutefois d'éliminer, par le processus dit de « renormalisation », l'« énergie de point zéro » $\hbar\omega/2$, qui serait ici bien gênante).

En utilisant l'expression (III.4) de la densité $\tilde{\rho}(\omega)$, on obtient ainsi la *« formule classique de Rayleigh-Jeans »*

$$u^{\mathrm{cl}}(\omega;\, T)=\frac{kT}{\pi^2 c^3}\,\omega^2\,. \tag{III.28}$$

Cette formule donne pour la densité spectrale d'énergie une fonction constamment croissante de la pulsation ω. Elle ne pouvait donc pas, lorsqu'elle constituait la seule prédiction théorique (avant la découverte de la formule de Planck et de la quantification de l'énergie), reproduire même de façon approximative les résultats expérimentaux, qui montraient une décroissance aux hautes fréquences. Elle souffrait d'ailleurs d'un défaut majeur : si l'on cherche à calculer la densité volumique totale d'énergie électromagnétique en intégrant l'expression (III.28) sur toutes les pulsations possibles, c'est-à-dire de 0 à $+\infty$, on trouve une intégrale divergente, ce qui est bien sûr un non-sens du point de vue physique (la densité volumique d'énergie est nécessairement finie. Voir formule (III.35) ci-dessous)[85].

Mais, maintenant que nous connaissons la formule quantique de Planck, nous devons nous attendre à ce qu'elle coïncide avec la formule classique (III.28) dans le domaine où la quantification de l'énergie des modes propres devient inappréciable, c'est-à-dire lorsque le quantum $\hbar\omega$ est très petit devant l'énergie caractéristique kT. Effectivement, si l'on compare (III.24) et (III.28), on voit que

$$u(\omega;\, T)\simeq u^{\mathrm{cl}}(\omega;\, T) \quad \text{pour } \hbar\omega \ll kT\,. \tag{III.29}$$

La loi de Rayleigh-Jeans apparaît donc comme une loi limite, valable dans le domaine des basses fréquences.

2. PROPRIÉTÉS DU RAYONNEMENT THERMIQUE

Les propriétés thermodynamiques du rayonnement électromagnétique contenu dans l'enceinte du corps noir se déduisent simplement de la formule de Planck (III.23).

a. Énergie totale : loi de Stefan-Boltzmann

L'énergie électromagnétique totale $E(T, V)$ présente dans l'enceinte de volume V lorsque celle-ci est maintenue à la température T peut être calculée à partir de (III.22) :

$$E(T,\, V)=V\int_0^{+\infty} u(\omega;\, T)\,\mathrm{d}\omega\,. \tag{III.30}$$

L'énergie totale est bien sûr proportionnelle au volume V, puisque la densité spectrale l'est pour toutes les valeurs de ω :

$$E(T,\, V)=V u_{\mathrm{tot}}(T)\,, \tag{III.31}$$

---

85. Le résultat classique (III.28) ne contient bien sûr pas la constante de Planck $\hbar$, qui n'était pas connue lorsque Rayleigh et Jeans écrivirent la formule qui porte leur nom. Soulignons d'ailleurs que, en l'absence de la constante fondamentale $\hbar$, il était *a priori* impossible de comprendre la décroissance exponentielle des courbes expérimentales vers les hautes fréquences (« loi de Wien » (III.25)) : avec pour seule grandeur pertinente la température T et en utilisant seulement les constantes fondamentales c (vitesse de la lumière) et k (constante de Boltzmann), on ne peut pas construire une expression ω_0 ayant les dimensions d'une pulsation et donc susceptible de donner l'échelle de variation de l'exponentielle figurant dans la loi de Wien (sous la forme $\exp[-\omega/\omega_0]$). On en conclut que, dans la théorie classique, *« il manque une constante fondamentale »*, sans laquelle on ne peut espérer résoudre la difficulté créée par la divergence de l'intégrale donnant l'énergie totale.

où la *densité volumique d'énergie* $u_{\text{tot}}(T)$ s'écrit, d'après (III.23),

$$u_{\text{tot}}(T)=\frac{\hbar}{\pi^2 c^3}\int_0^{+\infty} d\omega\,\frac{\omega^3}{e^{\hbar\omega/kT}-1}\,. \tag{III.32}$$

L'intégrale à calculer est ici parfaitement convergente. Pour déterminer sa dépendance par rapport à la température T, il suffit d'introduire la *variable d'intégration sans dimension*

$$x=\frac{\hbar\omega}{kT}\,. \tag{III.33}$$

La formule (III.32) prend alors la forme

$$u_{\text{tot}}(T)=\frac{\hbar}{\pi^2 c^3}\left(\frac{kT}{\hbar}\right)^4\int_0^{+\infty} dx\,\frac{x^3}{e^x-1}\,, \tag{III.34}$$

où l'on voit que *l'énergie totale est proportionnelle à la quatrième puissance de la température.* Ce résultat, qui avait été déduit de l'étude empirique des données expérimentales, est connu sous le nom de *«loi de Stefan-Boltzmann»;* il apparaît comme une simple conséquence de la loi de Planck qui permet en outre de trouver le coefficient de proportionnalité[86] : la densité volumique totale d'énergie vaut

$$u_{\text{tot}}=\frac{\pi^2}{15}\frac{k^4}{\hbar^3 c^3}\,T^4\,, \tag{III.35}$$

et l'énergie totale

$$E(T,\,V)=\frac{\pi^2}{15}\frac{k^4}{\hbar^3 c^3}\,VT^4\,. \tag{III.36}$$

b. Grand potentiel

Les photons étant des bosons, leur grand potentiel J se calcule à partir de la formule (II.6) :

$$J=-kT\sum_{(\lambda)}\ln\left[1+\overline{N}^P_\lambda\right], \tag{III.37}$$

que l'on peut transformer en une intégrale sur l'énergie individuelle ε :

$$\begin{aligned}J&=-kT\int_0^{+\infty} d\varepsilon\;\rho(\varepsilon)\ln\left[1+N^P(\varepsilon;\,T)\right]\\&=-kT\frac{V}{\pi^2\hbar^3 c^3}\int_0^{+\infty} d\varepsilon\;\varepsilon^2\ln\left[1+N^P(\varepsilon;\,T)\right]\end{aligned} \tag{III.38}$$

d'après (III.11). Comme au paragraphe II.A.2.c, une intégration par parties permet de mettre cette expression sous la forme (cf. note 39)

$$J=-\frac{V}{\pi^2\hbar^3 c^3}\int_0^{+\infty} d\varepsilon\,\frac{\varepsilon^3}{3}\,N^P(\varepsilon;\,T)\,. \tag{III.39}$$

---

86. Nous avons déjà rencontré des intégrales définies de forme analogue à celle qui figure dans (III.34) (note 36) :

$$\int_0^{+\infty}\frac{x^3\,dx}{e^x-1}=\Gamma(4)\zeta(4)=\frac{\pi^4}{15}\,.$$

On reconnaît alors, à un facteur près, la formule donnant l'énergie totale du gaz de photons (qui est évidemment, compte tenu de (III.5.b), identique à (III.31)-(III.32))[87] :

$$J = -\frac{1}{3} E(T, V) , \tag{III.40}$$

où l'énergie E est donnée par (III.36).

c. Pression, entropie, énergie libre

La *pression* p qui règne à l'intérieur de l'enceinte se déduit immédiatement du grand potentiel J :

$$p = -\left(\frac{\partial J}{\partial V}\right)_T \tag{III.41}$$

(puisque le potentiel chimique reste de toute façon constant : $\mu = 0$), ce qui donne explicitement[88]

$$p = \frac{1}{3}\frac{\pi^2}{15}\frac{k^4}{\hbar^3 c^3} T^4 = \frac{1}{3} u_{\text{tot}}(T) . \tag{III.42}$$

Dans le cadre classique, p est la *pression de radiation* associée au système d'ondes électromagnétiques établies dans la cavité. Au niveau microscopique, on peut la considérer comme due aux *chocs des photons sur les parois*[89].

On calcule de même l'*entropie* S du gaz de photons, c'est-à-dire du rayonnement présent dans le corps noir, par la formule habituelle

$$S = -\left(\frac{\partial J}{\partial T}\right)_V . \tag{III.43}$$

On obtient sans difficulté, à partir de (III.40) et (III.36),

$$S = \frac{4}{3}\frac{\pi^2}{15}\frac{k^4}{\hbar^3 c^3} VT^3 = \frac{4}{3}\frac{E}{T} . \tag{III.44}$$

---

87. Dans le cas des fermions ou des bosons massifs, la densité d'états individuels était en $\varepsilon^{1/2}$, et le facteur de proportionnalité entre J et E était $-2/3$ (formules (I.20) et (II.25)); c'est le même mécanisme (l'intégration par parties) qui fait apparaître ici $-1/3$ à partir d'une densité d'états en ε^2.

88. On peut aussi déduire ce résultat de la relation V (II.37) $J = -pV$.

Pour un gaz de photons, le nombre des variables macroscopiques d'équilibre est réduit d'une unité par rapport à un gaz ordinaire : parmi les variables grand-canoniques $\{\mu, T, V\}$, μ reste constamment nul; de même, N disparaît de l'ensemble des variables canoniques, puisqu'il n'est pas possible de fixer le nombre de photons du système. Restent donc seulement T et V. L'équation d'état du gaz, qui devrait être *a priori* une relation entre p, T et V, ne peut pas faire intervenir V, car c'est la seule variable extensive de l'ensemble et son extensivité ne peut donc pas être compensée ni équilibrée (dans le cas d'un gaz ordinaire, la présence de 2 variables extensives, N et V, permet une telle compensation : l'équation d'état est une relation entre p, T et N/V). C'est ce qui explique que, dans (III.42), p dépend seulement de T.

89. Le raisonnement est analogue à celui qui donne, en « théorie cinétique des gaz » (complément III.C, § III), la pression d'un gaz parfait classique. Cependant, les photons ont ici une vitesse toujours égale à c en module, et leur impulsion est proportionnelle à leur énergie (formule (III.6)); ceci aboutit simplement à modifier le facteur de proportionnalité entre p et la densité d'énergie moyenne (faire le calcul à titre d'exercice).

Notons au passage que cette entropie, comme celle d'un gaz de fermions (formule (I.60)) ou d'un gaz de bosons massifs (formule (II.57)), s'annule lorsque T tend vers zéro.

Enfin, l'*énergie libre* thermodynamique F du gaz de photons, donnée par

$$F = E - TS\,, \tag{III.45}$$

est elle aussi proportionnelle à l'énergie E : les formules (III.44) et (III.40) conduisent en effet à[90]

$$F = -\frac{1}{3}\,E = J\,. \tag{III.46}$$

90. L'égalité de F et J était prévisible dès lors que le potentiel chimique est nul, puisque de façon générale (formule V (III.20))

$$J = F - \mu N\,.$$

COMPLÉMENTS DU CHAPITRE VI

VI.A. Propriétés électroniques des métaux et des semi-conducteurs	Très important du point de vue physique. Les paragraphes I et II sont faciles, le paragraphe III (semiconducteurs) plus difficile, particulièrement les paragraphes III.4 et III.5.
VI.B. Les naines blanches	Étude d'un problème astrophysique simple car bien délimité. Facile. Le résultat final est très important. Le paragraphe II (rôle des forces gravitationnelles) et le paragraphe III.3 (gaz dégénéré de fermions relativistes) traitent des sujets dont l'intérêt n'est pas limité aux naines blanches.
VI.C. Potentiel chimique et chaleur spécifique d'un gaz de bosons libres au voisinage de la température de Bose	Établit un résultat annoncé dans le chapitre. Un peu technique, mais bonne illustration de la méthode grand-canonique et les calculs ne sont pas très compliqués.
VI.D. L'hélium à basse température	Sujet d'un très grand intérêt physique, mais assez complexe. Le paragraphe I, qui décrit les propriétés très particulières de l'hélium, n'est pas difficile. Le paragraphe II.1 non plus, et la notion qu'il présente (excitations collectives dans un système de particules en interaction) est très importante. Le paragraphe II.2 détaille les calculs du modèle de Landau; il est un peu plus difficile mais reste accessible. Le paragraphe II.3 est nettement plus difficile.
VI.E. Émission et absorption de rayonnement par un atome. Coefficients d'Einstein. **VI.F. Absorption et émission de rayonnement par un corps macroscopique. Loi de Kirchhoff**	Deux lois physiques importantes, déduites ici de la formule fondamentale de Planck. VI.E : facile; la méthode de raisonnement est intéressante, le résultat très important. VI.F : un peu plus difficile, car la définition des grandeurs à considérer est techniquement plus complexe; la notion de «corps noir» est présentée ici dans sa définition historique, et le lien avec la définition proposée dans le chapitre est précisé.
VI.G. Le rayonnement cosmique à trois kelvins	Facile. Décrit succinctement une découverte astronomique fondamentale et ses implications cosmologiques concernant la structure et l'histoire de l'Univers.

COMPLÉMENT VI.A

Propriétés électroniques des métaux et des semi-conducteurs

I. Métaux, isolants et semi-conducteurs
  1. Le courant électrique
  2. Les états électroniques dans les solides
  3. Les différents types de matériaux

II. Propriétés électroniques des métaux à l'équilibre
  1. Le modèle des électrons libres et indépendants
    a. Introduction du modèle
    b. Quelques ordres de grandeur
  2. Prédictions du modèle
    a. Chaleur spécifique électronique
    b. Compressibilité électronique
    c. Paramagnétisme de Pauli

III. Propriétés électroniques des semi-conducteurs à l'équilibre
  1. Les états électroniques dans un semi-conducteur
    a. États de valence et états de conduction
    b. États d'impuretés
  2. Notion de trou
  3. Densités de porteurs dans un semi-conducteur intrinsèque à l'équilibre
  4. Densités de porteurs dans un semi-conducteur extrinsèque à l'équilibre
    a. Population des bandes de valence et de conduction
    b. Population des états donneurs
    c. Population des états accepteurs
    d. Détermination du potentiel chimique
    e. Densités de porteurs aux températures intermédiaires
  5. La jonction *p-n* à l'équilibre
    a. Analyse qualitative
    b. Densités locales de porteurs en fonction du potentiel électrostatique
    c. Calcul du potentiel électrostatique
    d. Conclusion

I. MÉTAUX, ISOLANTS ET SEMI-CONDUCTEURS

1. LE COURANT ÉLECTRIQUE

Lorsqu'on impose une différence de potentiel entre deux de leurs points, la plupart des matériaux sont parcourus par un courant électrique. Ce courant correspond à un transport de charges, mises en mouvement par le champ électrique appliqué. Dans les solides, ce sont les électrons les moins liés aux noyaux atomiques qui participent au courant; ils proviennent des couches externes des atomes constituant le matériau. Sous l'action d'un champ électrique $\vec{E}$, ces « électrons de conduction » acquièrent une vitesse d'entraînement moyenne $\langle \vec{v} \rangle$; pour des champs pas trop élevés, celle-ci est proportionnelle à $\vec{E}$:

$$\langle \vec{v} \rangle = -\alpha \vec{E} \ ; \tag{A.1}$$

le coefficient de proportionnalité α, appelé *mobilité* dépend du matériau considéré, des impuretés qu'il contient et de la température.

Appelons n la densité d'électrons mobiles, c'est-à-dire leur nombre par unité de volume; comme chacun transporte la charge $(-e)$, la densité de courant s'écrit

$$\vec{j} = -ne\langle \vec{v} \rangle \ . \tag{A.2}$$

De (A.1) et (A.2), on déduit la *loi d'Ohm*

$$\vec{j} = \sigma \vec{E} \ , \tag{A.3.a}$$

où $$\sigma = ne\alpha \tag{A.3.b}$$

est la *conductivité*. Les valeurs de σ varient considérablement d'un matériau à l'autre, depuis quelques 10^8 siemens/m pour les métaux courants à basse température, jusqu'à des valeurs inférieures à 10^{-20} siemens/m pour les bons isolants.

La conductivité dépend des deux facteurs n et α. La mobilité α ne peut être analysée que dans le cadre d'une mécanique statistique des systèmes hors d'équilibre; nous ne la discuterons donc pas ici. La densité n des électrons de conduction, au contraire, est une propriété du matériau à l'équilibre; c'est ce dernier facteur, et plus particulièrement sa dépendance avec la température, que nous étudierons essentiellement dans ce complément.

2. LES ÉTATS ÉLECTRONIQUES DANS LES SOLIDES

Comme nous l'avons vu au complément I.F, les états électroniques dans les solides sont groupés en *bandes d'énergie permises.* Deux bandes permises sont séparées par une *bande interdite* que l'on désigne couramment par le mot anglais *«gap».* Un état électronique est repéré par un vecteur d'onde $\vec{k}$ et le numéro i de la bande à laquelle il appartient; on note $\varepsilon_i(\vec{k})$ l'énergie de cet état. Les électrons étant des fermions, à température nulle les états électroniques d'énergie inférieure à l'énergie de Fermi μ_0 sont tous occupés, ceux d'énergie supérieure sont tous vides (cf. VI, § I.B.1). La dernière bande qui contient des états occupés (par les électrons provenant des couches externes des atomes constitutifs) s'appelle la *bande de valence.*

Un électron qui se trouve dans l'état $(i, \vec{k})$, appelé «onde de Bloch», possède une vitesse moyenne[1]

$$\vec{v}_{\vec{k}}^{\,i} = \frac{1}{\hbar} \overrightarrow{\text{grad}}_{\vec{k}}\, \varepsilon_i(\vec{k}) \,. \tag{A.4}$$

La symétrie de l'équation de Schrödinger par rapport au renversement du temps impose que les états d'une même bande qui correspondent à des vecteurs d'onde opposés ont même énergie[2] :

$$\varepsilon_i(\vec{k}) = \varepsilon_i(-\vec{k}) \,. \tag{A.5}$$

A l'équilibre, quelle que soit la température, la probabilité d'occupation de ces deux états est donc la même; les vitesses (A.4) correspondantes sont alors opposées et la vitesse moyenne $\langle \vec{v} \rangle$ est nulle.

3. LES DIFFÉRENTS TYPES DE MATÉRIAUX

Suivant les matériaux, deux cas peuvent se présenter. Dans le premier, *la bande de valence est entièrement occupée* à température nulle (fig. 1). Comme nous venons de le voir, une bande de valence pleine ne peut transporter de courant. Pour obtenir un courant, il faut créer (à l'aide d'un champ électrique) une dissymétrie entre les populations d'états de $\vec{k}$ opposés. Ici, cela n'est possible que si l'on peut amener certains électrons à «franchir le gap» qui sépare la bande de valence de la bande d'énergie immédiatement supérieure (dite alors *«bande de conduction»*). Un tel effet nécessite une forte perturbation; il ne se produit pas si le champ électrique est peu intense. *Un matériau dont la bande de valence est totalement occupée est donc isolant à température nulle.*

---

1. $\vec{v}_{\vec{k}}^{\,i}$ est la vitesse de groupe associée à une onde de pulsation $\omega_i(\vec{k}) = \varepsilon_i(\vec{k})/\hbar$.

On trouvera une démonstration du résultat (A.4) dans l'appendice E de la référence 1.

2. En effet, le renversement du temps change «l'impulsion» $\hbar\vec{k}$ d'un électron en son opposée et l'énergie doit être invariante dans cette transformation.

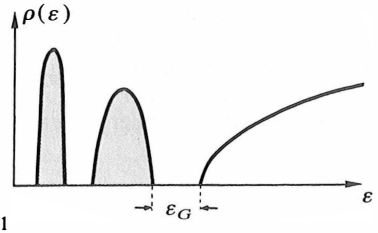


FIGURE 1
Densité d'états électroniques pour un isolant ou un semi-conducteur. A température nulle, tous les états de la bande de valence sont occupés, ceux de la bande de conduction sont tous vides.

Il se peut au contraire (deuxième cas) que *la bande de valence ne soit que partiellement remplie* à température nulle (fig. 2). A l'équilibre, un tel matériau ne transporte toujours pas de courant, mais il existe maintenant des états vides très voisins des états occupés, et l'on comprend que l'application d'un champ électrique, même faible, altère la distribution des états occupés et aboutit ainsi à une vitesse d'entraînement $\langle \vec{v} \rangle$ non nulle. La bande de valence est alors aussi la «bande de conduction»; *le matériau est conducteur, c'est un métal*[3].

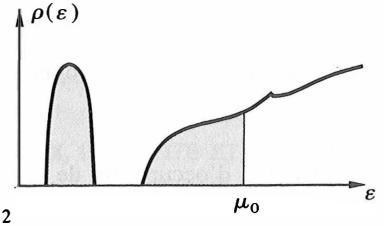


FIGURE 2
Densité d'états électroniques pour un métal. Le niveau de Fermi μ_0 est situé dans la bande de conduction. A température nulle, tous les états d'énergie inférieure à μ_0 sont occupés, ceux d'énergie supérieure sont vides.

La distinction entre conducteurs et isolants s'applique sans ambiguïté à $T=0$. A température non nulle, les choses sont un peu moins claires car la distribution d'équilibre des électrons est modifiée. Dans le cas d'un *métal,* ce sont les électrons d'énergie voisine du niveau de Fermi qui trouvent des états vides à proximité; ce sont donc eux qui participent à la conduction. La température de dégénérescence étant bien supérieure à la température ordinaire (voir paragraphe II.1.b ci-dessous), la densité n est assez peu sensible à la température et σ dépend de T principalement par la mobilité α. D'une façon générale, on constate que dans un métal σ décroît quand la température augmente.

Dans le cas d'un matériau dont la bande de valence est pleine à température nulle, certains états de la bande immédiatement supérieure, la bande de conduction, sont occupés à $T \neq 0$ (par des électrons laissant vides des états de la bande de valence), et un courant électrique est alors possible. Tout dépend en fait de la largeur de la

3. Cette distinction entre isolants et conducteurs, basée sur l'occupation de la bande de valence à température nulle, demande évidemment à être étayée sur des arguments plus solides. Pour cela, on pourra consulter les références 1 (chapitre 12) et 2 (chapitre 9).

bande interdite, « l'énergie de gap ». Si cette énergie est trop élevée (disons supérieure à 2 eV), le nombre n d'électrons excités dans la bande de conduction reste très faible aux températures usuelles et le matériau continue à se comporter comme un *isolant*. Pour des énergies de gap plus faibles (inférieures ou de l'ordre de 1 eV), les valeurs de n obtenues à température ordinaire ne sont plus négligeables et le matériau peut être parcouru par un courant électrique appréciable : on a alors affaire à un *semi-conducteur*. Évidemment, la distinction entre isolant et semi-conducteur n'est pas parfaitement tranchée (contrairement à celle qui sépare les métaux d'une part, les isolants ou semi-conducteurs d'autre part), et toutes les situations intermédiaires sont possibles. Dans un semi-conducteur, n augmente très vite (par un facteur exponentiel de Boltzmann) avec la température et σ est une fonction rapidement croissante de T.

Sauf à monter à de très hautes températures (auxquelles, le plus souvent, le matériau n'est plus à l'état solide), la distribution électronique des isolants est pratiquement gelée; nous ne nous intéresserons dans la suite qu'à la distribution électronique des métaux et des semi-conducteurs.

II. PROPRIÉTÉS ÉLECTRONIQUES DES MÉTAUX A L'ÉQUILIBRE

1. LE MODÈLE DES ÉLECTRONS LIBRES ET INDÉPENDANTS

a. Introduction du modèle

Un métal est constitué d'ions pratiquement immobiles et d'électrons délocalisés (c'est-à-dire dont la fonction d'onde s'étale sur tout le volume du cristal), l'ensemble étant électriquement neutre. L'approximation la plus simple consiste à négliger l'interaction des électrons avec les ions du réseau (électrons libres) et l'interaction des électrons entre eux (électrons indépendants). A cette approximation, les électrons de conduction forment donc un *gaz parfait de fermions* libres enfermé dans le volume V du métal. Ce modèle très simple permet de rendre compte, au moins semi-quantitativement, de la plupart des propriétés électroniques des métaux et c'est le seul que nous considérerons[4]. On peut lui appliquer directement les résultats du chapitre VI (§ 1). Nous allons préciser ici les ordres de grandeur et décrire quelques propriétés complémentaires.

b. Quelques ordres de grandeur

Dans ce modèle, le métal est caractérisé par le nombre d'électrons de conduction par unité de volume. C'est lui qui fixe toutes les autres grandeurs. Commençons donc par évaluer cette densité n d'électrons de conduction pour les métaux usuels.

Ce sont les électrons de valence de l'atome isolé qui deviennent les électrons de conduction du solide métallique. Appelons Z_v le nombre d'électrons de valence par atome, A la masse atomique de l'élément considéré et ρ la masse volumique du métal.

---

4. Il est assez surprenant que l'on puisse ainsi négliger des interactions de nature électrostatique qui ne sont pas faibles *a priori*.

L'approximation des *électrons indépendants* est du type des approximations de champ moyen : on considère que chaque électron est soumis à un potentiel moyen, créé par les ions localisés aux nœuds du réseau, mais aussi par l'ensemble des autres électrons, traité comme une distribution continue de charge; ce potentiel moyen a évidemment la périodicité du réseau. Cette approximation s'avère remarquablement bonne pour les états dont l'énergie est située dans un intervalle de largeur kT autour du niveau de Fermi. L. Landau a en effet montré que, si on se limite à ces états, on peut remplacer le système des électrons en interaction par un ensemble de « *quasi-particules* » pratiquement indépendantes (cf. compléments III.E, § V, et VI.D, § II.1.b); ce sont en fait ces quasi-particules que l'on traite ici, en continuant à les appeler, par commodité, « électrons ».

.../...

L'unité de volume du métal contient ρ/A moles, chaque mole comprend N_A atomes et chaque atome fournit Z_v électrons à la bande de conduction du solide; la densité électronique cherchée s'écrit donc

$$n = Z_v N_A \rho / A \,. \tag{A.6}$$

On obtient ainsi (tableau I) des densités de quelques 10^{22} électrons par cm^3, c'est-à-dire des valeurs environ mille fois plus grandes que pour un gaz classique dans les conditions normales. Remarquons que n varie légèrement avec la température, à cause du phénomène de dilatation thermique qui influence le facteur ρ.

TABLEAU I

Densité d'électrons de conduction et paramètres de Fermi associés, calculés dans le modèle des électrons libres et indépendants, pour quelques métaux usuels.

Élément	Z_v	n (10^{22}/cm^3)	r_s(Å)	r_s/a_0	K_F(Å^{-1})	v_F (10^6 m/s)	μ_0 (eV)	T_F (10^4 K)
Li	1	4,7	1,7	3,2	1,1	1,3	4,7	5,5
Na	1	2,6	2,1	3,9	0,9	1,1	3,2	3,8
K	1	1,4	2,6	4,9	0,7	0,9	2,1	2,5
Cu	1	8,5	1,4	2,7	1,3	1,6	7,0	8,2
Ag	1	5,8	1,6	3,0	1,2	1,4	5,5	6,4
Au	1	5,9	1,6	3,0	1,2	1,4	5,5	6,4
Mg	2	8,6	1,4	2,7	1,3	1,6	7,1	8,2
Fe	2	17,0	1,1	2,1	1,7	2,0	11,1	13,0
Al	3	18,1	1,1	2,1	1,7	2,0	11,7	13,6
Pb	4	13,2	1,2	2,3	1,6	1,8	9,5	11,0

On introduit souvent le rayon r_s de la sphère qui a pour volume $1/n$, c'est-à-dire le volume par électron de conduction :

$$r_s = \left(\frac{3}{4\pi n}\right)^{1/3} . \tag{A.7}$$

Ce rayon est toujours de l'ordre de quelques Å (deux à trois fois le rayon de Bohr a_0 pour la plupart des métaux, légèrement plus pour les métaux alcalins).

Des valeurs de n on déduit aisément, à l'aide des formules (I.30) et (I.31) du chapitre VI :

— *le vecteur d'onde de Fermi*

$$K_F = (3\pi^2 n)^{1/3} = \left(\frac{9\pi}{4}\right)^{1/3} \frac{1}{r_s} \,; \tag{A.8}$$

L'approximation des *électrons libres* consiste à remplacer en outre le réseau ionique par une distribution uniforme de charge; un électron donné se déplace alors dans un milieu où la densité de charge est nulle par compensation entre la distribution due aux ions et celle de l'ensemble des autres électrons. La distribution des ions peut être considérée comme continue si la longueur d'onde quantique des électrons est très supérieure au pas du réseau :

$$\lambda \gg a \,, \tag{1}$$

c'est-à-dire si le vecteur d'onde de Fermi K_F est très éloigné des bords de la première zone de Brillouin (complément I.F, § I). Lorsque c'est le cas, on améliore sensiblement les résultats si l'on remplace la masse de l'électron m_e par une « masse effective » m_e^* qui caractérise mieux la quasi-particule. Si la condition (1) n'est pas vérifiée, on ne peut plus négliger la structure périodique du réseau, et il faut tenir compte du fait que les fonctions d'onde électroniques ne sont pas véritablement des ondes planes mais des ondes de Bloch (complément I.F, § II).

Pour plus de détails sur ces approximations et sur les améliorations possibles, voir les chapitres 15 et 17 de la référence 1.

en l'écrivant sous la forme

$$K_F = \frac{3{,}63}{r_s/a_0}\ \text{Å}^{-1}\ , \tag{A.9}$$

on voit qu'il est de l'ordre de l'Å^{-1};

— *la vitesse de Fermi* (c'est-à-dire la vitesse d'un électron au niveau de Fermi)

$$v_F = \frac{\hbar K_F}{m_e} = \frac{4{,}20}{r_s/a_0} \times 10^6\ \text{m/s}\ ; \tag{A.10}$$

sa valeur est considérable, presque 1 % de la vitesse de la lumière (à température ordinaire, la vitesse moyenne d'une molécule d'azote n'est que de quelques centaines de m/s, celle d'une particule classique qui aurait la masse de l'électron serait de l'ordre de 10^5 m/s);

— *l'énergie de Fermi*

$$\mu_0 = \frac{\hbar^2 K_F^2}{2m_e} = \frac{\hbar^2}{2m_e a_0^2}(K_F a_0)^2\ , \tag{A.11}$$

où $\hbar^2/2m_e a_0^2$ est l'énergie de Rydberg; $K_F a_0$ étant voisin de l'unité, l'énergie de Fermi est de l'ordre du Rydberg, c'est-à-dire d'une dizaine d'eV; plus précisément, on obtient en utilisant (A.8)

$$\mu_0 = \frac{50{,}1}{(r_s/a_0)^2}\ \text{eV}\ ; \tag{A.12}$$

— *la température de Fermi*

$$T_F = \frac{\mu_0}{k} = \frac{5{,}8}{(r_s/a_0)^2} \times 10^5\ \text{K}\ , \tag{A.13}$$

de l'ordre d'au moins 10^4 K; *dans un métal, les électrons de conduction constituent un gaz de fermions dégénérés à toute température* où le métal est à l'état solide.

Le tableau I regroupe les valeurs numériques estimées à partir du modèle des électrons libres et indépendants pour quelques métaux usuels.

2. PRÉDICTIONS DU MODÈLE

a. Chaleur spécifique électronique

La capacité calorifique à volume constant d'un gaz de fermions libres a été calculée au chapitre VI (§ I.B.2.c.β). Pour $T \ll T_F$, elle est proportionnelle à la température :

$$C_V = \frac{\pi^2}{2} Nk \frac{T}{T_F}\ . \tag{A.14}$$

Ici, N est le nombre d'électrons de conduction dans le volume de métal considéré; pour une mole,

$$N = Z_v N_A\ , \tag{A.15}$$

de sorte que la *chaleur spécifique électronique* par mole de métal C_V est *proportionnelle à la température :*

$$C_V = \eta T\ , \tag{A.16.a}$$

avec

$$\eta = \frac{\pi^2}{2} \frac{RZ_v}{T_F}\ . \tag{A.16.b}$$

Cette prédiction est en bon accord avec les résultats expérimentaux; toutefois les coefficients η mesurés diffèrent parfois sensiblement des valeurs calculées à partir de (A.16.b) (tableau II), ce qui n'est guère étonnant vu la simplicité du modèle.

TABLEAU II

Coefficient du terme proportionnel à T dans la chaleur spécifique molaire de quelques métaux usuels : comparaison entre la valeur expérimentale et la valeur déduite du modèle des électrons libres et indépendants.

Élément	η calculé (mJ/mole K^2)	η expérimental (mJ/mole K^2)	η exp/η calc
Li	0,8	1,8	2,3
Na	1,1	1,5	1,3
K	1,7	2,0	1,2
Cu	0,5	0,7	1,3
Ag	0,6	0,7	1,1
Au	0,6	0,7	1,1
Mg	1,0	1,3	1,3
Fe	0,6	5,0	8,3
Al	0,9	1,3	1,4
Pb	1,5	2,9	1,9

Il ne faut cependant pas oublier que la principale contribution à la capacité calorifique du métal provient des vibrations du réseau (complément III.E). A basse température, cette contribution varie en T^3, de sorte que (formule (A.16) et formule (E.69) du complément III.E)

$$\frac{C_V^{\text{él}}}{C_V^{\text{vib}}} = \frac{5}{24\pi^2}\frac{Z_v T_D^3}{T^2 T_F}, \tag{A.17}$$

où T_D est la température de Debye du métal considéré. La capacité calorifique électronique n'est donc observable que pour des températures inférieures ou de l'ordre de

$$T_0 \simeq 0{,}145\left(\frac{Z_v T_D}{T_F}\right)^{1/2} T_D . \tag{A.18}$$

Les températures de Debye sont de quelques centaines de kelvins alors que T_F dépasse 10^4 K; T_0 est par conséquent de l'ordre de quelques centièmes de T_D, c'est-à-dire au plus quelques kelvins. On trouve par exemple $T_0 = 1{,}4$ K pour le sodium et 2,8 K pour le cuivre. La contribution électronique à la capacité calorifique ne se manifeste donc qu'à très basse température.

b. Compressibilité électronique

Le gaz électronique exerce une pression (formule VI.(I.26))

$$p = \frac{2}{3}\frac{E}{V}; \tag{A.19}$$

à température ordinaire ($T \ll T_F$), elle est pratiquement égale à la pression résiduelle de température nulle

$$p_0 = \frac{2}{3}\frac{E_0}{V} = \frac{2}{5}\frac{N}{V}\mu_0 = \frac{2}{5}\frac{\hbar^2}{2m_e}(3\pi^2)^{2/3}\left(\frac{N}{V}\right)^{5/3} . \tag{A.20}$$

A cette pression est associée une *compressibilité*

$$\varkappa \widehat{=} -\frac{1}{V}\frac{\partial V}{\partial p} = \frac{3}{2n\mu_0}, \tag{A.21}$$

et un *module de compression*

$$B \widehat{=} \frac{1}{\varkappa} = \frac{2}{3} n\mu_0 . \tag{A.22}$$

Il est intéressant de comparer la valeur déduite de (A.22) à la valeur expérimentale du module de compression (tableau III); l'ordre de grandeur est le même. Évidemment, il serait absurde d'expliquer la résistance à la compression d'un métal par la seule pression électronique : les interactions ioniques jouent aussi un rôle important. Mais le calcul précédent montre que la contribution électronique ne saurait en aucun cas être négligée dans une évaluation quantitative.

TABLEAU III

Module de compression B calculé pour des électrons libres et indépendants, comparé à la valeur expérimentale dans quelques métaux usuels.

Élément	B électrons libres (10^9 N/m^2)	B exp (10^9 N/m^2)
Li	24	11
Na	9	6
K	3	2,8
Cu	63	134
Ag	34	100
Al	225	76

c. Paramagnétisme de Pauli

Lorsqu'un métal est soumis à un champ magnétique $\vec{B}$, que nous supposerons uniforme, il acquiert une aimantation $\vec{M}$ (moment magnétique par unité de volume). On désigne sous le nom de *paramagnétisme de Pauli* la contribution des *spins des électrons de conduction* à cette aimantation[5].

Si $\mu_B \widehat{=} e\hbar/2m_e$ (positif) est le magnéton de Bohr, chaque électron contribue pour $-\mu_B$ à l'aimantation si son spin est parallèle ($\uparrow$) au champ appliqué (le signe moins provient de la charge négative $-e$ de l'électron), et pour $+\mu_B$ si son spin est antiparallèle ($\downarrow$) à $\vec{B}$. Par conséquent, si $n_\uparrow$ est la densité des électrons de spin $\uparrow$ et $n_\downarrow$ celle des électrons de spin $\downarrow$,

$$M = -\mu_B(n_\uparrow - n_\downarrow) . \tag{A.23}$$

Or

$$n_i = \frac{1}{V}\int \mathrm{d}\varepsilon \, \frac{\rho_i(\varepsilon)}{\mathrm{e}^{\beta(\varepsilon-\mu)}+1}, \quad i = \uparrow \text{ ou } \downarrow , \tag{A.24}$$

où $\rho_\uparrow(\varepsilon)$ est la densité d'états individuels des électrons de spin $\uparrow$ et $\rho_\downarrow(\varepsilon)$ celle des électrons de spin $\downarrow$. μ est le *potentiel chimique* du gaz d'électrons *en présence du champ* $\vec{B}$; il est déterminé par la condition

$$n = n_\uparrow + n_\downarrow . \tag{A.25}$$

5. Comme dans le cas d'un gaz électronique non dégénéré (complément III.A, § III.2), les électrons de conduction contribuent aussi à l'aimantation par un effet diamagnétique : c'est le *diamagnétisme de Landau* (voir exercice VI.18).

En l'absence de champ, $\rho_\uparrow$ et $\rho_\downarrow$ sont égaux chacun à la moitié de la densité électronique totale $\rho(\varepsilon)$. Le seul effet du champ magnétique est de déplacer l'énergie de chaque électron, de $+\mu_B B$ si son spin est $\uparrow$ et de $-\mu_B B$ dans le cas contraire. Le nombre d'états dont l'énergie est comprise entre ε et $\varepsilon + d\varepsilon$ en présence de $\vec{B}$ est donc égal au nombre d'états qui, en l'absence de champ, avaient une énergie comprise entre $\varepsilon - \mu_B B$ et $\varepsilon - \mu_B B + d\varepsilon$ pour les électrons de spin $\uparrow$, et entre $\varepsilon + \mu_B B$ et $\varepsilon + \mu_B B + d\varepsilon$ pour les électrons de spin $\downarrow$ (fig. 3) :

$$\rho_\uparrow(\varepsilon) = \frac{1}{2}\rho(\varepsilon - \mu_B B)\,, \tag{A.26}$$

$$\rho_\downarrow(\varepsilon) = \frac{1}{2}\rho(\varepsilon + \mu_B B)\,. \tag{A.27}$$

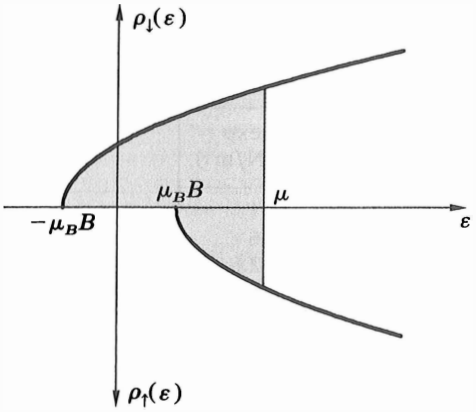


FIGURE 3
Densité d'états pour les électrons de spin $\uparrow$ et les électrons de spin $\downarrow$ en présence d'un champ magnétique $\vec{B}$.

La densité électronique $\rho(\varepsilon)$ étant connue, il suffit d'éliminer le potentiel chimique μ entre les relations (A.24) et (A.25) pour déterminer l'aimantation M en fonction de la densité électronique n.

Le calcul est particulièrement simple à *température nulle;* la fonction de Fermi est alors une fonction échelon, de sorte que

$$n_\uparrow(T=0) = \frac{1}{2V}\int_{\mu_B B}^{\mu_0} d\varepsilon\,\rho(\varepsilon - \mu_B B) = \frac{1}{2V}\int_0^{\mu_0 - \mu_B B} d\varepsilon\,\rho(\varepsilon)\,, \tag{A.28}$$

$$n_\downarrow(T=0) = \frac{1}{2V}\int_{-\mu_B B}^{\mu_0} d\varepsilon\,\rho(\varepsilon + \mu_B B) = \frac{1}{2V}\int_0^{\mu_0 + \mu_B B} d\varepsilon\,\rho(\varepsilon)\,. \tag{A.29}$$

En pratique, $\mu_B B$ est toujours très inférieur à μ_0 (pour $B = 1$ tesla, $\mu_B B \sim 10^{-4}$ eV) et donc

$$n_\uparrow(T=0) \simeq \frac{1}{2V}\int_0^{\mu_0} d\varepsilon\,\rho(\varepsilon) - \frac{\mu_B B}{2V}\rho(\mu_0)\,, \tag{A.30}$$

$$n_\downarrow(T=0) \simeq \frac{1}{2V}\int_0^{\mu_0} d\varepsilon\,\rho(\varepsilon) + \frac{\mu_B B}{2V}\rho(\mu_0)\,. \tag{A.31}$$

Ces expressions, reportées dans (A.25) et (A.23), montrent d'une part que le potentiel chimique μ_0 n'est pas déplacé (au premier ordre en $\mu_B B/\mu_0$) par la présence du champ magnétique, et que d'autre part

$$M = \frac{1}{V}\,\mu_B^2\,\rho(\mu_0)\,B\,. \tag{A.32}$$

La susceptibilité magnétique correspondante s'écrit

$$\chi_{\text{Pauli}} \widehat{=} \lim_{B\to 0}\frac{\partial M}{\partial B} = \frac{\mu_B^2}{V}\,\rho(\mu_0)\,, \tag{A.33}$$

ou encore, en remplaçant $\rho(\mu_0)$ par son expression (formule VI (I.28))

$$\rho(\mu_0) = \frac{3}{2}\,\frac{N}{\mu_0}\,, \tag{A.34}$$

$$\chi_{\text{Pauli}} = \frac{3}{2}\,\frac{n\mu_B^2}{\mu_0} = \frac{3}{2}\,\frac{n\mu_B^2}{kT_F}\,. \tag{A.35}$$

Cette expression, établie à température nulle, reste valable (à des corrections d'ordre $(T/T_F)^2$ près, cf. VI, § I.B.2) à toute température $T \ll T_F$, c'est-à-dire en pratique aux températures où le métal est à l'état solide[6]. Comme pour la capacité calorifique (VI, § I.B.2.c.β), χ est réduit d'un facteur $\sim T/T_F$ par rapport à la valeur classique $n\mu_B^2/kT$ (complément III.A, § II.3.a). Ici encore, ce facteur provient de la dégénérescence de la statistique de Fermi-Dirac : il représente le pourcentage de particules qui ne sont pas « gelées » par le principe d'exclusion et qui sont capables de répondre au champ appliqué. Comparée aux susceptibilités paramagnétiques des ions magnétiques, χ_{Pauli} est donc très faible. Le tableau IV compare les données expérimentales aux valeurs calculées à partir de (A.35) pour quelques alcalins[7].

TABLEAU IV

Susceptibilités magnétiques de Pauli pour le lithium, le sodium et le potassium. La première colonne donne la valeur calculée à partir du modèle des électrons libres et indépendants, la deuxième celle qu'on déduit des mesures expérimentales.

Élément	$\chi_{\text{Pauli}}/4\pi\varepsilon_0 c^2$ calculée	$\chi_{\text{Pauli}}/4\pi\varepsilon_0 c^2$ expérimentale
Li	$0{,}8\times10^{-6}$	$2{,}0\times10^{-6}$
Na	$0{,}7\times10^{-6}$	$1{,}1\times10^{-6}$
K	$0{,}5\times10^{-6}$	$0{,}8\times10^{-6}$

6. Cependant, en utilisant les formules générales, on peut montrer (exercice VI.17) qu'on retrouve, pour $T \gg T_F$, l'expression classique $n\mu_B^2/kT$.

7. On obtiendra rapidement des estimations de χ_{Pauli} en écrivant (A.35) sous la forme

$$\chi_{\text{Pauli}} = \left(\frac{\alpha}{2\pi}\right)^2 (a_o K_F).4\pi\varepsilon_0 c^2\,, \tag{1}$$

où $\alpha \widehat{=} e^2/4\pi\varepsilon_0\hbar c = 1/137$ est la constante de structure fine. En pratique, on utilise la grandeur sans dimension

$$\tilde{\chi}_{\text{Pauli}} \widehat{=} \frac{\chi_{\text{Pauli}}}{4\pi\varepsilon_0 c^2} = \left(\frac{\alpha}{2\pi}\right)^2 a_0 K_F \simeq \frac{2{,}6}{r_s/a_0}\times 10^{-6} \tag{2}$$

($\tilde{\chi}$ mesure directement la susceptibilité magnétique si on utilise, comme le font de nombreux ouvrages, le système d'unités C.G.S.).

III. PROPRIÉTÉS ÉLECTRONIQUES DES SEMI-CONDUCTEURS A L'ÉQUILIBRE

1. LES ÉTATS ÉLECTRONIQUES DANS UN SEMI-CONDUCTEUR

a. États de valence et états de conduction

Comme nous l'avons vu au paragraphe I, dans un semi-conducteur à $T=0$, tous les états de la bande de valence sont occupés, tous ceux de la bande de conduction sont vides. L'énergie de « gap » ε_G qui sépare le dernier état de la bande de valence du premier état de la bande de conduction est typiquement de l'ordre de l'électron-volt[8] : 1,12 eV pour le silicium, 0,67 eV pour le germanium, par exemple.

L'énergie d'un électron dans la bande de valence ou la bande de conduction est en général une fonction compliquée de son vecteur d'onde $\vec{k}$. Cependant, aux températures ordinaires ou inférieures, les électrons excités laissent vides les états du haut de la bande de valence pour venir occuper les états du bas de la bande de conduction, et les énergies $\varepsilon(\vec{k})$ de ces états peuvent être approchées par des fonctions quadratiques de $\vec{k}$ (complément I.F, § III.1). Dans les cas les plus simples, les seuls que nous considérerons ici, le maximum de la bande de valence et le minimum de la bande de conduction correspondent tous deux à $\vec{k}=\vec{0}$ (semi-conducteurs dits à « gap direct »). Nous supposerons en outre que les énergies ne dépendent que de $|\vec{k}|$; on écrira donc

$$\varepsilon_v(\vec{k}) = \varepsilon_{0v} - \frac{\hbar^2 k^2}{2m_v} \tag{A.36}$$

pour les états de valence qui seront affectés par l'excitation thermique, et de même

$$\varepsilon_c(\vec{k}) = \varepsilon_{0c} + \frac{\hbar^2 k^2}{2m_c} \tag{A.37}$$

pour les états de conduction mis en jeu (fig. 4 a); les paramètres m_v et m_c sont appelés les *« masses effectives »* des électrons dans les deux bandes[9], et l'énergie de gap est alors

$$\varepsilon_G = \varepsilon_{0c} - \varepsilon_{0v} \,. \tag{A.38}$$

b. États d'impuretés

Le plus souvent, les matériaux que l'on étudie contiennent des impuretés. Dans un semi-conducteur, celles-ci jouent un rôle très important (fig. 4 b). Si l'impureté comporte un électron de valence de plus que les atomes du semi-conducteur (par exemple une impureté de phosphore ou d'arsenic pentavalente qui remplace un atome de silicium ou de germanium quadrivalent), elle apporte un électron supplémentaire. A $T=0$, cet électron reste dans un état faiblement lié au noyau de l'impureté, d'énergie ε_d inférieure au minimum ε_{0c} de la bande de conduction mais très proche[10] (par

---

8. ε_G varie légèrement avec la température à cause de la dilatation thermique et de l'effet des vibrations du réseau sur la structure de bandes. Nous négligeons ici cette variation.

9. En général, les masses effectives sont des grandeurs tensorielles : le cristal étant anisotrope, l'énergie ε dépend de la direction du vecteur $\vec{k}$. C'est par souci de simplicité que nous nous limitons au cas particulier où m_v et m_c sont des scalaires.

10. On peut estimer l'énergie de liaison sur l'impureté (mesurée à partir du bas de la bande de conduction) en considérant qu'il s'agit en première approximation d'un atome hydrogénoïde; toutefois l'énergie d'interaction est réduite par la constante diélectrique ε_r du matériau et il faut faire intervenir la masse effective m_c à la place de la masse m_e de l'électron libre; on obtient ainsi

$$\varepsilon_{0c} - \varepsilon_d \simeq \frac{m_c}{m_e} \times \frac{1}{\varepsilon_r^2} \times 13{,}6 \text{ eV} \,.$$

.../...

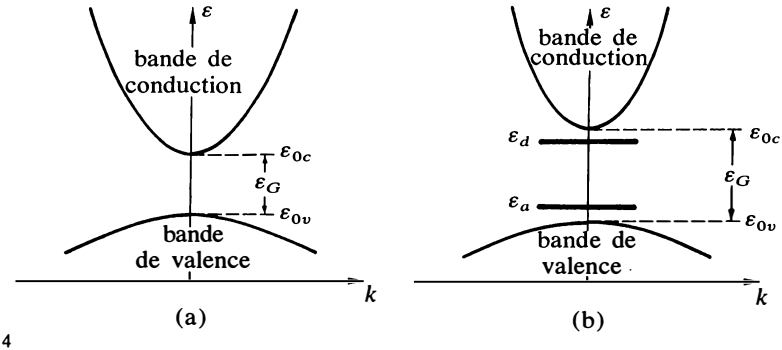


FIGURE 4

Énergies électroniques en fonction du vecteur d'onde $\vec{k}$, dans l'approximation de masse effective, pour un semi-conducteur à gap direct : a) semi-conducteur intrinsèque; b) semi-conducteur extrinsèque contenant des impuretés des deux types.

exemple $\varepsilon_{0c} - \varepsilon_d = 0{,}012$ eV pour une impureté de phosphore dans le germanium). A $T \neq 0$, l'électron de cet état peut facilement être excité dans la bande de conduction voisine; c'est pourquoi on appelle *«état donneur»* l'état lié sur une telle impureté. Nous supposerons qu'il y a N_d impuretés «donneuses» par unité de volume, et qu'à chacune d'elles correspond un état donneur pouvant accepter un seul électron.

De la même façon, des impuretés contenant un électron de valence de moins que les atomes du semi-conducteur (bore ou gallium trivalent dans du germanium ou du silicium par exemple) conduisent à des états liés d'énergie ε_a supérieure au maximum ε_{0v} de la bande de valence mais très proche ($\varepsilon_a - \varepsilon_{0v} = 0{,}06$ eV pour du gallium dans du silicium). A $T = 0$ (et en l'absence d'impuretés donneuses), ces états sont vides, mais à $T \neq 0$ ils peuvent facilement recevoir un électron excité à partir de la bande de valence : on les appelle *«états accepteurs»*. Si N_a est la densité d'impuretés «accepteuses», il y aura ainsi N_a états accepteurs par unité de volume pouvant recevoir chacun un électron.

Si le cristal contient un *excès d'états donneurs* ($N_d > N_a$), le semi-conducteur est de *«type n»* (pour «négatif»); si au contraire il contient un *excès d'états accepteurs* ($N_a > N_d$), il est *«de type p»* (pour «positif»; voir ci-dessous paragraphe 2).

Lorsque le semi-conducteur est pratiquement pur et que les impuretés jouent un rôle négligeable dans la distribution électronique, on parle de *«semi-conducteur intrinsèque»;* dans le cas contraire, il est *«extrinsèque»*.

2. NOTION DE TROU

Considérons un *état de valence,* de vecteur d'onde $\vec{k}$ et d'énergie $\varepsilon_v(\vec{k})$. Si le potentiel chimique des électrons dans le cristal est μ, le nombre moyen d'électrons occupant cet état à la température T est donné par la distribution de Fermi (formule VI (I.1)) :

$$\overline{N}_v(\vec{k}\,;\,T) = \frac{1}{e^{(\varepsilon_v(\vec{k})-\mu)/kT}+1}\,. \qquad \text{(A.39)}$$

Pour le germanium par exemple, $m_c \simeq m_e$ mais $\varepsilon_r \simeq 16$, ce qui donne

$$\varepsilon_{0c} - \varepsilon_d \simeq 0{,}05 \text{ eV} ;$$

l'énergie de liaison réelle est encore plus faible, par un facteur quatre à cinq suivant la nature de l'impureté.

Si l'on compare à la situation de température nulle où tous les états de valence sont occupés ($\overline{N}_v(\vec{k}\,;0)=1$), le nombre moyen d'*électrons «manquants»* dans cet état est

$$1-\overline{N}_v(\vec{k}\,;T)=1-\frac{1}{e^{(\varepsilon_v(\vec{k})-\mu)/kT}+1}=\frac{1}{e^{(-\varepsilon_v(\vec{k})+\mu)/kT}+1}\,. \tag{A.40}$$

Or ceci est la distribution de Fermi de particules qui auraient pour énergie $-\varepsilon_v(\vec{k})$ et pour potentel chimique $-\mu$. *On appelle «trous» ces particules fictives.* Plus exactement, on dit que dans chaque état électronique de valence, de vecteur d'onde $\vec{k}$ et d'énergie $\varepsilon_v(\vec{k})$, non occupé, il y a un trou

$$\begin{cases} \text{de charge } e'=e>0\,, \\ \text{de vecteur d'onde } \vec{k}\,'=-\vec{k}\,, \\ \text{d'énergie } \varepsilon'_v(\vec{k}\,')=-\varepsilon_v(\vec{k})\,, \\ \text{de potentiel chimique } \mu'=-\mu\,, \end{cases} \tag{A.41}$$

et on *réserve le nom d'«électrons» aux particules occupant les états de la bande de conduction.* Lorsqu'on applique un champ électrique au semi-conducteur, les trous sont mis en mouvement; ce sont en réalité les électrons de la bande de valence qui se déplacent : l'un d'eux vient occuper un trou, laissant derrière lui un trou dans son état précédent. Les trous évoluent donc en sens inverse des électrons, se comportant comme des particules chargées positivement. C'est ce que traduisent les deux premières relations de (A.41) qui ne se déduisent pas directement de (A.40) (voir référence 1, chapitre 12).

A température nulle, tous les états de conduction sont vides, tous les états de la bande de valence sont occupés : il n'y a ni trous ni «électrons». A $T\neq 0$, des «électrons» et des trous apparaissent; on appelle *«porteurs»* (sous-entendu de charge) l'ensemble des «particules» du système susceptibles de conduire le courant électrique, c'est-à-dire l'ensemble des «électrons» et des «trous» présents dans le semi-conducteur quand il se trouve dans un état excité.

3. DENSITÉS DE PORTEURS DANS UN SEMI-CONDUCTEUR INTRINSÈQUE A L'ÉQUILIBRE

Le nombre $n_c(T)$ d'électrons par unité de volume dans la bande de conduction et le nombre $p_v(T)$ de trous par unité de volume dans la bande de valence d'un semi-conducteur à la température T sont donnés par la distribution de Fermi. Si $\rho_c(\varepsilon)$ et $\rho_v(\varepsilon)$ sont les densités d'états de la bande de conduction et de la bande de valence, on a

$$n_c(T)=\frac{1}{V}\int_{\varepsilon_{0c}}^{\infty}d\varepsilon\,\rho_c(\varepsilon)\,\frac{1}{e^{(\varepsilon-\mu)/kT}+1}\,, \tag{A.42}$$

et

$$p_v(T)=\frac{1}{V}\int_{-\infty}^{\varepsilon_{0v}}d\varepsilon\,\rho_v(\varepsilon)\left[1-\frac{1}{e^{(\varepsilon-\mu)/kT}+1}\right]$$

$$=\frac{1}{V}\int_{-\infty}^{\varepsilon_{0v}}d\varepsilon\,\rho_v(\varepsilon)\,\frac{1}{e^{(\mu-\varepsilon)/kT}+1}\,. \tag{A.43}$$

Le potentiel chimique μ est déterminé par la conservation du nombre des véritables électrons. Dans un semi-conducteur intrinsèque, à chaque fois qu'un «électron» apparaît dans la bande de conduction il laisse un trou dans la bande de valence; la conservation des particules implique donc

$$n_c(T)=p_v(T)\,. \tag{A.44}$$

Cette relation détermine μ en fonction de la température; les formules (A.42) et (A.43) fixent ensuite sans ambiguïté les valeurs de n_c et p_v.

En pratique, dans tout le domaine de température utilisé,

$$kT\ll\varepsilon_G\,. \tag{A.45}$$

On peut alors montrer, et nous allons l'admettre dans un premier temps, que

$$\varepsilon_{0c} - \mu \gg kT\,,$$
$$\mu - \varepsilon_{0v} \gg kT \tag{A.46}$$

(ces conditions seront vérifiées *a posteriori* quand, grâce à elles, nous aurons calculé μ). Les facteurs de Fermi des expressions (A.42) et (A.43) peuvent alors être remplacés par des exponentielles :

$$\frac{1}{e^{(\varepsilon-\mu)/kT}+1} \simeq e^{-(\varepsilon-\mu)/kT} \quad \text{pour} \quad \varepsilon > \varepsilon_{0c}\,, \tag{A.47}$$

et

$$\frac{1}{e^{(\mu-\varepsilon)/kT}+1} \simeq e^{-(\mu-\varepsilon)/kT} \quad \text{pour} \quad \varepsilon < \varepsilon_{0v}. \tag{A.48}$$

Dans cette limite, les effets quantiques de la statistique ne se font pas sentir; on dit que le semi-conducteur est *non dégénéré*. En outre, les exponentielles (A.47) et (A.48) n'ont une valeur appréciable qu'au voisinage de ε_{0c} pour la première et de ε_{0v} pour la seconde. Dans ces domaines, les expressions (A.36) et (A.37) sont de bonnes approximations pour ε, et les densités d'états correspondantes s'écrivent (complément I.F, § III.1)

$$\rho_c(\varepsilon) = \frac{V\sqrt{2}}{\pi^2\hbar^3}\, m_c^{3/2}(\varepsilon - \varepsilon_{0c})^{1/2}\,, \tag{A.49}$$

$$\rho_v(\varepsilon) = \frac{V\sqrt{2}}{\pi^2\hbar^3}\, m_v^{3/2}(\varepsilon_{0v} - \varepsilon)^{1/2}\,. \tag{A.50}$$

En reportant (A.47), (A.48), (A.49) et (A.50) dans (A.42) et (A.43) on obtient[11]

$$n_c(T) \simeq \int_{\varepsilon_{0c}}^{\infty} d\varepsilon\, \frac{\sqrt{2}}{\pi^2\hbar^3}\, m_c^{3/2}(\varepsilon - \varepsilon_{0c})^{1/2} e^{-(\varepsilon-\mu)/kT}$$
$$= \frac{1}{4}\left(\frac{2m_c kT}{\pi\hbar^2}\right)^{3/2} e^{-(\varepsilon_{0c}-\mu)/kT}\,, \tag{A.51}$$

$$p_v(T) \simeq \int_{-\infty}^{\varepsilon_{0v}} d\varepsilon\, \frac{\sqrt{2}}{\pi^2\hbar^3}\, m_v^{3/2}(\varepsilon_{0v} - \varepsilon)^{1/2} e^{-(\mu-\varepsilon)/kT}$$
$$= \frac{1}{4}\left(\frac{2m_v kT}{\pi\hbar^2}\right)^{3/2} e^{-(\mu-\varepsilon_{0v})/kT}\,. \tag{A.52}$$

A partir de ces expressions, la relation (A.44) détermine le potentiel chimique :

$$\mu = \frac{\varepsilon_{0v} + \varepsilon_{0c}}{2} + \frac{3kT}{4} \ln \frac{m_v}{m_c} = \varepsilon_{0v} + \frac{1}{2}\,\varepsilon_G + \frac{3kT}{4} \ln \frac{m_v}{m_c}\,. \tag{A.53}$$

---

11. Le changement de variable

$$x \widehat{=} \frac{\varepsilon - \varepsilon_{0c}}{kT} \tag{1}$$

donne

$$n_c(T) = \frac{\sqrt{2}}{\pi^2\hbar^3}\,(m_c kT)^{3/2}\, e^{-(\varepsilon_{0c}-\mu)/kT} \int_0^{\infty} dx\, x^{1/2}\, e^{-x}\,. \tag{2}$$

Or

$$\int_0^{\infty} dx\, x^{1/2}\, e^{-x} = \Gamma\left(\frac{3}{2}\right) = \frac{1}{2}\sqrt{\pi}\,. \tag{3}$$

D'où le résultat (A.51); (A.52) s'obtient de façon analogue.

A température nulle, le potentiel chimique se situe exactement au milieu du gap[12]. A $T \neq 0$, μ se rapproche de la bande dont la masse effective est la plus petite; comme $\ln(m_v/m_c)$ reste toujours de l'ordre de l'unité, l'écart de μ par rapport à sa limite de température nulle est de l'ordre de kT. Puisque ε_G est de l'ordre de l'eV, dans tous les cas pratiques μ reste voisin du milieu du gap et les conditions (A.46) sont bien vérifiées; ceci justifie donc le calcul approché que nous venons de faire.

En reportant le résultat (A.53) pour μ dans (A.51) et (A.52) on obtient finalement les expressions des densités de porteurs en fonction de la température. Le plus simple est de remarquer que le produit des deux densités $n_c(T)$ et $p_v(T)$, qui sont égales, est indépendant de μ. Donc, pour un semi-conducteur intrinsèque tel que $\varepsilon_G \gg kT$, on a[13]

$$\boxed{n_c(T) = p_v(T) = \frac{1}{4}\left(\frac{2kT}{\pi\hbar^2}\right)^{3/2} (m_c m_v)^{3/4}\, e^{-\varepsilon_G/2kT}} \quad . \tag{A.54}$$

4. DENSITÉS DE PORTEURS DANS UN SEMI-CONDUCTEUR EXTRINSÈQUE A L'ÉQUILIBRE

Nous considérons maintenant un semi-conducteur qui contient des impuretés : N_d atomes donneurs et N_a atomes accepteurs par unité de volume.

a. Population des bandes de valence et de conduction

La densité n_c d'électrons dans la bande de conduction et la densité p_v de trous dans la bande de valence sont données, comme en (A.42) et (A.43), par

$$n_c(T) = \frac{1}{V}\int_{\varepsilon_{0c}}^{\infty} d\varepsilon\, \rho_c(\varepsilon)\, \frac{1}{e^{(\varepsilon-\mu)/kT}+1}\,, \tag{A.55}$$

et

$$p_v(T) = \frac{1}{V}\int_{-\infty}^{\varepsilon_{0v}} d\varepsilon\, \rho_v(\varepsilon)\, \frac{1}{e^{(\mu-\varepsilon)/kT}+1}\,. \tag{A.56}$$

Mais la relation (A.44) qui déterminait le potentiel chimique doit être modifiée pour tenir compte des populations des états d'impuretés. Calculons donc d'abord ces populations.

b. Population des états donneurs

Écrivons la grande fonction de partition Ξ_d du système des électrons liés aux N_d impuretés contenues dans un volume unité. Nous supposons les impuretés suffisamment diluées dans le cristal pour qu'on puisse négliger les interactions entre

---

12. On dit parfois que le potentiel chimique d'un système à température nulle est l'énergie qu'il faut lui fournir pour ajouter une particule. Ceci semblerait indiquer que μ est égal à ε_{0c}. Mais on peut dire aussi que c'est l'énergie qu'on lui retire lorsqu'on enlève une particule, c'est-à-dire ε_{0v}. Cette « définition » du potentiel chimique, qui ne donne pas la même valeur pour une particule de plus ou une particule de moins, n'est pas applicable ici; le résultat (A.53) montre que la bonne valeur est la moyenne arithmétique $(\varepsilon_{0v} + \varepsilon_{0c})/2$.

Remarquons aussi que, si nous nous étions placés d'emblée à température nulle, le potentiel chimique serait resté indéterminé : toute valeur comprise entre ε_{0c} et ε_{0v} donne en effet $n_c(0) = p_v(0) = 0$.

13. Pour $\varepsilon_G \simeq 0{,}7$ eV (cas du germanium) et $m_c \simeq m_v \simeq m_e$, on trouve ainsi à 300 K

$$n_c = p_v \simeq 3{,}3 \times 10^{13}\ \text{cm}^{-3}\,.$$

Les densités de porteurs sont donc beaucoup plus faibles que dans un métal (cf. paragraphe II.1.b ci-dessus).

électrons situés sur des sites différents. La grande fonction de partition Ξ_d se factorise alors sous la forme

$$\Xi_d = [\xi_d]^{N_d}, \tag{A.57}$$

où ξ_d est la grande fonction de partition associée à un site d'impureté déterminé. Chaque site peut accueillir zéro ou un électron, mais celui-ci peut se trouver dans l'un ou l'autre de ses deux états de spin (sans que son énergie ε_d en soit affectée). On a donc

$$\xi_d = 1 + 2\, e^{-(\varepsilon_d - \mu)/kT}, \tag{A.58}$$

et

$$\Xi_d = \left[1 + 2\, e^{-(\varepsilon_d - \mu)/kT}\right]^{N_d}. \tag{A.59}$$

On en déduit la densité moyenne $n_d(T)$ d'électrons qui restent piégés dans les états donneurs à la température T :

$$n_d(T) = kT \frac{\partial \ln \Xi_d}{\partial \mu} = \frac{N_d}{1 + \frac{1}{2} e^{(\varepsilon_d - \mu)/kT}}. \tag{A.60}$$

Notons la différence avec la distribution de Fermi; elle provient du fait qu'un site d'impureté ne peut pas recevoir deux électrons de spins antiparallèles[14].

c. Population des états accepteurs

La situation est symétrique pour les états accepteurs, à condition de raisonner sur les trous. En effet, une impureté trivalente dans un cristal quadrivalent fournit un électron de moins que les atomes voisins; elle porte donc une charge $+3e$ au lieu de $+4e$. Tout se passe comme si on avait ajouté une charge $-e$ au site correspondant du réseau. Cette charge, négative, repousse les électrons de son voisinage; on peut donc dire qu'elle attire les trous de la même façon qu'une impureté pentavalente attire les électrons. Un état accepteur d'énergie ε_a peut donc être considéré comme un *état lié d'énergie* $-\varepsilon_a$ *pour un trou;* chaque site d'impureté peut contenir zéro ou un trou dans l'un ou l'autre état de spin[15]. Si on se rappelle que le potentiel chimique des trous est égal à $-\mu$, on peut écrire la grande fonction de partition *pour les trous* dans les états accepteurs sous la forme

$$\Xi_a = \left[1 + 2\, e^{-(-\varepsilon_a + \mu)/kT}\right]^{N_a}. \tag{A.61}$$

La densité de trous $p_a(T)$ qui restent piégés dans les états accepteurs à la température T vaut donc

$$p_a(T) = kT \frac{\partial \ln \Xi_a}{\partial(-\mu)} = \frac{N_a}{1 + \frac{1}{2} e^{(\mu - \varepsilon_a)/kT}}. \tag{A.62}$$

d. Détermination du potentiel chimique

Le potentiel chimique est déterminé, comme toujours, par la conservation des particules. Comme les électrons et les trous ne peuvent apparaître (ou disparaître) que par paires électron-trou, la différence entre le nombre d'électrons et le nombre de

14. Si nous négligeons les interactions entre électrons, un niveau donneur pourrait aussi contenir deux électrons de spins antiparallèles. C'est la répulsion coulombienne entre ces deux électrons localisés sur un même site qui interdit pratiquement la double occupation du niveau donneur (voir exercice V.16).

15. C'est encore la répulsion coulombienne, entre les trous cette fois, qui interdit la présence de deux trous de spins antiparallèles sur la même impureté.

trous dans le cristal doit rester constante. Les N_d impuretés donneuses apportent N_d électrons par unité de volume, et les N_a impuretés accepteuses apportent N_a trous; on a donc

$$n_c(T)+n_d(T)-p_v(T)-p_a(T)=N_d-N_a\,. \tag{A.63}$$

En remplaçant les densités par leurs expressions (A.55), (A.56), (A.60) et (A.62), on obtient une équation pour $\mu(T)$:

$$\frac{1}{V}\int_{\varepsilon_{0c}}^{\infty} d\varepsilon\,\rho_c(\varepsilon)\,\frac{1}{e^{(\varepsilon-\mu)/kT}+1}-\frac{1}{V}\int_{-\infty}^{\varepsilon_{0v}} d\varepsilon\,\rho_v(\varepsilon)\,\frac{1}{e^{(\mu-\varepsilon)/kT}+1}$$
$$=\frac{N_d}{1+2\,e^{(\mu-\varepsilon_d)/kT}}-\frac{N_a}{1+2\,e^{(\varepsilon_a-\mu)/kT}}\,. \tag{A.64}$$

On reporte ensuite la solution de cette équation dans (A.55), (A.56), (A.60) et (A.62), ce qui donne les populations de porteurs libres (électrons dans la bande de conduction et trous dans la bande de valence) et de porteurs liés (dans les niveaux donneurs et accepteurs) en fonction de la température.

e. Densités de porteurs aux températures intermédiaires

Une analyse complète de l'équation (A.64) pour toutes les valeurs de T est assez compliquée et fastidieuse. Nous nous limiterons au cas, important en pratique, où le potentiel chimique μ reste éloigné des bords de bande et des niveaux d'impureté. Nous supposerons donc

$$\varepsilon_d-\mu\gg kT\,,$$
$$\mu-\varepsilon_a\gg kT\,. \tag{A.65}$$

Dans ces conditions, les formules (A.60) et (A.62) impliquent

$$n_d(T)\simeq 2N_d\,e^{-(\varepsilon_d-\mu)/kT}\ll N_d \tag{A.66}$$

et

$$p_a(T)\simeq 2N_a\,e^{-(\mu-\varepsilon_a)/kT}\ll N_a\,. \tag{A.67}$$

Il n'y a donc *pratiquement pas de porteurs liés dans les états d'impuretés.* Celles-ci sont pratiquement toutes ionisées (positivement pour les donneuses, négativement pour les accepteuses). Compte tenu de ce résultat, la condition (A.63) s'écrit

$$n_c(T)-p_v(T)\simeq N_d-N_a\,. \tag{A.68}$$

D'autre part, les approximations (A.51) et (A.52) sont encore valables et

$$n_c(T)p_v(T)=n_i^2(T)\,, \tag{A.69}$$

où

$$n_i(T)\widehat{=}\frac{1}{4}\left(\frac{2kT}{\pi\hbar^2}\right)^{3/2}(m_c m_v)^{3/4}\,e^{-\varepsilon_G/2kT} \tag{A.70}$$

est la densité de porteurs de chaque type dans le même semi-conducteur lorsqu'il est intrinsèque. La solution du système d'équations (A.68)-(A.69) est immédiate :

$$n_c(T)=\frac{1}{2}\left[(N_d-N_a)^2+4n_i^2\right]^{1/2}+\frac{1}{2}(N_d-N_a)\,, \tag{A.71}$$

$$p_v(T)=\frac{1}{2}\left[(N_d-N_a)^2+4n_i^2\right]^{1/2}-\frac{1}{2}(N_d-N_a)\,. \tag{A.72}$$

Pour discuter la validité de l'approximation (A.65), il nous faut calculer μ, ce qui se fait commodément à partir de la relation (A.68) : en divisant les deux membres par n_i et en utilisant les approximations (A.51) et (A.52), on obtient

$$\frac{\mu-\mu_i}{kT}=\mathrm{Argsh}\,\frac{N_d-N_a}{2n_i}\,, \tag{A.73}$$

où

$$\mu_i\widehat{=}\frac{\varepsilon_{0v}+\varepsilon_{0c}}{2}+\frac{3kT}{4}\ln\frac{m_v}{m_c} \tag{A.74}$$

est le potentiel chimique du même semi-conducteur lorsqu'il est intrinsèque (formule (A.53)). Comme nous l'avons vu au paragraphe 3, tant que $kT \ll \varepsilon_G$, μ_i reste proche du milieu du gap. La relation (A.73) montre qu'il en est de même pour μ, sauf pour des valeurs considérables[16] du rapport $(N_d - N_a)/n_i$, c'est-à-dire sauf pour des températures très basses (auxquelles $n_i \to 0$); l'approximation (A.65) est donc en général justifiée[17].

Les expressions de $n_c(T)$ et $p_v(T)$ se simplifient dans deux limites :

(i) A faible concentration d'impuretés :

$$|N_d - N_a| \ll n_i \,, \tag{A.75}$$

le comportement est très proche du cas intrinsèque : un développement au premier ordre de (A.71) et (A.72) donne

$$n_c(T) \simeq n_i + (N_d - N_a)/2 \,, \tag{A.76}$$

$$p_v(T) \simeq n_i - (N_d - N_a)/2 \,. \tag{A.77}$$

(ii) Si au contraire

$$|N_d - N_a| \gg n_i \,, \tag{A.78}$$

---

16. Par exemple, pour

$$\frac{N_d - N_a}{n_i} \simeq 10^4 \,,$$

on trouve que

$$\frac{\mu - \mu_i}{kT} = \text{Argsh}\, \frac{N_d - N_a}{2n_i} = \ln\left[\frac{N_d - N_a}{2n_i} + \sqrt{1 + \left(\frac{N_d - N_a}{2n_i}\right)^2}\right]$$

n'est encore que $\simeq 9{,}2$.

17. Essayons de préciser un peu ses limites de validité.

Pour $kT \ll \varepsilon_G$, μ_i reste voisin du milieu du gap et il suffit donc que

$$\left| kT\ \text{Argsh}\, \frac{N_d - N_a}{2n_i} \right| \ll \varepsilon_G \tag{1}$$

pour qu'il en soit de même pour μ, c'est-à-dire pour que la condition (A.65) soit vérifiée. Choisissons par exemple le cas $(N_d - N_a) > 0$ pour lequel le Argsh est positif (le raisonnement se transpose immédiatement au cas inverse). La condition (1) pose problème uniquement si $(N_d - N_a)/2n_i \gg 1$. Dans cette limite

$$\text{Argsh}\, \frac{N_d - N_a}{2n_i} \simeq \ln \frac{N_d - N_a}{n_i} \,; \tag{2}$$

la condition (1) est donc satisfaite si

$$\frac{N_d - N_a}{n_i} \ll e^{\varepsilon_G/kT} \,, \tag{3}$$

ou encore, en remplaçant n_i par sa définition (A.70), si

$$N_d - N_a \ll \frac{1}{4}\left(\frac{2}{\pi\hbar^2}\right)^{3/2} (m_c m_v)^{3/4} (kT)^{3/2} \,. \tag{4}$$

Le résultat (A.73) est donc valable pour des températures intermédiaires telles que

$$\frac{\pi\hbar^2}{(m_c m_v)^{1/2}} |N_d - N_a|^{2/3} \ll kT \ll \varepsilon_G \,, \tag{5}$$

ce qui suppose évidemment une densité d'impuretés non compensées $(N_d - N_a)$ pas trop importante $(|N_d - N_a| \ll (m_c m_v)^{3/4} (\varepsilon_G/\pi\hbar^2)^{3/2})$.

le comportement est dominé par les impuretés (régime extrinsèque); on a alors[18]

$$\left.\begin{aligned} n_c(T) &\simeq N_d - N_a \\ p_v(T) &\simeq \frac{n_i^2(T)}{N_d - N_a} \end{aligned}\right\} \text{ si } \quad N_d > N_a \tag{A.79}$$

c'est-à-dire pour les semi-conducteurs de type n, ou

$$\left.\begin{aligned} n_c(T) &\simeq \frac{n_i^2(T)}{N_a - N_d} \\ p_v(T) &\simeq N_a - N_d \end{aligned}\right\} \text{ si } \quad N_a > N_d \tag{A.80}$$

c'est-à-dire pour les semi-conducteurs de type p. Dans cette seconde limite, si par exemple $N_d > N_a$, tous les électrons en excès apportés par les impuretés se retrouvent dans la bande de conduction, la bande de valence ayant une densité de porteurs très faible. C'est l'inverse qui se produit si $N_a > N_d$: les porteurs libres sont presque exclusivement des trous dans la bande de valence.

Insistons sur le fait que ces résultats ne sont valables que si les conditions (A.65) sont vérifiées. Si elles ne l'étaient pas, il faudrait revenir à l'équation exacte (A.64).

5. LA JONCTION $p-n$ A L'ÉQUILIBRE

Les dispositifs électroniques à semi-conducteurs sont en général inhomogènes, c'est-à-dire que la densité d'impuretés qu'ils contiennent varie d'un point à l'autre du cristal. Le plus simple de ces dispositifs est la « jonction $p-n$ ». Elle est constituée d'un cristal semi-conducteur dans lequel la concentration d'impuretés varie selon une direction (choisie ici comme axe Ox) de la façon suivante[19] (fig. 5) :

$$N_a(x) = \begin{cases} N_a & \text{si } x < 0 \,, \\ 0 & \text{si } x > 0 \,; \end{cases}$$

$$N_d(x) = \begin{cases} 0 & \text{si } x < 0 \,, \\ N_d & \text{si } x > 0 \,. \end{cases} \tag{A.81}$$

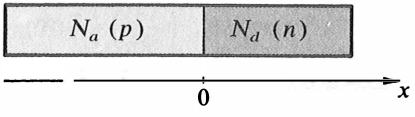


FIGURE 5

Principe de la jonction p-n. Du côté $x<0$, le semi-conducteur est de type p : il contient N_a impuretés accepteuses par unité de volume. Du côté $x>0$ il est de type n, avec N_d impuretés donneuses par unité de volume.

18. Dans les expressions (A.79) et (A.80), on n'a gardé que le premier terme non nul du développement en puissances de $n_i/(N_d - N_a)$. La neutralité électrique du matériau (condition (A.68)) n'y est donc assurée qu'à l'ordre dominant : pour un semi-conducteur de type n, par exemple, p_v est négligeable devant n_c.

19. En pratique, les concentrations d'impuretés ne changent pas de façon abrupte. L'approximation (A.81) est valable si la distance sur laquelle N_d et N_a varient reste faible devant la dimension de la « double couche » (voir ci-dessous paragraphe a).

Le semi-conducteur est donc de type p du côté $x<0$ et du type n du côté $x>0$. Nous supposerons que les valeurs de N_d, N_a et T sont telles que le régime extrinsèque (qui conduit aux formules (A.79) ou (A.80) dans un semi-conducteur homogène) est valable dans chacune des deux régions.

a. Analyse qualitative

Nous avons montré au paragraphe 4.e que, sauf si la température est trop basse, toutes les impuretés sont pratiquement ionisées : il y a N_d ions positifs par unité de volume du côté $x>0$ et N_a ions négatifs par unité de volume du côté $x<0$. Ces charges sont globalement compensées par les porteurs libres dans les bandes de conduction et de valence, mais il n'est pas possible que la neutralité électrique soit partout assurée localement. En effet, supposons que le cristal soit électriquement neutre en chaque point. Il y a alors une densité d'électrons de conduction égale à N_d du côté n et pratiquement nulle du côté p; inversement, la densité de trous de valence est égale à N_a du côté p et pratiquement nulle du côté n (relations (A.79) et (A.80)). Les électrons vont donc diffuser des régions à haute densité vers les régions à basse densité, c'est-à-dire du côté n vers le côté p, alors que les trous vont diffuser en sens inverse. Par suite de ce déplacement de charges, il apparaît un champ électrique qui s'oppose progressivement au mouvement des porteurs libres. Le phénomène se poursuit jusqu'à ce qu'un équilibre soit atteint, c'est-à-dire jusqu'à ce que l'effet du champ électrique compense exactement l'effet de diffusion, autrement dit jusqu'à ce que le potentiel chimique μ soit le même dans tout le cristal (cf. complément V.B, § II.2).

Il se constitue donc, de part et d'autre du plan $x=0$, deux couches de charge électrique non nulle, positive du côté $x>0$, négative du côté $x<0$; c'est ce qu'on appelle la *«double couche»* de charge de la jonction $p-n$. A cette distribution de charge est associé, via l'équation de Poisson, un potentiel électrostatique $\phi(x)$. Loin de la jonction, on doit retrouver le comportement d'un semi-conducteur homogène, avec $\phi(x)$ pratiquement constant; le potentiel $\phi(x)$ ne varie de façon appréciable que dans la double couche, et l'on s'attend à ce qu'il soit une fonction croissante de x[20].

b. Densités locales de porteurs en fonction du potentiel électrostatique

Nous supposons que $\phi(x)$ varie lentement à l'échelle microscopique. L'énergie de chaque état électronique est alors décalée de $-e\phi(x)$; elle varie donc avec x, c'est-à-dire d'un point à l'autre du cristal (fig. 6). Les densités locales de porteurs

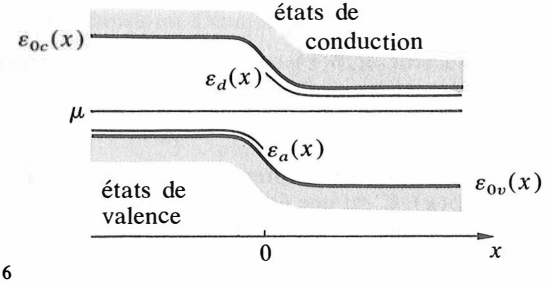


FIGURE 6

Niveaux d'énergie électroniques en fonction de la position dans une jonction *p-n*.

20. Le champ électrique $\vec{E}$ est en effet dirigé des charges positives vers les charges négatives, c'est-à-dire dans le sens des x négatifs. On a donc

$$\frac{d\varphi}{dx} = -E_x > 0 .$$

dans chacun des états sont obtenues en remplaçant simplement, dans les résultats du paragraphe 4.e pour un semi-conducteur homogène, μ par $\mu + e\phi(x)$ (cf. complément V.B, § I).

Loin de la jonction, la population des états liés aux impuretés est pratiquement nulle (formules (A.66) et (A.67)); elle l'est en réalité dans tout le cristal, puisque l'écart entre μ et les niveaux liés locaux croît encore lorsqu'on s'approche de la jonction (fig. 6); les impuretés sont donc ionisées pour toute valeur de x :

$$n_d(T; x) \simeq p_a(T; x) \simeq 0 . \tag{A.82}$$

Pour les porteurs libres, les formules (A.51) et (A.52) s'écrivent ici

$$n_c(T; x) \simeq \frac{1}{4}\left(\frac{2m_c kT}{\pi\hbar^2}\right)^{3/2} e^{-(\varepsilon_{0c} - e\varphi(x) - \mu)/kT} , \tag{A.83}$$

$$p_v(T; x) \simeq \frac{1}{4}\left(\frac{2m_v kT}{\pi\hbar^2}\right)^{3/2} e^{-(\mu + e\varphi(x) - \varepsilon_{0v})/kT} . \tag{A.84}$$

Loin de la jonction, on doit retrouver le comportement d'un semi-conducteur homogène, c'est-à-dire les résultats (A.79) et (A.80); on a donc

$$n_c(T; +\infty) \simeq \frac{1}{4}\left(\frac{2m_c kT}{\pi\hbar^2}\right)^{3/2} e^{-(\varepsilon_{0c} - e\varphi(+\infty) - \mu)/kT} \simeq N_d , \tag{A.85}$$

$$p_v(T; -\infty) \simeq \frac{1}{4}\left(\frac{2m_v kT}{\pi\hbar^2}\right)^{3/2} e^{-(\mu + e\varphi(-\infty) - \varepsilon_{0v})/kT} \simeq N_a . \tag{A.86}$$

Les relations (A.85) et (A.86) permettent d'éliminer μ des expressions (A.83) et (A.84) et d'écrire

$$n_c(T; x) \simeq N_d \, e^{-e[\varphi(+\infty) - \varphi(x)]/kT} , \tag{A.87}$$

$$p_v(T; x) \simeq N_a \, e^{-e[\varphi(x) - \varphi(-\infty)]/kT} . \tag{A.88}$$

c. Calcul du potentiel électrostatique

Il reste à calculer le potentiel $\phi(x)$ créé par la répartition inhomogène des charges dans le cristal. Il est donné par l'équation de Poisson

$$\frac{d^2\varphi}{dx^2} + \frac{1}{\varepsilon}\rho(x) = 0 , \tag{A.89}$$

où ε est la permittivité du semi-conducteur et $\rho(x)$ la densité locale de charge. A celle-ci contribuent les impuretés ionisées et les porteurs libres :

$$\rho(x) = e\left[N_d(x) - N_a(x) + p_v(T; x) - n_c(T; x)\right] . \tag{A.90}$$

On obtient ainsi, en remplaçant p_v et n_c par leurs expressions (A.87) et (A.88), une équation ne faisant plus intervenir que $\phi(x)$. C'est une équation différentielle du second ordre, non linéaire, qui doit être résolue numériquement, et dont nous n'allons donner qu'une *solution approchée.*

Calculons tout d'abord la variation totale de potentiel $\varphi(+\infty) - \phi(-\infty)$ dans la jonction. En multipliant membre à membre les égalités (A.85) et (A.86), on obtient :

$$e\left[\varphi(+\infty) - \varphi(-\infty)\right] = \varepsilon_G + kT \ln\left[\frac{16 N_d N_a}{(m_c m_v)^{3/2}}\left(\frac{\pi\hbar^2}{2kT}\right)^3\right] , \tag{A.91}$$

c'est-à-dire, en utilisant (A.70),

$$\frac{e}{kT}\left[\varphi(+\infty) - \varphi(-\infty)\right] = \ln\frac{N_d N_a}{n_i^2} . \tag{A.92}$$

D'autre part, nous avons supposé que la condition (A.78) est satisfaite pour chaque type d'impuretés, de sorte que

$$\frac{N_d N_a}{n_i^2} \gg 1 \,. \tag{A.93}$$

Il y a donc toujours des valeurs de x pour lesquelles les arguments des exponentielles de (A.87) et (A.88) sont grands et il n'est pas possible de les linéariser comme nous l'avons fait dans un problème analogue au complément III.H (§ II.2).

Pour résoudre de façon approchée l'équation de Poisson, nous allons supposer que la double couche s'étend de part et d'autre de $x=0$, entre les abscisses (à déterminer) $-d_p$ et $+d_n$. On délimite ainsi quatre régions (fig. 7).

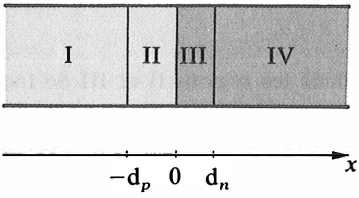


FIGURE 7

La jonction *p-n* et les quatre régions où l'on peut résoudre de façon approchée l'équation de Poisson. Les régions I et IV sont pratiquement homogènes et localement neutres. La double couche s'étend entre $-d_p$ et $+d_n$: la région II est chargée négativement, la région III positivement.

Dans la région I ($x < -d_p$), le semi-conducteur est pratiquement homogène, avec

$$\phi(x) \simeq \phi(-\infty) \quad \text{pour} \quad x < -d_p \,. \tag{A.94}$$

On en déduit, d'après (A.87) et (A.88) et compte tenu de (A.92)-(A.93),

$$\left.\begin{aligned} n_c(T;x) &\simeq 0 \\ p_v(T;x) &\simeq N_a \end{aligned}\right\} \quad \text{pour} \quad x < -d_p \,. \tag{A.95}$$

Ceci donne bien, dans cette région, une densité globale de charge ρ nulle, associée à un potentiel constant. De même dans la région IV,

$$\phi(x) \simeq \phi(+\infty) \quad \text{pour} \quad x > d_n \,, \tag{A.96}$$

ce qui s'accompagne de

$$\left.\begin{aligned} n_c(T;x) &\simeq N_d \\ p_v(T;x) &\simeq 0 \end{aligned}\right\} \quad \text{pour} \quad x > d_n \,. \tag{A.97}$$

Dans les deux régions intermédiaires II ($-d_p < x < 0$) et III ($0 < x < d_n$), au contraire, les variations de $e\phi(x)$ sont rapides (cf. (A.92) et (A.93)) et les arguments des exponentielles de (A.87) et (A.88) deviennent vite tous les deux grands et négatifs. De sorte que, sauf au voisinage immédiat des bords ($x \simeq -d_p$ et $x \simeq d_n$),

$$\left.\begin{aligned} n_c(T;x) &\simeq 0 \\ p_v(T;x) &\simeq 0 \end{aligned}\right\} \quad \text{pour} \quad -d_p < x < d_n \,. \tag{A.98}$$

L'équation de Poisson s'écrit alors

$$\frac{d^2\varphi}{dx^2} = e\frac{N_a}{\varepsilon} \quad (-d_p < x < 0) \tag{A.99}$$

dans presque toute la région II, et

$$\frac{d^2\varphi}{dx^2} = -e\frac{N_d}{\varepsilon} \quad (0 < x < d_n) \tag{A.100}$$

dans presque toute la région III. L'intégration des équations (A.99) et (A.100) est immédiate; on trouve ainsi

$$\phi(x) = \begin{cases} \phi(-\infty) & \text{pour} \quad x < -d_p \\ \phi(-\infty) + \dfrac{eN_a}{2\varepsilon}(x + d_p)^2 & \text{pour} \quad -d_p < x < 0 \\ \phi(+\infty) - \dfrac{eN_d}{2\varepsilon}(x - d_n)^2 & \text{pour} \quad 0 < x < d_n \\ \phi(+\infty) & \text{pour} \quad x > d_n \, . \end{cases} \tag{A.101}$$

Nous avons choisi les constantes d'intégration dans les régions II et III de façon que $\phi(x)$ et sa dérivée soient continus en $x = -d_p$ et $x = d_n$. Mais nous voulons aussi que $\phi(x)$ et $d\phi/dx$ soient continus en $x = 0$. Ces deux dernières conditions vont permettre de fixer les deux paramètres d_p et d_n qui étaient jusqu'ici indéterminés. La continuité de la dérivée du potentiel donne

$$N_a d_p = N_d d_n \, , \tag{A.102}$$

qui exprime simplement la neutralité électrique globale du cristal : l'accumulation de charges positives, de densité superficielle $N_d d_n$, du côté n de la jonction est exactement compensée par l'accumulation de charges négatives, de densité superficielle $N_a d_p$, du côté p. La continuité de $\phi(x)$ lui-même est assurée si

$$\frac{e}{2\varepsilon}[N_d d_n^2 + N_a d_p^2] = \varphi(+\infty) - \varphi(-\infty) \, , \tag{A.103}$$

ce qui donne, avec (A.102) et (A.92),

$$d_n = \left[\frac{N_a/N_d}{N_a + N_d} \times \frac{2\varepsilon kT}{e^2} \ln \frac{N_a N_d}{n_i^2}\right]^{1/2} , \tag{A.104}$$

$$d_p = \left[\frac{N_d/N_a}{N_a + N_d} \times \frac{2\varepsilon kT}{e^2} \ln \frac{N_a N_d}{n_i^2}\right]^{1/2} . \tag{A.105}$$

Les valeurs typiques[21] de d_n et d_p sont de plusieurs centaines d'angströms.

---

21. Par exemple, pour $\varepsilon_G \simeq 0{,}7$ eV avec $m_c \simeq m_v \simeq m_e$, on trouve à $T = 300$ K (note 13)

$$n_i \simeq 3{,}3 \times 10^{13} \text{ cm}^{-3} \, . \tag{1}$$

Pour des valeurs courantes $N_a = N_d \simeq 10^{16}$ cm^{-3}, on a donc (avec $\varepsilon = 16\varepsilon_0$)

$$d_n = d_p \simeq 1{,}6 . 10^{-7} \text{ m} \, . \tag{2}$$

La différence de potentiel d'un bord à l'autre de la double couche,

$$\Delta\varphi = \varphi(+\infty) - \varphi(-\infty) = \frac{kT}{e} \ln \frac{N_d N_a}{n_i^2} \, , \tag{3}$$

vaut alors $\simeq 0{,}3$ volt. D'où un champ électrique moyen dans la double couche

$$E = \frac{\Delta\varphi}{d_n + d_p} \simeq 10^6 \text{ V/m} \, . \tag{4}$$

Le potentiel $\phi(x)$ est maintenant complètement déterminé (quoique de façon approchée) : il suffit de reporter les expressions (A.104) et (A.105) de d_n et d_p dans les formules (A.101). On en déduit facilement la densité de charge en chaque point. Les résultats sont décrits graphiquement sur la figure 8.

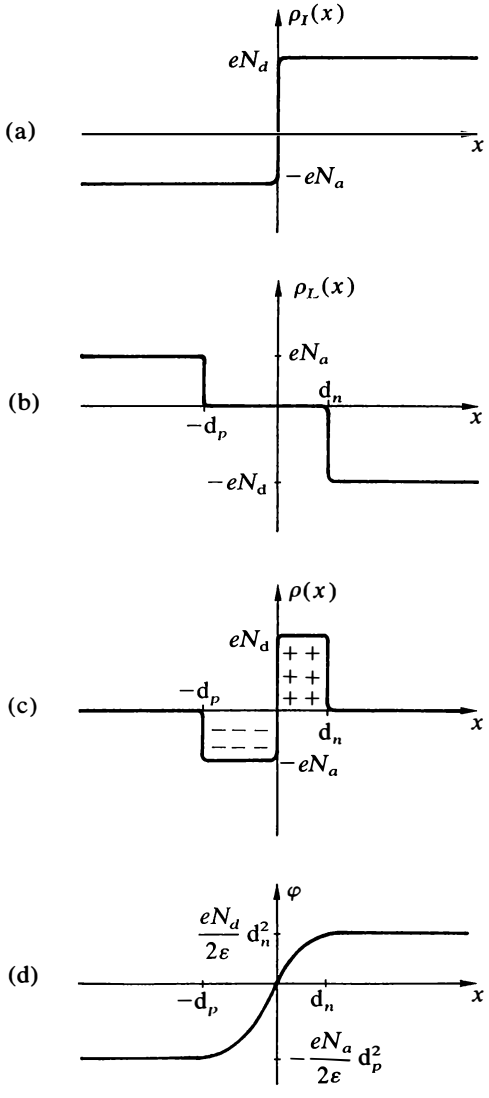


FIGURE 8

Densités de charge et potentiel dans une jonction *p-n* à l'équilibre.

a) Densité $\rho_I(x)$ de charges fixes (impuretés ionisées).

b) Densité de charge $\rho_L(x)$ des porteurs libres : $\rho_L(x) = e\,p_v(x)$ pour $x<0$, $\rho_L(x) = -e\,n_c(x)$ pour $x>0$.

c) Densité de charge totale $\rho(x)$.

d) Potentiel électrostatique $\varphi(x)$ (avec origine choisie en $x=0$).

d. Conclusion

Pour résumer, on peut se représenter schématiquement la jonction p-n de la façon suivante : en dehors de la double couche le semi-conducteur est électriquement neutre, les charges ioniques fixes dues aux impuretés étant compensées par les charges des porteurs libres, trous de valence du côté p, électrons de conduction du côté n; dans la double couche, au contraire, il n'y a pratiquement pas de porteurs libres, il reste seulement les charges fixes (non compensées) des impuretés, négatives du côté p, positives du côté n.

L'importance pratique de la jonction p-n vient de ses propriétés conductrices. Elle laisse en effet passer facilement le courant électrique du côté p vers le côté n, difficilement en sens inverse (fig. 9) : elle se comporte en *«redresseur»* de courant.

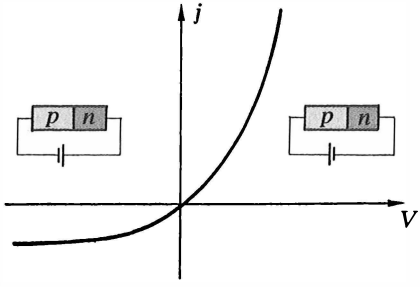


FIGURE 9

Densité de courant j en fonction de la différence de potentiel V appliquée à une jonction p-n.

La représentation schématique que nous avons donnée d'une jonction p-n à l'équilibre permet de comprendre simplement cet effet redresseur[22].

22. Une jonction $p-n$ peut être considérée comme constituée d'un échantillon de faible épaisseur d'un matériau peu conducteur, dans lequel se produit pratiquement toute la chute de potentiel (la double couche, qui contient très peu de porteurs libres), séparant deux conducteurs de faible résistance dans lesquels le potentiel varie très peu (les régions homogènes, qui conduisent facilement le courant grâce aux porteurs libres qu'elles contiennent : électrons du côté n, trous du côté p). Lorsqu'on applique une différence de potentiel V (inférieure à $\phi(+\infty)-\phi(-\infty)$) entre la région p et la région n, la chute de potentiel totale

$$\Delta V = \phi(+\infty)-\phi(-\infty)-V>0 \qquad (1)$$

se produit encore à la traversée de la zone peu conductrice, c'est-à-dire de la double couche (dont l'épaisseur est d'ailleurs modifiée par le potentiel supplémentaire).

Examinons alors le mouvement des porteurs induit par ce potentiel supplémentaire. Considérons par exemple les électrons de conduction. Lorsque $V=0$, dire qu'il n'y a pas de courant ne signifie pas qu'aucun électron ne traverse la jonction; simplement il y en a autant qui passent de la région n à la région p que de la région p à la région n. Lorsque $V\neq 0$, l'équilibre est rompu :

— les électrons de conduction sont en très petit nombre dans la région p, mais dès que l'un d'eux s'approche de la double couche, il est immédiatement «aspiré» par le champ électrique et passe dans la région n. Comme ces électrons ont été excités dans une région où le potentiel est constant, *leur nombre est pratiquement indépendant de V ainsi que le courant de particules* $J^{(e)}_{p\to n}$ *correspondant* (mesuré algébriquement sur l'axe Ox);

.../...

Références

1. N.W. Ashcroft et N.D. Mermin, *Solid State Physics,* Holt-Saunders, Tokyo (1981).
2. C. Kittel, *Introduction à la physique du solide,* Dunod, Paris (1972).

— les électrons de conduction sont en plus grand nombre dans la région n, mais la plupart sont repoussés par le champ électrique. Seuls ceux qui arrivent en d_n avec une énergie suffisante pour franchir la barrière de potentiel pourront passer dans la région p. La densité des électrons qui ont cette énergie minimale est proportionnelle à $\exp\{-(-e)(\phi(-\infty)+V-\phi(+\infty))/kT\}$; il en est de même du courant de particules correspondant :

$$J^{(e)}_{n\to p} \propto \exp\frac{e}{kT}\left[\varphi(-\infty)+V-\varphi(+\infty\right]. \tag{2}$$

Pour $V=0$, $J^{(e)}_{n\to p}$ et $J^{(e)}_{p\to n}$ se compensent exactement :

$$J^{(e)}_{p\to n}+J^{(e)}_{n\to p}(V=0)=0\ ; \tag{3}$$

on a donc

$$J^{(e)}_{n\to p}(V)=-J^{(e)}_{p\to n}\exp\frac{eV}{kT} \tag{4}$$

On peut raisonner de façon analogue pour les trous, en échangeant les régions n et p; on trouve ainsi :

$$J^{(t)}_{p\to n}(V)=-J^{(t)}_{n\to p}\exp\frac{eV}{kT} \tag{5}$$

avec $J^{(t)}_{n\to p}$ pratiquement indépendant de V.

La densité de courant électrique totale s'écrit donc

$$\begin{aligned} j &= -e\left(J^{(e)}_{p\to n}+J^{(e)}_{n\to p}\right)+e\left(J^{(t)}_{p\to n}+J^{(t)}_{n\to p}\right)\\ &= e\left(J^{(e)}_{p\to n}-J^{(t)}_{n\to p}\right)\left(\exp\frac{eV}{kT}-1\right) \end{aligned} \tag{6}$$

(noter que $J^{(t)}_{n\to p}$ est négatif). On en déduit la caractéristique $j(V)$ de la jonction (fig. 9).

COMPLÉMENT VI.B

Les naines blanches

I. Introduction. Principales données

II. Le rôle des forces gravitationnelles

III. Théorie élémentaire des naines blanches
1. Idée physique de base
2. Quelques ordres de grandeur
3. Gaz dégénéré d'électrons relativistes
4. Équilibre d'une naine blanche. Relation entre sa masse et son rayon

I. INTRODUCTION. PRINCIPALES DONNÉES

Les naines blanches sont des astres très particuliers[1]. Dans les diagrammes masse-luminosité, masse-rayon ou couleur-magnitude que tracent les astronomes pour mettre en évidence les corrélations entre différentes données d'observation (le diagramme couleur-luminosité est esquissé sur la figure 1), les naines blanches se situent très nettement en dehors de la « séquence principale » qui regroupe la plupart des étoiles analogues au Soleil : elles sont très peu brillantes pour leur couleur blanche ou bleue, et leur rayon est beaucoup plus petit que celui des étoiles de même masse. Aucune n'est visible à l'œil nu (on en a recensé à l'heure actuelle plus de 1500, et on estime qu'environ 3 % des étoiles de notre galaxie sont des naines blanches).

La première naine blanche connue a été découverte (F. Bessel, 1844) en analysant le mouvement d'une étoile très brillante et donc bien visible, Sirius : ce mouvement indiquait la présence d'un « compagnon » alors invisible, que l'on nomma « Sirius B ». Ce n'est que 18 ans plus tard que Sirius B fut aperçue pour la première fois à la lunette astronomique : c'est un astre de huitième grandeur seulement. Pourtant sa masse, déduite des trajectoires des deux compagnons, est à peu de chose près égale à celle du soleil ($\sim 2 \times 10^{30}$ kg); mais son rayon, évalué à partir de l'énergie rayonnée et de la température en surface (elle-même estimée grâce au spectre de la lumière émise), est d'environ 16000 km seulement (à comparer avec 700000 km pour le Soleil).

Les naines blanches sont des étoiles vieilles, parvenues au stade ultime de leur évolution : elles ont épuisé le combustible (hydrogène dans un premier stade, puis hélium dans un deuxième) qui alimente les réactions nucléaires fournissant l'énergie dans les étoiles plus jeunes, et se sont considérablement contractées sous l'effet des forces de gravitation. C'est ainsi que leur rayon R est de l'ordre de celui de la Terre pour une masse M de l'ordre de celle du Soleil; leur *masse volumique* ρ est donc *énorme.* Pour fixer les idées, nous prendrons comme *ordres de grandeur typiques*

$$R \sim 5000 \text{ km} , \tag{B.1.a}$$

$$M \sim 10^{30} \text{ kg} , \tag{B.1.b}$$

$$\rho \sim 5 \times 10^{6} \text{ g/cm}^3 , \tag{B.1.c}$$

1. Les caractéristiques des naines blanches et les problèmes qu'elles continuent de poser à l'heure actuelle sont exposés de façon assez précise dans un article de vulgarisation : G. Fontaine et F. Wesemael, « Les Naines Blanches », *La Recherche 165,* 464 (1985). On pourra consulter aussi : H.M. Van Horn, « The Physics of White Dwarfs », *Physics Today 32,* n° 1, p. 23 (1979); J.P. Cox, *Principles of Stellar Structure,* Gordon and Breach, Londres (1968), vol. II, chapitre 25.

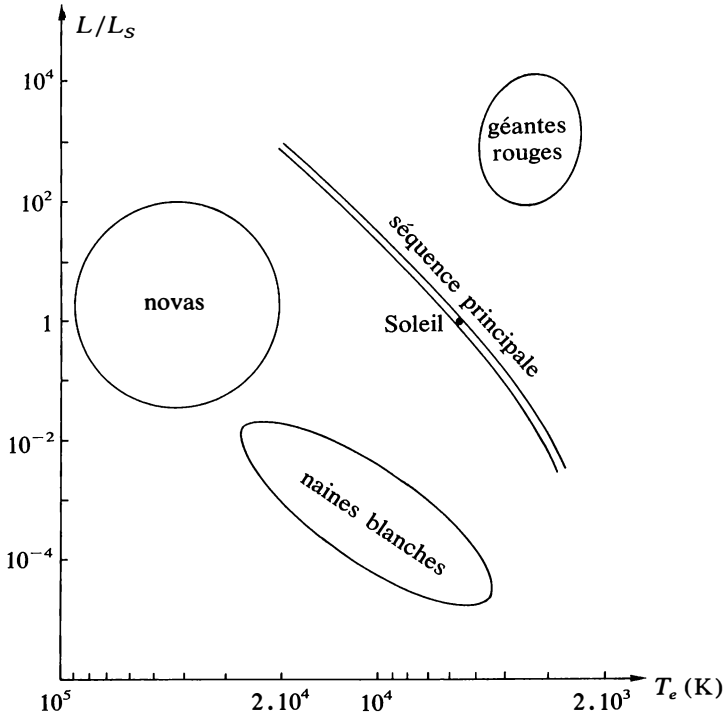


FIGURE 1

Diagramme de Hertzprung-Russell : en abscisse, l'inverse de la « température effective » T_e de l'étoile (déduite du spectre qu'elle émet), en ordonnée le rapport L/L_s de sa luminosité à celle du Soleil (noter la double échelle logarithmique). Les naines blanches se situent très nettement en dehors de la « séquence principale », où se trouvent le Soleil et Sirius *A*.

Quant à la température, elle est estimée à

$$T \sim 10^7 \text{ K} \tag{B.2}$$

dans les régions centrales d'une naine blanche, c'est-à-dire qu'elle est du même ordre que celle qui règne à l'intérieur des étoiles « normales ». La composition chimique est cependant très différente : le cœur d'une naine blanche ne contient plus d'hydrogène ni d'hélium; il est constitué essentiellement de « déchets nucléaires » produits par les réactions dont il a été le siège, principalement du carbone et de l'oxygène. Pour les évaluations d'ordres de grandeur, nous aurons surtout besoin de connaître la masse atomique par électron, c'est-à-dire le rapport A/Z du nombre de masse A au numéro atomique Z; nous prendrons

$$\frac{A}{Z} \sim 2 , \tag{B.3}$$

valeur de ce rapport pour le carbone 12 et l'oxygène 16, et dont les autres noyaux s'écartent peu. Lorsqu'il s'avèrera nécessaire de faire intervenir le nombre de masse A lui-même, nous l'évaluerons à

$$A \sim 15 . \tag{B.4}$$

II. LE ROLE DES FORCES GRAVITATIONNELLES

Les systèmes que nous avons étudiés jusqu'ici ont des dimensions linéaires de l'ordre du mètre et des masses de l'ordre du kilogramme. L'énergie potentielle associée aux forces gravitationnelles entre les constituants est alors totalement négligeable devant les autres formes d'énergie. Au contraire, les propriétés des corps célestes sont régies principalement par les forces gravitationnelles.

Il est facile de comprendre cette différence par un calcul d'ordres de grandeur. La force gravitationnelle exercée par un objet (1) de masse m_1 sur un objet (2) de masse m_2 est donnée par

$$\vec{F}_{21} = -G\frac{m_1 m_2}{r^2}\vec{e}_{21}\ , \tag{B.5}$$

où r est la distance entre les deux objets et $\vec{e}_{21}$ le vecteur unitaire porté par la droite qui les joint, dans le sens de (1) vers (2); cette force est *toujours attractive;* la constante gravitationnelle G a pour valeur

$$G = 6{,}67 \times 10^{-11}\ \mathrm{m^3\ kg^{-1}\ s^{-2}}\ . \tag{B.6}$$

L'énergie potentielle gravitationnelle E_G d'un corps de masse M et de dimensions linéaires de l'ordre de R vaut donc environ[2]

$$E_G \sim -G\frac{M^2}{R}\ . \tag{B.7}$$

On peut par exemple comparer numériquement $|E_G|$ à NkT, où N est le nombre de constituants microscopiques du système et T sa température : on trouve un rapport $|E_G|/NkT$ de l'ordre de 5×10^{-13} pour 1 m^3 d'eau à la température ordinaire, 3×10^{-14} pour 1 kg de cuivre à la même température, et voisin de 1 pour le Soleil[3].

---

2. Pour un corps *sphérique* de rayon R dont la masse M est répartie *uniformément,* on trouve facilement (en utilisant le théorème de Gauss)

$$E_G = -\frac{3}{5}G\frac{M^2}{R}\ .$$

Dans une étoile ou une naine blanche, la masse volumique est loin d'être uniforme, mais l'énergie gravitationnelle garde la forme (B.7) pour des raisons dimensionnelles, le coefficient numérique sans dimension étant toujours de l'ordre de 1.

3. Pour l'eau, l'énergie gravitationnelle vaut

$$|E_G| \sim 7 \times 10^{-11} \times \frac{10^6}{1} \sim 7 \times 10^{-5}\ \mathrm{J} \tag{1}$$

et le nombre de molécules

$$N \sim \frac{10^6}{18} \times 6 \times 10^{23} \sim 0{,}3 \times 10^{29}\ , \tag{2}$$

de sorte que

$$NkT \sim 1{,}4 \times 10^{8}\ \mathrm{J}\ . \tag{3}$$

Le calcul est analogue pour le cuivre, de masse volumique $\simeq 9$ g/cm^3 et de masse atomique 63,5 g; on trouve

$$|E_G| \sim 1{,}4 \times 10^{-9}\ \mathrm{J}\ , \tag{4.a}$$

$$NkT \sim 4 \times 10^{4}\ \mathrm{J}\ . \tag{4.b}$$

.../...

Les forces gravitationnelles ont une portée trop longue pour que les considérations d'extensivité (complément II.A, § II) soient applicables en leur présence. Effectivement, si l'on multiplie le volume du système par un facteur λ en maintenant sa densité constante, M est multipliée par λ et R par $\lambda^{1/3}$; l'énergie gravitationnelle E_G s'accroît alors par un facteur $\lambda^{5/3}$. L'énergie n'est donc pas extensive si E_G n'est pas négligeable, et *les propriétés du système dépendent de façon non triviale de sa taille.*

III. THÉORIE ÉLÉMENTAIRE DES NAINES BLANCHES

1. IDÉE PHYSIQUE DE BASE

Les forces gravitationnelles, purement attractives, ont toujours tendance à provoquer la contraction d'un objet tel qu'une étoile. Inversement, un gaz ou un plasma comme ceux qui constituent l'étoile pendant la majeure partie de sa vie a tendance à s'épandre dans tout le volume disponible. La stabilité (en réalité provisoire) d'une étoile comme le Soleil résulte d'un équilibre entre ces deux tendances contraires : l'effondrement s'arrête lorsque les processus physiques qui se produisent dans l'étoile (processus que la contraction contribue à déclencher ou à renforcer) créent une pression suffisante pour contrebalancer les forces gravitationnelles.

L'intensité de ces dernières est facile à évaluer. Imaginons en effet un accroissement (algébrique) $\mathrm{d}R$ du rayon de l'astre. Il provoque une augmentation de l'énergie gravitationnelle (B.7) de

$$\mathrm{d}E_G \sim G\frac{M^2}{R^2}\,\mathrm{d}R\,. \tag{B.8}$$

Pour faire apparaître une « pression gravitationnelle » P_G que l'on pourra comparer à la « pression cinétique » du gaz constituant l'astre, écrivons que l'augmentation (B.8) d'énergie correspond à un travail

$$-P_G\,\mathrm{d}V = -P_G \times 4\pi R^2\,\mathrm{d}R \tag{B.9}$$

de cette « pression gravitationnelle ». On en déduit

$$P_G \sim -\frac{G}{4\pi}\frac{M^2}{R^4}\,. \tag{B.10}$$

Dans le cas du Soleil,

$$M \simeq 2\times 10^{30}\ \mathrm{kg}\,, \quad R \simeq 7\times 10^{8}\ \mathrm{m}\,, \tag{5}$$

ce qui donne

$$|E_G| \sim 4\times 10^{41}\ \mathrm{J}\,. \tag{6}$$

Pour évaluer N, il faut savoir que la proportion (en masse) d'hydrogène est de 75 % et celle d'hélium de 25 % (nous négligeons les 2 % environ d'éléments lourds); comme les atomes sont à peu près totalement ionisés, nous multiplions par 2 le nombre d'atomes d'hydrogène (un proton et un électron) et par 3 celui des atomes d'hélium (un noyau et 2 électrons); nous obtenons ainsi (M étant exprimée en grammes)

$$N \sim \left[0{,}75\,\frac{M}{1}\times 2 + 0{,}25\,\frac{M}{4}\times 3\right]\times 6\times 10^{23}\,, \tag{7}$$

puis, avec $T \sim 10^7 K$,

$$NkT \sim 3\times 10^{41}\ \mathrm{J}\,. \tag{8}$$

Cette « pression » est *négative :* ceci reflète simplement le fait que les forces de gravitation *tendent toujours à provoquer la contraction du système,* et non pas son expansion comme une véritable pression.

Dans une étoile de la séquence principale (fig. 1), la pression contrebalançant P_G est d'origine thermique; elle est entretenue par l'énergie que dégagent les réactions nucléaires se déroulant dans le cœur de l'étoile. Mais la stabilité des naines blanches, dont les caractéristiques sont si radicalement particulières, a posé pendant longtemps un problème théorique préoccupant : quel phénomène physique est capable d'empêcher dans ce cas l'effondrement gravitationnel? Comme la masse M d'une naine blanche est analogue à celle d'une étoile ordinaire et que son rayon R est 50 à 100 fois plus petit, la « pression gravitationnelle » qu'il s'agit de compenser est considérablement plus forte (formule (B.10)). En outre, nous savons aujourd'hui que les réactions nucléaires y sont éteintes.

Cette énigme été résolue dès 1925 par Fowler, qui proposa d'appliquer aux électrons des naines blanches les résultats, alors tout récents, de la mécanique quantique. Nous allons effectivement constater que *c'est la « pression quantique »* (cf. VI, § I.B.1.c) *du gaz d'électrons, totalement dégénéré, qui empêche l'effondrement gravitationnel de la naine blanche et assure sa stabilité.*

2. QUELQUES ORDRES DE GRANDEUR

Explicitons cette idée en évaluant les principales caractéristiques du système. A la température 10^7 K (cf. (B.2)) correspond une énergie d'environ 1 000 eV. Dans ces conditions, *les électrons et les noyaux ne forment pas des atomes* ni même des ions, *mais deux « gaz » séparés* quoique intimement mélangés. Cependant, la naine blanche se refroidit progressivement, car elle perd de l'énergie par rayonnement et ne possède plus de combustible pour la renouveler. On pourrait penser que ce refroidissement lent devrait entraîner la recombinaison des électrons et des noyaux en atomes lorsque la température serait tombée à quelques milliers de kelvins. Il n'en est rien, car *c'est principalement la densité élevée qui est ici responsable de l'ionisation des atomes.* En effet, évaluons le nombre N_n/V de noyaux par unité de volume à l'intérieur d'une naine blanche :

$$\frac{N_n}{V} \sim \frac{\rho(\mathrm{g/cm^3})}{A} . N_A \sim 2 \times 10^{29}\ \mathrm{cm^{-3}} . \qquad \text{(B.11)}$$

Le volume disponible par noyau est donc de l'ordre de $5 \times 10^{-30}\ \mathrm{cm^3} = 5 \times 10^{-6}\ \text{Å}^3$, alors que le volume d'un atome[4] est de l'ordre de l'Å^3. Ainsi, pour que des atomes puissent se former au cours du refroidissement de la naine blanche, il faudrait que celle-ci retrouve un rayon 50 à 100 fois plus grand; or on ne peut imaginer aucun processus susceptible de produire une telle augmentation de volume, qui devrait s'effectuer contre les forces gravitationnelles.

---

4. L'ordre de grandeur du volume minimum d'un atome est fixé par la relation d'incertitude

$$\Delta x . \Delta p_x \geqslant \hbar . \qquad (1)$$

En effet, si un électron est confiné dans un volume de dimensions linéaires

$$\Delta x \sim a , \qquad (2)$$

les composantes de son impulsion ne peuvent pas être déterminées à mieux que $\Delta p_x \geqslant \hbar/a$, ce qui lui donne une énergie cinétique au moins de l'ordre de

$$E_c \sim \frac{3\hbar^2}{2ma^2} . \qquad (3)$$

.../...

Calculons maintenant la température de Fermi du gaz d'électrons[5]. Il nous faut pour cela le nombre d'électrons par unité de volume :

$$\frac{N_e}{V} \sim \frac{\rho(\mathrm{g/cm^3})}{A} . Z . N_A \sim 1{,}5 \times 10^{30}\ \mathrm{cm^{-3}} . \tag{B.12}$$

Nous en déduisons

$$T_F^{(e)} \widehat{=} \frac{\hbar^2}{2m_e k} \left(3\pi^2 \frac{N_e}{V}\right)^{2/3} \sim 5 \times 10^9\ \mathrm{K} . \tag{B.13}$$

Le gaz d'électrons est donc totalement dégénéré :

$$\frac{T}{T_F^{(e)}} \leqslant 2 \times 10^{-3} \ll 1 . \tag{B.14}$$

Or nous savons (VI, § I.B.1.c) que la pression d'un gaz dégénéré de fermions indépendants tend vers une limite non nulle lorsque la température devient très inférieure à la température de Fermi; l'ordre de grandeur de la *pression du gaz d'électrons* est donc

$$P_0^{(e)} = \frac{2}{15\pi^2\hbar^3} (2m_e)^{3/2} [kT_F^{(e)}]^{5/2} \sim 4{,}5 \times 10^{22}\ \mathrm{Pa} . \tag{B.15}$$

Pour estimer la *pression du gaz de noyaux,* nous pouvons utiliser la loi des gaz parfaits classiques. En effet, la température de dégénérescence $T_D^{(n)}$ des noyaux (température de Fermi ou de Bose suivant les cas) est nettement inférieure à T puisque leur densité (B.11) est à peine plus faible que celle des électrons mais leur masse environ $15 \times 1\,800$ fois plus grande; on trouve effectivement

$$T_D^{(n)} \sim 4{,}5 \times 10^4\ \mathrm{K} \implies T \gg T_D^{(n)} . \tag{B.16}$$

Comme son énergie potentielle E_p à la distance a d'un noyau de charge Ze vaut

$$|E_p| \sim \frac{Zq^2}{a} \quad (q^2 \widehat{=} e^2/4\pi\varepsilon_0) , \tag{4}$$

l'électron ne peut être lié que si

$$E_c \leqslant |E_p| , \tag{5}$$

c'est-à-dire

$$a \geqslant \frac{3}{2} \frac{1}{Z} \frac{\hbar^2}{mq^2} \sim \frac{a_0}{Z} , \tag{6}$$

où a_0 est le rayon de Bohr. Un atome neutre (dans lequel le Z-ième électron « voit » le noyau écranté par $(Z-1)$ électrons, c'est-à-dire une charge peu différente de $+e$) nécessite donc un volume minimum de

$$\frac{4}{3} \pi a_0^3 \sim 0{,}5\ \text{Å}^3 \tag{7}$$

et un ion carbone ou oxygène comportant un seul électron

$$\frac{4}{3} \pi \left(\frac{a_0}{Z}\right)^3 \sim 2 \times 10^{-4}\ \text{Å}^3 . \tag{8}$$

5. Nous traitons ici les électrons comme un gaz parfait, c'est-à-dire que nous négligeons leurs interactions. Cette approximation est tout à fait valable : un gaz d'électrons *dégénéré* est d'autant plus proche d'un gaz parfait que sa densité N/V est plus élevée (exercice VI.14).

On estime donc la pression cinétique $P^{(n)}$ du gaz de noyaux à[6]

$$P^{(n)} \sim \frac{N_n kT}{V} \sim 3 \times 10^{19} \text{ Pa} . \tag{B.17}$$

On constate effectivement qu'elle est très inférieure à celle du gaz d'électrons :

$$P^{(n)} \ll P_0^{(e)} . \tag{B.18}$$

C'est donc bien à ce dernier que revient le rôle crucial dans l'équilibre de la naine blanche[7].

Notons que la pression du gaz d'électrons est pratiquement *indépendante de la température* T dès lors que celle-ci est très inférieure à la température de Fermi $T_F^{(e)}$. Par conséquent, *le refroidissement progressif de la naine blanche ne modifiera pas ses conditions d'équilibre.*

Remarquons enfin que, pour des énergies de l'ordre de

$$\varepsilon_F^{(e)} = kT_F^{(e)} \sim 4{,}5 \times 10^5 \text{ eV} , \tag{B.19}$$

les électrons sont *relativistes :*

$$\varepsilon_F^{(e)} \sim m_e c^2 . \tag{B.20}$$

Ceci modifie peu l'évaluation (B.13) de la température de Fermi[8], de sorte que la conclusion (B.14) reste vraie. Mais il nous faut établir les formules qui décrivent un gaz dégénéré d'électrons relativistes avant de pouvoir préciser la condition d'équilibre d'une naine blanche.

3. GAZ DÉGÉNÉRÉ D'ÉLECTRONS RELATIVISTES

Nous considérons un gaz de fermions libres et indépendants (cf. note 5). S'ils sont non relativistes, leur énergie ε et leur impulsion $\vec{p}$ sont liées par la formule

$$\varepsilon = \frac{\vec{p}^{\,2}}{2m} \quad \text{(non relativiste);} \tag{B.21.a}$$

s'ils sont relativistes, cette relation doit être remplacée par

$$\varepsilon = \sqrt{\vec{p}^{\,2} c^2 + m^2 c^4} \quad \text{(relativiste).} \tag{B.21.b}$$

---

6. Nous négligeons encore les interactions entre les noyaux. Cette approximation est raisonnable, aux températures considérées ici, pour obtenir un *ordre de grandeur* de la pression. Mais elle cessera d'être valable lorsque la naine blanche se sera suffisamment refroidie : les noyaux constitueront peu à peu un liquide, puis un solide (comme les électrons forment toujours un gaz dégénéré, on aboutit à une structure semblable à celle d'un métal). Toutefois, ce refroidissement est extrêmement lent, et toutes les naines blanches que l'on a pu observer jusqu'ici sont encore très chaudes, comme l'indique (B.2).

7. La pression de radiation, qui joue un rôle important dans les étoiles ordinaires, est ici négligeable. Elle vaut en effet, d'après la formule VI (III.42),

$$P^{(r)} = \frac{1}{3} \frac{\pi^2}{15} \frac{(kT)^4}{\hbar^3 c^3} \sim 2{,}5 \times 10^{12} \text{ Pa} .$$

8. Nous établirons au paragraphe 3 l'expression de la température de Fermi dans le cas relativiste (voir note 12). Cette expression est à nouveau très simple dans la limite ultra-relativiste :

$$T_F^{(e)} = \frac{\hbar c}{k} \left(3\pi^2 \frac{N_e}{V} \right)^{1/3} \quad \text{si} \quad \varepsilon_F^{(e)} \gg m_e c^2 , \tag{1}$$

et elle donne encore

$$T_F^{(e)} \sim 8 \times 10^9 \text{ K} . \tag{2}$$

Nous allons mener les calculs en traitant le plus longtemps possible les deux cas en parallèle, pour pouvoir les comparer à chaque stade.

Reprenons l'expression du grand potentiel (formules V (IV.6) et V (IV.12))

$$J(T,\ V,\ \mu)=-kT\times 2\sum_{\vec{p}}\ln\left[1+\mathrm{e}^{-\beta\left[\varepsilon(\vec{p})-\mu\right]}\right]\ ; \tag{B.22}$$

le facteur 2 provient des deux états de spin possibles, et la somme porte sur les impulsions permises par les conditions aux limites périodiques dans une boîte de volume V (que nous prendrons pour simplifier cubique, d'arête L) : la composante de l'impulsion $\vec{p}$ sur chacun des trois axes de coordonnées doit être un multiple entier de $\hbar\times 2\pi/L$ (voir I, § III.B.2.a). La somme discrète peut être remplacée par une intégrale sur l'espace des impulsions, à condition de diviser le volume infinitésimal d^3p de cet espace par le volume $[\hbar.2\pi/L]^3$ d'une maille élémentaire de quantification :

$$J(T,\ V,\ \mu)=-kT\frac{2V}{(2\pi\hbar)^3}\int \mathrm{d}^3p\ \ln\left[1+\mathrm{e}^{-\beta\left[\varepsilon(\vec{p})-\mu\right]}\right]. \tag{B.23}$$

Comme l'énergie ε ne dépend que du module p de l'impulsion (qu'il ne faudra pas confondre avec la pression), on passe en coordonnées sphériques dans l'espace des $\vec{p}$, ce qui donne

$$J(T,\ V,\ \mu)=-kT\frac{8\pi V}{(2\pi\hbar)^3}\int_0^{+\infty} p^2\,\mathrm{d}p\ \ln\left[1+\mathrm{e}^{-\beta\left[\varepsilon(p)-\mu\right]}\right]. \tag{B.24}$$

Enfin, une intégration par parties aboutit à l'expression[9]

$$J(T,\ V,\ \mu)=-\frac{V}{3\pi^2\hbar^3}\int_0^{+\infty} p^3\,\mathrm{d}p\,\frac{\mathrm{d}\varepsilon}{\mathrm{d}p}\,N^F(\varepsilon;\ T,\ \mu)\ , \tag{B.25}$$

dans laquelle la dérivée $\mathrm{d}\varepsilon/\mathrm{d}p$ est donnée par l'une ou l'autre des formules :

$$\frac{\mathrm{d}\varepsilon}{\mathrm{d}p}=\frac{p}{m}\quad\text{(non relativiste)}, \tag{B.26.a}$$

$$\frac{\mathrm{d}\varepsilon}{\mathrm{d}p}=c^2\,\frac{p}{\sqrt{p^2c^2+m^2c^4}}\quad\text{(relativiste)}. \tag{B.26.b}$$

Pour un système suffisamment grand (limite thermodynamique), on calcule μ en fonction de N à partir de la condition habituelle[10] :

$$N=2\sum_{\vec{p}} N^F(\varepsilon;\ T,\ \mu)=\frac{8\pi V}{(2\pi\hbar)^3}\int_0^{+\infty} p^2\,\mathrm{d}p\,N^F(\varepsilon;\ T,\ \mu)\ . \tag{B.27}$$

---

9. La variable d'intégration étant ici p, et non pas ε, l'intégration par parties fait intervenir

$$\frac{\partial}{\partial p}\ln\left[1+\mathrm{e}^{-\beta\left[\varepsilon(p)-\mu\right]}\right]=\frac{\mathrm{d}\varepsilon}{\mathrm{d}p}.(-\beta)N^F(\varepsilon;\ T,\ \mu)\ .$$

Le reste du raisonnement est semblable à celui de la note 11 du chapitre VI.

10. Dans une naine blanche, les électrons et les noyaux sont évidemment accompagnés de photons, dont certains ont une énergie suffisante pour donner lieu à la création de paires électron-positron. Nous ne tenons pas compte ici de ce phénomène, qui n'affecte pas les ordres de grandeur que nous allons déterminer. Si nous voulions le faire, la relation (B.27) entre N et μ resterait valable pour les électrons, mais leur nombre N serait moins simple à évaluer car il serait couplé à celui des positrons : il faudrait analyser l'équilibre entre les électrons, les positrons et le gaz de photons, et faire intervenir les positrons dans la condition de neutralité électrique globale du système.

Les propriétés du sytème se déduisent alors des formules (B.25) et (B.27); en particulier, la *pression* P du gaz est telle que

$$P = -\frac{\partial J}{\partial V} = -\frac{J}{V} . \tag{B.28}$$

Dans le cas où le gaz est *totalement dégénéré,* c'est-à-dire lorsque la température T est très inférieure à la température de Fermi T_F[11], ces formules se simplifient considérablement, car le facteur de Fermi N_F se réduit alors à sa limite de température nulle (formule VI (I.4)): l'*« impulsion de Fermi »* p_F étant donnée par[12]

$$N = \frac{V}{\pi^2\hbar^3}\int_0^{p_F} p^2\,\mathrm{d}p = \frac{V}{3\pi^2\hbar^3}\,p_F^3 , \tag{B.29}$$

le grand potentiel s'écrit

$$J = -\frac{V}{3\pi^2\hbar^3}\int_0^{p_F} p^3\,\mathrm{d}p\,\frac{\mathrm{d}\varepsilon}{\mathrm{d}p} , \tag{B.30}$$

et la pression P_0 du gaz dégénéré

$$P_0 = \frac{1}{3\pi^2\hbar^3}\int_0^{p_F} p^3\,\mathrm{d}p\,\frac{\mathrm{d}\varepsilon}{\mathrm{d}p} . \tag{B.31}$$

Pour des particules non relativistes (formule (B.26.a)), on retrouve le résultat cité plus haut en (B.15) :

$$P_0 = \frac{1}{15\pi^2\hbar^3}\,\frac{p_F^5}{m} = \frac{2}{15\pi^2\hbar^3}\,(2m)^{3/2}(kT_F)^{5/2} \quad \text{(non relativiste).} \tag{B.32}$$

Si les particules sont relativistes (formule (B.26.b)), l'expression est un peu plus compliquée :

$$P_0 = \frac{c^2}{3\pi^2\hbar^3}\int_0^{p_F} \mathrm{d}p\,\frac{p^4}{\sqrt{p^2c^2+m^2c^4}} \quad \text{(relativiste).} \tag{B.33}$$

---

11. La température de Fermi T_F est définie ici par (cf. VI, § I.B.1.a)

$$kT_F \widehat{=} \mu_0 - mc^2 ,$$

où μ_0 est la limite du potentiel chimique μ lorsque la température T tend vers zéro (voir note 12).

12. Nous utilisons ici comme variable l'impulsion p, et non pas l'énergie ε. Le potentiel chimique μ_0 dans la limite de température nulle est relié à l'impulsion de Fermi p_F par

$$\mu_0 = \varepsilon(p_F) , \tag{1}$$

où la fonction $\varepsilon(p)$ a la forme (B.21.a) ou (B.21.b) suivant les cas. Pour des particules relativistes, on a donc

$$\mu_0 = \sqrt{p_F^2c^2 + m^2c^4} = kT_F + mc^2 , \tag{2}$$

où p_F est donnée par la relation (B.29). A la limite ultra-relativiste, la formule précédente devient

$$\mu_0 \simeq kT_F \simeq p_F c \quad \text{si} \quad p_F \gg mc , \tag{3}$$

c'est-à-dire

$$\mu_0 \simeq kT_F \simeq \hbar c\left(3\pi^2\,\frac{N}{V}\right)^{1/3} \quad \text{si} \quad p_F \gg mc . \tag{4}$$

Dans le domaine non relativiste, on prend évidemment

$$\mu_0 = kT_F = \frac{p_F^2}{2m} \quad \text{si} \quad p_F \ll mc . \tag{5}$$

En posant

$$x \widehat{=} \frac{p}{mc} , \tag{B.34}$$

on la met sous la forme

$$P_0 = \frac{m^4 c^5}{3\pi^2 \hbar^3} \int_0^{x_F} \mathrm{d}x \, \frac{x^4}{\sqrt{x^2+1}} \quad \text{(relativiste).} \tag{B.35}$$

Les formules non relativistes sont applicables tant que l'énergie cinétique $p^2/2m$ est petite devant l'énergie de masse mc^2, ce qui se traduit par la condition[13]

$$x_F \widehat{=} \frac{p_F}{mc} \ll 1 \implies \text{non relativiste.} \tag{B.36}$$

A l'autre extrême se situe le domaine ultra-relativiste, dans lequel au contraire l'énergie de masse est négligeable devant pc, c'est-à-dire

$$x_F \gg 1 \implies \text{ultra-relativiste.} \tag{B.37}$$

Dans ce domaine, on calcule aussi très simplement P_0[14] :

$$P_0 \simeq \frac{c}{12\pi^2\hbar^3} p_F^4 \left[1 - \frac{m^2c^2}{p_F^2}\right] \quad \text{si} \quad p_F \gg mc \ ; \tag{B.38}$$

nous conservons ici les deux premiers termes du développement, car le second nous sera nécessaire pour obtenir la relation entre la masse et le rayon de la naine blanche.

4. ÉQUILIBRE D'UNE NAINE BLANCHE. RELATION ENTRE SA MASSE ET SON RAYON

La condition d'équilibre de la naine blanche s'écrit dans notre modèle simple

$$P_0 + P_G = 0 , \tag{B.39}$$

où P_0 est la pression quantique du gaz d'électrons (§ 3) et P_G la « pression gravitationnelle » (B.10).

---

13. L'expression (B.35) permet de calculer les premières corrections relativistes, lorsque x_F n'est plus totalement négligeable devant 1 mais reste petit. On obtient en effet, en développant le dénominateur pour x petit,

$$P_0 = \frac{m^4c^5}{3\pi^2\hbar^3} \frac{x_F^5}{5} \left[1 - \frac{5}{14} x_F^2 + \frac{5}{24} x_F^4 + \ldots\right] ,$$

où le terme d'ordre le plus bas coïncide évidemment avec (B.32).

14. Il suffit de développer, par rapport à $1/x$, le dénominateur de la fonction à intégrer dans (B.35) pour obtenir

$$P_0 = \frac{m^4c^5}{3\pi^2\hbar^3} \left[\frac{1}{4} x_F^4 - \frac{1}{4} x_F^2 + \ldots\right] \tag{1}$$

(attention : pour démontrer que ce résultat est correct, il faut scinder l'intégrale en deux parties par une borne intermédiaire $x_0 > 1$, et n'utiliser le développement en $1/x$ que dans la deuxième intégrale, entre x_0 et x_F). D'ailleurs, l'intégrale (B.35) se calcule analytiquement sans difficulté pour x_F quelconque :

$$P_0 = \frac{m^4c^5}{3\pi^2\hbar^3} \left\{\frac{1}{4} x_F^3 \sqrt{x_F^2+1} - \frac{3}{8} x_F \sqrt{x_F^2+1} + \frac{3}{8} \ln\left[x_F + \sqrt{x_F^2+1}\right]\right\} . \tag{2}$$

Pour exploiter cette condition, exprimons P_0, comme P_G, en fonction du rayon R et de la masse totale M de la naine blanche. Les formules (B.32) et (B.38) se transforment facilement[15] en

$$P_0 \sim \frac{\hbar^2}{15\pi^{1/3}} \frac{1}{m} \left(\frac{9M}{8m_p}\right)^{5/3} \frac{1}{R^5} \quad \text{(non relativiste)}, \tag{B.40.a}$$

$$P_0 \sim \frac{\hbar c}{12\pi^{2/3}} \left(\frac{9M}{8m_p}\right)^{4/3} \frac{1}{R^4} \left[1 - \frac{m^2c^2}{\hbar^2\pi^{2/3}} \frac{R^2}{(9M/8m_p)^{2/3}}\right] \quad \text{(ultra-relativiste)}, \tag{B.40.b}$$

où m_p est la masse du proton et m celle de l'électron. Nous nous contenterons d'utiliser ces deux expressions limites; l'analyse du domaine intermédiaire nécessiterait seulement un calcul numérique (cf. note 14).

La condition d'équilibre (B.39) de la naine blanche prend donc la forme[16]

$$R \sim \frac{\hbar^2}{15\pi^{1/3}} \frac{1}{m} \left(\frac{9}{8m_p}\right)^{5/3} \cdot \frac{4\pi}{G} \cdot \frac{1}{M^{1/3}} \quad \text{(non relativiste)} \tag{B.41.a}$$

---

15. La relation (B.29) donne p_F en fonction de la densité N/V du gaz d'électrons :

$$p_F = \hbar \left(3\pi^2 \frac{N}{V}\right)^{1/3} . \tag{1}$$

Le volume V est relié de façon triviale à R et le nombre total N d'électrons a pour ordre de grandeur

$$N = \frac{M}{Am_p} \times Z \sim \frac{M}{2m_p} \tag{2}$$

(nous négligeons évidemment la masse m de l'électron devant celle du proton m_p). Un calcul algébrique simple conduit ensuite aux formules (B.40).

16. Il est facile de montrer que l'équilibre ainsi obtenu est *stable*. En effet, la pression gravitationnelle est de la forme

$$P_G = -\frac{A}{R^4} , \tag{1}$$

c'est-à-dire que son augmentation (algébrique) provoquée par un accroissement dR du rayon s'écrit

$$dP_G = +\frac{4A}{R^5} dR . \tag{2}$$

Pour la pression P_0 du gaz d'électrons, les expressions correspondantes sont

$$P_0 = \frac{B}{R^5} , \quad dP_0 = -\frac{5B}{R^6} dR \quad \text{(non relativiste)} \tag{3.a}$$

et

$$P_0 = \frac{C}{R^4} - \frac{D}{R^2} , \quad dP_0 = \left[-\frac{4C}{R^5} + \frac{2D}{R^3}\right] dR \quad \text{(ultra-relativiste)}. \tag{3.b}$$

L'équilibre sera stable si un accroissement $dR > 0$ crée un déséquilibre tendant à contracter le système :

$$dP_G + dP_0 < 0 \quad \text{si} \quad dR > 0 . \tag{4}$$

C'est effectivement le cas, tant à la limite non relativiste, car

$$-\frac{A}{R^4} + \frac{B}{R^5} = 0 \implies \frac{4A}{R^5} - \frac{5B}{R^6} < 0 , \tag{5.a}$$

que dans le domaine ultra-relativiste :

$$-\frac{A}{R^4} + \frac{C}{R^4} - \frac{D}{R^2} = 0 \implies \frac{4A}{R^5} - \frac{4C}{R^5} + \frac{2D}{R^3} < 0 . \tag{5.b}$$

ou $$R \sim \frac{\hbar}{mc}\,\pi^{1/3}\left(\frac{9M}{8m_p}\right)^{1/3}\left[1-\frac{G}{4\pi}\left(\frac{8m_p}{9}\right)^{4/3}\frac{12\pi^{2/3}}{\hbar c}\,M^{2/3}\right]^{1/2}$$

(ultra-relativiste), (B.41.b)

que nous écrirons de façon plus compacte

$$RM^{1/3} \sim R_0 M_0^{1/3} = \text{constante} \quad \text{(non relativiste)} \tag{B.42.a}$$

ou $$R \sim R_0\left(\frac{M}{M_0}\right)^{1/3}\left[1-\left(\frac{M}{M_0}\right)^{2/3}\right]^{1/2} \quad \text{(ultra-relativiste)}, \tag{B.42.b}$$

avec $$M_0 \widehat{=} \frac{1}{\pi}\left(\frac{\hbar c}{12}\cdot\frac{4\pi}{G}\right)^{3/2}\left(\frac{9}{8m_p}\right)^2 \tag{B.43}$$

et $$R_0 \widehat{=} \frac{\hbar}{mc}\,\pi^{1/3}\left(\frac{9M_0}{8m_p}\right)^{1/3}. \tag{B.44}$$

Numériquement, M_0 est de l'ordre de la masse du Soleil :

$$M_0 \sim 2\times 10^{30}\ \text{kg} \tag{B.45}$$

et R_0 voisin du rayon de la Terre :

$$R_0 \sim 6\,000\ \text{km}\,. \tag{B.46}$$

La relation (B.42.a) est valable lorsque $p_F \ll mc$, c'est-à-dire lorsque la densité N/V des électrons n'est pas trop grande; on montre facilement[17] que ceci signifie

$$p_F \ll mc \ \text{(non relativiste)} \implies \begin{cases} R \gg R_0\,, \\ M \ll M_0\,. \end{cases} \tag{B.47}$$

Au contraire, la relation (B.42.b) suppose que $p_F \gg mc$, ce que l'on peut traduire par[18]

$$p_F \gg mc \ \text{(ultra-relativiste)} \implies \begin{cases} R \ll R_0\,, \\ M \longrightarrow M_{0-}\,. \end{cases} \tag{B.48}$$

---

17. En reprenant les formules de la note 15, on met la condition

$$p_F \ll mc \tag{1}$$

sous la forme

$$R \gg \frac{\hbar}{mc}\,\pi^{1/3}\left(\frac{9M}{8m_p}\right)^{1/3} \tag{2}$$

Il suffit ensuite d'utiliser la relation (B.41.a) pour remplacer M en fonction de R ou R en fonction de M, ce qui conduit à

$$R \gg \frac{\hbar}{mc}\sqrt{\frac{\hbar c}{15}}\,\frac{9}{8m_p}\sqrt{\frac{4\pi}{G}} = \sqrt{\frac{4}{5}}\,R_0 \tag{3.a}$$

ou $$M^{2/3} \ll \frac{\hbar c}{15\pi^{2/3}}\left(\frac{9}{8m_p}\right)^{4/3}\cdot\frac{4\pi}{G} = \frac{4}{5}\,M_0^{2/3}\,. \tag{3.b}$$

18. Comme dans le note 17, la condition $p_F \gg mc$ s'exprime par

$$R \ll \frac{\hbar}{mc}\,\pi^{1/3}\left(\frac{9M}{8m_p}\right)^{1/3} \tag{1}$$

Si l'on remplace R en fonction de M selon la relation (B.42.b), on transforme facilement cette condition en

$$1-\left(\frac{M}{M_0}\right)^{2/3} \ll 1\,, \tag{2}$$

ce qui implique que M est très voisin de M_0. Mais alors le second membre de l'inégalité (1) est pratiquement égal à R_0. On trouve donc bien les conditions (B.48).

Nous aboutissons ainsi à une conclusion remarquable : *le rayon R d'une naine blanche est fonction de sa masse M*[19]. Tant que M est petite devant M_0, le rayon R, grand devant R_0, est inversement proportionnel à la puissance 1/3 de la masse M (formule (B.42.a))[20]. Pour des masses plus importantes, s'approchant de M_0, le rayon R peut être très faible (formule (B.42.b)). Mais, comme R est nécessairement réel, *il ne peut pas exister de naine blanche dont la masse serait supérieure à M_0* : c'est ce que l'on appelle la *« limite de Chandrasekhar »*. La relation masse-rayon, telle qu'elle découle du modèle très grossier que nous avons développé, est représentée sur la figure 2.

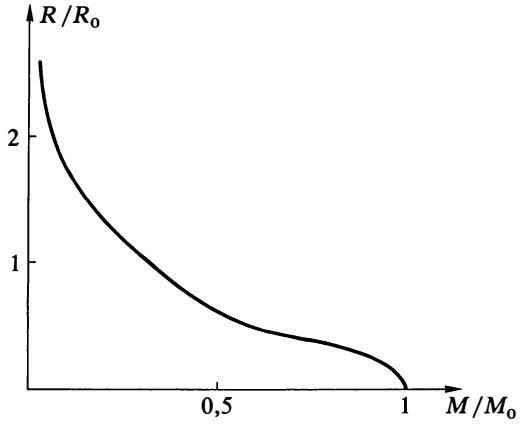


FIGURE 2
Relation entre la masse M et le rayon R d'une naine blanche dans le modèle simple présenté ici. Cette relation est donnée par la formule (B.42.a) pour les faibles valeurs de la masse et par (B.42.b) pour les masses voisines de M_0; les deux branches de courbe ont été simplement raccordées dans la région intermédiaire.

19. La masse d'une naine blanche est nettement inférieure à celle de l'étoile dont elle provient. Une étoile en effet perd constamment de la masse : le « vent stellaire » lui arrache continuellement de la matière. A certains moments cruciaux de la vie de l'étoile, cette perte de masse peut atteindre des proportions considérables : par exemple, au cours de la phase « nébuleuse planétaire » qui précède immédiatement la transformation en naine blanche, toute l'enveloppe externe de l'astre est éjectée. Typiquement, une étoile dont la masse initiale vaut 5 fois celle du Soleil se retrouve à la fin de son évolution avec seulement le cinquième ou le dixième de cette masse initiale.

20. Le domaine de validité de la relation non relativiste (B.42.a) est également limité du côté des grands rayons et des faibles masses. Il faut en effet que la densité du gaz d'électrons soit suffisante pour qu'il soit dégénéré. Si l'on prend pour la température T de la naine blanche l'ordre de grandeur (B.2), on doit imposer

$$T_F = \frac{\hbar^2}{2mk}\left(3\pi^2 \frac{N}{V}\right)^{2/3} \geqslant 10^7 \text{ K} . \qquad (1)$$

En utilisant à nouveau l'évaluation de N donnée dans la note 15 puis la relation (B.41.a), on transforme cette condition en

$$R \leqslant \frac{\hbar^{3/2}}{\sqrt{15}} \frac{1}{m^{3/4}} \frac{9}{8m_p} \left(\frac{4\pi}{G}\right)^{1/2} \left(\frac{1}{2k \times 10^7}\right)^{1/4} , \qquad (2)$$

.../...

Les observations effectuées sur les naines blanches connues actuellement vérifient de façon qualitative les résultats précédents. Mais la structure des naines blanches est évidemment plus complexe que nous ne l'avons indiqué, et leurs caractéristiques ne sont pas exactement les mêmes pour toutes (leur densité moyenne, par exemple, peut varier d'un facteur 10 de l'une à l'autre). Des calculs plus précis aboutissent à une masse limite de Chandrasekhar valant environ 1,4 fois la masse du Soleil, ce qui semble conforme aux observations[21]. La vertu du modèle très simple que nous avons présenté est de prendre en compte le phénomène physique essentiel : c'est la pression quantique du gaz d'électrons dégénéré qui contrebalance les forces gravitationnelles et régit ainsi l'équilibre de la naine blanche.

ce qui donne

$$R \leqslant 2 \times 10^4 \text{ km} . \qquad (3)$$

Cette limite étant supérieure à R_0, il peut exister des naines blanches dans lesquelles les électrons ne sont pas relativistes.

21. Une étoile ayant épuisé son combustible nucléaire et dont la masse résiduelle M est supérieure à la limite de Chandrasekhar se transforme, pense-t-on, en « étoile à neutrons » ou en « trou noir ». Voir par exemple : Ch. Ryter, « L'Étoile à Neutrons », *Bulletin de la Société Française de Physique 49,* p. 3 (juillet 1983); B. Carter et J.P. Luminet, « Les Trous Noirs : Maelströms Cosmiques », *La Recherche 94,* 944 (1978).

COMPLÉMENT VI.C

Potentiel chimique et chaleur spécifique d'un gaz de bosons libres au voisinage de la température de Bose

I. Domaine $T < T_B$: expression exacte de l'énergie et de la chaleur spécifique

II. Domaine $T > T_B$: développement de l'énergie au voisinage de T_B

III. Raccordement entre les deux domaines

Nous voulons montrer ici que *la chaleur spécifique à volume constant* d'un système thermodynamique de bosons libres *est continue* au passage de la température de Bose, mais que *sa dérivée ne l'est pas* (voir VI, fig. 10.b). Ceci nous amènera à préciser également le *comportement du potentiel chimique au voisinage de la température de Bose.*

I. DOMAINE $T < T_B$: EXPRESSION EXACTE DE L'ÉNERGIE ET DE LA CHALEUR SPÉCIFIQUE

Au-dessous de la température de Bose T_B, le fait que le potentiel chimique thermodynamique est constamment nul permet d'obtenir (VI, § II.B.2.a) une expression analytique explicite pour l'*énergie* du système :

$$E = AV \int_0^{+\infty} d\varepsilon \frac{\varepsilon^{3/2}}{e^{\beta\varepsilon} - 1} = 0{,}770\,1\ NkT \left(\frac{T}{T_B}\right)^{3/2} \quad \text{pour } T \leqslant T_B\ , \tag{C.1}$$

et donc aussi pour sa *chaleur spécifique à volume constant :*

$$c_V \widehat{=} \frac{N_A}{N} \left(\frac{\partial E}{\partial T}\right)_{N,V} = 1{,}925\ R \left(\frac{T}{T_B}\right)^{3/2} \quad \text{pour } T \leqslant T_B\ . \tag{C.2}$$

La *dérivée de la chaleur spécifique par rapport à la température* se calcule alors aisément; pour $T = T_B$, elle vaut

$$\left(\frac{\partial c_V}{\partial T}\right)_{N,V} \xrightarrow[T \to T_{B^-}]{} 2{,}888\ \frac{R}{T_B}\ . \tag{C.3}$$

II. DOMAINE $T > T_B$: DÉVELOPPEMENT DE L'ÉNERGIE AU VOISINAGE DE T_B

La situation est moins simple lorsque la température est supérieure à la température de Bose; c'est là que se situe en fait le problème que nous voulons résoudre.

L'énergie du système est ici donnée par (formule VI.(II.22))

$$E = AV \int_0^{+\infty} d\varepsilon \frac{\varepsilon^{3/2}}{e^{\beta(\varepsilon-\mu)}-1} \quad \text{pour } T \geqslant T_B , \tag{C.4}$$

où le potentiel chimique μ est déterminé par la relation VI.(II.13) :

$$AV \int_0^{+\infty} d\varepsilon \frac{\varepsilon^{1/2}}{e^{\beta(\varepsilon-\mu)}-1} = N \quad \text{pour } T \geqslant T_B . \tag{C.5}$$

Il n'est plus possible de calculer analytiquement ces intégrales, ni donc la chaleur spécifique c_V du système[1].

Mais nous savons que le potentiel chimique μ s'annule à la température de Bose (cf. VI, § II.C.1) :

$$\mu \longrightarrow 0_- \quad \text{quand } T \longrightarrow T_{B^+} . \tag{C.6}$$

Nous allons utiliser ce fait pour évaluer les intégrales (C.4) et (C.5) *au voisinage de la température de Bose,* en nous limitant *à l'ordre le plus bas en* μ.

Le calcul est simple pour l'énergie E (formule (C.4)), et donne[2]

$$E \simeq 0{,}770\,1\, NkT \left(\frac{T}{T_B}\right)^{3/2} + \frac{3}{2} N\mu \left(\frac{T}{T_B}\right)^{3/2} \quad \text{pour } T \geqslant T_B . \tag{C.7}$$

---

1. *Attention :* à N et V constants, *le potentiel chimique* μ *est fonction de la température* T à travers la relation implicite (C.5). Il faut en tenir compte si l'on veut calculer c_V à partir de l'expression (C.4) de l'énergie :

$$c_V = \frac{N_A}{N} \cdot \frac{AV}{kT^2} \int_0^{+\infty} d\varepsilon\, \varepsilon^{3/2} \left(\varepsilon - \mu - \beta \frac{d\mu}{d\beta}\right) \frac{e^{\beta(\varepsilon-\mu)}}{[e^{\beta(\varepsilon-\mu)}-1]^2} \quad \text{pour} \quad T \geqslant T_B .$$

2. La dérivée par rapport à μ du facteur de Bose se calcule facilement :

$$\frac{\partial}{\partial\mu} \frac{1}{e^{\beta(\varepsilon-\mu)}-1} = \beta \frac{e^{\beta(\varepsilon-\mu)}}{[e^{\beta(\varepsilon-\mu)}-1]^2} . \tag{1}$$

La formule (C.4) s'écrit donc, au *premier ordre en* μ,

$$E \simeq AV \int_0^{+\infty} d\varepsilon \frac{\varepsilon^{3/2}}{e^{\beta\varepsilon}-1} + \beta\mu AV \int_0^{+\infty} d\varepsilon \frac{\varepsilon^{3/2} e^{\beta\varepsilon}}{[e^{\beta\varepsilon}-1]^2} , \quad T \geqslant T_B . \tag{2}$$

La première intégrale est connue (VI, note 53). La seconde se transforme, par une intégration par parties, en

$$\int_0^{+\infty} d\varepsilon \frac{\varepsilon^{3/2} e^{\beta\varepsilon}}{[e^{\beta\varepsilon}-1]^2} = \frac{3}{2} \frac{1}{\beta} \int_0^{+\infty} d\varepsilon \frac{\varepsilon^{1/2}}{e^{\beta\varepsilon}-1} , \tag{3}$$

où l'on reconnaît à nouveau une intégrale déjà rencontrée (formule VI.(II.44)). En définitive, la formule (2) s'écrit explicitement

$$E \simeq 0{,}770\,1 NkT \left(\frac{T}{T_B}\right)^{3/2} + \frac{3}{2} \mu N \left(\frac{T}{T_B}\right)^{3/2} , \quad T \geqslant T_B . \tag{4}$$

L'évaluation de la formule (C.5) est moins facile, car la dérivée par rapport à μ de l'intégrale diverge en $\mu = 0$[3]. En analysant la difficulté de plus près, on trouve[4], toujours *à l'ordre le plus bas en* μ,

$$N \simeq N\left(\frac{T}{T_B}\right)^{3/2} - 1{,}357\ N\frac{kT}{(kT_B)^{3/2}}\sqrt{-\mu} \quad \text{pour} \quad T \geqslant T_B\ . \tag{C.8}$$

Cette relation s'inverse facilement, et donne le *comportement du potentiel chimique au voisinage de* $T = T_B$, *à* N *et* V *fixés :*

$$\mu \simeq -\frac{kT_B}{1{,}841}\left(\frac{T_B}{T}\right)^2\left[\left(\frac{T}{T_B}\right)^{3/2} - 1\right]^2 \quad \text{pour} \quad T \geqslant T_B\ . \tag{C.9}$$

On peut alors éliminer μ de la formule (C.7), pour obtenir l'expression (approchée) de l'énergie du système en fonction de T, N et V, pour T supérieure à T_B mais très voisine :

$$E \simeq 0{,}770\,1\ NkT\left(\frac{T}{T_B}\right)^{3/2} - 0{,}814\,6\ NkT_B\left(\frac{T_B}{T}\right)^{1/2}\left[\left(\frac{T}{T_B}\right)^{3/2} - 1\right]^2 \quad \text{pour} \quad T \geqslant T_B \tag{C.10}$$

(T_B étant donnée en fonction de N et V par la formule VI.(II.21)).

---

3. Si l'on dérive « naïvement » sous le signe somme, on obtient (cf. formule (1) de la note 2)

$$AV\beta\int_0^{+\infty} \mathrm{d}\varepsilon\ \frac{\varepsilon^{1/2}\,\mathrm{e}^{\beta(\varepsilon-\mu)}}{[\mathrm{e}^{\beta(\varepsilon-\mu)}-1]^2}\ . \tag{1}$$

Cette intégrale converge si $\mu \neq 0$, mais elle devient divergente à sa borne inférieure pour $\mu = 0$:

$$\frac{\varepsilon^{1/2}\,\mathrm{e}^{\beta\varepsilon}}{[\mathrm{e}^{\beta\varepsilon}-1]^2} \underset{\varepsilon\to 0}{\sim} \frac{1}{\beta^2}\frac{1}{\varepsilon^{3/2}}\ . \tag{2}$$

4. Écrivons l'intégrale à calculer sous la forme

$$AV\int_0^{+\infty} \mathrm{d}\varepsilon\ \frac{\varepsilon^{1/2}}{\mathrm{e}^{\beta(\varepsilon-\mu)}-1} = AV\int_0^{+\infty} \mathrm{d}\varepsilon\ \frac{\varepsilon^{1/2}}{\mathrm{e}^{\beta\varepsilon}-1} + AV\int_0^{+\infty} \mathrm{d}\varepsilon\ \varepsilon^{1/2}\left[\frac{1}{\mathrm{e}^{\beta(\varepsilon-\mu)}-1} - \frac{1}{\mathrm{e}^{\beta\varepsilon}-1}\right]. \tag{1}$$

Le premier terme est connu (note 2, formule (3)) :

$$AV\int_0^{+\infty} \mathrm{d}\varepsilon\ \frac{\varepsilon^{1/2}}{\mathrm{e}^{\beta\varepsilon}-1} = N\left(\frac{T}{T_B}\right)^{3/2} \tag{2}$$

C'est le deuxième qui pose problème, à cause de la divergence de l'intégrale à sa borne inférieure lorsque $\mu \to 0$ (note 3). Pour isoler la difficulté, nous scindons l'intégrale en deux :

$$AV\int_0^{+\infty} \mathrm{d}\varepsilon\ \varepsilon^{1/2}\left[\frac{1}{\mathrm{e}^{\beta(\varepsilon-\mu)}-1} - \frac{1}{\mathrm{e}^{\beta\varepsilon}-1}\right] = AV\left[\int_0^{\eta kT} + \int_{\eta kT}^{+\infty}\right]\mathrm{d}\varepsilon\ \varepsilon^{1/2}\left[\frac{1}{\mathrm{e}^{\beta(\varepsilon-\mu)}-1} - \frac{1}{\mathrm{e}^{\beta\varepsilon}-1}\right], \tag{3}$$

où η est un nombre sans dimension, choisi à l'avance tel que

$$\eta \ll 1\ . \tag{4}$$

Dans la formule décomposée (3), la deuxième intégrale ne présente plus de divergence quand $\mu \to 0$, puisque sa borne inférieure n'est plus nulle; elle est donc du premier ordre en μ :

$$AV\int_{\eta kT}^{+\infty} \mathrm{d}\varepsilon\ \varepsilon^{1/2}\left[\frac{1}{\mathrm{e}^{\beta(\varepsilon-\mu)}-1} - \frac{1}{\mathrm{e}^{\beta\varepsilon}-1}\right] = \mathcal{O}(\mu)\ . \tag{5}$$

.../...

III. RACCORDEMENT ENTRE LES DEUX DOMAINES

Dans l'expression (C.10), le deuxième terme s'annule en $T = T_B$. On retrouve le fait que *l'énergie est continue* en $T = T_B$:

$$E(T_B) = 0{,}770\,1\ NkT_B\ , \tag{C.11}$$

valeur que l'on obtient tant à partir de (C.1) que de (C.10).

Mais la dérivée première, par rapport à T, du second terme de (C.10) s'annule également pour $T = T_B$, de sorte que *la chaleur spécifique à volume constant est elle aussi continue :*

$$c_V(T) \simeq \frac{5}{2} \times 0{,}770\,1 R \left(\frac{T}{T_B}\right)^{3/2} - 0{,}814\,6\ R \left[\left(\frac{T}{T_B}\right)^{3/2} - 1\right]\left[\frac{5}{2} + \frac{1}{2}\left(\frac{T_B}{T}\right)^{3/2}\right]$$

$$\text{pour} \quad T \geqslant T_B\ , \tag{C.12}$$

ce qui donne

$$c_V(T_B) = 1{,}925\ R \tag{C.13}$$

comme cela avait été trouvé plus haut (formule (C.2)).

Finalement, c'est au niveau de la *dérivée seconde de* E par rapport à T, c'est-à-dire la *dérivée première de* c_V, que se manifeste une *discontinuité,* c'est-à-dire une différence entre les valeurs obtenues, pour $T = T_B$, à partir des expressions (C.1) et (C.10). On trouve en effet, en dérivant une nouvelle fois la formule (C.12),

$$\left(\frac{\partial c_V}{\partial T}\right)_{N,\ V} \xrightarrow[T \to T_{B+}]{} -0{,}777\,6\ \frac{R}{T_B}\ , \tag{C.14}$$

à comparer avec (C.3). On constate que les deux valeurs limites sont cette fois différentes, et qu'elles ont même des *signes opposés;* ceci explique la structure en pointe de la courbe représentée sur la figure 10 du chapitre VI.

Quant à la première intégrale de la décomposition (3), sa borne supérieure a été choisie (inégalité (4)) de telle sorte que

$$\beta\varepsilon \leqslant \eta \ll 1 \tag{6}$$

dans tout l'intervalle d'intégration. On peut donc y développer les deux exponentielles, ce qui donne

$$AV \int_0^{\eta kT} \mathrm{d}\varepsilon\ \varepsilon^{1/2}\left[\frac{1}{\mathrm{e}^{\beta(\varepsilon-\mu)}-1} - \frac{1}{\mathrm{e}^{\beta\varepsilon}-1}\right] \simeq AV\,\frac{\mu}{\beta}\int_0^{\eta kT} \mathrm{d}\varepsilon\ \frac{\varepsilon^{1/2}}{\varepsilon(\varepsilon-\mu)}\ . \tag{7}$$

μ étant négatif, l'intégrale est classique (on pose $y \widehat{=} \sqrt{\varepsilon}$) :

$$\begin{aligned} AV\frac{\mu}{\beta}\int_0^{\eta kT} \mathrm{d}\varepsilon\ \frac{1}{\sqrt{\varepsilon}(\varepsilon + |\mu|)} &= AV\frac{\mu}{\beta}\int_0^{\sqrt{\eta kT}} 2\ \mathrm{d}y\ \frac{1}{y^2 + |\mu|} \\ &= AV\frac{\mu}{\beta}\,\frac{2}{\sqrt{|\mu|}}\ \mathrm{Arctg}\ \sqrt{\frac{\eta kT}{|\mu|}} \\ &= AV\frac{\mu}{\beta}\,\frac{2}{\sqrt{|\mu|}}\left[\frac{\pi}{2} + \mathcal{O}\left(\sqrt{|\mu|}\right)\right] \\ &= -AV\pi kT\sqrt{|\mu|} + \mathcal{O}(\mu)\ . \end{aligned} \tag{8}$$

En regroupant (2), (5) et (8), on trouve finalement

$$AV\int_0^{+\infty} \mathrm{d}\varepsilon\ \frac{\varepsilon^{1/2}}{\mathrm{e}^{\beta(\varepsilon-\mu)}-1} \simeq N\left(\frac{T}{T_B}\right)^{3/2} - AV\pi kT\sqrt{|\mu|}\ , \quad T \geqslant T_B\ , \tag{9}$$

résultat dans lequel on peut remplacer AV en fonction de N et T_B (formule VI (II.21)).

COMPLÉMENT VI.D

L'hélium à basse température

I. Transitions de phase des deux isotopes de l'hélium
   1. Liquéfaction et solidification de l'hélium
   2. L'hélium 3 et l'effet Pomerantchuk
   3. L'hélium 4 et la « transition λ »
      a. Diagramme de phases de l'hélium 4
      b. Propriétés de l'hélium II

II. Interprétation théorique des propriétés de l'hélium II
   1. Considérations qualitatives et semi-quantitatives
      a. Comparaison avec le gaz de bosons sans interactions
      b. Excitations collectives et quasi-particules
      c. Le critère de superfluidité
   2. Mécanique statistique des quasi-particules : le modèle de Landau
      a. Relation de dispersion des excitations élémentaires
      b. Propriétés thermodynamiques du système
   3. Interprétation du modèle des deux fluides
      a. Définition de la densité normale
      b. Expression de la densité normale
      c. Contribution des phonons et des rotons à la densité normale
   4. Conclusion

Il existe *deux isotopes* stables de l'hélium : le plus courant, l'*hélium* 4, dont le noyau est constitué de deux protons et deux neutrons, est un *boson*[1]; l'*hélium* 3, beaucoup plus rare mais que l'on parvient néanmoins à produire en quantité suffisante pour l'étudier séparément, comporte un seul neutron, ce qui en fait un *fermion*[1]. Tous deux manifestent à basse température des propriétés très particulières.

I. TRANSITIONS DE PHASE DES DEUX ISOTOPES DE L'HÉLIUM

1. LIQUÉFACTION ET SOLIDIFICATION DE L'HÉLIUM

L'hélium est le seul corps qui ne se solidifie pas sous la pression atmosphérique. Il est gazeux à température ordinaire. L'hélium 4 se liquéfie à 4,2 K sous la pression atmosphérique, l'hélium 3 à 3,2 K. Ils restent ensuite liquides jusqu'aux plus basses températures que l'on ait pu atteindre, et donc vraisemblablement jusqu'au zéro absolu. Ce n'est qu'en appliquant des pressions supérieures à 25 atmosphères pour l'hélium 4, et 30 atmosphères pour l'hélium 3, que l'on arrive à les solidifier.

1. L'hélium, deuxième élément de la classification de Mendeleïeff, possède deux électrons, et donc deux protons dans son noyau. Lorsque ceux-ci sont accompagnés de deux neutrons (hélium 4), l'ensemble de ces six fermions constitue un boson. L'hélium 3 est au contraire un fermion, car il comporte un neutron de moins et donc un nombre impair de particules de spin 1/2. Dans l'état fondamental de l'atome d'hélium, le moment cinétique orbital des deux électrons est nul (couche 1s) ainsi que leur spin total pour vérifier le principe de Pauli; le spin du noyau d'hélium 4 est aussi nul, celui du noyau d'hélium 3 est égal à 1/2.

Cette propriété remarquable est d'*origine quantique;* on peut la comprendre semi-quantitativement de façon assez simple.

Comme tous les autres gaz rares, l'hélium reste toujours monoatomique. L'interaction entre deux atomes d'hélium est due aux forces de van der Waals; l'énergie potentielle correspondante a la forme que nous avons déjà rencontrée à plusieurs reprises (complément III.G, fig. 1); mais ici le minimum de ce potentiel, qui se produit pour une distance d'un ordre de grandeur habituel :

$$\rho_0 \simeq 3 \text{ Å} , \tag{D.1}$$

est *très peu profond :*

$$u_0 \simeq 7 \times 10^{-4} \text{ eV} . \tag{D.2}$$

Bien entendu, le problème est plus compliqué pour un cristal : chaque atome y est soumis à l'influence de plusieurs voisins, et la configuration d'équilibre classique (minimum de l'énergie potentielle totale) est beaucoup plus difficile à déterminer que pour un système de deux atomes seulement. Mais il nous suffira de considérer qu'un atome particulier du cristal est situé au voisinage immédiat du minimum de l'énergie potentielle due aux autres atomes; on s'attend à ce que la distance de l'atome considéré à ses voisins et la profondeur du puits de potentiel où il se trouve gardent les ordres de grandeur de l'interaction à deux corps.

Imaginons alors un cristal atomique de gaz rare. Les atomes n'y sont pas immobiles : ils vibrent autour de leur position d'équilibre, avec une fréquence dont l'ordre de grandeur est déterminé par leur *masse m* et la *courbure du potentiel* $u(\rho)$ en son minimum ρ_0 (cf. complément III.B, § II.2)[2] :

$$\omega \sim \sqrt{\frac{\varkappa}{m}} , \tag{D.3}$$

où $$\varkappa = \frac{\mathrm{d}^2 u}{\mathrm{d}\rho^2}(\rho_0) . \tag{D.4}$$

L'amplitude des vibrations croît avec la température. Lorsqu'elle n'est plus négligeable devant les distances interatomiques dans le cristal, celui-ci n'est plus stable et fond[3].

L'hélium est particulier parce que *sa masse est petite et le potentiel interatomique très faible.* Ces deux facteurs jouent dans le même sens pour empêcher la solidification de l'hélium. En effet, ρ_0 restant du même ordre de grandeur pour tous les atomes, c'est la profondeur u_0 du puits de potentiel qui régit principalement la valeur de la dérivée seconde (D.4). Or, même dans l'état fondamental, la fonction d'onde de l'oscillateur harmonique quantique a une *extension minimum non nulle,* donnée par (C.D.L., p. 502)

$$\Delta\rho = \sqrt{\frac{\hbar}{2m\omega}} . \tag{D.5}$$

2. Nous utilisons un modèle simplifié, dit « modèle d'Einstein » : nous supposons que les divers atomes du cristal vibrent de façon indépendante, ce qui n'est qu'une approximation assez grossière (voir complément III.E, § I).

3. Le *critère empirique de Lindemann,* bien vérifié pour de nombreux corps, considère que la fusion se produit lorsque l'amplitude des vibrations des constituants atteint environ 1/10 de la distance qui sépare les plus proches voisins.

On voit facilement, par un calcul d'ordre de grandeur[4], que cette « amplitude minimum des vibrations » est, pour l'hélium, comparable à la distance interatomique ρ_0; donc, en quelque sorte, *« le cristal d'hélium fond avant de s'être formé »*. Si l'on augmente considérablement la pression, on force les atomes à se rapprocher les uns des autres malgré la forte répulsion qui s'exerce entre eux dès que leurs couches électroniques viennent au contact; on « resserre » ainsi le potentiel u, ce qui augmente sa courbure et stabilise donc le cristal.

Les autres gaz rares se solidifient à la pression atmosphérique : les valeurs de u_0 sont comparables à celle de l'hélium, mais les masses sont beaucoup plus élevées. Quant à l'hydrogène, il peut aussi être solidifié à toute pression car, bien que plus légères que les atomes d'hélium, les molécules H_2 s'attirent beaucoup plus fortement.

2. L'HÉLIUM 3 ET L'EFFET POMERANTCHUK

Le *diagramme de phases* de l'hélium 3 est représenté sur la figure 1[5]. La courbe inférieure *(« courbe de vaporisation »)*, terminée par le point critique C, correspond à la transition liquide-gaz (complément III.G, § III). La courbe supérieure *(« courbe de fusion »)*, frontière entre les domaines de stabilité du liquide et du solide, est l'ensemble des points où ces deux phases coexistent à l'équilibre. On constate que le solide n'existe que pour des pressions supérieures à 29 atmosphères environ. La figure 2 (p. 884) donne, pour comparaison, l'allure d'un diagramme de phases « normal ».

On remarque en outre un comportement surprenant de la courbe de fusion : la fonction $p(T)$ correspondante présente un *minimum*, ce qui n'est observé pour aucun autre corps. Cette partie de la courbe est agrandie sur la figure 1.b. Imaginons alors

---

4. On peut évaluer la « constante de rappel » $\varkappa$ à partir des données (D.1)-(D.2) :

$$\varkappa \sim \frac{u_0}{\rho_0^2} \sim 10^{-3}\ \mathrm{J/m^2}\ . \qquad (1)$$

L'extension minimale $\Delta\rho$ qui s'en déduit par la formule (D.5) ($m \simeq 1{,}7 \times 10^{-27}$ kg $\times 3$ pour l'hélium 3 ou $\times 4$ pour l'hélium 4) vaut

$$\Delta\rho \sim 1\ \text{Å}\ . \qquad (2)$$

On peut aussi comparer l'énergie cinétique de vibration qu'aurait un atome d'hélium dans le cristal, à son énergie potentielle. D'après la relation d'incertitude

$$\Delta x\,.\,\Delta p \geqslant \hbar\ , \qquad (3)$$

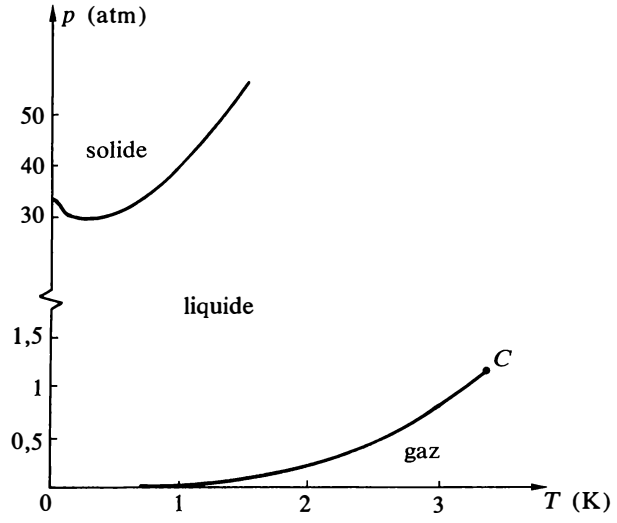
un atome d'hélium localisé à $\Delta x \sim 1$ Å près aurait une énergie cinétique minimale

$$\Delta E_c \sim \frac{(\Delta p)^2}{2m} \geqslant \frac{1}{2m}\left(\frac{\hbar}{\Delta x}\right)^2 \sim 6 \times 10^{-4}\ \mathrm{eV}\ , \qquad (4)$$

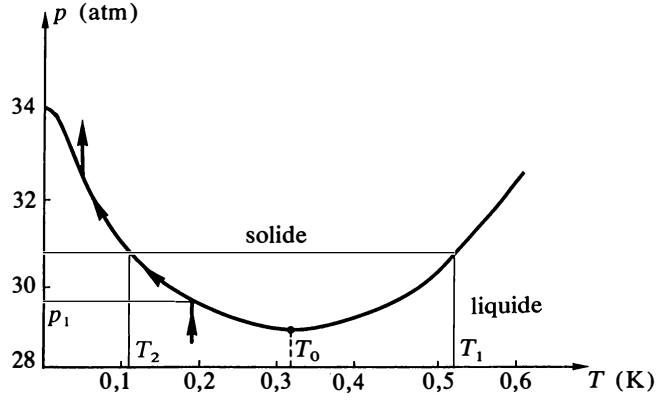
c'est-à-dire du même ordre que la profondeur (D.2) du puits de potentiel.

5. On a récemment mis en évidence, dans la région des températures inférieures au millikelvin, d'autres phases de l'hélium 3 possédant des propriétés intéressantes et curieuses (superfluidité, anisotropie, structure magnétique,...). Ces nouvelles phases sont dues à la formation de *« paires de Cooper »* : comme les électrons dans les métaux supraconducteurs, les atomes d'hélium 3 peuvent s'apparier à très basse température; un ensemble de deux fermions liés constituant un boson, on voit alors apparaître des phénomènes caractéristiques des systèmes de bosons, analogues à ceux que nous allons décrire plus loin pour l'hélium 4. Voir R. Combescot, « L'hélium 3 superfluide », *La Recherche 95*, p. 1064 (1978); G.E. Volovik, « Vortices in Rotating Superfluid ^{3}He », *Europhysics News* (Bulletin de la Société Européenne de Physique), vol. *15*, n° 6 (juin 1984), p. 13.

que, à *pression constante* (supérieure à 29 atmosphères), nous refroidissions de l'hélium 3 liquide, à partir d'une température de l'ordre de 1 K. Lorsque la température atteint la valeur T_1, un petit cristal d'hélium 3 apparaît. Si l'on continue à retirer de la chaleur au système, ce cristal grossit au détriment du liquide; tant que solide et liquide sont en présence à l'équilibre, la température reste égale à T_1. Après que la dernière goutte de liquide a disparu, la température du système devenu entièrement solide peut recommencer à décroître. Jusqu'ici, rien que de très habituel. Mais un phénomène analogue se produit en sens inverse pour $T = T_2$: *si l'on retire de la chaleur au système, le solide se transforme à nouveau progressivement en liquide.*



(a)



(b)

FIGURE 1
Diagramme de phases de l'hélium 3. La figure 1.b détaille la région du minimum de la courbe de fusion.

C'est le comportement en $T = T_2$ qui est exceptionnel : il implique que, *dans ce domaine de température et de pression, l'entropie de l'hélium* 3 *liquide est inférieure à celle du solide*[6], contrairement à ce qui se passe pour tous les autres corps.

On peut comprendre simplement, au moins de façon qualitative, l'origine du signe inhabituel de la différence d'entropie entre le liquide et le solide en s'appuyant sur les résultats du chapitre VI (§ I) concernant les gaz de *fermions*. L'entropie S_g d'un gaz de fermions sans interactions tend vers zéro avec la température T, proportionnellement à celle-ci :

$$S_g \underset{T \to 0}{\sim} \frac{\pi^2}{2} Nk \frac{T}{T_F} , \tag{D.6}$$

où T_F est la température de Fermi. Dans l'hélium 3 *liquide,* les interactions ne sont pas négligeables comme dans un gaz, et ceci d'autant moins que la température T est plus basse, c'est-à-dire l'énergie caractéristique kT plus faible. Mais le résultat (D.6) découle directement du *principe de Pauli,* qui régit de façon analogue l'hélium 3 liquide : dans un gaz de fermions, chaque état d'impulsion individuelle déterminée ne peut accueillir que deux particules, et encore *à condition que leurs spins soient « antiparallèles »;* pour s'entasser dans les états de plus basse énergie possible, lorsque la température devient très inférieure à T_F, les fermions du gaz doivent donc orienter leurs spins deux à deux de façon à former un état de spin nul. En d'autres termes, *l'expression* (D.6) *de l'entropie tient compte des degrés de liberté de spin du système.* Pour l'hélium 3, la température de Fermi vaut environ[7]

$$T_F \simeq 4{,}9 \text{ K} . \tag{D.7}$$

---

6. Lors d'une transition de phase du premier ordre comme le sont les transitions solide-liquide et liquide-gaz (voir complément III.G, § III.4), il faut fournir de la chaleur pour faire passer le corps considéré, *à T et p constantes* et de façon réversible, de la phase de basse température à celle de haute température. Cette *chaleur latente de transformation* explique pourquoi, si l'on chauffe le système à pression constante, sa température ne croît pas tant que coexistent les deux phases : la chaleur fournie sert à augmenter la proportion de la phase de haute température au détriment de celle de basse température. Ici, étant donné le domaine considéré, nous avons envisagé de refroidir le système, et non pas de le chauffer, mais cela ne change rien à l'argument : la « phase de basse température » (au voisinage de T_2) est le liquide, celle de « haute température » le solide.

La chaleur latente L est *proportionnelle à la différence entre les entropies* S_h *et* S_b du système dans la phase de haute température et celle de basse température (complément III.G, § III.4.b) :

$$L = T(S_h - S_b) . \tag{1}$$

Elle vérifie également la *« relation de Clapeyron »*

$$L = T(V_h - V_b) \left(\frac{dp}{dT}\right)_{éq} , \tag{2}$$

où V_h est le volume occupé par le système dans la phase de haute température et V_b celui de la phase de basse température, et $(dp/dT)_{éq}$ est la pente de la courbe d'équilibre (exercice III.19). Ici, cette pente est visiblement négative en T_2; mais le volume V_h du solide est inférieur au volume V_b de la même masse de liquide, de sorte que L est encore positive : d'après (1), l'entropie S_h du solide est donc supérieure, au voisinage de T_2, à celle S_b du liquide.

7. Il suffit d'utiliser la formule VI.(I.68) reliant la température de Fermi à la densité N/V, que l'on évalue à partir de la masse volumique de l'hélium 3 liquide (0,0826 g/cm^3) :

$$T_F = \frac{1}{k} \frac{\hbar^2}{2m} \left(3\pi^2 \frac{N}{V}\right)^{2/3} \simeq 4{,}9 \text{ K} .$$

Dans l'état solide au contraire, les atomes d'hélium 3 sont localisés dans l'espace aux nœuds d'un réseau cristallin. Les sites qu'ils occupent ainsi étant *discernables,* le principe de Pauli ne restreint plus leurs états de spin. Comme il s'agit de spins nucléaires, leurs interactions sont extrêmement faibles; même aux températures de quelques dixièmes de kelvins, les deux états de spin de chacun des atomes ont encore la même probabilité[8]. On a donc au total 2^N états de spin également probables, qui donnent à l'entropie une contribution

$$S_s = Nk \ln 2 \; ; \tag{D.8}$$

cette contribution reste présente dans l'entropie du solide jusqu'à des températures extrêmement basses (note 8). On comprend ainsi que l'entropie du liquide, qui ne comporte pas de terme de ce type et décroît régulièrement avec la température, puisse être, dans un certain domaine, inférieure à celle du solide[9]. Les valeurs numériques déduites des formules (D.6) et (D.8) sont effectivement du bon ordre de grandeur :

$$S_g \leqslant S_s \implies T \leqslant 0{,}14 T_F \simeq 0{,}67 \text{ K} \; . \tag{D.9}$$

Cet argument ne peut évidemment que rendre le phénomène plausible; une véritable explication théorique exigerait un raisonnement beaucoup plus élaboré. Pourtant l'effet avait été prédit théoriquement par Pomerantchuk en 1950, bien avant que l'expérience n'en confirme l'existence en précisant le domaine où il se produit.

L'effet Pomerantchuk est à la base d'une méthode d'*obtention de très basses températures*[10]; cette méthode permet d'amener des échantillons macroscopiques d'hélium 3 jusque dans le domaine du millikelvin.

---

8. Le moment magnétique associé à un spin nucléaire a pour ordre de grandeur le « magnéton nucléaire » (m_p est la masse du proton et e sa charge)

$$\mu_n \widehat{=} \frac{e\hbar}{2m_p} \simeq 5 \times 10^{-27} \text{ J/tesla} \; . \tag{1}$$

L'interaction entre deux tels moments distants de $\rho = 1$ Å donne lieu à une énergie potentielle d'environ

$$\frac{\mu_0}{4\pi} \frac{\mu_n^2}{\rho^3} \simeq 1{,}5 \times 10^{-11} \text{ eV} \; , \tag{2}$$

ce qui correspond à une température de 2×10^{-7} K. Le résultat est tellement petit que d'autres mécanismes d'interaction devraient être envisagés. Mais cette évaluation montre que les spins nucléaires ne commencent à se coupler qu'à des températures extrêmement basses.

9. On a pu réaliser expérimentalement des échantillons d'*hélium 3 polarisé* dans lesquels les spins sont au contraire tous parallèles. Les propriétés de tels échantillons diffèrent assez nettement de celles de l'hélium 3 non polarisé. En particulier, l'effet Pomerantchuk disparaît puisque la contribution (D.8) à l'entropie du solide n'existe plus si les spins sont orientés. Voir B. Castaing, « Les fluides quantiques polarisés », *Bulletin de la Société Française de Physique 48,* p. 19 (avril 1983).

10. On prépare d'abord un échantillon d'hélium 3 liquide à une température T inférieure à la position $T_0 \simeq 0{,}32$ K du minimum de la courbe de fusion. On comprime ensuite cet échantillon (trajet fléché de la figure 1.b). Lorsque son point représentatif atteint la courbe de fusion (en p_1), l'hélium 3 commence à se solidifier. On l'isole alors du thermostat, de façon que sa température puisse varier en même temps que la pression. Si celle-ci continue à augmenter, le point représentatif du système diphasé suit la courbe de fusion, jusqu'à ce que la solidification soit complète; comme la pente de la courbe est négative dans cette région, *l'augmentation de pression provoque une diminution de la température.*

3. L'HÉLIUM 4 ET LA « TRANSITION λ »

a. Diagramme de phases de l'hélium 4

La figure 3 représente le diagramme de phases de l'isotope 4 de l'hélium, lui aussi très différent d'un diagramme « normal » (fig. 2). Les courbes de vaporisation et

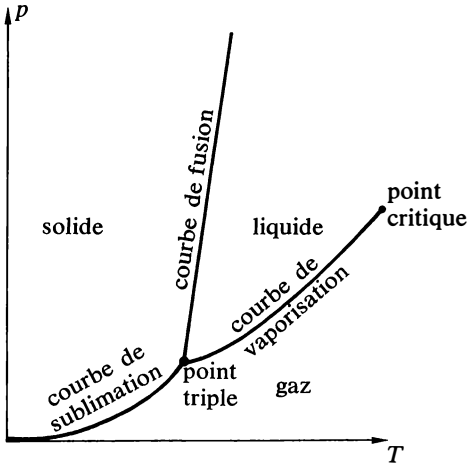


FIGURE 2
Allure du diagramme de phases pour un corps pur autre que l'hélium.

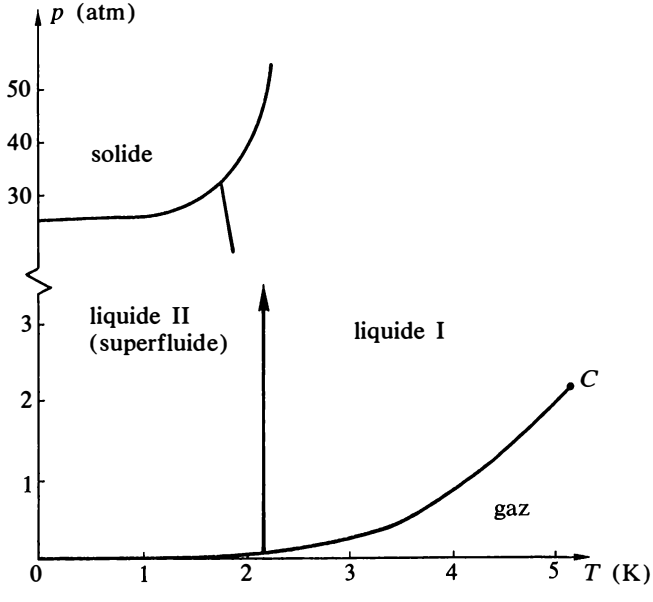


FIGURE 3
Diagramme de phases de l'hélium 4.

de fusion sont analogues à celles de l'hélium 3; les coordonnées du point critique et la valeur minimale de la pression permettant d'obtenir le solide [11] sont légèrement différentes, mais surtout *la courbe de fusion* est quasiment plate au voisinage de $T = 0$ [12].

La principale différence avec l'hélium 3, fondamentale celle-là, réside dans l'existence d'une *nouvelle transition de phase,* appelée « transition λ ». Elle fait passer d'une phase liquide de type habituel, dite *« hélium I »,* à *une autre phase liquide, l'« hélium II »,* possédant des propriétés étonnantes (voir paragraphe b). Cette transition n'est pas du même type que la vaporisation-liquéfaction ou la solidification-fusion : *les deux phases ne coexistent jamais,* elles ont *même volume molaire et même entropie molaire* au point de transition, de sorte qu'il n'y a *pas de chaleur latente.* Les points de transition constituent dans le plan (T, p) une courbe de transition, mais celle-ci trace seulement *la frontière entre les domaines de stabilité* des deux phases, ces deux domaines étant *strictement séparés* puisque les deux phases ne sont jamais en équilibre l'une avec l'autre. Une telle transition est dite *« du second ordre »* [13].

Sur la courbe de vaporisation, la transition λ de l'hélium 4 se produit à la température

$$T_\lambda = 2{,}17 \text{ K} \qquad \text{(D.10)}$$

(cette température varie peu avec la pression, comme le montre la figure 3). La transition se manifeste de façon spectaculaire sur la chaleur spécifique c_V (fig. 4), qui devient singulière lorsque T tend vers T_λ; c'est la forme de cette courbe qui a suggéré le nom de « transition λ ».

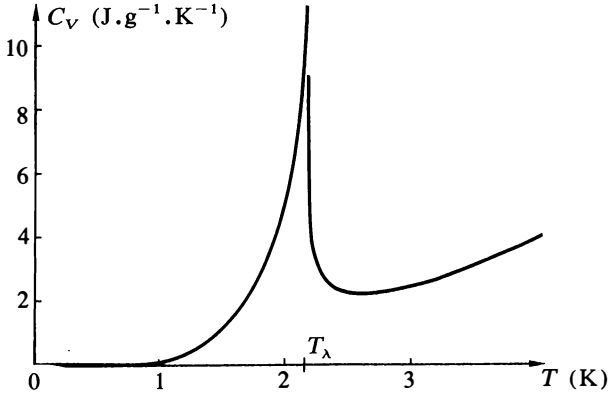


FIGURE 4
Chaleur spécifique c_v de l'hélium 4 liquide en fonction de la température.

11. On peut noter que cette différence va dans le sens prévu par l'argument du paragraphe 1 : l'hélium 4, plus lourd que l'hélium 3, est un peu plus facile à solidifier.

12. La figure 3 montre que sa pente est pratiquement nulle pour les températures inférieures à 1 K. On déduit alors de la relation de Clapeyron (note 6) que la chaleur latente est pratiquement nulle, c'est-à-dire que l'entropie du liquide et celle du solide sont pratiquement égales. En fait la courbe présente un minimum peu marqué à $T = 0{,}8$ K (cf. référence 5).

13. Les transitions du second ordre ont fait l'objet depuis une quinzaine d'années d'une étude très détaillée, tant théorique qu'expérimentale. Comme la transition ferro-paramagnétique (complément III.J), la transition λ peut être caractérisée par un ensemble de *« coefficients critiques »* définis à partir du comportement singulier de certaines grandeurs physiques à l'approche du point de transition. La transition λ et la transition ferro-paramagnétique appartiennent à des

.../...

b. Propriétés de l'hélium II

Pour présenter succinctement les principales propriétés de l'hélium II, nous allons d'abord introduire le *modèle phénoménologique des deux fluides,* qui en donne une description synthétique en les faisant apparaître comme conséquences d'un petit nombre d'hypothèses simples.

α. Le modèle des deux fluides

Ce modèle, dû à Tisza, représente l'hélium II comme un *mélange idéal de deux fluides,* de caractéristiques très différentes. Le premier est un *liquide normal,* présentant en particulier une certaine viscosité. L'autre composante, appelée *superfluide,* possède des propriétés extraordinaires : *sa viscosité est nulle,* c'est-à-dire qu'elle peut s'écouler sans opposer de résistance à travers des tubes extrêmement fins (de 0,1 mm de diamètre, ou même moins), et *son entropie est nulle.* Les deux fluides sont présents dans tout le volume du système; la masse volumique ρ de l'hélium II est ainsi la somme

$$\rho = \rho_n + \rho_s \qquad \text{(D.11)}$$

des masses volumiques ρ_n et ρ_s des deux fluides; de même, si l'hélium II est en mouvement par rapport au repère choisi, la densité volumique d'impulsion $\vec{j}$ est la somme des contributions des deux fluides :

$$\vec{j} = \rho_n \vec{v}_n + \rho_s \vec{v}_s \,, \qquad \text{(D.12)}$$

où $\vec{v}_n$ et $\vec{v}_s$ sont les vitesses de la composante normale et de la composante superfluide au point considéré. La masse volumique globale ρ dépend peu de la température. Par contre, les proportions ρ_n/ρ et ρ_s/ρ de fluide normal et de superfluide varient fortement avec la température : ρ_n/ρ tend vers zéro lorsque T tend vers zéro, et vers 1 lorsque T tend vers T_λ par valeurs inférieures (ρ_s/ρ étant égal à $1 - \rho_n/\rho$); une détermination expérimentale de ρ_n/ρ par des mesures de viscosité et de vitesse du « second son » (§ δ) donne la courbe représentée sur la figure 5. Pour $T > T_\lambda$, domaine de l'hélium I, ρ_s est identiquement nulle et ρ_n égale à ρ : l'hélium I est un liquide normal.

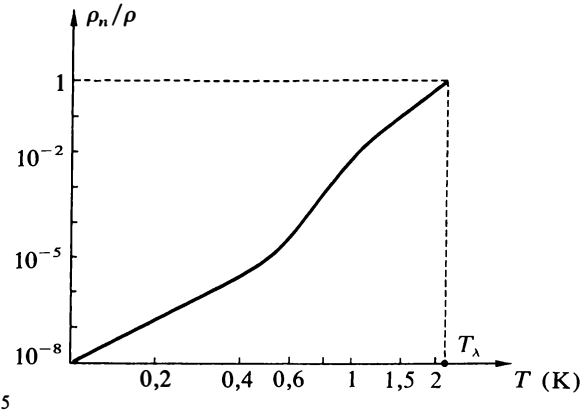


FIGURE 5

Variation avec la température de la proportion ρ_n/ρ de « fluide normal » dans l'hélium II, déterminée expérimentalement (noter la double échelle logarithmique). Pour $T \leqslant 0{,}5$ K, ρ_n/ρ varie approximativement comme T^4; pour $1{,}3 \leqslant T \leqslant T_\lambda$, son comportement empirique est en $(T/T_\lambda)^{5{,}3}$.

« classes d'universalité » différentes, c'est-à-dire que leurs coefficients critiques analogues ont des valeurs numériques différentes. L'existence et la caractérisation des classes d'universalité ont été récemment comprises dans le cadre de la théorie quantique des champs (méthode du « groupe de renormalisation » de K. Wilson). On a même pu calculer ainsi avec une grande précision les valeurs des coefficients critiques dans chaque classe, et obtenir des prédictions théoriques en excellent accord avec les mesures expérimentales (voir les références du complément III.J, note 49).

Avant de poursuivre, insistons sur le fait que *l'hélium II n'est pas un mélange :* il comprend une seule sorte d'atomes, indiscernables les uns des autres, et constitue une phase unique[14]. Mais il se comporte phénoménologiquement dans beaucoup de circonstances *comme si* c'était un mélange : le modèle des deux fluides permet d'interpréter et de prévoir de nombreuses propriétés de l'hélium II, dont nous allons donner quelques exemples marquants.

β. Viscosité de l'hélium II

Ce sont les caractéristiques très particulières de la viscosité de l'hélium II qui ont suggéré le modèle des deux fluides. Dans certaines circonstances en effet, cette viscosité est rigoureusement nulle : l'hélium II s'écoule à travers un tube capillaire sans opposer la moindre résistance, c'est-à-dire sans qu'il soit nécessaire d'exercer en amont du tube une pression supérieure à celle qui règne en aval (fig. 6)[15], comme il est indispensable pour tous les autres fluides. Dans ces circonstances, *le superfluide est le seul à franchir le tube,* ce qu'il fait sans frottements puisque sa viscosité est nulle par hypothèse; mais, au fur et à mesure de son écoulement, les proportions de fluide normal et de superfluide se rétablissent dans les deux récipients aux valeurs que déterminent la température T et la pression p extérieures. Au contraire une certaine viscosité, faible mais non nulle, se manifeste lorsqu'on déplace un objet dans l'hélium II : si par exemple (fig. 7) on fait glisser dans son plan une plaque en contact avec l'hélium II, elle entraîne les couches de liquide proches avec une vitesse qui décroît progressivement lorsqu'on s'éloigne de la plaque. *C'est alors le liquide normal qui entre en jeu* (le superfluide n'est pas entraîné par la plaque) : comme dans tout autre fluide, les différences de vitesse entre couches contiguës donnent lieu à des frottements visqueux, qui freinent le mouvement et dissipent de l'énergie mécanique en chaleur.

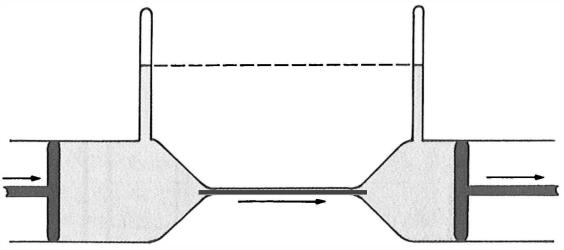


FIGURE 6
Écoulement superfluide de l'hélium II à travers un tube capillaire. Les deux tubes verticaux schématisent deux manomètres qui indiquent l'égalité des pressions en amont et en aval du capillaire.

14. Lorsque deux phases distinctes d'un corps pur unique sont en équilibre, il existe nécessairement une relation entre la température et la pression, c'est-à-dire que le point représentatif du système se situe, dans le plan (T, p), sur une courbe d'équilibre (complément III.G, § III.4.a). L'hélium II est au contraire observé dans tout un domaine du plan (T, p) (fig. 3).

15. En réalité, une viscosité non nulle apparaît brusquement si la vitesse d'écoulement dépasse un certain seuil (quelques mm/s), qui dépend des conditions de l'expérience. Nous reviendrons plus loin (§ II.1.c) sur cette propriété importante, qui sort du cadre très simple que nous décrivons ici.

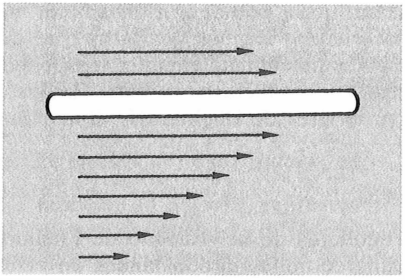


FIGURE 7
L'hélium II, comme tous les fluides, manifeste une certaine viscosité lorsqu'on déplace un objet avec lequel il est en contact.

γ. Effet fontaine et effet thermomécanique

Un autre exemple intéressant est fourni par deux effets d'origine commune, l'« effet fontaine » et l'« effet thermomécanique ». Construisons deux récipients (de volumes V_1 et V_2) que nous remplirons d'hélium II, et relions-les par une « superfuite », c'est-à-dire par un tube fin bourré d'une substance poreuse qui arrête le fluide normal mais laisse passer le superfluide (fig. 8). Nous avons alors un système global $\mathcal{S}$ constitué de *deux sous-systèmes $\mathcal{S}_1$ et $\mathcal{S}_2$ qui échangent seulement du superfluide, c'est-à-dire des particules mais pas d'entropie.* On montre (exercice II.20) que la valeur (la plus

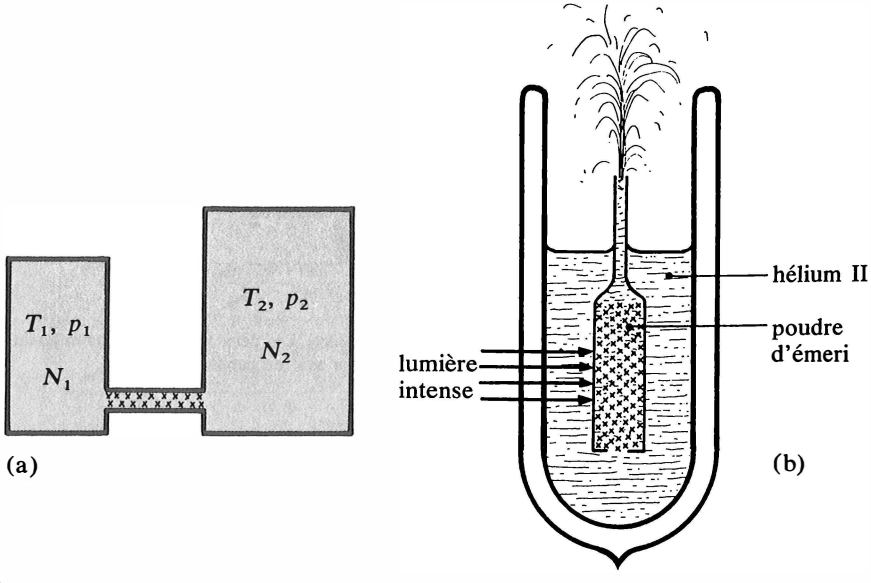


FIGURE 8
a) Deux échantillons d'hélium II reliés par une « superfuite » échangent des particules mais pas d'entropie.

b) L'effet fontaine : on chauffe, par exemple à l'aide d'un faisceau lumineux intense, l'hélium II remplissant les interstices de la poudre d'émeri; la surpression qui s'ensuit fait jaillir l'hélium jusqu'à des hauteurs de l'ordre de 30 cm.

probable) d'une variable interne s'obtient dans un tel cas en recherchant le *minimum de l'énergie E* du système global $\mathcal{S}$, *à entropie totale S, volume total V et nombre total N de particules fixés.* Il est facile d'en déduire[16] que la seule condition d'équilibre entre les sous-systèmes $\mathcal{S}_1$ et $\mathcal{S}_2$ est *l'égalité de leurs potentiels chimiques*[17] :

$$\mu_1 = \mu_2 \ . \tag{D.13}$$

Mais les deux sous-systèmes sont constitués du même corps pur, l'hélium II. Prenons comme variables la température T, la pression p et le nombre N de particules; le potentiel chimique, grandeur intensive, ne peut pas dépendre de N, seule variable extensive de l'ensemble (voir complément III.D, note 10) :

$$\mu = \mu(T, p) \ . \tag{D.14}$$

La condition d'équilibre (D.13) s'écrit donc alors

$$\mu(T_1, p_1) = \mu(T_2, p_2) \ . \tag{D.15}$$

L'équilibre est évidemment réalisé pour $T_1 = T_2$ et $p_1 = p_2$. Mais ce n'est pas la seule possibilité : *les deux sous-systèmes peuvent avoir des températures T_1 et T_2 différentes et des pressions p_1 et p_2 différentes, pourvu qu'elles vérifient la condition* (D.15). Notons que ces possibilités inhabituelles apparaissent parce qu'ici l'échange de particules ne s'accompagne d'aucun échange d'entropie, de sorte que *l'égalité des températures n'est pas imposée à l'équilibre* (voir note 16). On connaît (complément III.F, formules (F.13)) les dérivées de μ par rapport à T et p :

$$\frac{\partial \mu}{\partial T} = -\frac{S}{N} \ ; \quad \frac{\partial \mu}{\partial p} = \frac{V}{N} \ . \tag{D.16}$$

---

16. Le volume de chacun des deux sous-systèmes reste inchangé dans le contact :

$$\mathrm{d}V_1 = \mathrm{d}V_2 = 0 \ . \tag{1}$$

Il en est de même de leur entropie, puisque la superfuite ne laisse passer que la composante d'entropie nulle :

$$\mathrm{d}S_1 = \mathrm{d}S_2 = 0 \ . \tag{2}$$

La variable interne dont on cherche la valeur d'équilibre est par exemple N_1, nombre de particules du sous-système $\mathcal{S}_1$; le nombre N_2 de particules constituant $\mathcal{S}_2$ est simplement donné par

$$N_2 = N - N_1 \ , \tag{3}$$

où le nombre total de particules N est fixé.

Calculons alors, pour écrire ensuite qu'elle est nulle à l'équilibre, la variation d'énergie $\mathrm{d}E$ du système global lorsque la variable interne s'accroît de $\mathrm{d}N_1$:

$$\mathrm{d}E = \mathrm{d}E_1 + \mathrm{d}E_2 \ , \tag{4}$$

avec (formule (D.56) du complément II.D)

$$\mathrm{d}E_i = T_i \, \mathrm{d}S_i - p_i \, \mathrm{d}V_i + \mu_i \, \mathrm{d}N_i \ ; \quad i = 1,2 \ . \tag{5}$$

Compte tenu de (1), (2) et (3), on trouve

$$\mathrm{d}E = (\mu_1 - \mu_2)\mathrm{d}N_1 \ , \tag{6}$$

d'où la condition d'équilibre (D.13).

17. Pour un gaz de bosons libres sans interaction, le potentiel chimique est nul dans le domaine où se produit la condensation de Bose (VI, § II.B.1). Ici, le potentiel chimique de l'hélium II n'est pas nul : sur la courbe de vaporisation (la plupart des expériences sur l'hélium liquide se déroulent en présence de la vapeur), ce potentiel chimique est égal à celui du gaz (complément III.G, note 23); or ce dernier est négatif et grand, car la vapeur d'hélium, loin du point critique, est à une bonne approximation un gaz parfait classique. Nous verrons plus loin (§ II.2.b) que les quasi-particules qui portent dans l'hélium II l'énergie d'excitation ont, elles, un potentiel chimique nul; mais celui des atomes d'hélium, qui intervient ici, ne l'est pas.

Si donc les différences de température $T_2 - T_1 \widehat{=} \delta T$ et de pression $p_2 - p_1 \widehat{=} \delta p$ sont suffisamment petites pour être traitées au premier ordre, la condition (D.15) devient simplement

$$\frac{\delta p}{\delta T} = \frac{S}{V} . \tag{D.17}$$

Noter que ce rapport est positif : les différences de température et de pression sont de même signe.

Donc, si l'on maintient les deux récipients de la figure 8.a à des pressions différentes, leurs températures seront différentes à l'équilibre, selon la loi que nous venons d'expliciter : c'est l'« effet thermomécanique ». Si au contraire on établit entre eux une différence de température (en chauffant l'un d'eux par un procédé sélectif), le superfluide va migrer du récipient froid vers le récipient chaud pour créer la différence de pression nécessaire à l'équilibre : c'est l'« effet fontaine ». Son nom vient de ce que la surpression dans le récipient chaud peut être mise en évidence de façon spectaculaire par un jet de liquide si l'on perce un trou de dimensions convenables dans la paroi supérieure du récipient (fig. 8.b).

δ. Le second son

Dans tout fluide se propagent des ondes de densité et de pression, sinusoïdales s'il s'agit d'un son pur. Dans l'hélium II, ces ondes ordinaires sont appelées *« premier son »;* en un point donné, la masse volumique (globale) ρ varie sinusoïdalement, en fonction du temps, autour de sa valeur d'équilibre, et cette perturbation se propage de proche en proche; dans ce son ordinaire, les densités ρ_n et ρ_s du fluide normal et du superfluide oscillent *en phase,* de sorte que c'est l'ensemble du liquide qui entre en mouvement pour créer les variations locales de densité et de pression. Mais il peut exister dans l'hélium II un *autre type d'ondes,* dans lesquelles ρ_n *et* ρ_s *oscillent en opposition de phase,* de sorte que leur somme ρ reste constante : c'est le *« second son »,* qui consiste donc en des *variations locales des proportions de fluide normal et de superfluide, à densité globale constante.* Comme les rapports ρ_n/ρ et ρ_s/ρ dépendent de la température, le second son se traduit concrètement par des oscillations locales de la température, comme d'ailleurs de l'entropie puisque seul le fluide normal la transporte. Ainsi, *le second son est une onde de température et d'entropie sans variation de la masse volumique.* Il a effectivement été observé expérimentalement, et sa vitesse mesurée en fonction de la température, ce qui fournit une détermination de ρ_n/ρ indépendante des mesures de viscosité (réf. 3, p. 516).

L'existence et la nature du second son permettent de comprendre une autre propriété remarquable de l'hélium II : c'est un *supraconducteur de la chaleur.* Dans un matériau ordinaire en effet, la conduction de la chaleur se fait par diffusion : schématiquement, lorsqu'une région est portée à une température plus élevée que ses voisines, l'agitation thermique des particules qui s'y trouvent est augmentée; ces particules, au hasard des chocs et de l'orientation de leur vitesse, gagnent progressivement les régions plus froides où elles transmettent par collision une partie de leur énergie aux autres particules. Ce processus chaotique est lent, même pour les bons conducteurs de la chaleur que sont les métaux[18]. Dans l'hélium II au contraire,

---

18. Les processus de diffusion ne sont pas caractérisés par une vitesse unique, car la distance d parcourue est proportionnelle à la racine carrée du temps t (complément I.D, § III). L'*« équation de la chaleur »,* qui décrit la diffusion, donne

$$\frac{d^2}{t} \sim \frac{\varkappa M}{\rho c_V} , \tag{1}$$

où M est la masse molaire, ρ la masse volumique, c_V la chaleur spécifique (molaire) et $\varkappa$ la conductivité thermique, qui s'exprime en watts par mètre et par kelvin. Pour le cuivre à

.../...

les fluctuations locales de température, et donc la chaleur, se propagent de façon cohérente et ordonnée, sous la forme du second son, à une vitesse de plusieurs dizaines de mètres par seconde (20 m/s au-dessus de 1 K; cette vitesse croît très rapidement lorsque T décroît, pour atteindre ~ 200 m/s vers 0,1 K)[19],[20].

On peut construire, dans le cadre du modèle à deux fluides, une *hydrodynamique* de l'hélium II (réf. 1, p. 335). On peut aussi étudier d'autres propriétés de ce liquide extraordinaire[21], dont certaines ne sont comprises que si l'on complète le modèle en faisant intervenir les caractéristiques quantiques du superfluide. Nous nous contenterons ici de l'image simplifiée que nous venons de présenter.

II. INTERPRÉTATION THÉORIQUE DES PROPRIÉTÉS DE L'HÉLIUM II

Le modèle des deux fluides que nous avons résumé au paragraphe précédent a la vertu de regrouper à l'intérieur d'un cadre unique et simple la plupart des propriétés de l'hélium II. Mais il ne dit rien sur la nature physique des deux fluides qu'il introduit dans ses hypothèses, et ne permet pas une étude statistique du système. Ces questions ont fait l'objet de nombreux travaux théoriques; il reste encore des problèmes non résolus, mais on peut considérer que les aspects essentiels des phénomènes ont été compris.

---

température ordinaire, par exemple, $\varkappa \simeq 3{,}8$ W/cm K, $c_V \simeq 24$ J/K par mole, $\rho \simeq 9$ g/cm^3 et $M = 63{,}5$ g par mole. On trouve ainsi

$$\frac{d^2}{t} \sim 1 \text{ cm}^2/\text{s} , \qquad (2)$$

c'est-à-dire qu'il faut environ 100 s pour que la chaleur diffuse sur 10 cm, et près de 3 h pour qu'elle diffuse sur 1 m.

19. La supraconductivité de l'hélium II pour la chaleur se manifeste elle aussi de façon spectaculaire. Lorsqu'on refroidit un échantillon d'hélium au-dessus du point λ, en pompant rapidement sa vapeur, on le voit bouillir. Mais dès qu'on atteint le point λ, l'ébullition cesse brusquement : l'hélium II obtenu est trop bon conducteur de la chaleur pour laisser s'établir les inhomogénéités de température qui provoquent la formation de bulles dans un liquide ordinaire; la vaporisation de l'hélium II a lieu uniquement en surface, le refroidissement qu'elle produit se transmettant quasi-instantanément au reste du liquide.

20. Notons que la supraconductivité de la chaleur nécessite la présence simultanée des deux fluides. Dans l'effet fontaine ou thermomécanique (fig. 8), les deux sous-systèmes peuvent avoir des températures différentes car ils n'échangent que du superfluide.

21. L'hélium II a par exemple la faculté de s'écouler hors du vase qui le contient, en grimpant le long des parois et ruisselant ensuite à l'extérieur. Ceci s'explique si l'on sait que l'hélium est un agent mouillant extrêmement efficace et universel : les forces d'attraction entre atomes d'hélium sont tellement faibles (cf. § 1) qu'elles sont toujours inférieures à celles qu'exercent les atomes d'une paroi quelconque sur ceux d'hélium. L'hélium forme donc une mince pellicule (quelques centaines d'Å) s'élévant le long des parois au-dessus du niveau du liquide. Le fluide normal est bloqué sur la paroi par sa viscosité, mais le superfluide circule librement dans cette mince pellicule et peut ainsi amorcer un siphon qui videra progressivement le vase (réf. 3, p. 522).

En outre, la composante superfluide peut être le siège de courants fermés se maintenant indéfiniment après que la cause qui leur a donné naissance a été supprimée (réf. 1, pp. 326 et 333 et réf. 3, p. 527).

1. CONSIDÉRATIONS QUALITATIVES ET SEMI-QUANTITATIVES

a. Comparaison avec le gaz de bosons sans interactions

Il paraît clair que les propriétés de l'hélium II, qui se manifestent pour l'isotope 4 de l'hélium et pas pour l'isotope 3 (voir cependant note 5), sont liées au fait que l'hélium 4 est un *boson.* Reprenons alors les résultats du chapitre VI (§ II) sur les gaz de bosons indépendants, et voyons s'ils peuvent nous aider à comprendre le comportement de l'hélium II.

Nous savons qu'un gaz de bosons sans interactions est le siège d'une *transition de phase.* La température de Bose à laquelle elle se produit a pour expression

$$T_B = \frac{2\pi\hbar^2}{mk}\left[\frac{N}{V}\cdot\frac{1}{2{,}612\,(2s+1)}\right]^{2/3} \tag{D.18}$$

Si nous calculons sa valeur pour l'hélium 4, en évaluant N/V à partir de la densité de l'hélium liquide (0,15 g/cm^3 au voisinage du point λ), nous obtenons

$$T_B \simeq 3{,}2 \text{ K}\,, \tag{D.19}$$

qui est comparable à T_λ. Mais cette coïncidence est fortuite, ou plus exactement presque inévitable : à partir de $\hbar$, k, N/V et m, on peut construire une seule grandeur ayant les dimensions d'une température (chapitre VI, note 31); il suffit donc que les coefficients numériques sans dimension soient de l'ordre de 1 pour que T_λ et T_B aient des valeurs voisines. Si l'on y regarde de près, *les deux transitions de phase sont très différentes :* celle qui se produit dans un gaz sans interactions[22] est du premier ordre alors que la transition λ est du second ordre; cela signifie que le gaz de bosons condensé comporte *deux phases,* de sorte que sa pression est fixée si sa température l'est, alors que l'hélium II est *une phase unique* pouvant exister dans toute une plage du plan (T, p).

Comparons ensuite la phase condensée du gaz de bosons avec la composante superfluide de l'hélium II. Le condensat de bosons indépendants a un volume nul, ce qui n'est pas très réaliste; mais il a aussi une entropie nulle, et cet aspect est intéressant. Il permet de comprendre comment *un nombre macroscopique de bosons peuvent donner une contribution nulle à l'entropie :* lorsque ces bosons sont indépendants, ils s'accumulent dans l'état fondamental individuel; si maintenant leurs interactions ne sont plus négligeables et leur volume propre n'est plus nul, ils se trouveront encore dans un état quantique global déterminé de façon unique. Nous verrons cependant plus loin (§ c) que, dans un gaz de bosons sans interactions, le condensat n'est pas superfluide; mais nous verrons également (§ 2.a) comment cette difficulté disparaît dans un modèle plus réaliste.

Analysons enfin la dépendance de la chaleur spécifique c_V par rapport à la température T. Pour un gaz de bosons sans interactions, la courbe représentant cette

22. On a tenté de réaliser expérimentalement des gaz de bosons pratiquement parfaits, dans lesquels la condensation de Bose devrait apparaître «à l'état pur». On a d'abord essayé l'hydrogène atomique polarisé, mais c'est un corps instable, difficile à produire et à manipuler. Tout récemment, on a pu étudier des gaz constitués par des atomes d'hélium 4 dissous dans un verre poreux et inerte; il semble bien que la transition de phase change de nature lorsque la dilution du gaz devient suffisante pour que les interactions entre atomes d'hélium soient négligeables. Voir B. Castaing : référence de la note 9; S. Balibar, «Un gaz quantique presque parfait», *La Recherche 162* (janvier 1985), p. 96.

dépendance est celle de la figure 10 du chapitre VI; il faut la comparer à la figure 4 ci-dessus. Nous constatons aussitôt que le comportement de c_V à la transition de phase n'est pas le même dans les deux cas. De plus, *la chaleur spécifique de l'hélium II varie expérimentalement comme T^3 lorsque T tend vers zéro, alors que celle d'un gaz de bosons sans interactions est en $T^{3/2}$* (formule VI.(II.49)). Ceci indique clairement que *le modèle du gaz sans interactions décrit mal les excitations du fluide* à partir de son niveau fondamental.

Pour résumer, l'étude du gaz de bosons sans interactions permet de comprendre l'existence d'une transition de phase mais pas sa nature; elle suggère que la composante d'entropie nulle existe parce qu'un ensemble macroscopique de particules peut se trouver de façon certaine dans son état quantique fondamental, mais elle n'explique pas la superfluidité de cette composante; elle montre que les excitations au-dessus de l'état fondamental ne peuvent être décrites comme affectant individuellement et séparément les particules du fluide.

b. Excitations collectives et quasi-particules

Les atomes constituant l'hélium II ne sont donc pas indépendants. Ceci n'est pas surprenant : même si leurs interactions sont extrêmement faibles, la valeur (D.2) de l'énergie potentielle entre deux atomes d'hélium correspond à une température u_0/k d'environ 10 K, et nous nous situons ici nettement au-dessous.

Dans un gaz où les interactions sont négligeables, on peut utiliser l'approximation qui consiste à traiter les particules comme des sous-systèmes indépendants : lorsque l'énergie du gaz augmente, c'est l'énergie individuelle d'une ou plusieurs particule(s) qui a augmenté. Nous avons indiqué au paragraphe a que cette approximation ne donne pas de bons résultats dans le cas qui nous occupe.

La situation inverse où les interactions sont fortes a déjà été rencontrée dans l'analyse des vibrations d'un cristal (complément III.E, § II) ou d'une molécule polyatomique (complément III.B, § II.2.b). Dans ce cas, chaque atome vibrant autour de sa position d'équilibre constitue à une bonne approximation un oscillateur harmonique; mais ces divers oscillateurs sont couplés par des termes du même ordre que leurs énergies individuelles. On remplace alors (en mécanique classique aussi bien qu'en mécanique quantique) le système des N oscillateurs réels couplés par l'ensemble des N *modes normaux* de vibration du système, qui sont des *oscillateurs harmoniques indépendants mais fictifs*. Chaque mode normal correspond à une *oscillation collective* particulière des particules réelles. L'énergie du système est ici la somme des énergies des modes normaux, et non plus des particules réelles : augmenter l'énergie du système consiste à augmenter l'énergie d'un ou plusieurs modes normaux collectifs, et non plus l'énergie individuelle d'une ou plusieurs particules. Dans le cas du solide, les modes normaux peuvent être repérés par un vecteur d'onde $\vec{K}$ (quantifié), c'est-à-dire par une impulsion $\hbar\vec{K}$. Les excitations élémentaires se comportent alors comme des particules; ce ne sont pas de vraies particules, mais des quanta d'excitation d'un mode de vibration collectif du système des véritables particules; on les appelle des *«quasi-particules»*, les phonons dans ce cas particulier (complément III.E, § V).

L'étude des phonons d'un cristal nous a montré que leur contribution à la chaleur spécifique du solide se comporte comme T^3 à basse température. Ceci suggère fortement que *dans l'hélium II également,* à température basse mais non nulle où l'énergie (moyenne) E est supérieure à celle du niveau fondamental E_0, l'énergie supplémentaire $E - E_0$ est portée par des *excitations élémentaires de mouvements collectifs de l'ensemble des atomes;* ces excitations élémentaires sont des *quasi-particules pratiquement indépendantes,* de sorte que l'énergie des premiers états excités du système est de la forme

$$E_{\{n_{\vec{p}}\}} = E_0 + \sum_{\vec{p}} n_{\vec{p}}\,\varepsilon(\vec{p}\,)\,, \tag{D.20}$$

où $n_{\vec{p}}$ est le nombre (entier) de quasi-particules d'impulsion $\vec{p}$ et d'énergie $\varepsilon(\vec{p}\,)$.

c. Le critère de superfluidité

L'image suivante se dégage donc des paragraphes a et b précédents : l'état fondamental dans lequel se trouve le système à température nulle est un condensat d'atomes d'hélium[23]; à température non nulle apparaissent des quasi-particules décrivant des excitations collectives du système; en première approximation, ces quasi-particules constituent un gaz sans interactions. Il est tentant d'identifier le condensat avec la composante superfluide de l'hélium II et le gaz de quasi-particules avec sa composante normale; nous allons voir dans la suite que cette identification semble effectivement correcte.

Cherchons pour commencer à quelle condition le condensat peut s'écouler de façon superfluide dans un tube. A l'échelle microscopique, la viscosité est due à des échanges d'énergie et d'impulsion entre le fluide et les parois, et entre les différentes couches de fluide. *Le condensat manifestera une viscosité nulle* si ces échanges ne peuvent pas se produire, c'est-à-dire (puisque le condensat ne peut que gagner de l'énergie) *si le contact avec les parois du tube ne peut exciter aucune quasi-particule;* dès que cette excitation deviendra possible, l'écoulement cessera d'être superfluide.

Considérons donc un condensat d'atomes d'hélium se déplaçant à la vitesse $\vec{v}$ par rapport à un objet macroscopique de masse M, par exemple la paroi d'un tube. Comme il s'agit d'analyser les *excitations éventuelles du condensat,* il est préférable de se placer *dans le référentiel où il est au repos.* Dans ce référentiel, l'objet considéré est animé de la vitesse $-\vec{v}$; il a donc une impulsion $\vec{P} = -M\vec{v}$ et une énergie $E = Mv^2/2$. Supposons qu'il se crée à son contact une excitation élémentaire d'impulsion $\vec{p}$ et d'énergie ε. Les lois de conservation de l'énergie et de l'impulsion donnent alors les relations

$$\frac{1}{2} Mv^2 = \frac{1}{2} Mv'^2 + \varepsilon \,, \tag{D.21.a}$$

$$-M\vec{v} = M\vec{v}' + \vec{p} \,, \tag{D.21.b}$$

où $\vec{v}'$ est la vitesse de l'objet après le processus d'excitation. Il est facile de montrer[24] que ces relations impliquent

$$|\vec{v}| \geqslant \frac{\varepsilon}{|\vec{p}|} \,. \tag{D.22}$$

---

23. Cet état fondamental est l'état propre d'énergie la plus basse de l'hamiltonien total du système, comprenant ici l'énergie d'interaction entre les atomes d'hélium.

24. Éliminons $\vec{v}'$ entre les deux relations. La deuxième donne

$$\vec{v}' = -\left(\vec{v} + \frac{\vec{p}}{M}\right) . \tag{1}$$

En reportant cette expression dans la première relation, on trouve

$$\vec{v} \,.\vec{p} = -\varepsilon - \frac{\vec{p}^{\,2}}{2M} \,. \tag{2}$$

Mais ε est nécessairement positif (il s'agit d'exciter le fluide *au-dessus* de son niveau fondamental); $\vec{v} \,.\vec{p}$ est donc négatif, et l'on peut écrire

$$|\vec{v} \,.\vec{p}| \geqslant \varepsilon \,. \tag{3}$$

Comme $|\vec{v} \,.\vec{p}|$ est au plus égal au produit des modules des deux vecteurs, on a finalement

$$|\vec{v}| \,.\, |\vec{p}| \geqslant |\vec{v} \,.\vec{p}| \geqslant \varepsilon \,, \tag{4}$$

d'où l'on tire (D.22).

L'excitation envisagée ne pourra donc pas avoir lieu si la vitesse d'écoulement v du condensat vérifie la condition

$$v < \frac{\varepsilon}{|\vec{p}|} . \tag{D.23}$$

Cette dernière inégalité est appelée le *« critère de superfluidité »*. Pour qu'il existe des vitesses non nulles donnant lieu à un écoulement superfluide, *il faut que le second membre de l'inégalité ait une valeur minimum non nulle* lorsqu'on envisage toutes les excitations élémentaires possibles.

Cette condition n'est pas vérifiée pour un gaz de bosons libres sans interactions. En effet, les excitations sont dans ce cas portées par les particules elles-mêmes, et la relation entre leur énergie et leur impulsion s'écrit

$$\varepsilon = \frac{\vec{p}^{\,2}}{2m} , \tag{D.24}$$

de sorte que

$$\frac{\varepsilon}{|\vec{p}|} = \frac{|\vec{p}|}{2m} . \tag{D.25}$$

Ce rapport a pour valeur minimum zéro comme l'impulsion $|\vec{p}|$[25]. Les particules seront donc excitées quelle que soit la vitesse de l'écoulement, et l'échange d'énergie avec la paroi se manifestera par une résistance visqueuse : le condensat d'un gaz parfait de bosons n'est pas superfluide. Voilà une raison supplémentaire d'introduire, dans le cas de l'hélium II, des excitations élémentaires qui ne soient pas les véritables particules du système.

2. MÉCANIQUE STATISTIQUE DES QUASI-PARTICULES : LE MODÈLE DE LANDAU

a. Relation de dispersion des excitations élémentaires

Les quasi-particules décrivant les excitations élémentaires de l'hélium II peuvent à une bonne approximation être considérées comme indépendantes; nous sommes ainsi ramenés à l'étude d'un *gaz parfait de quasi-particules*. Il nous suffira de connaître l'énergie ε de chaque quasi-particule en fonction de son impulsion $\vec{p}$, c'est-à-dire la « relation de dispersion » des excitations élémentaires, pour pouvoir calculer les propriétés thermodynamiques du système. Le fluide étudié étant isotrope, l'énergie ε sera seulement fonction du module p de l'impulsion (qu'il faudra éviter de confondre avec la pression si celle-ci apparaît dans les formules).

La relation de dispersion $\varepsilon(p)$ peut être obtenue expérimentalement par diffusion inélastique de neutrons lents : lorsque, dans une telle diffusion, une quasi-particule d'impulsion $\vec{p}$ et d'énergie ε est créée dans l'hélium II, l'impulsion et l'énergie du neutron diffusé diffèrent de celles du faisceau incident précisément de $\vec{p}$ et ε. La courbe ainsi déterminée est représentée sur la figure 9. En réalité, Landau avait deviné sa forme avant qu'elle ne soit mesurée, en montrant qu'elle permettait d'expliquer la dépendance en température de la chaleur spécifique de l'hélium II (voir paragraphe b).

Pour les petites valeurs de p, ε est simplement proportionnelle à p; c'est le spectre de *phonons* que l'on attendait :

$$\varepsilon = c_1 p \text{ (phonons)}, \tag{D.26}$$

où c_1 est la vitesse du (premier) son dans l'hélium II :

$$c_1 \simeq 240 \text{ m/s} . \tag{D.27}$$

---

25. On pourrait objecter que les valeurs possibles pour l'impulsion $\vec{p}$ sont en réalité quantifiées, puisque le gaz est enfermé dans une boîte de dimensions finies. Mais l'impulsion minimale $\hbar . 2\pi/L$ donnerait une vitesse maximale d'écoulement très faible (0,5 μm/s pour $L = 10$ cm).

Mais le comportement de la fonction $\varepsilon(p)$ devient ensuite assez curieux : elle présente un maximum suivi d'un minimum, avant de croître à nouveau pour les grandes valeurs de p. La portion de courbe comprise entre le maximum et le minimum est une *« région de dispersion anormale » :* la dérivée $d\varepsilon/dp$, qui n'est autre que la vitesse de groupe des quasi-particules, y est négative (dans cette région, les excitations se déplacent donc dans le sens opposé à leur impulsion). Le minimum de la courbe a pour coordonnées p_m et Δ telles que

$$\frac{p_m}{\hbar} \simeq 1{,}9 \text{ Å}^{-1} , \tag{D.28.a}$$

$$\frac{\Delta}{k} \simeq 8{,}7 \text{ K} . \tag{D.28.b}$$

Notons que *le critère de superfluidité est vérifié :* le rapport ε/p est représenté sur la figure 9 par la pente de la droite joignant l'origine à un point de la courbe; sa valeur minimum est obtenue lorsque cette droite est tangente à la courbe; le point de contact de cette tangente étant voisin du minimum, on peut estimer la vitesse critique prédite à

$$v_c \simeq \frac{\Delta}{p_m} \simeq 60 \text{ m/s} . \tag{D.29}$$

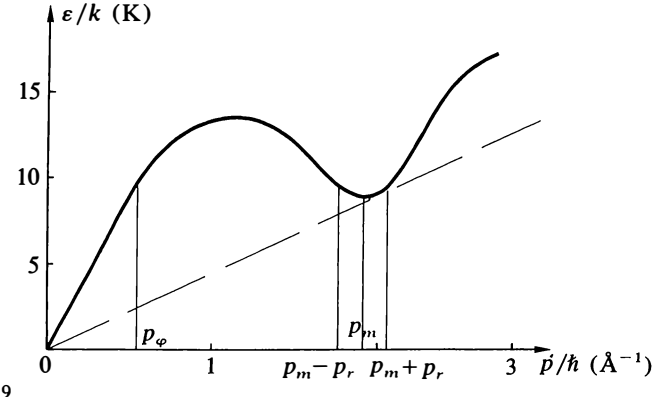


FIGURE 9

Relation de dispersion des excitations élémentaires dans l'hélium II, déterminée expérimentalement par diffusion inélastique de neutrons lents. On a tracé en tirets la tangente de pente minimale issue de l'origine, qui intervient dans le critère de superfluidité. On a également indiqué les limites de validité des expressions approchées (D.26) et (D.30) de la relation de dispersion.

Ce résultat est cependant *beaucoup trop grand :* expérimentalement, l'écoulement de l'hélium II cesse d'être superfluide dès que la vitesse atteint quelques mm/s, au mieux quelques cm/s dans les conditions les plus favorables. Cette différence quantitative est vraisemblablement due au fait que l'écoulement de l'hélium II dans un capillaire est un mouvement plus complexe que nous ne l'avons supposé[26].

26. Les valeurs expérimentales de la vitesse critique indiquent qu'il doit exister des excitations de rapport ε/p beaucoup plus petit que ceux de la figure 9. Mais alors pourquoi ne se manifestent-elles pas dans la chaleur spécifique? Voir réf. 1, p. 365, pour quelques arguments relativement simples concernant ce problème.

Nous voulons étudier l'hélium II à des températures T de l'ordre de 1 à 2 K, puisqu'elles doivent être inférieures à T_λ, donné en (D.10). Nous pourrons donc ne prendre en compte que les excitations de plus basse énergie. Celles-ci comprendront évidemment les *phonons* de faible impulsion, pour lesquels la relation de dispersion prend la forme simple (D.26); les phonons domineront le comportement du système à très basse température (pour $T \lesssim 0{,}5$ K, où la chaleur spécifique varie en T^3). Mais il faut également inclure les excitations qui se situent au voisinage du minimum de la courbe $\varepsilon(p)$, et que Landau a appelées *« rotons » :* leur énergie est relativement élevée (formule (D.28.b)), mais la densité d'états, proportionnelle comme toujours à $\mathrm{d}p/\mathrm{d}\varepsilon$ (cf. I, § III.B.2), est très grande dans cette région puisqu'elle devient même infinie en p_m. La relation de dispersion des rotons s'obtient en développant la fonction $\varepsilon(p)$ autour de son minimum :

$$\varepsilon = \Delta + \frac{(p - p_m)^2}{2\alpha} \quad \text{(rotons)}; \tag{D.30}$$

le paramètre α a les dimensions d'une masse et vaut environ

$$\alpha \simeq 0{,}16m \simeq 1{,}06 \times 10^{-27} \text{ kg} . \tag{D.31}$$

Soulignons que α n'est la masse d'aucun objet physique : la relation (D.30) entre l'impulsion et l'énergie d'un roton, qui est une quasi-particule, a une forme différente de celle qu'elle prend pour une véritable particule.

b. Propriétés thermodynamiques du système

α. Contribution des phonons

La mécanique statistique des phonons est *analogue à celle des photons* (VI, § III) : la relation de dispersion a même forme, la vitesse du son c_1 remplaçant ici la vitesse de la lumière c, et le nombre de phonons comme celui des photons n'est pas fixé de sorte que leur potentiel chimique est nul. Il y a cependant *deux différences :* en premier lieu, un photon peut prendre deux états de polarisation distincts, alors qu'un phonon en possède trois dans un solide et seulement un dans l'hélium II (seul le mode longitudinal existe dans un fluide); deuxièmement, la relation de dispersion (D.26) n'est valable pour un phonon que dans un domaine limité d'impulsion et d'énergie, alors qu'elle l'est pour un photon dans toute la gamme des énergies.

Les *propriétés des phonons* ont été calculées dans le complément III.E (§ III); il suffit, par rapport aux résultats obtenus alors, de diviser par trois la densité d'états (pour éliminer les deux modes transversaux)[27]. La température de Debye vaut ici[28]

$$T_D \simeq 20 \text{ K} , \tag{D.32}$$

de sorte que l'on se trouve dans le domaine où

$$T \ll T_D . \tag{D.33}$$

---

27. Il n'y a pas lieu de conserver ici l'« énergie de point zéro » $\hbar\omega/2$ qui figure dans le spectre de l'oscillateur harmonique quantique : comme les phonons décrivent les excitations de l'hélium II au-dessus de son état fondamental, leur énergie s'annule en même temps que leur impulsion, comme pour les photons.

28. L'expression de la température de Debye n'est pas affectée par la réduction du nombre de polarisations possibles (formule (E.61) du complément III.E) :

$$T_D = \frac{\hbar}{k} c_1 \left(6\pi^2 \frac{N}{V}\right)^{1/3}$$

Le résultat (D.32) découle alors des valeurs numériques suivantes : $\rho = 0{,}146$ g/cm^3 (pratiquement indépendante de T); $m = 6{,}65 \times 10^{-27}$ kg; $c_1 = 240$ m/s.

En particulier, la contribution des phonons à la capacité calorifique de l'hélium II a pour expression

$$c_V^{(\varphi)} = \frac{4}{5}\pi^4 Nk\left(\frac{T}{T_D}\right)^3 . \tag{D.34}$$

β. Contribution des rotons

Les rotons sont des *bosons de spin nul dont le nombre n'est pas conservé.* Le nombre moyen de rotons occupant l'état individuel d'impulsion $\vec{p}$ est donc donné par la distribution de Bose-Einstein pour un potentiel chimique μ nul :

$$\bar{n}^{(r)}(\vec{p}) = \frac{1}{e^{\beta\varepsilon(\vec{p})} - 1} . \tag{D.35}$$

Mais la valeur minimale Δ de $\varepsilon(\vec{p})$ est grande devant kT dans le domaine qui nous intéresse (cf. (D.28.b)), de sorte que

$$e^{\beta\varepsilon(\vec{p})} \gg 1 . \tag{D.36}$$

Le nombre d'occupation moyen de tous les états est donc très petit :

$$\bar{n}^{(r)}(\vec{p}) \simeq e^{-\beta\varepsilon(\vec{p})} \ll 1 . \tag{D.37}$$

Il s'ensuit que le gaz de rotons obéit à la *statistique de Maxwell-Boltzmann :* c'est un *gaz parfait classique de potentiel chimique nul*[29].

La *fonction de partition* d'un gaz de N_r rotons s'écrit donc

$$Z_r = \frac{1}{N_r!}[z_r]^{N_r} , \tag{D.38}$$

où la fonction de partition z_r d'un roton a pour expression[30]

$$z_r = \sum_{\vec{p}} e^{-\beta\varepsilon(\vec{p})} = \frac{V}{(2\pi\hbar)^3}\int d^3p\ e^{-\beta[\Delta + (p-p_m)^2/2\alpha]} . \tag{D.39}$$

Le nombre N_r de rotons est une *variable interne;* il est déterminé, à chaque température, par la condition

$$\frac{\partial F_r}{\partial N_r} = 0 . \tag{D.40}$$

---

29. Pour des particules libres, la statistique de Maxwell-Boltzmann est obtenue à partir de la distribution de Bose-Einstein ou de Fermi-Dirac lorsque le potentiel chimique μ devient grand et négatif (V, § IV.C). Ici, le potentiel chimique est constamment nul, mais c'est l'existence d'une énergie minimale grande qui conduit à un résultat analogue au précédent.

30. Le système étant enfermé dans une boîte de volume V, les impulsions permises $\vec{p}$ sont quantifiées. Lorsqu'on remplace la somme discrète par une intégrale, il faut diviser le volume infinitésimal d^3p de l'espace des impulsions par le volume de la maille élémentaire de quantification dans cet espace, égal à (chapitre I, note 19)

$$\left(\hbar\frac{2\pi}{L}\right)^3 = \frac{(2\pi\hbar)^3}{V} . \tag{1}$$

On peut aussi, puisque le gaz de rotons est *classique,* écrire directement la fonction de partition d'un roton sous la forme (III, § V.A.2)

$$z_r = \frac{1}{(2\pi\hbar)^3}\int d^3r \int d^3p\ e^{-\beta\varepsilon(\vec{p})} , \tag{2}$$

qui redonne l'expression (D.39).

Avec l'énergie libre F_r déduite de (D.38) :

$$F_r = -N_r kT [\ln z_r - \ln N_r + 1] , \tag{D.41}$$

la condition (D.40) donne simplement

$$N_r = z_r . \tag{D.42}$$

On trouve donc finalement

$$F_r = -N_r kT , \tag{D.43}$$

où N_r a la même expression (D.39) que z_r. Un calcul simple[31] conduit à la formule approchée

$$N_r(T, V) \simeq \frac{2}{(2\pi)^{3/2}} \frac{V}{\hbar^3} p_m^2 (\alpha kT)^{1/2} e^{-\Delta/kT} . \tag{D.44}$$

Les *propriétés du gaz de rotons* se déduisent de son énergie libre par les formules habituelles. C'est évidemment l'exponentielle de $-\Delta/kT$ qui régit la variation des diverses grandeurs en fonction de la température. On trouve ainsi

$$E_r = -\frac{\partial}{\partial\beta} N_r = N_r(T, V) \left[\frac{kT}{2} + \Delta\right] \tag{D.45}$$

pour l'énergie du gaz,

$$S_r = -\frac{\partial F_r}{\partial T} = k N_r(T, V) \left[\frac{3}{2} + \frac{\Delta}{kT}\right] \tag{D.46}$$

pour son entropie, et

$$C_V^{(r)} = \frac{\partial E_r}{\partial T} = k N_r(T, V) \left[\frac{3}{4} + \frac{\Delta}{kT} + \left(\frac{\Delta}{kT}\right)^2\right] \tag{D.47}$$

pour la contribution des rotons à la capacité calorifique de l'hélium II.

γ. *Comparaison avec l'expérience*

Dans le modèle de Landau, la chaleur spécifique de l'hélium II est la somme des deux contributions calculées au paragraphe α et au paragraphe β :

$$c_V = c_V^{(\varphi)} + c_V^{(r)} . \tag{D.48}$$

---

31. Pour calculer

$$N_r = z_r = \frac{V}{(2\pi\hbar)^3} \int d^3p \; e^{-\beta[\Delta + (p - p_m)^2/2\alpha]} , \tag{1}$$

on passe en coordonnées sphériques dans l'espace des $\vec{p}$:

$$N_r = \frac{4\pi V}{(2\pi\hbar)^3} e^{-\beta\Delta} \int p^2 \, dp \; e^{-\beta(p - p_m)^2/2\alpha} . \tag{2}$$

Nous n'avons pas précisé jusqu'ici les bornes d'intégration, c'est-à-dire le domaine dans lequel la relation de dispersion approchée (D.30) peut être considérée comme valable. Mais la gaussienne à intégrer ici décroît très rapidement lorsque p s'éloigne de p_m, de sorte que l'on ne change pratiquement pas le résultat en prenant pour bornes $-\infty$ et $+\infty$. Pour la même raison, on peut remplacer le facteur p^2 par p_m^2 et le sortir de l'intégrale. On obtient ainsi

$$N_r \simeq \frac{4\pi V}{(2\pi\hbar)^3} p_m^2 e^{-\beta\Delta} \int_{-\infty}^{+\infty} dp \; e^{-\beta(p - p_m)^2/2\alpha} \quad \text{pour } \frac{\beta p_m^2}{2\alpha} \gg 1 , \tag{3}$$

c'est-à-dire finalement

$$N_r \simeq \frac{4\pi V}{(2\pi\hbar)^3} p_m^2 e^{-\beta\Delta} \sqrt{\frac{2\pi\alpha}{\beta}} . \tag{4}$$

Les formules (D.34) et (D.47), compte tenu des valeurs des paramètres données au paragraphe a, reproduisent bien les résultats expérimentaux (fig. 4)[32]. On constate que la contribution des rotons est négligeable pour $T \leqslant 0{,}6$ K, domaine où l'on observe la loi en T^3 caractéristique des phonons. Inversement, les rotons dominent pour $T \geqslant 1$ K : malgré le nombre élevé de phonons qui sont excités à de telles températures, c'est la dépendance exponentielle de $c_V^{(r)}$ qui apparaît essentiellement. Le modèle cesse d'être applicable pour $T \geqslant 1{,}8$ K : on est alors trop près de la transition de phase, que l'on n'a pas prise en compte, et on ne peut pas espérer reproduire la divergence qu'elle induit sur c_V.

3. INTERPRÉTATION DU MODÈLE DES DEUX FLUIDES

Nous sommes maintenant persuadés que l'hélium II est composé d'un condensat superfluide et d'un gaz de quasi-particules pratiquement indépendantes, essentiellement des phonons et des rotons. Nous allons préciser le lien entre ces deux composantes et celles que le modèle des deux fluides a introduites pour décrire phénoménologiquement les propriétés de l'hélium II; c'est Landau qui a développé les arguments que nous allons présenter.

a. Définition de la densité normale

La tâche n'est pas facile. Les deux « composantes » qu'il s'agit de séparer sont constituées des *mêmes atomes d'hélium indiscernables,* et *ce ne sont pas deux phases distinctes.* Le « fluide normal » est un ensemble de *quasi-particules,* pour lesquelles *la notion de masse n'est pas définie :* la relation entre leur énergie ε et leur impulsion $\vec{p}$, donnée par la courbe de la figure 9, n'a pas la forme simple qu'elle prend pour de véritables particules :

$$\varepsilon \neq \frac{\vec{p}^{\,2}}{2m} + u \ . \tag{D.49}$$

On ne peut donc pas envisager de compter leur nombre par unité de volume et de multiplier par leur masse pour calculer leur contribution à la masse volumique totale. Quant au condensat qui forme le « superfluide » du modèle de Tisza, sa structure est encore moins facile à expliciter.

Considérons la composante que nous connaissons le mieux, c'est-à-dire le gaz de quasi-particules. Nous savons définir pour chacune d'elles son impulsion $\vec{p}$ et son énergie ε; c'est donc à partir de ces grandeurs qu'il nous faut raisonner. Parmi les relations de base (D.11)-(D.12) du modèle des deux fluides, la première peut seulement nous servir à définir la densité superfluide ρ_s si nous sommes parvenus à définir par ailleurs la densité normale ρ_n :

$$\rho_s = \rho - \rho_n \tag{D.50}$$

(la masse volumique globale ρ est, elle, directement accessible à l'expérience). Pour que la deuxième relation, (D.12), puisse permettre de distinguer entre les deux types de fluide, *il faut que les vitesses $\vec{v}_n$ et $\vec{v}_s$ soient différentes.* Si c'est le cas, *on peut*

32. Nous ne tenterons pas d'évaluer la précision avec laquelle ces résultats sont reproduits. Il y a plusieurs façons de le faire, suivant le point de vue que l'on privilégie. Pour donner une idée, indiquons que si on ne fixe pas *a priori* les valeurs des paramètres et qu'on les détermine de façon à reproduire au mieux la chaleur spécifique c_V jusqu'à 1,8 K environ, on trouve

$$c_1 = 226 \text{ m/s} \ , \quad \frac{\Delta}{k} = 9 \text{ K} \ , \quad \frac{p_m}{\hbar} = 2 \text{ Å}^{-1} \ , \quad \alpha = 0{,}3m \ ,$$

que l'on peut comparer à (D.27), (D.28) et (D.31).

isoler la contribution de la composante normale en se plaçant dans le référentiel où la composante superfluide est au repos. Pour la clarté de l'exposé, nous affecterons ici un indice zéro aux grandeurs mesurées dans ce référentiel $\mathcal{R}_0$. Nous écrirons donc

$$\vec{v}_{s0} = \vec{0} \,, \tag{D.51}$$

de sorte que la densité d'impulsion dans $\mathcal{R}_0$ prend la forme

$$\vec{j}_0 = \rho_n \vec{v}_{n0} \tag{D.52}$$

(la masse volumique ρ_n est, en mécanique non relativiste, indépendante du référentiel).

Nous définissons donc ρ_n grâce à la relation (D.52) : dans le référentiel $\mathcal{R}_0$ où la composante superfluide est au repos, ρ_n est le coefficient de proportionnalité entre la densité volumique d'impulsion $\vec{j}_0$ et la vitesse $\vec{v}_{n0}$ de la composante normale. Notons que cette définition est bien adaptée à notre propos : d'une part, elle fait appel à la densité d'impulsion, et l'impulsion est une caractéristique bien définie des quasi-particules; d'autre part, elle demande la connaissance de cette grandeur dans le référentiel où le superfluide, c'est-à-dire le condensat, est au repos, et c'est précisément dans ce référentiel que nous connaissons la relation de dispersion pour les quasi-particules; pour être fidèles à la convention proposée plus haut, nous écrirons ici la relation représentée sur la figure 9

$$\varepsilon_0 = \varepsilon_0(\vec{p}_0) \,. \tag{D.53}$$

b. Expression de la densité normale

Il est facile d'imaginer une situation concrète où les deux composantes de l'hélium II ont des vitesses différentes : c'est le cas lors de l'écoulement superfluide à travers un capillaire (fig. 6). Dans le *référentiel du laboratoire* $\mathcal{R}$ (nous n'ajoutons pas d'indice aux grandeurs correspondantes), la composante superfluide s'écoule à la vitesse

$$\vec{v}_s \neq \vec{0} \,, \tag{D.54}$$

alors que la composante normale est bloquée par sa viscosité :

$$\vec{v}_n = \vec{0} \,. \tag{D.55}$$

Dans le *référentiel $\mathcal{R}_0$ où le superfluide est au repos,* ces vitesses deviennent

$$\vec{v}_{s0} = \vec{0} \,, \tag{D.56.a}$$

$$\vec{v}_{n0} = -\vec{v}_s \,. \tag{D.56.b}$$

Nous voulons calculer *l'impulsion par unité de volume dans le référentiel* $\mathcal{R}_0$, qui a pour expression

$$\vec{j}_0 = \frac{1}{V} \sum_{\vec{p}_0} \vec{p}_0 \bar{n}(\vec{p}_0) = \frac{1}{(2\pi\hbar)^3} \int \vec{p}_0 \, \bar{n}(\vec{p}_0) \, \mathrm{d}^3 p_0 \,, \tag{D.57}$$

où $\bar{n}(\vec{p}_0)$ est le nombre moyen de quasi-particules ayant l'impulsion $\vec{p}_0$. Mais *attention :* ce n'est pas dans ce référentiel que nous connaissons la distribution d'équilibre des quasi-particules. *Cette distribution est régie par l'énergie ε de la quasi-particule dans le référentiel $\mathcal{R}$ du laboratoire,* puisque c'est dans ce référentiel que s'établit l'équilibre entre le gaz de quasi-particules, globalement au repos, et le thermostat (comprenant les parois du tube) :

$$\bar{n}(\vec{p}_0) = \frac{1}{\mathrm{e}^{\beta \varepsilon(\vec{p}_0)} - 1} \,, \tag{D.58}$$

où $\varepsilon(\vec{p}_0)$ est l'énergie, dans $\mathcal{R}$, d'une quasi-particule ayant l'impulsion $\vec{p}_0$ dans $\mathcal{R}_0$.

Il nous faut donc effectuer un *changement de référentiel* pour l'impulsion et l'énergie d'une quasi-particule. Comme celle-ci n'a pas de masse, nous devons à nouveau utiliser des moyens indirects. Considérons le fluide dans son ensemble, de masse volumique ρ. Appelons e_0 et $\vec{j}_0$ ses densités volumiques d'énergie et d'impulsion

dans le référentiel $\mathcal{R}_0$; sa densité volumique d'énergie e dans le référentiel $\mathcal{R}$ est donnée par la transformation de Galilée[33] :

$$e = e_0 + \frac{1}{2}\rho \vec{v}_e^{\,2} + \vec{j}_0 \cdot \vec{v}_e \,, \tag{D.59}$$

où $\vec{v}_e = \vec{v}_s$ est la vitesse de $\mathcal{R}_0$ par rapport à $\mathcal{R}$. Lorsqu'une quasi-particule supplémentaire est créée dans le volume unité auquel s'applique la relation (D.59), son impulsion $\vec{p}_0$ et son énergie ε_0 dans $\mathcal{R}_0$ s'ajoutent à $\vec{j}_0$ et e_0, respectivement; l'accroissement correspondant de e est égal à l'énergie ε de la quasi-particule dans $\mathcal{R}$, que nous cherchons. Mais cette création de quasi-particule ne modifie pas la masse volumique totale ρ du fluide (seulement la proportion ρ_n/ρ de fluide normal), ni la vitesse d'écoulement $\vec{v}_s = \vec{v}_e$ de la composante superfluide. La relation cherchée s'écrit donc

$$\varepsilon = \varepsilon_0 + \vec{p}_0 \cdot \vec{v}_e = \varepsilon_0 + \vec{p}_0 \cdot \vec{v}_s \,. \tag{D.60}$$

On en déduit aussitôt, d'après (D.58), le nombre moyen de quasi-particules ayant, à l'équilibre thermique avec les parois du tube, l'impulsion $\vec{p}_0$ dans $\mathcal{R}_0$:

$$\bar{n}(\vec{p}_0) = n\left(\varepsilon_0(\vec{p}_0) + \vec{p}_0 \cdot \vec{v}_s, \beta\right) , \tag{D.61}$$

où nous avons introduit la fonction

$$n(\varepsilon, \beta) \widehat{=} \frac{1}{e^{\beta\varepsilon} - 1} \,. \tag{D.62}$$

Revenons maintenant à l'expression (D.57) de la *densité volumique d'impulsion* $\vec{j}_0$. Comme c'est la vitesse $\vec{v}_{n0}$ de la composante normale qui nous intéresse, nous utilisons (D.56.b) pour écrire

$$\vec{j}_0 = \frac{1}{(2\pi\hbar)^3} \int \vec{p}_0 \, n\left(\varepsilon_0(\vec{p}_0) - \vec{p}_0 \cdot \vec{v}_{n0}, \beta\right) \mathrm{d}^3 p_0 \,. \tag{D.63}$$

Pour de faibles vitesses d'écoulement, nous pouvons traiter $\vec{p}_0 \cdot \vec{v}_{n0}/\varepsilon_0$ au premier ordre :

$$n\left(\varepsilon_0 - \vec{p}_0 \cdot \vec{v}_{n0}, \beta\right) = n\left(\varepsilon_0, \beta\right) - \vec{p}_0 \cdot \vec{v}_{n0} \frac{\partial n}{\partial \varepsilon} \,. \tag{D.64}$$

Lorsqu'on reporte ce développement dans la formule (D.63), le terme d'ordre zéro est nul par symétrie, car ε_0 dépend seulement du module de $\vec{p}_0$:

$$\int \vec{p}_0 \, n(\varepsilon_0, \beta) \mathrm{d}^3 p_0 = \vec{0} \,. \tag{D.65}$$

---

33. Un objet de masse m a pour énergie

$$E_0 = \frac{1}{2} m \vec{v}_0^{\,2} + u \tag{1}$$

dans le référentiel $\mathcal{R}_0$ où sa vitesse est $\vec{v}_0$, et

$$E = \frac{1}{2} m \vec{v}^{\,2} + u \tag{2}$$

dans le référentiel $\mathcal{R}$. La loi de composition des vitesses :

$$\vec{v} = \vec{v}_0 + \vec{v}_e \tag{3}$$

donne aussitôt

$$E = E_0 + \frac{1}{2} m \vec{v}_e^{\,2} + m \vec{v}_0 \cdot \vec{v}_e \,, \tag{4}$$

où l'on reconnaît en $m\vec{v}_0$ l'impulsion de l'objet dans $\mathcal{R}_0$. Ici, l'objet est un volume unité du fluide, de masse ρ et d'impulsion $\vec{j}_0$ dans $\mathcal{R}_0$.

Reste le terme du premier ordre en $\vec{v}_{n0}$:

$$\vec{j}_0 = -\frac{1}{(2\pi\hbar)^3}\int \vec{p}_0(\vec{p}_0\cdot\vec{v}_{n0})\frac{\partial n}{\partial \varepsilon}\,\mathrm{d}^3p_0\ , \tag{D.66}$$

que l'on peut mettre sous la forme[34],[35]

$$\vec{j}_0 = -\frac{4\pi}{3(2\pi\hbar)^3}\vec{v}_{n0}\int_0^{+\infty} p^4\frac{\partial n}{\partial \varepsilon}\,\mathrm{d}p\ . \tag{D.67}$$

Il suffit alors d'utiliser la définition (D.52) de ρ_n et la forme simple (D.62) de $n(\varepsilon, \beta)$ pour obtenir l'expression générale de la densité normale :

$$\rho_n = \frac{4\pi}{3(2\pi\hbar)^3}\beta\int_0^{+\infty} p^4\frac{e^{\beta\varepsilon(p)}}{[e^{\beta\varepsilon(p)}-1]^2}\,\mathrm{d}p\ . \tag{D.68}$$

Cette expression est calculable dès que la relation de dispersion $\varepsilon(p)$ est connue.

c. Contribution des phonons et des rotons à la densité normale

Pour évaluer, de façon approchée, la densité normale (D.68), nous allons ajouter les contributions des phonons et des rotons (voir la discussion du paragraphe 2.a) :

$$\rho_n = \rho_n^{(\varphi)} + \rho_n^{(r)}\ . \tag{D.69}$$

Dans la contribution des *phonons,* la relation de dispersion prend la forme très simple (D.26) :

$$\rho_n^{(\varphi)} = \frac{4\pi}{3(2\pi\hbar)^3}\beta\int_0^{p_\varphi} p^4\frac{e^{\beta c_1 p}}{[e^{\beta c_1 p}-1]^2}\,\mathrm{d}p\ ; \tag{D.70}$$

la borne supérieure p_φ de l'intégrale peut être déterminée d'après la figure 9 : c'est la limite de validité de l'approximation linéaire (D.26) de la relation de dispersion. Mais la fonction à intégrer décroît très rapidement quand p croît, de sorte que l'intégrale

---

34. On veut calculer

$$\vec{j}_0 = -\frac{1}{(2\pi\hbar)^3}\int \vec{p}\,(\vec{p}\cdot\vec{v}_{n0})\frac{\partial n}{\partial \varepsilon}\,\mathrm{d}^3p\ . \tag{1}$$

Repérons $\vec{p}$ par son module p et par ses angles polaires θ_p et φ_p, en choisissant $\vec{v}_{n0}$ comme axe origine de l'angle θ_p :

$$\vec{p}\cdot\vec{v}_{n0} = p\,v_{n0}\cos\theta_p\ . \tag{2}$$

On a alors

$$\mathrm{d}^3p = p^2\sin\theta_p\,\mathrm{d}p\,\mathrm{d}\theta_p\,\mathrm{d}\varphi_p\ , \tag{3}$$

et l'on sait que $\partial n/\partial\varepsilon$ dépend seulement de p. Pour calculer les composantes de $\vec{j}_0$ perpendiculaires à $\vec{v}_{n0}$, on doit intégrer celles de $\vec{p}$, c'est-à-dire $p\sin\theta_p\cos\varphi_p$ ou $p\sin\theta_p\sin\varphi_p$; l'intégration sur l'angle azimutal φ_p donne alors zéro. La seule composante non nulle est donc suivant $\vec{v}_{n0}$, le vecteur $\vec{p}$ étant alors remplacé par sa composante $p\cos\theta_p$. On peut donc écrire

$$\vec{j}_0 = -\frac{1}{(2\pi\hbar)^3}\vec{v}_{n0}\int p^2\cos^2\theta_p\frac{\partial n}{\partial\varepsilon}\,p^2\sin\theta_p\,\mathrm{d}p\,\mathrm{d}\theta_p\,\mathrm{d}\varphi_p\ . \tag{4}$$

L'intégration sur les angles donne simplement $4\pi/3$, d'où la formule (D.67).

35. Le fait que $\vec{j}_0$ est colinéaire à $\vec{v}_{n0}$ n'est pas dû au développement limité (D.64) de $n(\varepsilon, \beta)$. En effet, dans la formule générale (D.63) donnant le vecteur $\vec{j}_0$, deux vecteurs seulement figurent au second membre, $\vec{p}_0$ et $\vec{v}_{n0}$. Comme on somme sur toutes les directions possibles de $\vec{p}_0$, c'est forcément $\vec{v}_{n0}$ qui fixe la direction de $\vec{j}_0$.

est pratiquement indépendante de sa borne supérieure, que l'on peut sans inconvénient considérer comme infinie. Introduisons la variable sans dimension

$$x \widehat{=} \beta c_1 p \ . \tag{D.71}$$

La formule devient

$$\rho_n^{(\varphi)} \simeq \frac{4\pi}{3(2\pi\hbar)^3} \beta \frac{1}{(\beta c_1)^5} \int_0^{+\infty} x^4 \frac{e^x}{[e^x - 1]^2} \, dx \ . \tag{D.72}$$

L'intégrale définie restant à calculer a pour valeur $4\pi^4/15$. Donc

$$\rho_n^{(\varphi)} \simeq \frac{2\pi^2}{15} \frac{(kT)^4}{\hbar^3 c_1^5} \ . \tag{D.73}$$

Cette contribution varie en T^4 comme l'énergie totale des phonons.

Dans le cas des *rotons,* la relation de dispersion est plus compliquée (formule (D.30)). Mais on peut utiliser le fait que, dans la région des rotons,

$$e^{\beta\varepsilon(p)} \gg 1 \tag{D.74}$$

pour simplifier l'expression à calculer :

$$\rho_n^{(r)} \simeq \frac{4\pi}{3(2\pi\hbar)^3} \beta \int_{p_m - p_r}^{p_m + p_r} p^4 e^{-\beta[\Delta + (p - p_m)^2/2\alpha]} \, dp \ ; \tag{D.75}$$

ici aussi, p_r délimite le domaine de validité de l'approximation (D.30) (voir figure 9). Mais, comme la fonction à intégrer devient rapidement négligeable quand p s'écarte de p_m, tout se passe pratiquement comme si les bornes étaient encore 0 et $+\infty$; on peut en outre, comme pour l'évaluation de N_r au paragraphe 2.b.β, remplacer le facteur p^4 par p_m^4 et le sortir de l'intégrale, ce qui donne finalement

$$\rho_n^{(r)} \simeq \frac{2}{3} \frac{1}{(2\pi)^{3/2}} \frac{\alpha^{1/2} p_m^4}{\hbar^3} (kT)^{-1/2} e^{-\Delta/kT} \ . \tag{D.76}$$

La *densité normale totale* est la somme des deux contributions (D.73) et (D.76). Aux basses températures (jusqu'à 0,5 K environ), les phonons dominent et la proportion ρ_n/ρ de fluide normal croît comme T^4. Un changement de régime se produit ensuite, l'influence des rotons s'affirmant rapidement pour devenir prépondérante lorsque la température dépasse 0,8 K. Ces prédictions sont bien confirmées par les déterminations expérimentales de ρ_n/ρ fondées sur la mesure de la viscosité et de la vitesse du second son (fig. 5); le changement de comportement vers 0,5 K y est en particulier clairement visible. Le modèle n'est cependant valable qu'à température suffisamment basse : au-dessus de $T \simeq 1{,}5$ K, les valeurs qu'il donne sont systématiquement inférieures aux valeurs expérimentales.

4. CONCLUSION

En définitive le modèle de Landau, fondé sur la notion de quasi-particule, fournit une interprétation théorique très convaincante des propriétés de l'hélium II, qui sont elles-mêmes bien résumées dans le modèle phénoménologique des deux fluides : la composante superfluide de l'hélium II apparaît comme un condensat d'atomes d'hélium décrit par une fonction d'onde unique, ce qui lui confère une entropie nulle; la composante « normale » est constituée par un gaz de quasi-particules (phonons et rotons) pratiquement indépendantes les unes des autres. Le modèle reproduit bien la dépendance en température de la chaleur spécifique de l'hélium II et de la proportion ρ_n/ρ de fluide normal.

Cependant, bien des questions restent encore sans réponse satisfaisante. Les plus importantes sont sans doute celles qui concernent la transition de phase hélium I-hélium II : malgré des progrès récents (cf. note 13), sa nature continue à poser problème. Mais nous avons aussi mentionné en passant la difficulté liée aux valeurs des vitesses critiques auxquelles l'écoulement de l'hélium II cesse d'être superfluide

(§ 2.a). En outre, on aimerait comprendre plus clairement la structure physique du condensat (c'est-à-dire connaître sa fonction d'onde) et l'origine de la forme assez surprenante du spectre des excitations élémentaires (fig. 9). Des considérations théoriques dues principalement à Feynman (réf. 1, p. 348) ont permis des progrès intéressants sur ces problèmes fondamentaux.

Références

1. D.L. Goodstein, *States of Matter,* Prentice Hall, Englewood Cliffs, N.J. (1975), paragraphe 5.2.
2. K. Huang, *Statistical Mechanics,* John Wiley, New York (1963), chapitre 18.
3. M.W. Zemansky, *Heat and Thermodynamics,* McGraw-Hill Kogakusha, Tokyo (1968), chapitres 12 et 15.
4. R.B. Hallock, « Superfluid Helium », Resource Letter SH1, *American Journal of Physics 50* (3), 202 (1982) : liste de références classées et commentées.
5. J. Wilks, *An Introduction to Liquid Helium,* Clarendon Press, Oxford (1970).

COMPLÉMENT VI.E

Émission et absorption de rayonnement par un atome. Coefficients d'Einstein

I. Définition des coefficients d'Einstein
1. Coefficient d'émission spontanée
2. Coefficients d'Einstein pour l'absorption et l'émission induite

II. Relations entre les coefficients d'Einstein
1. Méthode de raisonnement
2. Analyse de l'équilibre entre les atomes et le gaz de photons du corps noir
3. Discussion succincte des résultats

I. DÉFINITION DES COEFFICIENTS D'EINSTEIN

Considérons un atome ou une molécule, et étudions les transitions entre deux de ses niveaux d'énergie E_1 et E_2 ($E_2 > E_1$). Ces transitions s'accompagnent de l'émission ou l'absorption d'un photon de pulsation

$$\omega_{21} = \frac{E_2 - E_1}{\hbar} . \tag{E.1}$$

Nous noterons g_1 et g_2 les degrés de dégénérescence de ces deux niveaux, c'est-à-dire le nombre d'*états* quantiques indépendants qui leur sont associés[1].

1. COEFFICIENT D'ÉMISSION SPONTANÉE

Si l'atome se trouve à un instant t dans le niveau supérieur E_2, il peut « redescendre » dans le niveau E_1 en émettant spontanément un photon de pulsation ω_{21}. Comme tous les processus quantiques, cette *émission spontanée* se produit avec une certaine *probabilité*. Celle-ci est indépendante de l'instant t considéré (c'est-à-dire du temps qui s'est écoulé depuis que l'atome a été porté dans le niveau excité E_2). On peut donc définir une *probabilité de transition par unité de temps* A_{12} du niveau E_2 au niveau E_1 (cf. IV, § I.A.2), que l'on appelle *« coefficient d'émission spontanée »*.

Pour préciser, supposons que l'on observe un grand nombre d'atomes identiques. Soit $N_2(t)$ le nombre de ceux qui, à l'instant t, se trouvent dans le niveau E_2. Entre

1. Une raie atomique présente souvent une structure relativement compliquée (« stucture fine », « structure hyperfine ») : chacun des deux niveaux E_1 et E_2 est en réalité scindé en plusieurs niveaux proches. En outre, le photon émis ou absorbé lors du passage d'un état déterminé de l'un des niveaux à un état déterminé de l'autre possède un état de polarisation bien précis.

Nous ignorons ici ces détails et traitons globalement les g_1 états du niveau E_1 et les g_2 états du niveau E_2 : nous supposons que l'analyse du rayonnement émis ou absorbé ne cherche pas à atteindre cette précision. En particulier, le rayonnement électromagnétique qui baigne l'atome et y induit des transitions (§ 2) est non polarisé et sa distribution spectrale est très large devant les intervalles caractéristiques du spectre atomique.

Il en irait tout autrement pour l'étude d'un laser ou d'un maser, où ces hypothèses seraient totalement inadaptées. Il faudrait alors considérer séparément chacun des états atomiques et définir les coefficients d'Einstein pour la transition entre deux états, et non plus deux niveaux.

t et l'instant $t+\mathrm{d}t$ infiniment voisin, le nombre $\mathrm{d}N^{(s)}_{2\to 1}$ d'atomes qui passent du niveau E_2 au niveau E_1 par le processus d'émission spontanée s'écrit, par *définition* du coefficient d'émission spontanée A_{12},

$$\mathrm{d}N^{(s)}_{2\to 1} \mathrel{\widehat{=}} A_{12}N_2(t)\,\mathrm{d}t\,. \tag{E.2}$$

Le coefficient A_{12} est *indépendant du temps;* c'est une *caractéristique de l'atome considéré* et de ses niveaux E_2 et E_1.

2. COEFFICIENTS D'EINSTEIN POUR L'ABSORPTION ET L'ÉMISSION INDUITE

Supposons maintenant que l'atome soit soumis à un rayonnement électromagnétique dont la *densité d'énergie* par unité de volume, dans la bande infinitésimale de pulsation $(\omega,\ \omega+\mathrm{d}\omega)$ est $u(\omega)\mathrm{d}\omega$.

Si l'atome se trouve dans le niveau inférieur E_1, il peut être excité vers le niveau E_2 par *absorption* d'un photon du rayonnement ambiant, de pulsation ω_{21}. La probabilité par unité de temps correspondante est proportionnelle au nombre de photons disponibles, c'est-à-dire à la densité spectrale d'énergie $u(\omega_{21})$ pour la *«pulsation de résonance»* ω_{21}. Le *coefficient d'Einstein pour l'absorption* B_{21} est alors défini par[2]

$$\mathrm{d}N^{(a)}_{1\to 2} \mathrel{\widehat{=}} B_{21}u(\omega_{21})\,N_1(t)\,\mathrm{d}t\ : \tag{E.3}$$

$N_1(t)$ est le nombre d'atomes de l'échantillon observé qui se trouvent à l'instant t dans le niveau E_1, et $\mathrm{d}N^{(a)}_{1\to 2}$ le nombre de ceux qui passent de E_1 à E_2 par absorption d'un photon, entre t et $t+\mathrm{d}t$.

Mais l'expérience montre que la présence d'un rayonnement de pulsation appropriée produit également le phénomène inverse du précédent : l'*«émission induite»*, ou *«stimulée»*, de photons par les atomes qui se trouvent dans le niveau supérieur E_2. Ce phénomène, prévu par Einstein en 1917, se décrit comme en (E.3) : il provoque, entre t et $t+\mathrm{d}t$, le passage de $\mathrm{d}N^{(i)}_{2\to 1}$ atomes du niveau E_2 au niveau E_1, selon la formule

$$\mathrm{d}N^{(i)}_{2\to 1} \mathrel{\widehat{=}} B_{12}u(\omega_{21})\,N_2(t)\,\mathrm{d}t\,, \tag{E.4}$$

qui définit le *coefficient d'émission induite* B_{12}.

Comme A_{12}, les coefficients B_{21} et B_{12} sont *indépendants du temps* et *caractéristiques de l'atome considéré.*

II. RELATIONS ENTRE LES COEFFICIENTS D'EINSTEIN

1. MÉTHODE DE RAISONNEMENT

Puisque les coefficients d'Einstein sont des constantes caractéristiques de l'atome, nous pouvons nous placer pour les étudier dans une *situation très particulière :* un *très grand nombre d'atomes* sont enfermés dans une enceinte maintenue à la *température* T par un *thermostat;* cette enceinte contient évidemment aussi un *gaz de photons,* qui interagit avec les atomes par les processus d'absorption

2. L'atome peut en fait absorber des photons de pulsation ω légèrement différente de ω_{21}, avec une probabilité qui décroît très rapidement lorsque $|\omega-\omega_{21}|$ croît : la raie atomique que nous étudions a une certaine *largeur* $\Delta\omega$, qui dépend de l'environnement de l'atome (température, pression, ...) mais ne peut pas être inférieure à une valeur minimale non nulle, dite « largeur naturelle » de la raie, caractéristique intrinsèque des niveaux E_2 et E_1. Nous supposons ici que la fonction $u(\omega)$, qui décrit la répartition de l'énergie du rayonnement incident en fonction de la pulsation, est pratiquement constante sur tout la largeur $\Delta\omega$ de la plage d'absorption (cf. note 1). C'est pourquoi la définition (E.3) fait intervenir simplement ω_{21}.

et d'émission (spontanée ou induite) que nous avons décrits au paragraphe I; nous supposons que l'ensemble a atteint l'*équilibre macroscopique.*

Comme nous connaissons par ailleurs la population d'équilibre des niveaux E_1 et E_2 (distribution canonique) et la densité spectrale d'énergie $u_P(\omega; T)$ du rayonnement à l'équilibre thermique (loi de Planck), l'analyse de l'équilibre entre les atomes et le rayonnement va nous fournir des *relations entre les coefficients d'Einstein*[3]. Insistons sur le fait que les relations ainsi obtenues seront *valables dans n'importe quelle autre situation;* en particulier, elles sont applicables aux *situations hors d'équilibre* dans lesquelles se produisent le plus souvent les processus d'émission et d'absorption de lumière.

2. ANALYSE DE L'ÉQUILIBRE ENTRE LES ATOMES ET LE GAZ DE PHOTONS DU CORPS NOIR

Revenons donc à la situation d'équilibre macroscopique à la température T, et examinons ses implications.

Les populations N_1 et N_2 des niveaux E_1 et E_2 sont indépendantes du temps; il faut donc qu'il y ait constamment, pour chaque niveau, *compensation entre son dépeuplement et son peuplement.* On pourrait penser qu'il est nécessaire de considérer à la fois tous les niveaux de l'atome qui sont couplés par des transitions radiatives, puisqu'ils concourent tous au peuplement et au dépeuplement de chacun d'eux. Mais nous avons vu au chapitre IV (§ II.A) que l'invariance par renversement du temps des lois qui régissent les processus microscopiques a pour conséquence la *relation de «bilan détaillé» :* pour chaque couple de niveaux tels que E_1 et E_2, le peuplement de l'un d'eux à partir de l'autre est exactement compensé par son dépeuplement vers ce même niveau[4]. Il est donc légitime de considérer seulement les deux niveaux E_1 et E_2, et d'écrire que le nombre d'atomes qui passent, durant l'intervalle de temps infinitésimal $(t, t+dt)$, du niveau E_1 au niveau E_2 est égal au nombre de ceux qui passent de E_2 à E_1 durant le même temps :

$$B_{21}u(\omega_{21})N_1=\left[A_{12}+B_{12}u(\omega_{21})\right]N_2 \ . \tag{E.5}$$

A cette relation il faut adjoindre celle qui donne le rapport N_1/N_2 des populations des deux niveaux à l'équilibre :

$$\frac{N_1}{N_2}=\frac{g_1\exp[-E_1/kT]}{g_2\exp[-E_2/kT]}\ , \tag{E.6}$$

et l'expression de la densité spectrale d'énergie du gaz de photons à l'équilibre thermique :

$$u(\omega)=u_P(\omega;\,T)=\frac{\hbar}{\pi^2c^3}\frac{\omega^3}{e^{\hbar\omega/kT}-1}\ . \tag{E.7}$$

---

3. Ce raisonnement est fondé sur l'hypothèse que *l'équilibre est possible* entre les atomes et le rayonnement, l'ensemble étant maintenu à la température T. L'expérience montre que cette hypothèse est effectivement vérifiée. Le calcul théorique des coefficients d'Einstein, qui est possible dans le cadre de l'électrodynamique quantique, confirme la validité de cette hypothèse, car les résultats obtenus satisfont aux relations que nous allons démontrer ici.

4. Plus exactement, le bilan détaillé s'applique à l'ensemble du système atomes + photons : pendant un intervalle de temps donné, il y a autant d'atomes qui passent du niveau E_2 au niveau E_1 en émettant chacun un photon de pulsation ω_{21}, que d'atomes qui passent de E_1 à E_2 en absorbant un tel photon.

On obtient ainsi[5]

$$[B_{21}g_1\, e^{\hbar\omega_{21}/kT} - B_{12}g_2]\,\frac{\hbar}{\pi^2c^3}\,\frac{\omega_{21}^3}{e^{\hbar\omega_{21}/kT}-1} = A_{12}g_2\ . \tag{E.8}$$

Cette égalité doit être vérifiée *quelle que soit la température T;* on en déduit les deux relations suivantes entre les coefficients d'Einstein[6] :

$$B_{21}g_1 = B_{12}g_2\ , \tag{E.9.a}$$

$$\frac{A_{12}}{B_{12}} = \frac{\hbar}{\pi^2c^3}\,\omega_{21}^3 = \frac{8\pi\hbar}{\lambda_{21}^3} \tag{E.9.b}$$

(où λ_{21} est la longueur d'onde associée à la transition).

Il est intéressant de noter que *le rôle du phénomène d'émission induite* (découvert par Einstein en 1917) *est crucial* dans l'établissement de l'équilibre avec le gaz de photons : sans le terme correspondant, il serait impossible d'identifier les deux membres de la condition (E.8) pour toute température T.

3. DISCUSSION SUCCINCTE DES RÉSULTATS

Les relations (E.9) sont très importantes en physique atomique. La première montre que *l'émission induite est du même ordre de grandeur que l'absorption :* le rapport g_2/g_1 est en effet toujours de l'ordre de l'unité. La deuxième relation permet de *comparer l'émission induite et l'émission spontanée.* La comparaison fait intervenir la densité d'énergie du rayonnement $u(\omega_{21})$, qui varie selon les conditions expérimentales :

$$\frac{dN_{2\to 1}^{(i)}}{dN_{2\to 1}^{(s)}} = \frac{B_{12}u(\omega_{21})}{A_{12}} = u(\omega_{21})\,\frac{\lambda_{21}^3}{8\pi\hbar}\ . \tag{E.10}$$

On constate cependant que, *à densité d'énergie électromagnétique constante, l'importance relative de l'émission induite par rapport à l'émission spontanée croît comme le cube de la longueur d'onde.*

Le phénomène d'émission induite et ses propriétés sont à la base du fonctionnement des masers et lasers, que l'on peut décrire schématiquement comme suit. Par un procédé de « pompage » approprié, on prépare des atomes ou molécules dans un état excité (avec nos notations, un des états du niveau E_2). Lorsqu'ils reçoivent des photons de pulsation égale à celle d'une des transitions qu'ils sont susceptibles d'effectuer pour se désexciter, cette transition est stimulée et ils émettent alors un photon présentant les mêmes caractéristiques (pulsation, polarisation, direction *et phase*) que les photons incidents. Ainsi se crée (la désexcitation spontanée de quelques atomes suffit à enclencher le processus) un faisceau de photons parfaitement cohérent, dont l'intensité s'amplifie à la traversée du milieu « actif » (dans lequel le pompage

---

5. Compte tenu de (E.6) et (E.7), la relation (E.5) s'écrit

$$[B_{21}g_1\, e^{-E_1/kT} - B_{12}g_2\, e^{-E_2/kT}]\,\frac{\hbar}{\pi^2c^3}\,\frac{\omega_{21}^3}{e^{\hbar\omega_{21}/kT}-1} = A_{12}g_2\, e^{-E_2/kT}\ , \tag{1}$$

c'est-à-dire, si l'on multiplie les deux membres par $\exp E_2/kT$ et qu'on utilise l'expression (E.1) de ω_{21},

$$[B_{21}g_1\, e^{\hbar\omega_{21}/kT} - B_{12}g_2]\,\frac{\hbar}{\pi^2c^3}\,\frac{\omega_{21}^3}{e^{\hbar\omega_{21}/kT}-1} = A_{12}g_2\ . \tag{2}$$

6. Il suffit, après multiplication des deux membres de l'égalité (E.8) par $[\exp(\hbar\omega_{21}/kT)-1]$, d'identifier les termes en $\exp(\hbar\omega_{21}/kT)$ et les termes constants.

renouvelle la population de l'état excité). On comprend que, dans ce processus d'amplification cohérente, le phénomène d'émission spontanée est une gêne : il réduit le nombre d'atomes excités susceptibles de participer par émission induite à l'amplification du faisceau, et fournit des photons certes analogues, mais émis au hasard et sans lien de cohérence avec ceux du faisceau lorsque ce dernier est déjà constitué. La relation (E.10) permet alors de comprendre qualitativement pourquoi le maser a été plus facile à réaliser que le laser : dans le domaine des micro-ondes[7], l'émission induite domine l'émission spontanée pour des densités d'énergie que l'on peut obtenir sans difficulté; dans le domaine visible ou proche du visible, l'émission spontanée devient beaucoup plus importante[8]. La réalisation de lasers de longueur d'onde encore plus courte (rayons X mous) achoppe à l'heure actuelle sur des difficultés considérables, dont l'une provient de cette croissance très rapide du rapport entre l'émission spontanée et l'émission induite.

Références

1. R. Loudon, *The Quantum Theory of Light,* Clarendon Press, Oxford (1973).

2. D.C. O'Shea, W.R. Callen et W.T. Rhodes, *Introduction aux lasers et à leurs applications,* Eyrolles, Paris (1980).

3. J.T. Verdeyen, *Laser Electronics,* Prentice Hall, Englewood Cliffs, N.J. (1981).

7. Le « maser à ammoniac », par exemple, fonctionne sur une longueur d'onde de 1,2 cm associée à l'« inversion » de la molécule NH_3, c'est-à-dire au passage du plan des trois hydrogènes d'un côté à l'autre de l'atome d'azote (C.D.L., p. 453). Le « maser à hydrogène » utilise la transition entre les deux sous-niveaux de structure hyperfine du niveau fondamental $1s_{1/2}$ de l'atome d'hydrogène (couplage entre le spin de l'électron et celui du proton — C.D.L., p. 1222); la fréquence associée à cette transition est connue expérimentalement avec une très grande précision; elle correspond à une longueur d'onde voisine de 21 cm. Il existe évidemment beaucoup d'autres systèmes atomiques permettant la construction de masers, dans le domaine des ondes centimétriques ou millimétriques.

8. Lorsque la longueur d'onde λ_{21} varie de 10 cm (micro-ondes) à 0,5 μm (visible), le facteur λ_{21}^3 de la formule (E.10) décroît d'un facteur 10^{16}.

COMPLÉMENT VI.F

Absorption et émission de rayonnement par un corps macroscopique. Loi de Kirchhoff

I. Définitions
1. Émissivité en un point
2. Pouvoir absorbant en un point
3. Le corps noir

II. Relation entre émission et absorption
1. Analyse de l'équilibre statistique
2. Conséquence : loi de Kirchhoff
3. Le rayonnement du corps noir
4. Rayonnement thermique d'un corps quelconque

L'expérience courante nous apprend qu'un corps suffisamment chaud émet de la lumière : filament de lampe, « fer rouge » ou « fer chauffé à blanc », ... Nous nous proposons d'étudier ici le phénomène d'*émission de rayonnement thermique*[1] par un corps macroscopique quelconque, liquide ou solide, et de montrer que ce phénomène est relié à la capacité d'*absorption* que manifeste ce même corps pour un rayonnement reçu de l'extérieur.

I. DÉFINITIONS

Considérons un corps $\mathcal{C}$ porté à la température T. Nous choisissons un point M de sa surface extérieure s, et un élément de surface ds autour de ce point; la normale à la surface s en M, orientée vers l'extérieur, nous servira d'axe de référence.

1. ÉMISSIVITÉ EN UN POINT

La *puissance* d$\mathcal{P}$ *rayonnée par l'élément de surface* ds *au point* M, *dans l'angle solide infinitésimal* dΩ *autour de la direction de vecteur unitaire* $\vec{K}$, *et dans la bande infinitésimale de pulsations* $(\omega, \omega + \mathrm{d}\omega)$, s'écrit (fig. 1)

$$\mathrm{d}\mathcal{P} \widehat{=} \eta(\omega, \vec{K}; M, T)\,\mathrm{d}\omega.\mathrm{d}\Omega.\mathrm{d}s \ ; \qquad \text{(F.1)}$$

cette relation définit l'*émissivité* η du corps $\mathcal{C}$, qui a pour dimensions celles d'une énergie par unité de surface ($\mathrm{J/m^2}$). Elle dépend de la nature du corps considéré, du point M choisi sur sa surface, de la température T à laquelle il est porté, de la

1. De multiples autres causes peuvent amener un corps à émettre du rayonnement électromagnétique : une antenne parcourue par un courant alternatif de haute fréquence émet des ondes radio; un gaz soumis à une décharge électrique émet un spectre de raies caractéristique des atomes ou molécules qui le composent; une anticathode métallique bombardée par un faisceau d'électrons produit des rayons X; une source radioactive émet un rayonnement γ; de façon plus générale, un corps qui reçoit de l'énergie de l'extérieur, ou qui est le siège d'une réaction chimique libérant de l'énergie, peut en réémettre une partie sous forme de luminescence.

.../...

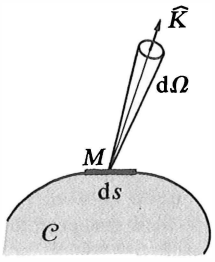


FIGURE 1
On analyse l'émission par l'élément de surface ds du corps $\mathcal{C}$ situé au voisinage du point M en isolant les vecteurs unitaires $\hat{K}$ dont la direction appartient à un angle solide infinitésimal $d\Omega$, et les pulsations comprises entre ω et $\omega + d\omega$.

pulsation ω du rayonnement observé et de la direction dans laquelle il est émis[2]; si θ et φ ($0 \leqslant \theta \leqslant \pi$, $0 \leqslant \varphi < 2\pi$) sont les angles qui repèrent la direction du vecteur unitaire $\hat{K}$ (par rapport à la normale à s au point M et à un axe perpendiculaire à cette normale servant d'origine pour φ), l'angle solide élémentaire $d\Omega$ a pour expression

$$d\Omega = \sin\theta \, d\theta \, d\varphi \, . \tag{F.2}$$

2. POUVOIR ABSORBANT EN UN POINT

Supposons maintenant que le corps $\mathcal{C}$ que nous étudions reçoive un rayonnement électromagnétique produit par des sources extérieures. Plus précisément, nous noterons

$$\varpi(\omega', \hat{K}'; M) d\omega' . d\Omega' . ds \tag{F.3}$$

la *puissance reçue en M par l'élément de surface* ds, *dans la bande infinitésimale de pulsations* $(\omega', \omega' + d\omega')$, *et qui arrive dans l'angle solide* $d\Omega'$ *autour de la direction de vecteur unitaire* $\hat{K}'$. La grandeur ϖ a mêmes dimensions que l'émissivité η (énergie par unité de surface).

Cette puissance reçue au point M par ds est en partie absorbée par le corps $\mathcal{C}$; le reste est renvoyé vers l'extérieur, par réflexion ou diffusion (nous supposons le corps suffisamment épais pour qu'aucun rayonnement ne puisse le traverser). On définit le *pouvoir absorbant* $\mathcal{A}$ du corps au point M comme *la fraction de la puissance reçue qui est absorbée.* Le pouvoir absorbant est donc un *nombre sans dimension;* il dépend de la nature du corps, du point M et de la température T, mais aussi de la pulsation ω' et de la direction $\hat{K}'$ du rayonnement incident[3] :

$$\mathcal{A} = \mathcal{A}(\omega', \hat{K}'; M, T) \, . \tag{F.4}$$

Ces diverses causes particulières ou extérieures sont supposées absentes dans les situations que nous envisageons ici. Le rayonnement thermique est universel; il est dû à l'énergie emmagasinée dans tout corps macroscopique maintenu à une température non nulle.

2. L'émissivité peut également, dans certains cas, dépendre de l'*état de polarisation* du rayonnement émis. Pour simplifier, nous négligeons ici cet effet.

3. Le pouvoir absorbant peut également, dans certains cas particuliers, dépendre de la polarisation du rayonnement reçu.

3. LE CORPS NOIR

On appelle ainsi un corps dont *le pouvoir absorbant est égal à* 1 *pour toutes les pulsations* ω' *et toutes les directions* $\hat{K}'$[4] :

$$\mathcal{A}_N(\omega', \hat{K}'; M, T) = 1 \quad \text{pour tous } \omega' \text{ et } \hat{K}' . \tag{F.5}$$

Le corps noir absorbe donc intégralement la puissance du rayonnement incident au point M. On l'appelle ainsi précisément parce que, éclairé par un rayonnement quelconque, il n'en réfléchit ni diffuse aucune composante, ce qui le fait apparaître noir[5].

Une surface enduite de noir de fumée se comporte approximativement comme un corps noir, au moins pour les radiations visibles et proches du visible. Mais la réalisation pratique la plus conforme à la définition du corps noir consiste à percer un trou de faible dimension dans une enceinte parfaitement close par ailleurs (fig. 2) :

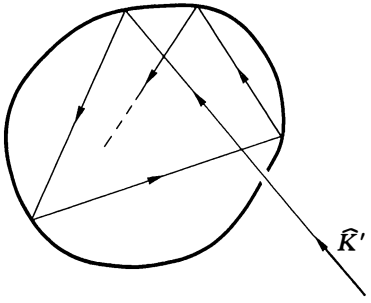


FIGURE 2

Réalisation pratique du corps noir par un trou percé dans une enceinte fermée : un rayonnement quelconque, dirigé vers le trou, est « piégé » à l'intérieur de l'enceinte.

un rayonnement quelconque dirigé vers ce trou sous n'importe quelle incidence pénètre dans la cavité et n'a pratiquement aucune chance d'en ressortir après réflexion et diffusion sur les faces internes. C'est parce qu'on se réfère toujours à ce prototype que l'on dit *« le »* corps noir[6].

4. Il peut arriver qu'un corps se comporte comme un corps noir seulement dans un certain domaine de température.

5. Nous verrons plus loin que le corps noir émet lui-même du rayonnement (formule (F.11)). Il apparaît véritablement noir à l'œil si sa température n'est pas trop élevée : il n'émet alors pratiquement pas dans le domaine visible; si on l'éclaire avec de la lumière visible, il l'absorbe donc entièrement et ne renvoie aucune couleur.

6. Dans le chapitre VI, nous avons défini le « corps noir » comme un gaz de photons en équilibre thermique avec un thermostat de température T. Effectivement, dans une enceinte fermée maintenue à la température T se créent spontanément des photons qui constituent progressivement un gaz à l'équilibre thermique avec les parois; si l'on perce un trou dans l'enceinte pour observer ce gaz, on retrouve la situation définie ici comme réalisant le « corps noir ». Le lien entre les propriétés du gaz de photons et celles du rayonnement s'échappant par le trou apparaîtra au paragraphe II.3.

On parle parfois aussi de « corps blanc », dont le pouvoir absorbant serait identiquement nul, et de « corps gris » pour les cas intermédiaires.

II. RELATION ENTRE ÉMISSION ET ABSORPTION

1. ANALYSE DE L'ÉQUILIBRE STATISTIQUE

Plaçons d'abord le corps $\mathcal{C}$ que nous voulons étudier dans une *situation très particulière,* mais que nous savons analyser : nous le mettons *en équilibre thermique avec un gaz de photons* à la température T. Il suffit pour cela de l'introduire à l'intérieur d'une enceinte fermée maintenue à la température T et d'attendre que l'équilibre se rétablisse. Dans cette situation, le corps $\mathcal{C}$ manifeste à la fois son émissivité et son pouvoir absorbant, qui prennent là leur valeur correspondant à la température T.

Dans cette situation d'équilibre, la puissance totale rayonnée par le corps $\mathcal{C}$ est égale à la puissance totale qu'il absorbe[7] :

$$\int \mathrm{d}\omega \, \mathrm{d}\Omega \, \mathrm{d}s \left[\eta(\omega, \hat{K}; M, T) - \mathcal{A}(\omega, -\hat{K}; M, T) \varpi(\omega, -\hat{K}; M, T) \right] = 0 \, . \quad \text{(F.6)}$$

Mais cette compensation n'est pas seulement globale; elle se produit *séparément en chaque point de la surface, pour chaque pulsation et chaque direction.* En effet, supposons que la fonction $(\eta - \mathcal{A}\varpi)$ à intégrer dans la condition d'équilibre (F.6) soit non nulle, disons positive, sur un petit domaine δ_1 de valeurs des variables ω, θ, φ et M (la température est fixée). La contribution de ce domaine δ_1 à l'intégrale doit être compensée par celle d'un autre domaine δ_2 dans lequel la fonction est négative. Mais il est possible de modifier l'émissivité et le pouvoir absorbant du corps $\mathcal{C}$ dans le domaine δ_2 sans toucher au domaine δ_1 : on peut par exemple recouvrir un élément de surface $\mathrm{d}s_2$ de δ_2 par un écran opaque bloquant les échanges entre cet élément de surface et le rayonnement qui règne dans la cavité; ou bien on y place une lame absorbant sélectivement une bande étroite de pulsations $(\omega, \omega + \delta\omega)$; ou bien encore on déforme légèrement la surface de l'élément $\mathrm{d}s_2$ de façon à changer en ce point la dépendance de η et $\mathcal{A}$ par rapport à la direction $\hat{K}$ d'émission ou d'absorption. Des interventions de ce genre vont évidemment détruire la compensation supposée entre les contributions des domaines δ_1 et δ_2, sans remettre en cause l'équilibre entre le corps $\mathcal{C}$ et le rayonnement. La condition d'équilibre (F.6) ne peut donc être vérifiée que par annulation de la fonction à intégrer, en chaque point du domaine d'intégration.

Finalement, si nous notons $\varpi_P(\omega', \hat{K}'; M, T)$ la puissance fournie au point M par le *gaz de photons à l'équilibre thermique,* pour la pulsation ω' et selon la direction $\hat{K}'$, le bilan d'équilibre entre le corps $\mathcal{C}$ et ce gaz de photons s'écrit[8]

$$\eta(\omega, \hat{K}; M, T) = \mathcal{A}(\omega, -\hat{K}; M, T) \varpi_P(\omega, -\hat{K}; M, T) \, . \quad \text{(F.7)}$$

---

7. Un vecteur unitaire $\hat{K}$ ayant été choisi, nous considérons le rayonnement émis dans sa direction, et le rayonnement reçu depuis cette même direction, c'est-à-dire selon le vecteur unitaire $-\hat{K}$.

8. Ceci est l'expression, dans la situation d'équilibre particulière examinée ici, de la *« relation de bilan détaillé »*, qui est une propriété générale des distributions d'équilibre (IV, § II.A) : la probabilité pour qu'un photon de pulsation ω et de direction $\hat{K}$ soit émis est égale à la probabilité pour qu'un photon de même pulsation et de direction $\hat{K}' = -\hat{K}$ soit absorbé. Le changement de $\hat{K}$ en $-\hat{K}$, compréhensible vu la géométrie du problème (note 7), provient de l'application du renversement du temps, qui est à l'origine de la relation de bilan détaillé.

En reprenant le calcul qui nous a conduits à la loi de Planck (VI, § III.A.2 et § III.B.1.a), on trouve facilement[9]

$$\varpi_P(\omega, -\vec{K}; M, T)=\frac{c}{4\pi} u_p(\omega; T) \cos\theta , \tag{F.8}$$

où θ est l'angle entre le vecteur unitaire $\vec{K}$ et la normale en M à la surface s, et $u_P(\omega; T)$ la densité spectrale d'énergie du gaz de photons, donnée par la loi de Planck VI.(III.23).

---

9. On obtient le nombre $dN_P(\vec{K}'; T)$ de photons ayant, à la température T, le vecteur d'onde $\vec{K}'$ à d^3K' près en multipliant le nombre de modes propres de la cavité ayant ce vecteur d'onde, égal (chapitre VI, note 67) à

$$2\times\frac{V}{(2\pi)^3} d^3K' , \tag{1}$$

par le nombre d'occupation moyen de chacun de ces modes propres :

$$dN_P(\vec{K}'; T)=2\frac{V}{(2\pi)^3}\frac{d^3K'}{e^{\hbar\omega'/kT}-1} . \tag{2}$$

Ces photons se déplacent tous à la vitesse c dans la direction de $\vec{K}'$, chacun transportant une énergie $\hbar\omega'=\hbar|\vec{K}'|c$.

Ceux qui frappent l'élément de surface ds (centré en M) entre t et $t+dt$ sont ceux qui étaient situés à l'instant t dans le cylindre de base ds et de génératrices parallèles à $\vec{K}'$ et ayant pour longueur $c\,dt$ (fig. 3). L'énergie reçue par ds pendant l'intervalle de temps dt vaut donc

$$\varpi_P(\omega', \vec{K}'; M, T)dt\, ds\, d\Omega'\, d\omega'=\hbar\omega'.dN_P(\vec{K}'; T)\frac{c\, dt\, ds\, \cos\theta}{V} , \tag{3}$$

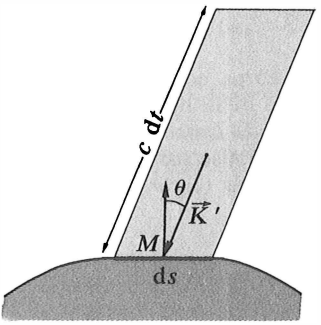


FIGURE 3

Les photons qui frappent, entre t et $t+dt$, l'élément de surface ds situé en M sont ceux qui se trouvaient à l'instant t dans le cylindre de base ds et de génératrices $c\,dt$ parallèles au vecteur d'onde $\vec{K}'$.

où θ est l'angle entre $-\vec{K}'$ et la normale en M à la surface s. Comme

$$d^3K'=K'^2\, dK'\, d\Omega'=\frac{\omega'^2\, d\omega'}{c^3} d\Omega' , \tag{4}$$

on en déduit

$$\varpi_P(\omega', \vec{K}'; M, T)=\frac{\hbar}{4\pi^3}\frac{\omega'^3}{c^2}\frac{1}{e^{\hbar\omega'/kT}-1}\cos\theta=\frac{c}{4\pi} u_P(\omega'; T)\cos\theta . \tag{5}$$

2. CONSÉQUENCE : LOI DE KIRCHHOFF

Bien que démontrée dans des conditions particulières, la relation (F.7) a une portée générale. Écrivons-la

$$\frac{\eta(\omega, \widehat{K}; M, T)}{\mathcal{A}(\omega, -\widehat{K}; M, T)} = \varpi_P(\omega, -\widehat{K}; M, T) . \tag{F.9}$$

Sous cette forme, on constate que *le premier membre ne fait intervenir que des caractéristiques intrinsèques du corps étudié :* l'émissivité η et le pouvoir absorbant $\mathcal{A}$ sont indépendants des conditions dans lesquelles est placé le corps $\mathcal{C}$, et en particulier du rayonnement auquel il est soumis[10]; ils restent les mêmes lorsque le corps émet ou absorbe du rayonnement dans une *situation hors d'équilibre.* Dans le second membre, au contraire, figure une *fonction universelle* de la température, de la pulsation ω et de la direction (formule (F.8)).

On en tire une première conséquence importante *(loi de Kirchhoff) : le rapport entre l'émissivité et le pouvoir absorbant est indépendant du corps considéré,* et du point choisi sur sa surface. Exprimée sous forme plus qualitative, cette loi affirme qu'*un bon absorbeur de rayonnement est aussi un bon émetteur,* ceci étant vrai pour chaque point particulier, chaque pulsation et chaque direction.

3. LE RAYONNEMENT DU CORPS NOIR

Le corps noir, au pouvoir absorbant égal à 1 dans tout le spectre, est un cas particulièrement intéressant. La relation (F.7) devient dans ce cas

$$\eta_N(\omega, \widehat{K}; T) = \varpi_P(\omega, -\widehat{K}; T) . \tag{F.10}$$

L'émissivité du corps noir est *indépendante de la nature du matériau* qui le constitue; elle sert de référence parce qu'elle est caractérisée par la *fonction universelle* (F.8) déduite des propriétés d'un gaz de photons en équilibre thermique; ainsi, les propriétés du gaz de photons peuvent être directement mesurées en étudiant l'émissivité du corps noir. Si l'on explicite selon (F.8) le second membre de (F.10), on obtient la *loi du rayonnement du corps noir :*

$$\eta_N(\omega, \widehat{K}; T) = \frac{c}{4\pi} u_P(\omega; T) \cos\theta . \tag{F.11}$$

On constate tout d'abord que l'émissivité du corps noir est proportionnelle à $\cos\theta$, où θ est l'angle entre la direction d'observation et la normale à la surface du corps noir. Ce résultat est connu sous le nom de *« loi de Lambert ».*

Calculons ensuite la puissance $\Pi_N(\omega; T)\, d\omega\, ds$ émise par l'élément de surface ds du corps noir dans la bande infinitésimale de pulsations $(\omega, \omega + d\omega)$. Il suffit pour cela d'intégrer sur toutes les directions, ce qui donne[11]

$$\Pi_N(\omega; T) = \frac{1}{4} c\, u_P(\omega; T) . \tag{F.12}$$

---

10. Ce n'est que dans des conditions extrêmes, pour de très fortes intensités de rayonnement, que celui-ci peut modifier les propriétés du corps $\mathcal{C}$: avec des faisceaux laser de forte puissance, on peut par exemple provoquer la fusion du matériau. Nous excluons évidemment de telles conditions dans la présente étude.

11. En effet, d'après la définition (F.1),

$$\Pi_N(\omega)\, d\omega\, ds = d\omega\, ds \int d\Omega\, \eta_N(\omega, \widehat{K}; T)$$

$$= d\omega\, ds \frac{c}{4\pi} u_P(\omega; T) \int_0^{2\pi} d\varphi \int_0^{\pi/2} \cos\theta\, \sin\theta\, d\theta$$

(noter que l'intégration sur θ porte sur la moitié de l'intervalle habituel, car l'émission est observée seulement vers l'extérieur du corps noir). La formule (F.12) en découle aussitôt.

Donc, *une mesure de la puissance émise par unité de surface du corps noir donne directement la densité volumique d'énergie du gaz de photons en équilibre thermique;* c'est pourquoi la loi de Planck, destinée initialement à décrire le rayonnement du corps noir, a pu être déduite au chapitre VI des propriétés d'un gaz de photons.

La *puissance totale émise par unité de surface* obéit à la *«loi de Stefan-Boltzmann»* :

$$\int_0^{+\infty} \mathrm{d}\omega \, \Pi_N(\omega; T) = \sigma T^4 , \tag{F.13}$$

avec (voir formule VI.(III.35))

$$\sigma = \frac{\pi^2}{60} \frac{k^4}{c^2 \hbar^3} . \tag{F.14}$$

4. RAYONNEMENT THERMIQUE D'UN CORPS QUELCONQUE

Considérons maintenant un corps $\mathcal{C}$ quelconque, maintenu à la température T, et un point M de sa surface. Nous saurons calculer l'émissivité de $\mathcal{C}$ en M si nous connaissons son pouvoir absorbant $\mathcal{A}(\omega', \hat{K}'; M, T)$; on a en effet, d'après (F.7) et (F.10),

$$\eta(\omega, \hat{K}; M, T) = \mathcal{A}(\omega, -\hat{K}; M, T) \, \eta_N(\omega, \hat{K}; T) . \tag{F.15}$$

Donc, *l'émissivité d'un corps quelconque est égale à celle du corps noir modulée par le pouvoir absorbant du corps étudié.* Si l'on explicite l'émissivité η_N du corps noir selon la formule (F.11), on obtient (cf. formule VI.(III.23))

$$\eta(\omega, \hat{K}; M, T) = \mathcal{A}(\omega, -\hat{K}; M, T) \frac{\hbar}{4\pi^3 c^2} \frac{\omega^3}{\mathrm{e}^{\hbar\omega/kT} - 1} \cos\theta . \tag{F.16}$$

Si le pouvoir absorbant $\mathcal{A}$ du corps $\mathcal{C}$ au point M est *indépendant de la direction du vecteur* $\hat{K}$, la puissance émise est proportionnelle à $\cos\theta$, c'est-à-dire que le corps vérifie la loi de Lambert. Dans ce cas, le calcul de la puissance émise dans la bande de pulsations $(\omega, \omega + \mathrm{d}\omega)$ est analogue à celui du paragraphe 3, et donne

$$\begin{aligned} \Pi(\omega; T) &= \mathcal{A}(\omega; M, T) \frac{c}{4} u_P(\omega; T) \\ &= \mathcal{A}(\omega; M, T) \frac{\hbar}{4\pi^2 c^2} \frac{\omega^3}{\mathrm{e}^{\hbar\omega/kT} - 1} . \end{aligned} \tag{F.17}$$

Si en outre le pouvoir absorbant est pratiquement *indépendant de la pulsation* dans le domaine où u_p est non négligeable, on trouve à nouveau la loi de Stéfan-Boltzmann :

$$\int_0^{+\infty} \mathrm{d}\omega \, \Pi(\omega; T) = \mathcal{A}\sigma T^4 . \tag{F.18}$$

Les formules de ce paragraphe supposent seulement que le corps $\mathcal{C}$ a une température T bien définie au voisinage du point M dont on étudie l'émission. Il n'est par contre *pas nécessaire qu'il y ait équilibre entre le corps* $\mathcal{C}$ *et le rayonnement :* ces formules s'appliquent à l'émission thermique d'un rayonnement hors d'équilibre,

et couvrent donc un domaine d'application beaucoup plus vaste que ne le laisserait penser la méthode utilisée pour les démontrer (§ 1)[12]. Notons que, parmi tous les corps portés à une même température T, *c'est le corps noir qui émet le plus d'énergie,* pour chaque pulsation et dans chaque direction (en effet, le pouvoir absorbant d'un corps quelconque est par définition inférieur ou au plus égal à 1).

12. Une hypothèse implicite sous-tend cette démonstration : l'équilibre est possible entre un corps quelconque et un gaz de photons, l'ensemble étant en contact avec un thermostat de température fixe T. On ne pourrait évidemment pas écrire la relation (F.7) si cette hypothèse n'était pas vérifiée. Mais l'expérience nous apprend qu'un système macroscopique possède toujours un état d'équilibre.

COMPLÉMENT VI.G

Le rayonnement cosmique à trois kelvins

I. Introduction : notions succinctes sur le rayonnement à trois kelvins et l'expansion de l'Univers
   1. Observation du rayonnement cosmique
   2. L'expansion de l'Univers

II. Origine du rayonnement fossile à trois kelvins
   1. Découplage du rayonnement et de la matière
   2. Évolution du gaz de photons découplé

I. INTRODUCTION : NOTIONS SUCCINCTES SUR LE RAYONNEMENT A TROIS KELVINS ET L'EXPANSION DE L'UNIVERS

1. OBSERVATION DU RAYONNEMENT COSMIQUE

Le rayonnement cosmique fut découvert en 1964 par deux radio-astronomes, A. Penzias et R. Wilson. Ils disposaient d'une antenne radio à très faible bruit de fond et voulaient l'utiliser pour détecter le rayonnement émis par notre galaxie dans le domaine des longueurs d'onde centimétriques, principalement hors du plan de la galaxie. Les caractéristiques du récepteur, dont le bruit de fond intrinsèque avait été soigneusement étudié et réduit, devaient permettre une mesure précise des signaux recherchés. Pour calibrer leur instrument et s'assurer que ses performances atteignaient bien le niveau prévu, Penzias et Wilson le réglèrent d'abord sur une longueur d'onde de 7,35 cm, où l'émission galactique devait être négligeable. A leur grande surprise, ils reçurent un signal beaucoup plus fort qu'ils ne l'attendaient. Ce signal se révéla indépendant de la direction d'observation, et constant dans le temps, sans variations journalières ni saisonnières. Son intensité était égale à celle qu'émettrait sur cette longueur d'onde un corps noir maintenu à une température d'environ 3 K[1].

Avant d'admettre et de publier un résultat aussi étonnant, Penzias et Wilson reprirent minutieusement l'analyse des diverses sources de bruit de fond dans l'appareillage, afin de réexaminer leur évaluation initiale. Au cours de cette période de vérifications, ils apprirent que des astrophysiciens théoriciens avaient prédit l'existence d'un rayonnement cosmique fossile, vestige d'une étape passée de l'évolution de l'Univers, qui devait présenter la distribution spectrale d'un corps noir de quelques kelvins de température. Ayant pris contact avec ces théoriciens, Penzias et Wilson se décidèrent finalement à publier, en 1965, un article décrivant leurs observations; cet article était associé à un autre, signé de R. Dicke, P. Peebles, P. Roll et D. Wilkinson, qui donnait dans la même revue l'interprétation cosmologique de ces observations[2].

---

1. Si l'on se fixe une longueur d'onde λ (c'est-à-dire une pulsation $\omega = 2\pi c/\lambda$), à chaque intensité de rayonnement on peut faire correspondre une température T : celle du corps noir qui émettrait cette même intensité sur la longueur d'onde choisie. Ceci ne signifie évidemment pas que le rayonnement reçu provient effectivement d'un corps noir; pour le savoir, il faut faire varier la longueur d'onde d'observation et chercher si l'intensité varie ou non suivant la loi de Planck, c'est-à-dire si la température associée aux diverses longueurs d'onde reste ou non la même.

2. L'idée que l'Univers doit être baigné par un rayonnement de corps noir de quelques kelvins provenant de son évolution a été émise pour la première fois en 1948 par R. Alpher et R. Herman, exploitant une théorie de G. Gamow. Plusieurs autres physiciens ont par la suite exprimé à nouveau cette idée, de façon apparemment indépendante.

Restait à rechercher, en mesurant son intensité à différentes longueurs d'onde, si le rayonnement ainsi découvert présente bien la distribution caractéristique du corps noir. Ce n'est pas tâche aisée. En effet, pour une température de 3 K, la longueur d'onde correspondant au maximum d'émission du corps noir est de 0,1 cm, c'est-à-dire beaucoup plus courte que celles qui étaient accessibles à l'instrument de Penzias et Wilson. Il est évidemment crucial, pour comparer avec le corps noir, d'établir si l'intensité du rayonnement étudié passe bien par un maximum aux alentours de 0,1 cm, pour décroître ensuite lorsque la longueur d'onde continue à décroître (c'est-à-dire lorsque la pulsation croît), comme le prévoit la loi de Planck. Or les mesures s'avèrent très difficiles, dans ce domaine des ondes submillimétriques puis infrarouges, notamment parce que l'atmosphère terrestre y est pratiquement opaque. Il faut donc utiliser des ballons sondes, des fusées ou des satellites artificiels. On peut cependant faire également appel à des méthodes indirectes : il existe, dans l'espace interstellaire, certaines molécules ou radicaux tels que le cyanogène CN; l'étude des raies d'absorption par lesquelles ils se manifestent permet dans certains cas d'obtenir des renseignements sur leur environnement, et donc en particulier sur le rayonnement cosmique qui les entoure.

Il paraît établi à l'heure actuelle que l'Univers est rempli d'un rayonnement distribué selon la loi de Planck du corps noir[3]; sa température est voisine de 3 K, et il paraît réparti dans les diverses directions de façon isotrope[4].

2. L'EXPANSION DE L'UNIVERS

Sans entrer dans les détails (voir réf. 1, 3, 4), rappelons que la lumière qui nous provient des galaxies lointaines est *décalée vers le rouge,* ce qui indique, selon les lois de l'effet Doppler, que ces galaxies s'éloignent de la nôtre. De façon plus précise, E. Hubble énonça en 1929 la loi empirique qui porte son nom, et selon laquelle *la vitesse de récession des galaxies est proportionnelle à leur distance.*

Notre galaxie ne joue évidemment aucun rôle privilégié dans l'Univers. La loi de Hubble signifie par conséquent que deux objets quelconques[5] de l'Univers

3. La température observée dans le domaine des radiofréquences, c'est-à-dire sur la partie décroissante de la courbe de Planck vers les hautes fréquences, est actuellement de $2{,}72 \pm 0{,}06$ K.

4. En réalité, on s'attend à ce que la Terre soit en mouvement par rapport au référentiel du rayonnement cosmique, ce mouvement étant dû principalement à celui de notre galaxie. Si c'est le cas, la température T observée pour le rayonnement cosmique doit varier avec l'angle θ d'observation selon la loi caractéristique de l'effet Doppler :

$$T = T_0 \left[1 + \frac{v}{c} \cos \theta \right], \tag{1}$$

v étant la vitesse de la Terre. On a pu effectivement mettre en évidence une variation de cette forme, de l'ordre de un pour mille en valeur relative, ce qui donne

$$v \sim 300 \text{ km/s} . \tag{2}$$

Si l'on retranche cette anisotropie attendue (ce qu'autorise sa forme très spécifique en $\cos \theta$), le rayonnement cosmologique est parfaitement isotrope à la précision des mesures, c'est-à-dire

$$\frac{\delta T}{T} < 10^{-4} . \tag{3}$$

5. Ces objets ne doivent pas être tout à fait quelconques : au mouvement d'expansion universel qu'il s'agit de caractériser se superposent d'autres mouvements relatifs (rotation des galaxies sur elles-mêmes, mouvements à l'intérieur de chaque amas de galaxies, ...) qu'il faut éliminer ici.

s'éloignent l'un de l'autre avec une vitesse relative proportionnelle à la distance qui les sépare. Il en résulte que, durant un intervalle de temps donné, *toutes les distances de l'Univers s'accroissent dans les mêmes proportions*[6],[7]. Il suffit donc, pour caractériser l'expansion de l'Univers, de connaître l'évolution, en fonction du temps t, de la distance $R(t)$ entre deux objets typiques. Cette distance caractéristique $R(t)$ est souvent appelée abusivement la « taille de l'Univers ».

Il est important pour la suite de noter que *les longueurs d'onde du rayonnement électromagnétique sont affectées de la même façon que les distances intergalactiques.* Dans le cadre relativiste, ceci est une conséquence directe du formalisme. Pour s'en convaincre sans avoir recours à ce formalisme, il suffit de calculer le décalage vers le rouge du rayonnement émis par une galaxie lorsqu'il est reçu en un point d'une autre galaxie[8].

---

6. La vitesse relative v de deux objets séparés par la distance R est donnée par

$$v = HR\,, \tag{1}$$

où H est la *« constante de Hubble »* (bien qu'évoluant au cours de l'histoire de l'Univers, cette « constante » est indépendante des objets choisis, c'est-à-dire de R). On a donc

$$\frac{\mathrm{d}R}{\mathrm{d}t} = HR\,, \tag{2}$$

c'est-à-dire

$$\frac{\mathrm{d}R}{R} = H\,\mathrm{d}t\ : \tag{3}$$

l'accroissement relatif de la distance de deux objets est indépendant du choix de ces objets.

7. Pour les vitesses de récession des galaxies les plus lointaines connues à l'heure actuelle, l'approximation non relativiste n'est plus valable. De façon générale, c'est dans le cadre de la relativité que doivent être exprimés tous les raisonnements cosmologiques. C'est par souci de simplicité que nous continuons à utiliser le langage non relativiste.

8. Raisonnons encore, pour simplifier, dans le cadre non relativiste.

Nous recevons en O des signaux lumineux émis par un point M dont la distance $R(t)$ à O croît avec le temps. Un signal émis à l'instant t par M est reçu en O à l'instant t' tel que

$$c\,.\,(t'-t) = R(t)\,. \tag{1}$$

S'il s'agit d'une onde sinusoïdale de période τ à l'émission, sa période τ' à la réception en O peut être calculée en écrivant la relation (1) pour l'instant $t+\tau$:

$$c\,.\,(t'+\tau'-t-\tau) = R(t+\tau)\,. \tag{2}$$

Il suffit ensuite de combiner les égalités (1) et (2) pour trouver aussitôt

$$c\,.\,(\tau'-\tau) = \tau\frac{\mathrm{d}R}{\mathrm{d}t}\,. \tag{3}$$

C'est la formule de l'effet Doppler non relativiste, que l'on peut transcrire sous la forme

$$\lambda' = \lambda\left(1+\frac{v}{c}\right) \quad \text{pour} \quad \frac{v}{c} \ll 1\,, \tag{4}$$

où $\lambda' = c\tau'$ et $\lambda = c\tau$ sont les longueurs d'onde à la réception et à l'émission, respectivement, et $v = \mathrm{d}R/\mathrm{d}t$ la vitesse de récession du point M.

Pendant le même intervalle de temps $t'-t$, la distance R entre les points O et M devient, au premier ordre en v/c,

$$R(t') = R(t) + v\,.\,(t'-t)\ ; \tag{5}$$

.../...

II. ORIGINE DU RAYONNEMENT FOSSILE A TROIS KELVINS

L'existence et les propriétés du rayonnement cosmique découvert par Penzias et Wilson s'expliquent essentiellement par les deux phénomènes physiques que nous allons maintenant décrire dans leurs grandes lignes.

1. DÉCOUPLAGE DU RAYONNEMENT ET DE LA MATIÈRE

L'expansion de l'Univers a pour conséquence son refroidissement graduel. A partir des valeurs fantastiquement élevées qui ont dû régner aussitôt après la « grande explosion » (« big bang » en anglais) qui a donné naissance à l'Univers, sa température a progressivement décrû. Lorsqu'elle atteint environ 3 000 K se produit le premier des deux phénomènes cruciaux qui nous intéressent ici : le rayonnement, qui jusque-là était en équilibre thermique avec les particules matérielles, cesse pratiquement d'interagir avec elles et en devient indépendant. Dans le « modèle standard » d'évolution de l'Univers, on calcule que ce moment crucial se situe $0{,}68 \times 10^6$ ans après l'explosion initiale.

On peut comprendre qualitativement les raisons physiques de ce phénomène. Un peu avant, lorsque par exemple la température était de 100 000 K, l'Univers contenait essentiellement des photons, des électrons et des noyaux atomiques « nus » (principalement des protons, et, dans une moindre proportion, des particules α, noyaux d'hélium 4). La température était trop élevée ($kT \sim 10$ eV) pour que les électrons et les noyaux puissent former des atomes, autrement que de manière transitoire et labile. L'interaction entre les photons et les particules chargées (surtout les électrons, les plus légères d'entre elles) est suffisamment intense, et la densité de ces dernières était alors suffisamment forte, pour que les photons soient sans arrêt diffusés, émis et absorbés. Malgré son expansion, l'Univers était alors à chaque instant en équilibre; sa température T était constamment bien définie, bien que décroissante au cours du temps. L'énergie des photons, c'est-à-dire la pulsation du rayonnement, était donc distribuée suivant la loi de Planck correspondant à cette température T.

La diminution de la température a ensuite permis la *formation d'atomes* à partir des électrons et des noyaux. Ce processus a entraîné une *chute rapide de la section efficace moyenne d'interaction entre les photons et les particules matérielles* (principalement à cause de la disparition des électrons libres), de sorte que *l'Univers est devenu transparent aux photons.* Une évaluation quantitative des caractéristiques du phénomène[9] situe ce découplage au moment où la température est descendue à 3 000 K.

---

en effet, d'après l'égalité (1),

$$v.(t'-t) \ll c.(t'-t) = R(t) . \tag{6}$$

On a donc ici aussi

$$R(t') = R(t)\left[1+\frac{v}{c}\right] \quad \text{pour} \quad \frac{v}{c} \ll 1 . \tag{7}$$

Les distances R et les longueurs d'onde λ sont donc multipliées par le même facteur pendant le même intervalle de temps.

9. Le libre parcours moyen ℓ d'un photon est de l'ordre de

$$\ell \sim \frac{1}{\rho\sigma} ,$$

où ρ est la densité de particules matérielles et σ la section efficace moyenne d'interaction (en

.../...

2. ÉVOLUTION DU GAZ DE PHOTONS DÉCOUPLÉ

Au moment du découplage, la densité volumique d'énergie du rayonnement est distribuée dans le spectre des pulsations selon la loi de Planck :

$$u(\omega)=u_P(\omega;\ T)=\frac{\hbar}{\pi^2c^3}\frac{\omega^3}{e^{\hbar\omega/kT}-1}\ , \tag{G.1}$$

où T est la température (3000 K environ) à ce moment-là. Cette distribution va ensuite évoluer sous l'influence de l'expansion de l'Univers.

Considérons les photons situés, à cet instant t, dans le volume $V=R^3$, et dont la pulsation est ω à $d\omega$ près; leur nombre est égal à

$$dN=R^3\times\frac{1}{\hbar\omega}\,u_P(\omega;\ T)\,d\omega=\frac{R^3}{\pi^2c^3}\frac{\omega^2\,d\omega}{e^{\hbar\omega/kT}-1}\ . \tag{G.2}$$

Comme il n'y a plus d'absorption ni d'émission de photons, *ce nombre va rester constant;* mais, à cause de l'expansion de l'Univers, ces photons en nombre constant vont occuper un volume plus grand, et acquérir une longueur d'onde plus grande, c'est-à-dire une pulsation plus petite. Pour préciser, examinons la situation à un instant t' ultérieur. Toutes les longueurs de l'Univers ont été multipliées, entre t et t', par le même facteur f : l'arête R du volume cubique choisi est ainsi devenue

$$R'=fR\ , \tag{G.3}$$

et la longueur d'onde des photons considérés

$$\lambda'=f\lambda\ , \tag{G.4}$$

de sorte que leur pulsation vaut à l'instant t'

$$\omega'=\frac{2\pi c}{\lambda'}=\frac{\omega}{f}\ . \tag{G.5}$$

Donc, l'énergie contenue à cet instant dans le volume $V'=R'^3$ et dans la bande de pulsations $(\omega',\ \omega'+d\omega')$, que nous écrirons $R'^3u'(\omega')d\omega'$, est donnée par[10]

$$R'^3u'(\omega')d\omega'=\hbar\omega'\ dN=\hbar\omega'\frac{R^3}{\pi^2c^3}\frac{\omega^2\,d\omega}{e^{\hbar\omega/kT}-1}\ . \tag{G.6}$$

La densité volumique d'énergie $u'(\omega')d\omega'$ à l'instant t', pour la bande de pulsations $(\omega',\ \omega'+d\omega')$, s'écrit donc

$$u'(\omega')d\omega'=\hbar\omega'\frac{1}{f^3}\frac{1}{\pi^2c^3}\frac{f^2\omega'^2\times f\,d\omega'}{e^{\hbar f\omega'/kT}-1}\ . \tag{G.7}$$

---

réalité, $\rho\sigma$ est une somme de termes, un pour chaque type de particules matérielles). La densité ρ décroît régulièrement au cours du temps par suite de l'expansion de l'Univers; la section efficace σ chute brutalement pendant la période de formation des atomes. Le découplage entre les photons et la matière se produit au moment où ℓ devient de l'ordre de ct, où t est l'âge de l'Univers : les photons existant à cet instant t n'ont pratiquement plus aucune chance d'interagir avec la matière. Ce découplage est irréversible, ℓ ne repassant plus jamais au-dessous de la limite cruciale ct. On peut dire de façon plus imagée que, après le découplage, un photon peut traverser tout l'Univers sans interagir.

10. Les photons contenus à l'instant t' dans le volume V' ne sont pas vraiment ceux qui se trouvaient dans le volume V à l'instant t : ceux-ci s'éloignaient à la vitesse c dans toutes les directions. Mais la distribution initiale est homogène et isotrope, et l'expansion préserve ces propriétés, de sorte que le nombre dN de photons dans le volume R'^3 à l'instant t' est bien le même qu'à l'instant t dans le volume R^3.

Il s'ensuit que *la distribution spectrale de l'énergie est encore à l'instant t' celle du corps noir :*

$$u'(\omega')=\frac{\hbar}{\pi^2 c^3}\frac{\omega'^3}{e^{\hbar f\omega'/kT}-1}=u_P(\omega';\,T')\,, \tag{G.8}$$

la température correspondante T' étant telle que

$$T'=\frac{T}{f}\,. \tag{G.9}$$

Ainsi, après son découplage d'avec la matière, le rayonnement cosmique évolue en conservant la distribution d'un corps noir dont la température décroît régulièrement, dans la même proportion que s'accroissent les distances au cours de l'expansion de l'Univers (depuis le moment du découplage, le facteur f d'expansion est très voisin de 1 000).

C'est en se fondant sur ce raisonnement que divers auteurs furent amenés à prédire l'existence, dans l'Univers actuel, d'un rayonnement fossile de quelques kelvins. La découverte de Penzias et Wilson, qui confirme parfaitement cette prédiction, est généralement considérée comme l'argument le plus solide en faveur du « modèle cosmologique standard », qui reconstitue l'histoire de l'Univers à partir de la « grande explosion » initiale.

Références

1. S. Weinberg, *Les Trois Premières Minutes de l'Univers,* Seuil, Paris (1978).
2. R.A. Muller, « Le Rayonnement Thermique Cosmologique et le Nouveau Vent d'Éther », *Pour la Science 9,* p. 81 (juillet 1978).
3. P.J.E. Peebles, *Physical Cosmology,* Princeton University Press, Princeton (1971).
4. S. Weinberg, *Gravitation and Cosmology,* John Wiley, New York (1972).

Exercices VI

GAZ PARFAITS DE FERMIONS OU DE BOSONS

VI.1. On considère un gaz parfait constitué de N particules en équilibre à la température T dans un récipient de volume V; soit p la pression du gaz. Montrer que $pV > NkT$ si les particules sont des fermions et que $pV < NkT$ si ce sont des bosons.

VI.2. Gaz parfait de fermions à deux dimensions. On étudie un gaz constitué de N fermions indépendants ($N \gg 1$) de masse m et de spin 1/2, libres mais astreints à se déplacer sur une surface d'aire s.

1) Écrire la relation qui détermine le potentiel chimique μ en fonction de N, s et de la température T.

2) On suppose d'abord que l'approximation de *température nulle* est valable.

a) Calculer le potentiel chimique μ_0 et la température de Fermi $T_F \widehat{=} \mu_0/k$.

b) Quelle est la vitesse maximale d'une particule du gaz? Quelle condition doit satisfaire la densité superficielle de particules N/s pour que les formules non relativistes utilisées ici soient valables? Préciser numériquement cette condition pour des électrons, en estimant leur nombre par Å^2 ; quel est l'ordre de grandeur correspondant de la température de Fermi ?

c) Calculer l'énergie E_0 du gaz. Soit T_{cl} la température d'un gaz parfait classique de N atomes à 2 dimensions qui aurait cette même énergie E_0; calculer le rapport T_{cl}/T_F.

3) On se place maintenant à *température non nulle.*

a) En utilisant la relation entre le facteur de Fermi $\overline{N}_\lambda^F$ et la dérivée $\partial (\ln \xi_\lambda^F)/\partial \varepsilon_\lambda$, où ξ_λ^F est la grande fonction de partition de l'état individuel (λ), calculer le potentiel chimique μ en fonction de μ_0 et de T. Tracer la courbe représentant μ en fonction de T, à N/s fixé. Préciser la façon dont μ tend vers μ_0 lorsque T tend vers zéro.

b) Démontrer les relations

$$J = -E ,$$

où J et E sont le grand potentiel et l'énergie du gaz, et

$$C_s = \frac{2E}{T} - Nk \frac{T_F}{T} \frac{1}{1 - e^{-T_F/T}} ,$$

où C_s est la capacité calorifique à surface s et nombre de particules N constants.

c) En utilisant la formule VI (I.45) et le résultat de la question *a* ci-dessus, calculer l'énergie E à basse température, jusqu'au 2^e ordre en T/T_F. En déduire le comportement de la capacité calorifique C_s dans le domaine $T \ll T_F$.

d) Que donne le développement de basse température du chapitre VI (§ I.B.2.a) si on l'applique au calcul du potentiel chimique μ? Peut-on comprendre cette anomalie à partir de l'expression exacte de μ (question *a*)?

VI.3. Gaz parfait de bosons à deux dimensions. On étudie un gaz constitué de N bosons indépendants de masse m et de spin nul ($N \gg 1$), libres mais astreints à se déplacer sur une surface d'aire s.

1) Écrire la relation qui détermine le potentiel chimique μ en fonction de N, s et de la température T, dans l'hypothèse où le gaz est « normal ».

a) Montrer que l'intégrale qui figure dans cette relation diverge lorsque μ tend vers zéro. En déduire qu'il existe, pour toute valeur de N/s et de T, une valeur de μ négative vérifiant la relation. En conclure que *le phénomène de condensation de Bose ne se produit pas à deux dimensions.*

b) En utilisant la relation entre le facteur de Bose $\overline{N}_\lambda^B$ et la dérivée $\partial(\ln \xi_\lambda^B)/\partial\varepsilon_\lambda$, où ξ_λ^B est la grande fonction de partition de l'état individuel (λ), calculer le potentiel chimique μ en fonction de T et N/s. Tracer la courbe représentant la fugacité $e^{\beta\mu}$ en fonction de T, et comparer avec la courbe correspondante pour un gaz de bosons libres à trois dimensions.

c) On pose

$$T_0 \widehat{=} \frac{2\pi\hbar^2}{mk}\frac{N}{s}\,.$$

Montrer, en comparant la longueur d'onde thermique

$$\Lambda \widehat{=} \left(\frac{2\pi\hbar^2}{mkT}\right)^{1/2}$$

à la distance moyenne entre particules voisines, que des effets quantiques sont attendus pour $T \lesssim T_0$.

d) Évaluer numériquement T_0 dans le cas de l'hélium 4 (on utilisera les données du complément VI.A, note 28, et on estimera s/N en prenant la puissance 2/3 du volume par atome dans l'hélium 4 liquide). Calculer le nombre d'occupation moyen de l'état fondamental individuel pour $T = 1K$; est-il d'ordre macroscopique comme à trois dimensions?

2) Démontrer les relations

$$J = -E\,,$$

où J et E sont le grand potentiel et l'énergie du gaz, et

$$C_s = \frac{2E}{T} - Nk\frac{T_0}{T}\frac{1}{e^{T_0/T}-1}\,,$$

où C_s est la capacité calorifique à surface s et nombre de particules N constants.

3) Donner une expression approchée pour l'énergie E et pour la capacité calorifique C_s dans chacun des deux domaines $T \gg T_0$ et $T \ll T_0$.

VI.4. Gaz parfait dégénéré de fermions relativistes (solution : complément VI.B, § III.3).

On considère un gaz constitué de N fermions indépendants ($N \gg 1$) de masse m et de spin 1/2, libres et relativistes : lorsque l'un d'eux a l'impulsion $\vec{p}$, son énergie s'écrit

$$\varepsilon(\vec{p}) = \sqrt{\vec{p}^{\,2}c^2 + m^2c^4}\,.$$

On note V le volume du récipient contenant le gaz; la température T est supposée suffisamment faible pour que l'on se contente de l'*approximation de température nulle.*

1) On pourrait utiliser la densité d'états individuels trouvée à l'exercice I.23, mais il est plus commode de prendre pour variable le module p de l'impulsion. Montrer que la somme sur les états individuels s'explicite alors sous la forme

$$\sum_{(\lambda)} \rightsquigarrow 2\frac{V}{(2\pi\hbar)^3}\int_0^{+\infty} 4\pi p^2\,dp\,.$$

2) On désigne par μ_0 le potentiel chimique du gaz à température nulle et par p_F l'impulsion d'une particule d'énergie μ_0. Calculer p_F en fonction de la densité N/V.

3) Soit J_0 et E_0 le grand potentiel et l'énergie du gaz à température nulle.

a) Montrer que J_0 peut s'écrire

$$J_0 = \frac{2V}{(2\pi\hbar)^3} \int_0^{p_F} 4\pi p^2 \, dp \left[\varepsilon(p) - \mu_0 \right] .$$

b) Vérifier que cette expression conduit bien à la relation générale (cf. formule V (II.23))

$$J_0 = E_0 - N\mu_0 .$$

Pourquoi J_0 n'est-il pas simplement proportionnel à E_0 comme dans le cas non relativiste (cf. formule VI (I.20))?

c) En posant

$$x \widehat{=} \frac{p}{mc}$$

et en effectuant une intégration par parties, établir l'expression suivante pour le grand potentiel :

$$J_0 = -V \frac{m^4 c^5}{3\pi^2 \hbar^3} I(x_F)$$

avec $I(x_F) \widehat{=} \int_0^{x_F} \frac{x^4 \, dx}{\sqrt{1+x^2}}$, $x_F \widehat{=} \frac{p_F}{mc}$.

d) Écrire l'expression qui en découle pour la pression P_0 du gaz à température nulle. Montrer qu'on a toujours

$$I(x_F) < \frac{x_F^4}{4} .$$

En déduire une borne supérieure pour la pression P_0 d'un gaz de densité N/V donnée.

4) Examiner le comportement des résultats précédents dans les domaines non relativiste $p \ll mc$ (cf. VI, § I.B.1) et ultra-relativiste $p \gg mc$.

VI.5. On considère un gaz parfait monoatomique de particules non relativistes.

1) Qu'il s'agisse de particules « classiques », de fermions ou de bosons, montrer que le grand potentiel J est de la forme

$$J = VT^{5/2} f\left(\frac{\mu}{T}\right) .$$

En déduire des expressions analogues pour l'entropie et le nombre (moyen) de particules :

$$S = VT^{3/2} f_1\left(\frac{\mu}{T}\right) ,$$

$$N = VT^{3/2} f_2\left(\frac{\mu}{T}\right) .$$

2) Montrer que, dans une transformation fermée ($N = \text{c}^{\text{te}}$) et isentropique ($S = \text{c}^{\text{te}}$) d'un tel système, le rapport μ/T reste constant. En déduire que le produit $VT^{3/2}$ reste également constant, ainsi que $pT^{-5/2}$ et $pV^{5/3}$.

3) Écrire la relation liant p et V dans une transformation fermée isentropique lorsque les particules constituant le gaz sont ultra-relativistes.

VI.6. On veut montrer que les isothermes de la figure 11.a du chapitre VI ont une pente continue (nulle) à l'amorce du palier de transition ($V = V_B$).

Écrire, dans le domaine où le gaz de bosons est « normal », les relations donnant la pression p en fonction de T et μ et le potentiel chimique μ en fonction de T, N, V. En déduire l'expression de $(\partial p/\partial \mu)_T$ et $(\partial V/\partial \mu)_{T,N}$ et analyser le comportement de ces dérivées lorsque $V \to V_B$ (c'est-à-dire $\mu \to 0_-$). Le résultat cherché en découle.

VI.7. Chaleur latente et formule de Clapeyron pour la condensation de Bose. On considère un gaz parfait de bosons libres dans le domaine où se produit le phénomène de condensation de Bose (VI, § II.B et § II.C.2). Le nombre N de particules est fixé. *A température T et pression p constantes,* on diminue de δN_0 le nombre de particules condensées.

1) Calculer l'accroissement de volume δV qui s'ensuit, puis ceux de l'énergie et de l'entropie du système, δE et δS. Vérifier que ces accroissements satisfont bien à la relation (complément III.L, formule (L.12))

$$\delta E = T\delta S - p\delta V . \tag{1}$$

2) La chaleur latente (molaire) de « décondensation » peut être définie par

$$L \mathrel{\widehat{=}} N_A T \frac{\delta S}{\delta N_0} , \tag{2}$$

où N_A est le nombre d'Avogadro. Montrer qu'elle vérifie la relation

$$L = Tv \frac{\mathrm{d}p}{\mathrm{d}T} , \tag{3}$$

où $$v \mathrel{\widehat{=}} N_A \frac{V}{N - N_0} \tag{4}$$

est le volume molaire de la phase « normale » et $\mathrm{d}p/\mathrm{d}T$ la pente de la courbe tracée sur la figure 12.a du chapitre VI. Comparer l'égalité (3) avec la relation de Clapeyron pour la transition liquide-vapeur (cf. exercice III.19).

VI.8. Dans un récipient de volume V est contenu un gaz parfait de N bosons libres, possédant des degrés de liberté internes. Pour simplifier, on admet que chaque particule a un seul état d'excitation, d'énergie ε_1 positive (l'énergie de l'état fondamental étant choisie égale à zéro).

1) Donner l'expression, en fonction de la température et du potentiel chimique, du nombre de particules qui se trouvent dans l'état interne fondamental et du nombre de celles qui se trouvent dans l'état excité. Écrire la condition déterminant la température critique T_B de la condensation de Bose.

2) En supposant $\varepsilon_1 \gg kT_B$, montrer que T_B est donnée par la formule

$$T_B = T_B^0[1 - 0{,}255\ \mathrm{e}^{-\varepsilon_1/kT_B^0}],$$

où T_B^0 est la température critique calculée sans tenir compte des degrés de liberté internes.

$$\left(\text{On rappelle que } \int_0^\infty \frac{\sqrt{x}}{\mathrm{e}^x - 1}\,\mathrm{d}x = 2{,}612\,\frac{\sqrt{\pi}}{2}.\right)$$

3) Application à l'hélium 4 : est-il nécessaire de tenir compte de la structure interne des atomes d'hélium lorsqu'on étudie la transition superfluide?

VI.9. N particules de spin 1/2 et de masse m, contenues dans un récipient de volume V à la température T, peuvent se lier deux à deux pour former de nouvelles particules : des « paires » de spin nul. L'énergie de liaison d'une paire est $-\varepsilon_0 (\varepsilon_0 > 0)$. On néglige

toute structure interne (autre que le spin) et on suppose que les particules et les paires forment deux gaz parfaits sans interactions.

1) Quelle relation existe-t-il, à l'équilibre, entre le potentiel chimique des particules isolées et celui des paires?

2) Montrer que les paires sont susceptibles de se condenser dans l'état fondamental (condensation de Bose) et calculer la température de condensation T_c (on supposera $kT_c \ll \varepsilon_0$).

PHOTONS

VI.10. Que deviendraient la loi de Planck pour la densité spectrale d'énergie et la loi de Stefan-Boltzmann pour la densité totale si le gaz de photons était confiné dans un espace à deux dimensions?

VI.11. On considère un gaz de photons en équilibre avec un thermostat de température T. On note (ℓ) un état microscopique du gaz, d'énergie E_ℓ et nombre de photons N_ℓ et (λ) un état individuel possible pour un photon, l'énergie correspondante étant ε_λ. L'application du postulat fondamental de la mécanique statistique donne la probabilité pour que le gaz soit dans l'état (ℓ) sous la forme (VI, § III.3.b)

$$P_\ell^P = \frac{1}{A} e^{-\beta E_\ell} , \tag{1}$$

où A est une constante de normalisation. On veut retrouver directement à partir de cette formule la distribution des photons sur les états individuels (λ) et la loi de Planck.

1) En caractérisant un état (ℓ) du gaz par la configuration $\{N_\lambda\}$ correspondante (appendice II, § IV.3), montrer que la «constante de normalisation» A peut être factorisée :

$$A = \prod_{(\lambda)} a_\lambda , \tag{2}$$

et calculer a_λ en fonction de ε_λ et T.

2) Montrer que le nombre d'occupation moyen $\overline{N}_\lambda^P$ de l'état individuel (λ) est donné par

$$\overline{N}_\lambda^P = -\frac{1}{\beta} \frac{\partial}{\partial \varepsilon_\lambda} \ln A = -\frac{1}{\beta} \frac{\partial}{\partial \varepsilon_\lambda} \ln a_\lambda , \tag{3}$$

et l'énergie moyenne $\overline{E}$ du gaz par

$$\overline{E} = -\frac{\partial}{\partial \beta} \ln A = -\sum_{(\lambda)} \frac{\partial}{\partial \beta} \ln a_\lambda . \tag{4}$$

Retrouver ainsi la formule VI (III.18) et la loi de Planck.

VI.12. Une molécule A peut absorber un photon de pulsation ω pour se trouver dans un état d'excitation interne A^*. Inversement, la molécule A^* peut se désexciter en émettant un photon d'énergie $\hbar\omega$. On a donc la réaction

$$A^* \rightleftarrows A + \text{photon } \omega .$$

1) Si l'on considère A et A^* comme des molécules de type différent, quelle relation existe-t-il à l'équilibre entre leurs potentiels chimiques μ_A et μ_{A^*}?

2) On enferme N_0 molécules A dans une enceinte à basse température T_0 ($\ll \hbar\omega/k$). On porte ensuite l'enceinte à la température $T \gg T_0$. En supposant que les

deux espèces se comportent comme des gaz parfaits, déduire de la question 1 les nombres $N_A(T)$ et $N_{A^*}(T)$ de molécules des deux types à l'équilibre.

VI.13. La température de surface du Soleil est $T_0 = 5\,500$ K. La Terre reçoit son énergie du Soleil. En admettant que la Terre et le Soleil se comportent comme des corps noirs, estimer la température de la Terre.

On connaît le rayon du Soleil : $R_S = 7.10^8$ m, le rayon de la Terre : $R_T = 6{,}4.10^6$ m et la distance Terre-Soleil : $d = 1{,}5.10^{11}$ m.

ÉLECTRONS DANS LES MÉTAUX

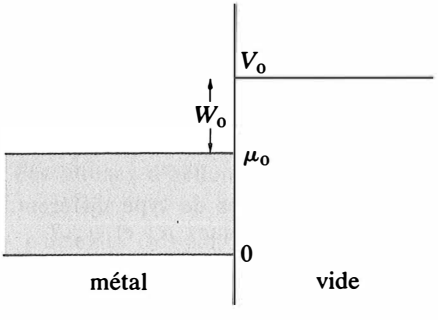
VI.14. Un échantillon de métal, de volume V, est constitué d'un gaz de N électrons et d'ions pratiquement immobiles, chargés positivement de façon à compenser la charge négative du gaz électronique. Si Ze est la charge d'un ion et a la distance moyenne entre un électron et l'ion le plus proche, on peut estimer que l'énergie potentielle d'un électron dans le champ électrique des ions est de l'ordre de $-Ze^2/4\pi\varepsilon_0 a$. L'énergie d'interaction coulombienne des électrons entre eux, ramenée à un électron, est, en module, du même ordre de grandeur.

1) Estimer l'énergie potentielle du gaz d'électrons et la comparer à son énergie cinétique. En déduire que l'approximation de gaz parfait est d'autant meilleure que la densité électronique N/V est plus élevée.

2) Peut-on dire que l'énergie potentielle du gaz électronique est négligeable devant son énergie cinétique dans les métaux usuels (on utilisera les données du tableau I du complément VI.A)?

VI.15. L'énergie de Fermi (potentiel chimique des électrons à température nulle) de l'aluminium est $\mu_0 = 11{,}7$ eV. Montrer que, à 900 K, $|\mu - \mu_0|/\mu_0 < 10^{-4}$. Pour quelle température aurait-on $|\mu - \mu_0|/\mu_0 \simeq 10^{-3}$? (on rappelle que la température de fusion de l'aluminium est 933 K et sa température d'ébullition 2 740 K).

VI.16. Effet thermoionique. Dans un métal, les électrons de conduction sont quasiment libres et indépendants. Si on choisit comme zéro l'énergie minimale d'un électron de la bande de conduction, l'énergie d'un électron de vitesse nulle dans le vide est égale à V_0 (voir figure).



Soit μ_0 le niveau de Fermi des électrons du métal. Le gaz d'électrons étant toujours très fortement dégénéré, son potentiel chimique à la température ordinaire est pratiquement égal à μ_0.

1) Calculer le « travail d'extraction » W_0 du métal, c'est-à-dire l'énergie minimum qu'il faut fournir, à température nulle, pour extraire un électron du métal (W_0 est de l'ordre de l'électron-volt : 2,35 eV pour le sodium, 4,25 eV pour l'aluminium).

2) Montrer que, dans le métal, le nombre d'électrons, par unité de volume, dont la quantité de mouvement vaut $\vec{p}$ à d^3p près s'écrit

$$dn(\vec{p}) = \frac{2}{(2\pi\hbar)^3} \frac{d^3p}{e^{\beta(\vec{p}^2/2m - \mu_0)} + 1} .$$

3) On suppose la surface du métal plane et perpendiculaire à l'axe Ox.

a) Montrer que, pour qu'un électron puisse s'échapper du métal, il faut que la composante p_x de sa quantité de mouvement soit supérieure à $\sqrt{2mV_0}$.

b) Calculer le nombre d'électrons qui quittent ainsi le métal (maintenu à la température T) par unité de surface et par seconde (en pratique, on applique un petit champ électrique qui entraîne les électrons émis; on évite de cette manière la formation d'une charge d'espace au voisinage de la surface du métal).

c) Calculer la densité de courant j_x ainsi créée (« courant thermoionique »). Donner son expression en fonction de la température T et du travail d'extraction W_0 (« loi de Richardson-Dushman »). Évaluer l'ordre de grandeur du courant à la température ordinaire.

Montrer que l'étude de la variation de j_x avec T permet de mesurer W_0.

VI.17. Paramagnétisme de Pauli. On considère un métal, en équilibre à la température T, contenant N électrons par unité de volume. On suppose les électrons indépendants et on note $\rho(\varepsilon)$ la densité d'états d'un électron d'énergie ε. On soumet le métal à un champ magnétique uniforme $\vec{B}$ et on se propose de déterminer sa susceptibilité paramagnétique

$$\chi_P(T) \widehat{=} \lim_{B \to 0} \frac{M}{B} ,$$

où M est l'aimantation acquise par le système des spins électroniques.

1) μ_B étant le magnéton de Bohr, chaque électron contribue à M pour $-\mu_B$ si son spin est parallèle ($\uparrow$) au champ appliqué, pour $+\mu_B$ si son spin est antiparallèle ($\downarrow$) à $\vec{B}$.

a) Montrer que les densités d'états des électrons ayant un état de spin déterminé sont

$$\rho_\uparrow(\varepsilon) = \frac{1}{2} \rho(\varepsilon - \mu_B B) , \quad \rho_\downarrow(\varepsilon) = \frac{1}{2} \rho(\varepsilon + \mu_B B) .$$

b) Donner l'expression générale, en fonction du potentiel chimique μ des électrons et de la température T, des nombres moyens $N_\uparrow$ et $N_\downarrow$ d'électrons de spin $\uparrow$ et $\downarrow$ à l'équilibre. Exprimer M en fonction des mêmes grandeurs et écrire l'équation donnant μ à partir de N et T.

2) Dans la limite $B \to 0$, on peut développer $\rho_\uparrow$ et $\rho_\downarrow$ au premier ordre en B.

a) En déduire que, à cet ordre, le potentiel chimique est le même qu'en l'absence de champ magnétique(*).

b) Montrer que

$$\chi_P(T)=\mu_B^2\int\frac{\mathrm{d}\rho}{\mathrm{d}\varepsilon}(\varepsilon)f(\varepsilon,\,T)\,\mathrm{d}\varepsilon=-\mu_B^2\int\rho(\varepsilon)\frac{\partial f}{\partial\varepsilon}(\varepsilon,\,T)\,\mathrm{d}\varepsilon\ ,$$

où $f(\varepsilon, T)$ est la fonction de Fermi à la température T en l'absence de champ magnétique.

3) On considère tout d'abord le domaine des basses températures où le gaz d'électron est dégénéré ($T\ll T_F$).

a) Calculer la susceptibilité à température nulle $\chi_P(0)$.

b) Montrer que, pour $T\ll T_F$,

$$\chi_P(T)=\chi_P(0)\left[1+\frac{\pi^2}{6}k^2T^2\left(\frac{\rho''}{\rho}-\left(\frac{\rho'}{\rho}\right)^2\right)+\ldots\right],$$

où ρ et ses dérivées ρ' et ρ'' sont prises au niveau de fermi μ_0.

c) Donner l'expression explicite de $\chi_P(0)$ dans le cas d'un gaz d'électrons libres.

4) Montrer que, dans le domaine des hautes températures ($T\gg T_F$), on retrouve l'expression classique

$$\chi_P(T)\simeq N\mu_B^2/kT\ .$$

VI.18. Diamagnétisme de Landau. Lorsque, au complément VI.A (§ II.2.c), nous avons étudié les propriétés magnétiques d'un métal, nous n'avons considéré que l'effet paramagnétique provenant du couplage des spins électroniques avec le champ extérieur. Mais il existe aussi un effet diamagnétique induit par le mouvement des électrons. C'est cet effet, appelé « diamagnétisme de Landau », que nous nous proposons d'étudier ici dans le cas d'un gaz de N électrons libres indépendants, en négligeant pour simplifier le couplage entre les spins et le champ.

Rappelons tout d'abord (complément III.A, note 15) que, dans une boîte de volume V, les niveaux d'énergie d'un électron libre soumis à un champ magnétique uniforme $\vec{B}_0$ sont de la forme

$$\varepsilon_n(k_z)=\frac{\hbar^2k_z^2}{2m}+(2n+1)\mu_BB_0\ ,$$

avec $\begin{cases} n=0,\ 1,\ 2,\ 3,\ \ldots \\ k_z \text{ multiple entier positif, négatif ou nul de } 2\pi/L_z. \end{cases}$

Chacun de ces niveaux est dégénéré, le nombre d'états de vecteur d'onde compris entre k_z et $k_z+\mathrm{d}k_z$, pour n fixé, étant (compte tenu de la dégénérescence de spin)

$$\rho\left(\varepsilon_n(k_z)\right)\mathrm{d}k_z=\frac{eB_0V}{2\pi^2\hbar}\,\mathrm{d}k_z\ .$$

1) On connaît l'expression

$$M=-\frac{1}{V}\left(\frac{\partial F}{\partial B_0}\right)_{T,\,V,\,N}$$

(*) Ce résultat est d'ailleurs évident : la symétrie du problème impose que μ n'est pas modifié lorsqu'on change B en $-B$.

de l'aimantation à partir de l'énergie libre; en déduire son expression à partir du grand potentiel $J(T, V, \mu, B_0)$. Montrer que la susceptibilité magnétique est reliée à J par

$$\chi \widehat{=} \lim_{B_0 \to 0} \frac{\partial M}{\partial B_0} = -\frac{1}{V} \lim_{B_0 \to 0} \left(\frac{\partial^2 J}{\partial B_0^2} \right)_{T, V, \mu}$$

2) Calculer la grande fonction de partition ξ_λ pour un état individuel (λ). En déduire l'expression générale du grand potentiel $J(T, V, \mu, B_0)$.

3) En utilisant la relation mathématique (formule d'Euler)

$$\int_0^\infty f(x)\,\mathrm{d}x = \sum_{n=0}^{\infty} \alpha f\left(\left(n+\frac{1}{2}\right)\alpha\right) - \frac{\alpha^2}{24}\frac{\mathrm{d}f}{\mathrm{d}x}(x=0) + \mathcal{O}(\alpha^3)\,,$$

établir le développement limité de J jusqu'au deuxième ordre en B_0. En déduire l'expression de la susceptibilité diamagnétique χ_d sous forme d'une intégrale définie faisant intervenir la température et le potentiel chimique.

4) *a*) Calculer χ_d pour un gaz fortement dégénéré ($T \ll T_F$), en ne retenant que le terme significatif d'ordre le plus bas en T/T_F. Montrer que la valeur obtenue est, au signe près, égale au tiers de la susceptibilité paramagnétique à température nulle (formule (A.35) du complément VI.A).

b) En utilisant les résultats de l'exercice VI.17 (question 2.*b*), montrer que ce résultat est indépendant de la température : pour un gaz d'électrons libres, on a toujours $\chi_d(T) = -\chi_P(T)/3$ quelle que soit T, où $\chi_P(T)$ est la susceptibilité paramagnétique due au couplage des spins électroniques avec le champ.

VI.19. Écrantage d'une charge dans un métal. Dans le modèle des électrons libres et indépendants, un métal est constitué d'ions positifs et d'électrons qu'on suppose répartis uniformément. Si n_0 est le nombre d'électrons par unité de volume, la neutralité électrique du métal impose que la densité de charge ionique (que nous supposerons, dans toute la suite, invariable) vaut $n_0 e$. Si on introduit, en un point du métal (choisi comme origine), une impureté ionisée de charge Ze (Z pouvant être positif ou négatif), la distribution des électrons n'est plus uniforme et dépend de la distance r à l'impureté. On cherche à déterminer la densité électronique d'équilibre $n(r)$ au point $\vec{r}$ en présence de l'impureté, à la température T.

1) Soit $\phi(r)$ le potentiel électrostatique créé au point $\vec{r}$ par la nouvelle distribution des charges. On choisira $\phi(\infty) = 0$. Écrire l'équation de Poisson qui relie $\phi(r)$ et $n(r)$. Que vaut $\phi(r)$ pour $r \to 0$?

2) Considérons un électron au point $\vec{r}$. A son énergie cinétique s'ajoute alors une énergie potentielle $-e\phi(r)$.

a) Exprimer $n(r)$ en fonction de la température T et du potentiel chimique μ des électrons.

b) En examinant la limite $r \to \infty$, montrer que μ est le même qu'en l'absence de l'impureté.

c) Montrer qu'on obtient ainsi un système fermé d'équations qui permet, en principe, de déterminer $\phi(r)$ et $n(r)$ en fonction de T.

3) Dans un métal, on a $T \ll T_F$ et on obtient une bonne approximation en résolvant le problème à $T = 0$.

a) Montrer alors que, pour les distances r telles que $|e\phi(r)| \ll \mu_0$, la solution cherchée est de la forme

$$\varphi(r) = \frac{A}{r} e^{-r/\ell_0} .$$

Déterminer A et ℓ_0, et calculer la densité $n(r)$ correspondante.

b) Estimer numériquement ℓ_0 (appelée « longueur d'écran de Thomas-Fermi »). Au vu de cette estimation, que peut-on dire de la validité du modèle?

4) On se place maintenant dans la limite inverse ($T \gg T_F$), ce qui correspond par exemple au cas des électrons dans la bande de conduction d'un semi-conducteur. Montrer que, pour les distances r telles que $|e\phi(r)| \ll kT$ cette fois, on a encore une solution de la forme

$$\varphi(r) = \frac{A}{r} e^{-r/\ell} .$$

Déterminer ℓ (appelée « longueur d'écran de Debye »). Calculer la densité $n(r)$ correspondante.

Solution : Complément V.B, § II.2.b.

ÉLECTRONS DANS LES SEMI-CONDUCTEURS

VI.20. On veut étudier les propriétés d'un semi-conducteur de type n (complément VI.A, § III.1.b), par exemple un cristal de germanium (4 électrons externes) dopé au phosphore (5 électrons externes) : dans ce cristal, N atomes de germanium sont remplacés par N atomes de phosphore; ce « dopage » apporte donc N électrons supplémentaires, mais aussi N états supplémentaires localisés sur les « pièges » que constituent les atomes de phosphore. A température nulle, la bande de conduction du cristal est vide et sa bande de valence totalement occupée; les N états « pièges », dont l'énergie (comptée à partir du fond de la bande de conduction) vaut $-\Delta$ ($\Delta > 0$), sont également tous occupés à $T = 0$. A température non nulle, certains électrons quittent les pièges pour passer dans la bande de conduction, où on les considérera comme libres; on suppose que Δ est très inférieure au « gap » entre la bande de conduction et la bande de valence et que, dans le domaine de température choisi, aucune transition d'électrons ne se produit entre les deux bandes.

Le volume V du cristal et la température T sont fixés. On suppose valable la limite thermodynamique, de sorte qu'on peut négliger les fluctuations du nombre d'électrons qui restent piégés et du nombre d'électrons libres dans la bande de conduction.

1) Montrer que le nombre N_p des électrons piégés est donné par la condition

$$\mu_p(N_p, T, V) = \mu_e(N - N_p, T, V) ,$$

où μ_p est le potentiel chimique des électrons piégés et $\mu_e(N_e, T, V)$ celui des électrons libres lorsque leur nombre est N_e.

2) On considère d'abord les électrons libres de la bande de conduction. On suppose qu'ils constituent un gaz parfait auquel s'applique l'approximation de Maxwell-Boltzmann.

a) Écrire la fonction de partition Z_e de N_e électrons libres. En déduire leur potentiel chimique $\mu_e(N_e, T, V)$.

b) Quelle condition doit satisfaire μ_e pour que l'approximation de Maxwell-Boltzmann soit valable?

3) On considère ensuite les électrons piégés. Chaque piège peut accueillir au maximum un électron, dans l'un ou l'autre état de spin.

a) On raisonne en premier lieu comme si le nombre N_p d'électrons piégés était fixé. Combien de façons distinctes y a-t-il de les placer dans les N pièges? Calculer leur fonction de partition Z_p et en déduire leur potentiel chimique μ_p en fonction de N_p.

b) On peut aussi raisonner dans le cadre grand-canonique. Calculer la grande fonction de partition Ξ_p des électrons piégés, et en déduire leur nombre N_p en fonction du potentiel chimique μ_p.

c) Montrer que les deux méthodes donnent bien la même relation entre μ_p et N_p.

4) *a*) Montrer que le nombre N_p d'électrons piégés est déterminé par

$$\frac{(N-N_p)^2}{N_p}=\frac{V}{\Lambda^3}\,\mathrm{e}^{-\Delta/kT},$$

où $$\Lambda \widehat{=} \left(\frac{2\pi\hbar^2}{mkT}\right)^{1/2}.$$

b) Que devient N_p à basse température et à haute température? Peut-on comprendre qualitativement ces deux comportements limites?

c) Calculer N_p et le potentiel chimique μ des électrons à température basse mais non nulle. Comment se situe alors μ par rapport à l'énergie $-\Delta$ des états donneurs?

d) Le potentiel chimique μ peut-il passer au-dessous de $-\Delta$ lorsque la température augmente? Examiner si la condition écrite à la question 2.b est vérifiée à toute température (les transitions à partir de la bande de valence restant toujours négligeables).

VI.21. Semi-conducteur intrinsèque. Un semi-conducteur intrinsèque contient des électrons (dans les états occupés de la bande de conduction) et des trous (dans les états vides de la bande de valence) en nombre égal. L'énergie d'un électron de vecteur d'onde $\vec{k}$ est donnée par la relation

$$\varepsilon_e(\vec{k})=E_g+\hbar^2\vec{k}^{\,2}/2m_e,$$

où E_g est le «gap», c'est-à-dire la différence d'énergie entre le haut de la bande de valence et le bas de la bande de conduction. On peut considérer un trou comme une particule libre (fermion de masse m_t), d'énergie

$$\varepsilon_t(\vec{k})=\hbar^2\vec{k}^{\,2}/2m_t.$$

1) Électrons et trous peuvent apparaître ou disparaître par paires, par exemple suivant la réaction

$$\text{électron} + \text{trou} \rightleftarrows \text{photon}.$$

En déduire que le potentiel chimique des trous est opposé à celui des électrons.

2) A basse température, il y a peu d'électrons et de trous; le nombre moyen d'occupation de chaque état est donc très inférieur à 1 et on peut utiliser l'approximation de Maxwell-Boltzmann. Calculer le nombre moyen n d'électrons par unité de volume et leur potentiel chimique μ. Application numérique :

$$E_g=0{,}7\ \text{eV}, \quad T=300\ \text{K}, \quad m_t\simeq m_e\simeq 10^{-30}\ \text{kg}.$$

3) Vérifier que la condition de validité de l'approximation de Maxwell-Boltzmann est satisfaite, et ce d'autant mieux que la température est plus basse.

VI.22. Semi-conducteur extrinsèque. Soit un semi-conducteur de type n, contenant N_D états donneurs par unité de volume (par exemple un cristal de germanium dopé au phosphore). Les électrons peuvent se trouver soit dans la bande de conduction, où ils se comportent comme des particules libres de masse effective m^* et de spin 1/2, soit piégés sur une impureté dans un état donneur, avec une énergie égale à $-\varepsilon_D (\varepsilon_D > 0)$ comptée à partir du bas de la bande de conduction. Les N_D pièges sont discernables et peuvent retenir chacun *au plus un* électron, dans l'un ou l'autre état de spin. Ce semi-conducteur est maintenu à température constante T, suffisamment basse pour que l'on n'ait pas à tenir compte des électrons provenant de la bande de valence.

On supposera la statistique de Maxwell-Boltzmann applicable aux électrons de conduction et on vérifiera sur les résultats obtenus qu'il en est bien ainsi.

1) Donner l'expression de n, nombre moyen d'électrons par unité de volume dans la bande de conduction, en fonction de la température T et du potentiel chimique μ.

2) Donner l'expression de n_D, nombre moyen par unité de volume d'électrons piégés sur les donneurs, en fonction de μ et de T. Pourquoi n'obtient-on pas une distribution de Fermi-Dirac?

3) Déterminer n et μ en fonction de T dans les domaines des basses et des hautes températures : $T \ll T_1$ et $T \gg T_1$, où la température T_1 est définie par la relation

$$4N_D = \frac{1}{\Lambda^3(T_1)} \, e^{-\varepsilon_D / kT_1} ,$$

$\Lambda(T)$ étant la longueur d'onde thermique de de Broglie. Interpréter les résultats.

Déterminer l'ordre de grandeur de T_1 pour $N_D \simeq 10^{20}$ atomes/m^3, $\varepsilon_D = 0{,}012$ eV et $m^* \simeq 0{,}5 m_e$ (où m_e est la masse de l'électron). Comparer à la température T_2 à partir de laquelle les électrons provenant de la bande de valence deviennent plus nombreux que ceux des états donneurs (on prendra $\varepsilon_G \simeq 0{,}8$ eV pour l'énergie de gap).

4) On place le cristal dans un champ magnétique B (le moment magnétique d'un électron étant μ_B). Calculer la contribution des électrons de conduction et celle des électrons piégés à l'aimantation, en fonction du nombre n d'électrons dans la bande de conduction. Vérifier que (si l'on admet que le moment magnétique d'un électron est indépendant de l'état dans lequel il se trouve) l'aimantation totale ne dépend pas de n. Calculer la susceptibilité magnétique du système.

VI.23. On considère un semi-conducteur extrinsèque contenant N_D atomes donneurs par unité de volume. En utilisant les résultats des exercices VI.21 et VI.22, étudier l'allure de la variation du potentiel chimique des électrons en fonction de la température (on situera sa position par rapport au bas de la bande de conduction, par rapport aux niveaux donneurs, et par rapport au haut de la bande de valence).

APPLICATIONS SIMPLES A L'ASTROPHYSIQUE

VI.24. On cherche à évaluer le potentiel chimique des diverses particules constituant l'Univers à un stade de son évolution où ces particules sont en équilibre entre elles. Pour simplifier, on va raisonner comme s'il existait seulement, outre des photons notés γ, des neutrons n, des protons p, des électrons e^- et des neutrinos ν, ainsi que leurs antiparticules $\bar{n}$, $\bar{p}$, e^+ et $\bar{\nu}$. On suppose que ces particules constituent des gaz parfaits (quantiques) et qu'elles interagissent seulement à travers des réactions du type

$$a + b \rightleftarrows c + d + e \, . \qquad (1)$$

Ces réactions conservent la charge électrique, le « nombre baryonique » et le « nombre leptonique » : si q_i est la charge de la particule i (on prend pour unité la charge e du proton), on a

$$q_c + q_d + q_e = q_a + q_b \ , \tag{2}$$

et des égalités analogues pour le nombre baryonique et le nombre leptonique. Les caractéristiques des diverses particules sont données dans le tableau ci-après.

	γ	n	$\bar{n}$	p	$\bar{p}$	e^-	e^+	ν	$\bar{\nu}$
charge électrique	0	0	0	+1	−1	−1	+1	0	0
nombre baryonique	0	+1	−1	+1	−1	0	0	0	0
nombre leptonique	0	0	0	0	0	+1	−1	+1	−1

1) Montrer (cf. exercice III.23) que, à l'équilibre, les potentiels chimiques des particules prenant part à la réaction (1) doivent vérifier

$$\mu_c + \mu_d + \mu_e = \mu_a + \mu_b \ . \tag{3}$$

2) Montrer que, si la réaction (1) est permise par les lois de conservation, la réaction

$$a + b \rightleftarrows c + d + e + \gamma \tag{4}$$

l'est aussi. En déduire que le potentiel chimique μ_γ des photons est nul.

3) Montrer qu'un système constitué d'une particule a et de son anti-particule $\bar{a}$ peut « communiquer » avec un système constitué seulement de photons : par exemple,

$$a + \bar{a} \rightleftarrows 2\gamma \tag{5}$$

est possible. En déduire que les potentiels chimiques de a et de $\bar{a}$ sont opposés :

$$\mu_{\bar{a}} = -\mu_a \ . \tag{6}$$

4) Montrer que la réaction

$$e^- + p \longrightarrow n + \nu \tag{7}$$

est permise. En déduire que les potentiels chimiques de toutes les particules énumérées plus haut peuvent être obtenus à partir de 3 d'entre eux. Montrer qu'inversement l'existence de 3 lois de conservation interdit de réduire à moins de 3 le nombre de potentiels chimiques indépendants.

5) Choisissons par exemple μ_p, μ_{e^-} et μ_ν comme potentiels chimiques indépendants. On peut les déterminer si l'on connaît la charge globale Q, le nombre baryonique total B et le nombre leptonique total L (par unité de volume) dans l'Univers.

a) Montrer en effet que, si N_a désigne le nombre de particules a par unité de volume, on a

$$Q = N_p - N_{\bar{p}} - N_{e^-} + N_{e^+} \ . \tag{8}$$

Écrire les relations analogues donnant B et L.

Indiquer (sans nécessairement expliciter les formules) comment ces relations déterminent μ_p, μ_{e^-} et μ_ν à partir de Q, B, L. On admettra que la correspondance entre $\{Q, B, L\}$ et $\{\mu_p, \mu_{e^-}, \mu_\nu\}$ est bijective.

b) Montrer, à partir de l'égalité (6), que les trois potentiels chimiques changent de signe si l'on change de signe les trois densités Q, B et L.

c) La charge globale de l'Univers est nulle. Son nombre baryonique global ne l'est pas, puisque nous constatons aujourd'hui l'existence de matière (constituée de n et de p), alors que l'antimatière ($\bar{n}$ et $\bar{p}$) est pratiquement absente. Mais B, qui vaut environ $10^{-7}\ \mathrm{cm}^{-3}$, est 10^9 ou 10^{10} fois plus petit que le nombre N_γ de photons par unité de volume (exercice VI.27, question 1); on peut donc à une bonne approximation considérer que B est nul. Le nombre leptonique par unité de volume L est difficile à évaluer, car il est impossible actuellement de mesurer la densité de neutrinos; il est raisonnable de penser que L est lui aussi très petit devant N_γ. On prend donc

$$B \simeq L \simeq Q = 0 \ . \tag{9}$$

Montrer que ces valeurs impliquent la *nullité des potentiels chimiques* de toutes les particules considérées.

VI.25. Dans une enceinte à température très élevée (comme une étoile), le nombre des électrons n'est pas fixé du fait de la réaction de création de paires

$$\gamma + \gamma \rightleftarrows e^+ + e^-$$

(e^+ et e^- désignent respectivement un positron et un électron, γ un photon). Par contre, la différence $N_- - N_+ = N_0$ entre le nombre d'électrons N_- et le nombre de positrons N_+ est fixée puisque les électrons et les positrons ne peuvent apparaître ou disparaître que par paires. On supposera que N_0/V ne dépasse pas la densité électronique dans un atome.

1) Quelle relation existe-t-il, à l'équilibre, entre les potentiels chimiques μ_+ et μ_- des positrons et des électrons?

2) On se place tout d'abord dans le domaine de basse température $T \ll mc^2/k$ (mais suffisamment élevée cependant pour que des atomes ne puissent pas se former). Dans ce domaine, les électrons et les positrons forment deux gaz (que l'on supposera parfaits) bien décrits par la statistique de Maxwell-Boltzmann, et leur énergie individuelle reste voisine de mc^2 : $\varepsilon(\vec{p}) \simeq mc^2 + p^2/2m$.

a) Exprimer les densités $n_+ = N_+/V$ et $n_- = N_-/V$ des positrons et des électrons à l'équilibre en fonction des potentiels chimiques μ_+ et μ_-.

b) Déduire de la question 1 que
$n_+ n_- = f(T)$,
où $f(T)$ est une fonction de la température que l'on explicitera. Déterminer n_+ et n_- en fonction de $n_0 = N_0/V$ et de T.

3) On considère maintenant l'autre domaine extrême où $kT \gg mc^2$. Dans ce domaine, le nombre de paires électron-positron formées est très grand devant le nombre N_0 d'électrons excédentaires (on pourra vérifier cette affirmation *a posteriori* sur le résultat de la question *b*). On peut donc considérer qu'on a alors, avec une bonne précision, $N_+ = N_-$.

a) Montrer que ceci implique $\mu_+ = \mu_- = 0$.

b) Dans cette limite, les électrons et les positrons sont dégénérés et pratiquement tous ultra-relativistes ($\varepsilon(\vec{p}) \simeq cp$). En déduire les nombres $N_+(T)$ et $N_-(T)$ de positrons et d'électrons dans l'enceinte, sachant que

$$\int_0^\infty \frac{x^2\,\mathrm{d}x}{e^x + 1} = \frac{3}{2}\zeta(3) \simeq 1{,}803 \ .$$

c) Calculer l'énergie du gaz de positrons et celle du gaz d'électrons en équilibre dans l'enceinte, sachant que

$$\int_0^\infty \frac{x^3\,dx}{e^x+1}=\frac{7\pi^4}{120}\,.$$

Comparer avec l'énergie du gaz de photons dans l'enceinte.

VI.26. Ordres de grandeur concernant les étoiles à neutrons. Lorsqu'une étoile a épuisé son combustible nucléaire, elle s'effondre sous l'effet des forces gravitationnelles. Si sa masse est supérieure à la limite de Chandrasekhar (complément VI.B, § III.4), la pression du gaz dégénéré d'électrons qu'elle contient n'est pas suffisante pour arrêter le processus. L'effondrement se poursuit alors jusqu'à ce que l'étoile devienne un trou noir. Mais il peut arriver que l'échauffement qui accompagne l'effondrement conduise à une explosion, donnant lieu à une supernova; l'étoile éjecte alors de la matière, et sa masse résiduelle peut descendre au-dessous de la limite de Chandrasekhar. Dans ce cas l'astre, trop comprimé pour pouvoir se stabiliser en une naine blanche, se transforme en une « étoile à neutrons ».

1) La différence de masse $m_n - m_p$ entre le neutron et le proton est telle que $(m_n - m_p)c^2 \simeq 1{,}3$ MeV.

a) Calculer (cf. § III.2 du complément VI.B) la masse volumique ρ qu'il faut atteindre pour que l'énergie de Fermi des électrons dépasse cette valeur.

b) On admet que les protons et les neutrons de l'astre sont pratiquement au repos (comment pourrait-on s'assurer que cette hypothèse est vérifiée?). Lorsque la masse volumique atteint l'ordre de grandeur calculé en *a*, la réaction

$$e^- + p \longrightarrow n + \nu$$

devient possible, les neutrinos s'échappant de l'étoile. Montrer qu'en outre la désintégration du neutron

$$n \longrightarrow p + e^- + \bar{\nu}$$

est bloquée par le principe de Pauli (appliqué aux électrons).

2) Le processus qui vient d'être esquissé aboutit à la formation d'un gaz de neutrons dégénéré; c'est essentiellement la pression quantique de ce gaz qui stabilise l'astre.

a) Montrer que les ordres de grandeur correspondants s'obtiennent à partir des formules du complément VI.B (§ III.4) en remplaçant la masse m de l'électron par celle du neutron m_n.

b) En déduire que la masse limite M_0 reste pratiquement la même, mais que le rayon d'une étoile à neutrons est environ 1 000 fois plus petit que celui d'une naine blanche de même masse. Quel est l'ordre de grandeur de la masse volumique d'une étoile à neutrons?

VI.27. Ordres de grandeur concernant le rayonnement fossile (complément VI.G).

1) Montrer que le nombre total de photons d'un corps noir est, comme l'entropie, proportionnel à T^3. Évaluer le nombre de photons par unité de volume dans le rayonnement cosmique à 3 K (avec les notations de la note 36 du chapitre VI, on donne $\Gamma(3)\zeta(3) = 2 \times 1{,}202\ldots$).

2) Pour $T = 3$ K, quelle est la longueur d'onde correspondant au maximum de la distribution de Planck?

3) Les premières mesures sur le rayonnement fossile ont été effectuées à des longueurs d'onde comprises entre 1 cm et 75 cm environ. Comparer, dans ce domaine, la prédiction de la loi de Planck à celle de la formule de Rayleigh-Jeans VI (III.24). Pensez-vous que ces premières mesures pouvaient déjà permettre d'établir que le rayonnement cosmique est distribué comme celui d'un corps noir?

4) On observe parfois, dans la lumière qui nous parvient de certaines étoiles, des raies sombres dues à l'absorption sélective de molécules présentes dans l'espace interstellaire. C'est ainsi que le radical cyanogène CN se manifeste par une raie de 3874 Å de longueur d'onde : en absorbant un photon de cette longueur d'onde, la molécule CN passe de son état fondamental à un état électronique excité. En réalité, la rotation de la molécule (complément III.B, § II.3.a) a pour effet de scinder chaque niveau électronique en plusieurs sous-niveaux proches. On constate, en analysant la structure de la raie, que le cyanogène interstellaire absorbe à 3874,608 Å, ce qui correspond à une transition à partir de son véritable état fondamental (nombre quantique de rotation $J=0$), mais aussi à 3873,998 Å et à 3875,763 Å; ces deux dernières composantes de la raie ne peuvent avoir comme point de départ que le premier sous-niveau rotationnel excité, situé à $4{,}70\times10^{-4}$ eV au-dessus du fondamental et 3 fois dégénéré ($J=1$).

a) Si les molécules CN sont à l'équilibre thermique avec le rayonnement cosmique, la population du premier niveau rotationnel excité est-elle appréciable?

b) Imaginer comment la mesure de l'intensité relative des diverses composantes de la raie d'absorption du cyanogène permet de déterminer la température du rayonnement qui emplit l'espace interstellaire; comment se placent ces mesures par rapport au maximum du spectre calculé à la question 2?

DIVERS

VI.28. Équilibre liquide-vapeur de l'hélium 3 en présence de champ magnétique. On se propose d'étudier l'équilibre liquide-vapeur de l'hélium 3 aux très basses températures ($T<0{,}1$ K), en présence de champ magnétique. Les atomes de ^{3}He ont un spin 1/2 (porté par leur noyau) et sont donc des fermions. Aux températures considérées, la vapeur est suffisamment diluée pour être traitée comme un *gaz parfait classique,* mais on doit tenir compte des *effets quantiques dans le liquide.* Le moment magnétique d'un atome de ^{3}He vaut $m_N=9{,}6.10^{-27}$ J/T et on supposera que, dans le domaine étudié, le volume molaire du liquide est une constante $v_L=37$ cm^3.

1) *Étude de la vapeur.* On considère N_V atomes de ^{3}He en équilibre à la température T dans un récipient de volume V en présence d'un champ magnétique uniforme $\vec{B}$.

a) Calculer la fonction de partition du système. En déduire son énergie libre et sa pression. Montrer que cette dernière est indépendante du champ appliqué.

b) Calculer le moment magnétique moyen $\mathcal{M}_V$ d'une mole de vapeur ainsi que la susceptibilité magnétique molaire

$$\chi_V(T)=\lim_{B\to 0}\frac{\mathcal{M}_V}{B}.$$

Évaluer numériquement $\chi_V(T)$ à 0,1 K.

c) Calculer le degré de polarisation du gaz, c'est-à-dire le rapport $(N_\uparrow-N_\downarrow)/N_V$, où $N_\uparrow$ est le nombre d'atomes dont le spin est orienté dans le sens de $\vec{B}$ et $N_\downarrow$ celui des atomes orientés en sens inverse. Pour quelle valeur de B le gaz est-il polarisé à 99 % à $T=0{,}1$ K?

d) Calculer le potentiel chimique μ_V de la vapeur en fonction des variables T, p, B. Montrer que

$$\left(\frac{\partial \mu_V}{\partial B}\right)_{T,\,p} = -\frac{\mathcal{M}_V}{N_A},$$

où N_A est le nombre d'Avogadro.

2) *Étude du liquide.* Pour traiter le liquide, on utilise un modèle très simple dans lequel l'interaction entre atomes est prise en compte en ajoutant à l'énergie cinétique et magnétique de chacun d'eux une énergie de cohésion $-\varepsilon_0(\varepsilon_0 > 0)$ constante, la même pour tous. Mis à part ce décalage d'énergie, les N_L atomes d'hélium 3 qui constituent le liquide se déplacent comme des particules libres et indépendantes, soumises au champ $\vec{B}$, dans le volume $V_L = v_L N_L / N_A$.

a) Montrer que, pour $T < 0{,}1$ K, l'hélium 3 liquide constitue un système de fermions fortement dégénéré. On obtient donc une bonne approximation en calculant ses propriétés à $T = 0$. Dans toute la suite, on se limitera à cette approximation.

b) Calculer le moment magnétique moyen $\mathcal{M}_L$ d'une mole de liquide, ainsi que la susceptibilité magnétique molaire χ_L. Évaluer numériquement χ_L et montrer qu'aux températures considérées, $\chi_L \ll \chi_V$.

c) Montrer qu'il existe une valeur B_0 de B au-dessus de laquelle le liquide est totalement polarisé. Que vaut le potentiel chimique μ_L pour $B > B_0$? Évaluer numériquement B_0, comparer à la valeur trouvée en 1-*c*, et montrer qu'en pratique on est toujours dans le cas $B < B_0$ (ce que l'on supposera dans toute la suite).

d) Calculer le grand potentiel du liquide et en déduire sa pression p_L en fonction de μ_L et de B. Montrer que

$$\left(\frac{\partial \mu_L}{\partial B}\right)_{P_L,\,T} = -\frac{\mathcal{M}_L}{N_A}.$$

Remarquons que nous n'obtenons ici que la pression cinétique (cf. complément III.C, § IV); pour connaître la véritable pression du liquide, il faudrait ajouter la pression dynamique (négative), que ce modèle trop simpliste ne permet pas de calculer.

3) *Étude de l'équilibre liquide-vapeur.*

a) On considère tout d'abord l'équilibre en champ nul ($B = 0$).

α) Déterminer la pression $p_0^s(T)$ de la vapeur en présence du liquide (pression de vapeur saturante); montrer que c'est une fonction croissante de T.

β) Soit $L_0(T)$ la chaleur latente de vaporisation par mole à la température T :

$$L_0(T) \widehat{=} T\left[S_V(T) - S_L(T)\right],$$

où $S_V(T)$ et $S_L(T)$ sont les entropies par mole de la vapeur et du liquide (en présence l'un de l'autre) à la température T. Montrer qu'aux basses températures on a

$$L_0(T) = \alpha + \gamma T + \mathcal{O}(T^2).$$

Déterminer α et γ. Montrer que, à cet ordre, $L_0(T)$ vérifie la relation de Clapeyron

$$L_0(T) = T(v_V - v_L)\frac{\mathrm{d}p_0^s}{\mathrm{d}T},$$

où v_V et v_L sont les volumes molaires de la vapeur et du liquide en équilibre à la température T.

γ) L'expérience montre que, à basse température, L_0 suit bien une loi linéaire du type précédent, avec $\alpha = 20$ J/mole. En déduire la valeur de ε_0. Montrer qu'il est justifié de traiter le gaz classiquement pour $T < 0{,}1$ K.

b) On considère maintenant l'équilibre en présence de champ magnétique. La pression de vapeur saturante est alors fonction des deux variables T et B : $p^s(T, B)$.

α) Montrer que la variation de $p^s(T, B)$ avec B est donnée par une formule analogue à celle de Clapeyron :

$$\left(\frac{\partial p^s}{\partial B}\right)_T = \frac{\mathcal{M}_V - \mathcal{M}_L}{v_V - v_L} .$$

β) En déduire $p^s(T, B)$ pour les faibles champs magnétiques en écrivant un développement limité de $p^s(T, B)$ jusqu'au premier ordre significatif en B. Placer la courbe représentant la pression de vapeur saturante $p^s(T, B)$ à B donné par rapport à la courbe en champ nul. Si on établit le champ en maintenant la température constante, est-ce le liquide qui s'évapore ou le gaz qui se liquéfie?

VI.29. Ondes de spin et « magnons » dans un ferromagnétique

N.B. La question 3 peut être abordée même si les questions 1 et 2 n'ont pas été résolues.

On considère un système de spins 1/2, fixes dans l'espace, interagissant par l'hamiltonien de Heisenberg (complément III.J, § I.3). Pour simplifier, nous prendrons d'abord une *chaîne linéaire,* fermée sur elle-même, de N tels spins, la distance entre voisins étant a. Lorsque ce système est plongé dans un champ magnétique $\vec{B}$ dirigé suivant Oz, l'hamiltonien s'écrit

$$H = -g\mu_B B \sum_{q=1}^{N} S_q^z - J \sum_{q=1}^{N} \vec{S}_q . \vec{S}_{q+1} , \tag{1}$$

avec $N+1 \equiv 1$ et $J>0$ (ainsi que g).

On note $|+\rangle$ et $|-\rangle$, pour chaque spin, les états propres de S^z :

$$S^z |\pm\rangle = \pm \frac{1}{2} |\pm\rangle , \tag{2}$$

et on rappelle les formules suivantes :

$$S^{\pm} \widehat{=} S^x \pm i S^y , \tag{3}$$

$$S^+ |+\rangle = 0 , \qquad S^+ |-\rangle = |+\rangle ,$$
$$S^- |+\rangle = |-\rangle , \quad S^- |-\rangle = 0 , \tag{4}$$

$$\vec{S}_1 . \vec{S}_2 = S_1^z S_2^z + \frac{1}{2} (S_1^+ S_2^- + S_1^- S_2^+) . \tag{5}$$

1) Considérons en premier lieu l'état du système, que nous noterons $|0\rangle$, dans lequel les N spins se trouvent tous dans leur état $|+\rangle$.

a) Montrer que $|0\rangle$ est un état propre de l'hamiltonien H et déterminer l'énergie E_0 correspondante.

b) Montrer que, dans cet état, chacun des termes des deux sommes figurant dans la formule (1) prend sa valeur maximale possible (il pourra être commode d'exprimer $\vec{S}_q . \vec{S}_{q+1}$ à partir de $(\vec{S}_q + \vec{S}_{q+1})^2$). En déduire que $|0\rangle$ est l'état fondamental du système.

2) Considérons maintenant les états du système dans lesquels l'un des N spins est dans l'état $|-\rangle$ et les $(N-1)$ autres dans l'état $|+\rangle$. Nous noterons $|q\rangle$ l'état du système dans lequel c'est le spin numéro q qui est dans l'état $|-\rangle$.

a) Montrer que

$$H\,|q\rangle = [E_0 + g\mu_B B + J]\,|q\rangle - \frac{J}{2}\,[\,|q-1\rangle + |q+1\rangle\,]\,. \tag{6}$$

b) L'état $|q\rangle$ n'est donc pas état propre de H, mais on doit pouvoir construire des combinaisons linéaires des N états $|q\rangle$ qui soient vecteurs propres de H. Posons

$$|u_k\rangle \widehat{=} \frac{1}{\sqrt{N}} \sum_{q=1}^{N} e^{ikqa}\,|q\rangle\,. \tag{7}$$

Montrer que cet état est normé si les $|q\rangle$ le sont. Montrer (cf. complément III.E, § II.1.b) que les valeurs permises de k sont de la forme

$$k = n\,\frac{2\pi}{Na}\,, \quad \text{avec} \quad -\frac{N}{2} < n < \frac{N}{2}\,, \tag{8}$$

et qu'elles sont donc au nombre de N.

c) Montrer que $|u_k\rangle$ est vecteur propre de H avec la valeur propre

$$E(k) = E_0 + g\mu_B B + J(1 - \cos ka)\,. \tag{9}$$

d) Pour interpréter physiquement l'« onde de spin » $|u_k\rangle$, on établira les propriétés suivantes de cet état :

α) le spin total vaut $N/2 - 1$;

β) la probabilité pour que l'état individuel $|-\rangle$ soit trouvé sur le site (q) est indépendante de q et vaut $1/N$;

γ) si $\vec{S}^{\perp}$ est la composante d'un spin $\vec{S}$ dans le plan perpendiculaire à Oz, la valeur moyenne quantique du produit scalaire

$$\vec{S}^{\perp}_{q} \cdot \vec{S}^{\perp}_{q'} = \frac{1}{2}\,(S^{-}_{q} S^{+}_{q'} + S^{+}_{q} S^{-}_{q'}) \tag{10}$$

est égale à

$$\langle u_k | \vec{S}^{\perp}_{q} \cdot \vec{S}^{\perp}_{q'} | u_k \rangle = \frac{1}{N} \cos\, k(q'-q)\,a\,. \tag{11}$$

Ainsi, chaque spin possède une petite composante transverse, c'est-à-dire perpendiculaire au champ magnétique, dont le module est égal, en valeur moyenne, à $1/\sqrt{N}$; les composantes transverses de deux spins distants de $d = (q'-q)\,a$ forment un angle égal à kd.

3) On considère maintenant un cristal ferromagnétique à *trois dimensions,* à *basse température* et en l'absence de champ extérieur. On admet que ses états excités de faible énergie sont assimilables à ceux d'un *gaz parfait de quasi-particules,* les *« magnons »*. Comme les phonons (complément III.E, § V), les magnons sont des *bosons de potentiel chimique nul.* Leurs états individuels possibles sont les ondes de spin, caractérisées par un vecteur d'onde $\vec{k}$ et une énergie individuelle de la forme

$$\varepsilon(\vec{k}) = Ak^2\,, \tag{12}$$

où A est une constante.

a) Montrer que la forme (12) de la « relation de dispersion » pour les ondes de spin est celle qui découle de la formule (9) lorsqu'on se limite aux grandes longueurs d'onde, c'est-à-dire aux petites valeurs de k. Ce sont les seules qui contribuent de façon significative dans le domaine des basses températures. On peut donc sans inconvénient, comme dans le modèle de Debye (complément III.E, § III.2.b), considérer que k varie de zéro à l'infini et étendre à tout cet intervalle la relation (12).

b) Calculer à partir de (12) la densité $\rho(\varepsilon)$ d'états individuels pour les magnons.

c) Écrire l'expression du nombre moyen de magnons à la température T, et montrer que ce nombre varie comme $T^{3/2}$. Comme chaque magnon réduit de 1 le spin total du système (cf. 2.*d*.α), en déduire que l'aimantation $M(T)$ d'un ferromagnétique se comporte à basse température selon la « loi de Bloch » :

$$M(T) = M(0)[1 - \alpha T^{3/2}] , \tag{13}$$

où α est une constante. Comparer avec le résultat obtenu dans l'approximation de champ moyen (formule (J.36) du complément III.J).

d) Écrire l'expression de l'énergie du gaz de magnons à la température T. En déduire que la capacité calorifique du système de spins se comporte comme $T^{3/2}$ à basse température.

Appendices

APPENDICE I

Notions élémentaires sur les probabilités

I. Éléments de calcul des probabilités
  1. La notion de probabilité
    a. Ensemble discret d'événements
    b. Ensemble continu. Densité de probabilité
  2. Les deux lois de composition fondamentales
    a. La « loi d'addition »
    b. La « loi de multiplication »
  3. Évaluation concrète des probabilités élémentaires
  4. Valeurs moyennes et distributions statistiques
    a. Les états d'un système et leurs probabilités
    b. Définition générale des valeurs moyennes
    c. Distribution statistique d'une grandeur physique
    d. Principales caractéristiques d'une distribution statistique

II. Deux exemples de distributions statistiques
  1. La distribution binomiale
    a. Énoncé du problème
    b. Solution
    c. Caractéristiques de la distribution binomiale
  2. La distribution gaussienne
    a. Définition
    b. Caractéristiques
    c. La distribution gaussienne comme limite

Cet appendice présente les définitions et résultats principaux concernant les probabilités. Il ne saurait être question ici d'un exposé complet ni rigoureux : nous nous limiterons aux notions élémentaires qui sont couramment utilisées dans ce livre, et nous traiterons en particulier de façon totalement cavalière les questions de convergence. Les compléments I.B à I.D développent quelques arguments un peu plus techniques sur ces mêmes sujets. Pour plus de détails, nous renvoyons aux ouvrages spécialisés cités en référence à la fin de l'appendice.

I. ÉLÉMENTS DE CALCUL DES PROBABILITÉS

La présentation que nous adoptons est délibérément concrète : nous nous appuierons sur des exemples simples et des arguments physiques.

1. LA NOTION DE PROBABILITÉ

a. Ensemble discret d'événements

Pour introduire la notion de probabilité, commençons par le cas où l'ensemble des événements possibles est fini.

α. Un exemple simple : pile ou face

Il est probablement inutile de décrire en détail ce jeu de hasard simple, qui consiste à jeter en l'air une pièce de monnaie après que chacun des deux joueurs en présence ait lié son sort à l'une des deux faces de la pièce. Malgré sa simplicité, le jeu de pile ou face contient les caractéristiques essentielles de toute situation analysable en termes de probabilités.

Si les joueurs pouvaient déterminer à l'avance laquelle des deux faces de la pièce va finalement être visible, il n'y aurait plus de jeu. Pourtant, on devrait être

capable, en principe, de calculer tout le mouvement de la pièce à partir des lois de la mécanique (newtonienne, bien sûr!). Mais ceci est illusoire, car des modifications imperceptibles affectant les conditions de lancement ou le point de chute peuvent changer radicalement le résultat. On parle alors de *hasard,* parce que l'information que l'on peut rassembler avant l'expérience est totalement insuffisante pour en prédire le résultat. Il est dès lors plus simple et plus efficace de renoncer à obtenir une telle prédiction par les lois physiques qui décrivent en principe le système, et de raisonner sur les caractéristiques du hasard lui-même : ici, on dira (nous discuterons ce point au paragraphe 3) que les deux joueurs ont *a priori* autant de chances de gagner.

D'autre part, on peut répéter le lancement de la pièce autant de fois que l'on veut, avec cette propriété remarquable que chaque nouveau lancement est indépendant des précédents : son résultat est *a priori* tout aussi aléatoire que si aucun autre lancement ne l'avait précédé[1].

Supposons la pièce non truquée et le lanceur impartial. Quand nous disons que les deux joueurs ont autant de chances de gagner, cela ne signifie pas qu'il ne va pas y avoir de gagnant. C'est évident pour le premier lancement de la pièce : il faut bien que ce soit pile *ou* face. Mais nous savons que ce sera vrai aussi, dans la plupart des cas, après plusieurs lancements : en général, si l'on joue N fois de suite à pile ou face, on n'obtient pas exactement $N/2$ fois pile et $N/2$ fois face (même si N est pair, même si N est grand). Ce n'est donc pas par une simple égalité de leurs gains que se manifeste concrètement l'égalité des chances des deux joueurs, mais de la façon suivante : si N_1 est le nombre de fois où l'on obtient pile, et N_2 celui où l'on obtient face, lorsqu'on joue N $(=N_1+N_2)$ coups de pile ou face, alors les *proportions relatives* N_1/N et N_2/N de piles et de faces *deviennent égales lorsque N devient très grand* : N_1 ni N_2 ne sont égaux à $N/2$, mais N_1/N et N_2/N tendent tous deux vers 1/2, c'est-à-dire que les différences $|N_1-N/2|$ et $|N_2-N/2|$ croissent moins vite que N quand celui-ci devient grand. C'est ce fait d'expérience qui, généralisé, conduit à la notion de *probabilité.*

β) *Généralisation. Définition des probabilités*

Imaginons donc une expérience qui possède les deux caractéristiques principales du jeu de pile ou face, à savoir :

(i) *Son résultat n'est pas connu à l'avance* avec certitude. Il appartient à un ensemble discret de résultats possibles, que l'on appelle souvent *«événements»,* et que nous noterons e_m, où m est un indice entier :

$$\{e_m;\ m=1,\ 2,\ \ldots\ .\} \tag{1}$$

est donc l'ensemble des résultats possibles.

(ii) *L'expérience peut être répétée,* dans les mêmes conditions, autant de fois que l'on veut.

Alors, si N_m est le nombre de fois où l'événement e_m se produit lorsqu'on effectue N expériences, la *probabilité* P_m de e_m est *définie* par

$$P_m \mathrel{\widehat{=}} \lim_{N\to\infty} \frac{N_m}{N}\,. \tag{2}$$

1. On entend parfois le «bon sens» affirmer que si pile, par exemple, est déjà «sorti» n fois de suite, il a moins de chances que face de le faire lors du $(n+1)$-ième coup. Si l'on veut bien s'en tenir à des raisonnements objectifs — et ne pas invoquer une quelconque «fatalité» — on se convaincra facilement que les lancements successifs sont bien indépendants, et l'on gardera cette conviction aussi ferme dans toutes les situations analogues.

Soulignons que ceci ne signifie *pas* que N_m devient égal à NP_m lorsque N devient très grand (voir paragraphe II.1.c.γ).

Il ressort de la définition (2) que

$$0 \leqslant P_m \leqslant 1 \quad \forall m \,. \tag{3}$$

Un événement de probabilité nulle est dit *impossible;* un événement de probabilité 1 est dit *certain.*

b. Ensemble continu. Densité de probabilité

La notion de probabilité peut être étendue à des cas où les résultats possibles forment un ensemble continu, comme par exemple les diverses positions d'un point (qui pourra représenter une particule) sur un axe, ou dans un plan, ou dans l'espace.

Supposons qu'on ait ainsi affaire à une *«variable aléatoire»* pouvant prendre *a priori* une valeur quelconque sur un intervalle continu (x_1, x_2), fini ou infini, de l'ensemble des réels. Ceci signifie que l'expérience permettant de déterminer la valeur de cette variable (c'est-à-dire de la mesurer) est du même type que celles du paragraphe a précédent, à la seule différence que l'ensemble des résultats possibles est continu. Il suffit alors de reprendre la définition des probabilités en la modifiant selon la technique habituelle de passage du discret au continu.

On commence par définir la probabilité, que nous noterons $\delta P(x)$, pour que la variable appartienne à l'intervalle $[x, x+\delta x]$: soit $\delta N(x)$ le nombre de fois que, dans N expériences identiques de mesure, la valeur de la variable est trouvée comprise entre x et $x+\delta x$; alors

$$\delta P(x) \mathrel{\widehat{=}} \lim_{N \to \infty} \frac{\delta N(x)}{N} \,. \tag{4}$$

S'il existe une fonction $w(x)$ telle que

$$\lim_{\delta x \to 0_+} \frac{\delta P(x)}{\delta x} = w(x) \,, \tag{5}$$

cette fonction est appelée *densité de probabilité* au point x. La *probabilité infinitésimale* $\mathrm{d}P(x)$ pour que la variable considérée appartienne à l'intervalle infiniment petit $[x, x+\mathrm{d}x]$ s'écrit alors

$$\mathrm{d}P(x) = w(x)\,\mathrm{d}x \qquad (\mathrm{d}x > 0) \,. \tag{6}$$

Elle est définie de la manière habituelle :

$$\mathrm{d}P(x) \mathrel{\widehat{=}} \lim_{N \to \infty} \frac{\mathrm{d}N(x)}{N} \,, \tag{7}$$

$\mathrm{d}N(x)$ étant le nombre de fois où, au cours de N mesures, le résultat est compris entre x et $x+\mathrm{d}x$.

La densité de probabilité $w(x)$ est partout positive ou nulle. Notons que, si x a des dimensions (par exemple celles d'une longueur), la probabilité $\mathrm{d}P(x)$ reste sans dimension d'après (7), mais d'après (6) $w(x)$ a les dimensions de l'inverse de x.

Les définitions précédentes se généralisent aisément aux situations où l'on doit traiter simultanément plusieurs variables aléatoires continues. Supposons par exemple que celles-ci soient au nombre de trois : x, y, z. La densité de probabilité (si elle existe) est alors une fonction $w(x, y, z)$ de ces trois variables. Elle caractérise la probabilité infinitésimale

$$\mathrm{d}P(x, y, z) = w(x, y, z)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z \tag{8.a}$$

pour que, simultanément, la première variable appartienne à l'intervalle $[x, x+\mathrm{d}x]$, la deuxième à $[y, y+\mathrm{d}y]$ et la troisième à $[z, z+\mathrm{d}z]$. On peut considérer x, y, z

comme les composantes cartésiennes d'un vecteur $\vec{r}$ repérant un point dans un espace à trois dimensions, et on note souvent

$$dP(\vec{r}) = w(\vec{r})\,d^3r \tag{8.b}$$

la probabilité infinitésimale $dP(x, y, z)$ pour que l'extrémité du vecteur aléatoire $\vec{r}$ appartienne au volume infinitésimal

$$d^3r \widehat{=} dx.dy.dz \tag{9}$$

de l'espace.

2. Les deux lois de composition fondamentales

Lorsque les probabilités d'un (ou plusieurs) ensemble(s) d'événements sont connues, on peut les utiliser pour en déduire celles d'autres événements, éventuellement plus complexes, dérivés des précédents. C'est le *calcul des probabilités* qui permet cette déduction. Il nous suffira ici de connaître les deux lois simples suivantes, qui se démontrent aisément à partir de la définition (2).

a. La «loi d'addition»

α. Énoncé

Considérons deux événements e_1 et e_2, de probabilités respectives $P(e_1)$ et $P(e_2)$. Supposons que ces événements s'excluent l'un l'autre; ceci signifie que, si l'un d'eux se produit, on est sûr que l'autre ne se produit pas dans la même expérience. Alors la probabilité $P(e_1 \text{ ou } e_2)$ pour que *l'un ou l'autre* de ces événements se réalise est la *somme de leurs probabilités*[2] :

$$\boxed{P(e_1 \text{ ou } e_2) = P(e_1) + P(e_2) \qquad \text{si } e_1 \text{ et } e_2 \text{ s'excluent.}} \tag{10}$$

Par exemple, la loi d'addition (10) permet de calculer la probabilité pour obtenir un nombre pair lorsqu'on jette un dé : c'est la somme des probabilités $P(2) + P(4) + P(6)$, où $P(n)$ est la probabilité d'obtenir n.

La loi d'addition s'étend sans difficulté aux variables aléatoires continues. Ainsi, la probabilité pour qu'une variable, de densité de probabilité $w(x)$, soit comprise entre a et b (avec $a < b$ et $[a, b] \in (x_1, x_2)$) s'écrit

$$P(a \leqslant x \leqslant b) = \int_a^b w(x)\,dx\,. \tag{11.a}$$

En effet, les diverses valeurs de x s'excluent évidemment les unes les autres, et l'on somme donc les probabilités $dP(x) = w(x)dx$ pour toutes les valeurs de x appartenant

---

2. Soit N_1 le nombre de fois que l'événement e_1 se produit dans un grand nombre N d'expériences, et N_2 celui où e_2 se produit. Si e_1 et e_2 sont incompatibles, il n'y a pas d'expérience qui contribue à la fois à N_1 et à N_2. Alors, par définition, l'événement composé « e_1 ou e_2 » s'est produit $(N_1 + N_2)$ fois dans ce lot de N expériences. Donc

$$P(e_1 \text{ ou } e_2) = \lim_{N \to \infty} \frac{N_1 + N_2}{N} = P(e_1) + P(e_2)\,. \tag{a}$$

Dans le cas général où les deux événements ne sont pas incompatibles, il faut, pour obtenir $P(e_1 \text{ ou } e_2)$, retrancher de $(N_1 + N_2)$ le nombre de fois où e_1 et e_2 se sont produits tous les deux (car les expériences correspondantes sont comptées dans N_1 *et* dans N_2). La formule générale s'écrit donc

$$P(e_1 \text{ ou } e_2) = P(e_1) + P(e_2) - P(e_1 \text{ et } e_2)\,. \tag{b}$$

à l'intervalle $[a, b]$ considéré. De même, dans le cas d'une variable aléatoire vectorielle de densité de probabilité $w(\vec{r})$, la probabilité pour que le point qui lui est associé appartienne à un domaine D de l'espace correspondant est donnée par

$$P(\vec{r} \in D) = \int_D w(\vec{r})\, \mathrm{d}^3 r \,. \tag{11.b}$$

β. Conséquence : normalisation des probabilités

Supposons que nous connaissions *tous les résultats possibles* e_m d'une expérience, avec leurs probabilités P_m, et qu'ils soient définis de telle sorte qu'*ils s'excluent mutuellement.* Lorsqu'on réalise l'expérience, la probabilité pour que le résultat appartienne à cet ensemble exhaustif d'événements incompatibles vaut évidemment 1. On déduit alors de la loi d'addition la relation suivante, appelée *condition de normalisation* des probabilités :

$$\boxed{\sum_m P_m = 1} \,. \tag{12}$$

Lorsqu'on écrit cette relation, il faut bien prendre garde que ses deux conditions de validité soient remplies dans le cas particulier considéré : la sommation doit porter sur tous les événements possibles, et il faut qu'ils s'excluent l'un l'autre.

Pour un événement e quelconque, de probabilité P, on définit son *contraire* (parfois appelé son *complémentaire*) e' comme étant l'événement « e ne se produit pas ». La probabilité P' du contraire e' vaut d'après ce qui précède

$$P' = 1 - P \,. \tag{13}$$

Lorsqu'on jette un dé, par exemple, la somme des probabilités associées aux six faces vaut toujours 1, même si le dé est pipé; l'événement contraire de « le dé donne 6 » est « le dé donne 1 ou 2 ou 3 ou 4 ou 5 ».

La condition de normalisation s'applique également aux cas de variables continues. En effet si, dans les formules (11), on somme sur toutes les valeurs possibles (qui constituent pour x un intervalle (x_1, x_2) fini ou infini de l'ensemble des réels, et pour $\vec{r}$ un domaine D_0, borné ou non, de l'espace à trois dimensions), on doit obtenir[3]

$$\int_{x_1}^{x_2} w(x)\, \mathrm{d}x = 1 \,, \tag{14.a}$$

$$\int_{D_0} w(\vec{r})\, \mathrm{d}^3 r = 1 \,. \tag{14.b}$$

Bien que $w(x)$ ne soit définie que sur (x_1, x_2), et $w(\vec{r})$ sur D_0, on les étend parfois à l'ensemble des réels et à tout l'espace, respectivement, en les définissant comme nulles à l'extérieur de (x_1, x_2), ou de D_0; dans les formules précédentes, l'intégrale porte alors sur $(-\infty, +\infty)$ dans (14.a) et sur tout l'espace à trois dimensions dans (14.b).

b. La « loi de multiplication »

Considérons maintenant deux événements e et f *statistiquement indépendants* (on dit aussi tout simplement « indépendants »). Ceci signifie que la probabilité

3. Une densité de probabilité est partout positive ou nulle, mais, contrairement aux probabilités elles-mêmes, elle n'est pas nécessairement bornée : c'est seulement son intégrale qui doit l'être pour que la condition de normalisation (14) soit satisfaite, et pour que les formules (11) donnent des probabilités inférieures ou égales à 1.

d'observer l'un d'eux ne dépend pas du résultat concernant l'autre. Par exemple, si l'on joue à pile ou face avec deux pièces distinctes, la probabilité d'obtenir pile avec la première ne dépend pas du résultat affiché par l'autre. Ou bien, si l'on joue avec une seule pièce, la probabilité de tirer pile dans un lancement donné ne dépend pas des résultats des essais précédents. Dans un cas de ce genre, la probabilité $P(e \text{ et } f)$ d'observer *à la fois les deux événements* e et f est le *produit de leurs probabilités* $P(e)$ et $P(f)$[4] :

$$\boxed{P(e \text{ et } f) = P(e).P(f) \quad \text{si } e \text{ et } f \text{ indépendants}} \quad . \tag{15}$$

La loi de multiplication s'étend sans difficulté aux cas de variables aléatoires continues. Supposons en effet qu'on ait affaire à deux telles variables, x de densité de probabilité $u(x)$ et y de densité de probabilité $v(y)$. Si ces deux variables sont statistiquement indépendantes, la probabilité infinitésimale $\mathrm{d}P(x, y)$ pour que la première soit comprise entre x et $x + \mathrm{d}x$ et que, simultanément, la seconde soit comprise entre y et $y + \mathrm{d}y$ s'obtient par le produit

$$\mathrm{d}P(x, y) = u(x)\mathrm{d}x \,.\, v(y)\mathrm{d}y \,, \tag{16}$$

de sorte que la densité de probabilité $w(x, y)$ associée au système des deux variables est simplement

$$w(x, y) = u(x)v(y) \quad \text{si } x \text{ et } y \text{ indépendantes.} \tag{17}$$

---

4. Supposons que l'on effectue un grand nombre N d'expériences « doubles », c'est-à-dire dans lesquelles on observe à la fois un résultat de chacun des deux ensembles auxquels appartiennent e et f (dans le cas du jeu de pile ou face avec deux pièces, on lance N fois les deux pièces simultanément). Soit $N(e)$ le nombre de fois que l'on obtient e (associé à n'importe quel résultat de l'autre ensemble), et N (e et f) le nombre de fois que les deux événements e et f apparaissent simultanément. Bien entendu, $N(e \text{ et } f)$ est compté dans $N(e)$, d'après les définitions que nous venons de donner.

Si $P(e)$ n'est pas nulle, on peut choisir N suffisamment grand pour que $N(e)$ soit grand lui aussi ($N(e) \to \infty$ en même temps que N). $N(e \text{ et } f)$ est alors le nombre de fois où f apparaît parmi les $N(e)$ expériences qui donnent e par ailleurs. D'après l'hypothèse d'indépendance statistique, le rapport $N(e \text{ et } f)/N(e)$ tend vers la probabilité $P(f)$ de l'événement f, puisque celle-ci est indépendante du résultat trouvé dans l'autre partie de l'expérience :

$$P(f) = \lim_{N(e)\to\infty} \frac{N(e \text{ et } f)}{N(e)} = \lim_{N\to\infty} \frac{N(e \text{ et } f)}{N} \frac{N}{N(e)} \,. \tag{a}$$

On obtient ainsi la formule cherchée, puisque la relation précédente implique

$$P(f) = \frac{P(e \text{ et } f)}{P(e)} \,. \tag{b}$$

Si $P(e)$ est nulle, il en est de même pour $N(e \text{ et } f)$ et donc pour $P(e \text{ et } f)$, de sorte que la formule (15) reste valable.

Dans le cas où e et f ne sont pas indépendants, on définit la *probabilité conditionnelle* $P(f/e)$ de trouver f sachant qu'on a e par ailleurs,

$$P(f/e) \widehat{=} \lim_{N(e)\to\infty} \frac{N(e \text{ et } f)}{N(e)} \,. \tag{c}$$

La formule (15) se généralise alors sous la forme

$$P(e \text{ et } f) = P(e).P(f/e) \,. \tag{d}$$

Dans ces notations, e et f sont statistiquement indépendants si et seulement si

$$P(f/e) = P(f) \,. \tag{e}$$

3. ÉVALUATION CONCRÈTE DES PROBABILITÉS ÉLÉMENTAIRES

Les deux lois que nous venons d'énoncer, et plus généralement le calcul des probabilités, permettent de calculer les probabilités d'événements composés, à partir de celles d'événements plus simples. Mais il faut un point de départ : dans chaque situation concrète, il existe un ensemble d'événements, que nous qualifierons d'*élémentaires,* dont il faut se donner *a priori* les probabilités pour pouvoir calculer toutes les autres. Comment détermine-t-on ces probabilités élémentaires?

Pour le comprendre, reprenons l'exemple de pile ou face. Qu'est-ce qui nous permet de dire que les deux joueurs ont autant de chances de gagner (ce qui implique la probabilité 1/2 pour chacun des deux résultats possibles)? C'est sans doute la constatation qu'aucun des deux résultats n'est *a priori* privilégié par rapport à l'autre, à partir du moment où l'on admet l'impossibilité d'une prédiction fondée sur les lois de la mécanique : *la situation est symétrique* pour les deux événements possibles, et c'est ce qui conduit à leur attribuer des *probabilités égales.*

Mais ceci n'est plus vrai si par exemple la pièce a été truquée, de façon à privilégier l'une de ses deux faces. Nous ne pouvons plus dans ce cas évaluer *a priori* les probabilités des deux résultats possibles, et celles-ci ne sont plus égales. Il nous faut les *mesurer,* en étudiant *expérimentalement* le comportement de la pièce : nous la lancerons un grand nombre N de fois, et compterons le nombre N_1 de cas où elle donne pile, et le nombre N_2 ($= N - N_1$) où elle donne face; nous en déduirons, par la définition (2), une évaluation des probabilités P_1 et P_2 des deux résultats possibles[5], et pourrons ensuite utiliser ces valeurs pour effectuer des prédictions sur des expériences ultérieures où cette même pièce serait impliquée[6].

D'une façon générale, il y a essentiellement *deux manières d'évaluer les probabilités élémentaires.* La première est applicable dans les situations où il y a *symétrie* entre tous les événements élémentaires, auquel cas ces événements sont tous *également probables* (c'est-à-dire que leurs probabilités sont égales, et donc déterminées par la condition de normalisation)[7]; la seconde est complètement générale, mais exige une *étude statistique préalable* du système considéré, étude qui conduit à une *mesure des probabilités* à partir de leur définition[5], [6].

4. VALEURS MOYENNES ET DISTRIBUTIONS STATISTIQUES

a. Les états d'un système et leurs probabilités

Nous adoptons désormais la terminologie utilisée habituellement en physique : nous étudions un *système,* qui se trouve dans un *état* appartenant à un *ensemble d'états possibles,* et sur lequel on peut mesurer diverses *grandeurs physiques ou observables.*

---

5. Le nombre N d'expériences est en pratique toujours fini, même s'il est grand. Il en résulte une *incertitude expérimentale* sur la mesure des probabilités, incertitude qui décroît bien sûr lorsque N croît, et qui peut être évaluée de façon objective (voir réf. 2, chapitres 6, 7, 8).

6. Il est des cas où la situation n'est pas symétrique (la pièce est truquée), où l'on n'a pas la possibilité d'effectuer la mesure des probabilités, et où pourtant on veut en avoir une estimation pour pouvoir les utiliser. Dans de tels cas, si l'on n'a aucune information sur la manière dont la symétrie est brisée (c'est-à-dire si l'on ne sait pas si c'est pile ou face qui est privilégié), on ne peut que raisonner comme si la situation était symétrique, et choisir, comme estimation, des probabilités égales : si l'on faisait autrement, on risquerait d'aggraver l'erreur commise, en affectant une probabilité plus grande à un résultat qui est en réalité défavorisé.

7. La théorie de l'information permet d'étendre cette méthode d'évaluation *a priori* des probabilités à des situations où il n'y a pas vraiment symétrie entre les événements élémentaires, et où donc leurs probabilités ne sont pas égales (complément V.D, § I).

La manière dont est défini l'état d'un système dépend de la nature de ce système et de la description qu'on veut en donner. Dans l'exemple de pile ou face, il y a seulement deux états possibles parce qu'on ne s'intéresse qu'à la face de la pièce qui est finalement visible, et pas à la position précise de cette pièce. Pour les systèmes que nous étudions en mécanique statistique, il s'agit principalement d'états quantiques (I, § III.A), quelquefois d'états au sens de la mécanique newtonienne (I, § III.C). Dans tous les cas, la définition adoptée doit être telle que la donnée de l'état suffise à déterminer toutes les caractéristiques du système auxquelles on peut s'intéresser.

Dans certains cas, les états possibles du système constituent un ensemble dénombrable, c'est-à-dire qu'on peut les repérer par un ou plusieurs indices discrets, que nous symbolisons comme ci-dessus par m. Dans un tel cas, nous noterons simplement (m) un état possible du système, et P_m la probabilité pour que le système se trouve dans cet état. Nous supposons que les divers états possibles s'excluent l'un l'autre[8], de sorte que les probabilités P_m sont normalisées :

$$\sum_m P_m = 1 \ , \tag{18.a}$$

la somme portant sur tous les états possibles.

Il se peut aussi que les états du système soient caractérisés par une ou plusieurs variables aléatoires continues (« indices continus »), que nous symbolisons par x. Il faut alors se donner la densité de probabilité $w(x)$ correspondante, elle aussi normalisée :

$$\int_{-\infty}^{+\infty} w(x)\,\mathrm{d}x = 1 \tag{18.b}$$

(si x représente plusieurs variables, $\mathrm{d}x$ est le produit de leurs différentielles et l'intégrale ci-dessus est multiple).

Si l'on connaît l'état du système, toutes ses caractéristiques sont déterminées, par définition même de la notion d'état. Par conséquent, *les probabilités associées aux divers états possibles du système constituent la donnée de base,* à partir de laquelle on peut calculer la probabilité de n'importe quel résultat ou propriété du système.

b. Définition générale des valeurs moyennes

Considérons une propriété quelconque du système étudié, que nous notons f, et soit f_m la valeur qui lui est associée[9] lorsque le système se trouve dans l'état (m). La valeur moyenne de f, notée $\bar{f}$, est par définition[10]

$$\boxed{\bar{f} \mathrel{\widehat{=}} \sum_m P_m f_m} \tag{19.a}$$

8. On pourrait penser que ceci ne s'applique pas aux états quantiques, dont on peut former des superpositions linéaires. Nous indiquons au chapitre I (§ III.A) comment éviter cette difficulté en choisissant une base particulière dans l'espace des états quantiques (la solution générale du problème est donnée dans le complément I.H).

9. Nous supposons ici que f possède une valeur unique dans l'état (m). Ce n'est pas nécessairement le cas si (m) est un état quantique et f une observable; f_m est alors la valeur moyenne (au sens de la mécanique quantique) de l'observable f dans l'état (m).

10. C'est effectivement $\bar{f}$ que l'on trouve si l'on effectue la moyenne, au sens ordinaire, des divers résultats $f(1)$, $f(2)$, ..., $f(N)$ obtenus dans une série d'un grand nombre N de mesures successives de la propriété f, sur des systèmes tous préparés de la même façon :

$$\bar{f} = \lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} f(n) \ .$$

($\bar{f}$ est quelquefois appelée « espérance mathématique » de f). Si les états du système forment un continuum, la propriété considérée aura une valeur $f(x)$ déterminée par la donnée de la (ou des) variables(s) aléatoires(s) caractérisant un état, et sa valeur moyenne s'écrira

$$\bar{f} \widehat{=} \int w(x) f(x) \, dx \, . \tag{19.b}$$

La définition (19) implique que *l'opération valeur moyenne est linéaire :* si λ et μ sont des réels quelconques,

$$\overline{\lambda f + \mu g} = \sum_m P_m (\lambda f_m + \mu g_m) = \lambda \bar{f} + \mu \bar{g} \, . \tag{20}$$

Étant donné une grandeur physique A, on peut calculer la valeur moyenne d'une puissance α *a priori* quelconque de cette grandeur :

$$\overline{A^\alpha} = \sum_m P_m (A_m)^\alpha \tag{21.a}$$

ou $$\overline{A^\alpha} = \int w(x) \left[A(x) \right]^\alpha dx \tag{21.b}$$

(à condition que la somme ou l'intégrale converge). Soulignons que, sauf cas exceptionnel, *la valeur moyenne $\overline{A^\alpha}$ n'est pas égale à la valeur moyenne $\bar{A}$ élevée à la puissance α.*

c. Distribution statistique d'une grandeur physique

Considérons un système, dont nous supposons connus les états possibles ainsi que leurs probabilités, et soit A une grandeur physique associée à ce système. On appelle *distribution statistique* de la grandeur A l'ensemble des probabilités des diverses valeurs possibles de cette grandeur.

Examinons d'abord le cas où les états (m) du système forment un ensemble discret, ainsi que les valeurs possibles A_i $(i = 1, 2, \ldots)$ de la grandeur considérée. Les probabilités $P(A_i)$ de ces valeurs, qui constituent la distribution statistique de A, peuvent être déduites des probabilités P_m des divers états (m). Supposons en effet que la valeur A_m de la grandeur A dans l'état (m)[11] soit égale à A_i. Il existe en général *plusieurs états (m) donnant la même valeur A_i.* Il suffit alors d'appliquer la loi d'addition (10) pour obtenir :

$$P(A_i) = \sum_{m \in M_i} P_m \, , \tag{22}$$

où M_i est l'ensemble des états (m) tels que $A_m = A_i$.

Le raisonnement est analogue dans le cas où les états du système forment un continuum. Si les valeurs possibles A_i de A sont encore discrètes, la formule (22) se généralise immédiatement :

$$P(A_i) = \int_{D_i} w(x) \, dx \, , \tag{23}$$

11. Si (m) est un état quantique et A une observable, le raisonnement reste valable à condition que les états (m) soient états propres de A. Si ce n'est pas le cas, la distribution statistique de A s'obtient par un calcul légèrement différent (voir I, § III.A.2.a et complément I.H, § I).

où D_i est le domaine de valeurs de x pour lesquelles $A(x)=A_i$. Si maintenant la grandeur considérée peut prendre un ensemble continu de valeurs, que nous noterons simplement A, sa distribution statistique est caractérisée par une densité de probabilité $W(A)$. La probabilité $W(A)\mathrm{d}A$ pour que la grandeur soit comprise entre A et $A+\mathrm{d}A$ s'obtient par une formule analogue aux précédentes :

$$W(A)\,\mathrm{d}A=\int_{\mathrm{d}D(A)} w(x)\,\mathrm{d}x\ , \tag{24}$$

où $\mathrm{d}D(A)$ est le domaine (infinitésimal) de valeurs de x telles que

$$A\leqslant A(x)\leqslant A+\mathrm{d}A \tag{25}$$

(ce dernier cas est étudié plus en détail dans le complément I.B).

La distribution statistique d'une grandeur physique est évidemment normalisée : si ses valeurs forment un ensemble discret,

$$\sum_i P(A_i)=1; \tag{26.a}$$

si elles sont continues,

$$\int_{-\infty}^{+\infty} W(A)\,\mathrm{d}A=1\ . \tag{26.b}$$

d. Principales caractéristiques d'une distribution statistique

α. Valeur la plus probable

La valeur la plus probable de la grandeur A est celle qui a la probabilité la plus grande, ou qui maximise la densité de probabilité $W(A)$. La valeur la plus probable n'est pas forcément unique : il peut y avoir plusieurs valeurs de A ayant la probabilité maximale, ou plusieurs maximums égaux de $W(A)$.

β. Valeur moyenne

La valeur moyenne $\overline{A}$ de la grandeur A étudiée a été définie ci-dessus (§ b) comme

$$\overline{A}=\sum_m P_m A_m \tag{27.a}$$

ou
$$\overline{A}=\int_{-\infty}^{+\infty} w(x)\,A(x)\,\mathrm{d}x \tag{27.b}$$

suivant les cas. Il est facile de voir qu'on peut également la calculer à partir de la distribution statistique de A :

$$\overline{A}=\sum_i P(A_i)\,A_i\ , \tag{28.a}$$

ou
$$\overline{A}=\int_{-\infty}^{+\infty} W(A)\,A\ \mathrm{d}A\ . \tag{28.b}$$

Dans la formule (27.a), par exemple, il suffit de regrouper les termes pour lesquels $A_m=A_i$ et de mettre A_i en facteur pour trouver, compte tenu de (22), le terme $A_iP(A_i)$ de (28.a). Le raisonnement peut être étendu au cas continu (voir complément I.B).

γ. Écart quadratique moyen

Une mesure de la grandeur A peut donner *a priori* n'importe quelle valeur permise, même parmi celles qui sont très différentes de la valeur moyenne $\overline{A}$. Ceci dépend de la distribution statistique : si les probabilités ne sont appréciables que pour des valeurs de A très voisines entre elles, presque toutes les mesures donneront des résultats très proches de $\overline{A}$; si au contraire les probabilités sont réparties presque également sur un large intervalle de valeurs possibles de A, les résultats de mesure pourront différer considérablement de $\overline{A}$. On dit dans ce dernier cas que la distribution statistique de la grandeur A présente une grande *dispersion,* ou qu'elle donne lieu à des *fluctuations* de grande amplitude.

Pour caractériser quantitativement la dispersion d'une distribution statistique, on pourrait songer à prendre la moyenne des écarts entre les résultats de mesure et la valeur moyenne $\overline{A}$. Mais une telle moyenne est nécessairement nulle, par définition même de $\overline{A}$. Pour éviter cette compensation entre valeurs positives et négatives de l'écart par rapport à la moyenne, on les élève au carré avant d'en prendre la moyenne. On définit ainsi l'*écart quadratique moyen* ΔA par

$$\boxed{(\Delta A)^2 \widehat{=} \overline{(A-\overline{A})^2}} \tag{29}$$

(ΔA est aussi appelé « écart-type », et $(\Delta A)^2$ « variance » de la distribution statistique). Cette définition s'explicite sous la forme

$$(\Delta A)^2 = \sum_m P_m (A_m - \overline{A})^2 = \sum_i P(A_i)(A_i - \overline{A})^2 \tag{30.a}$$

dans le cas de probabilités discrètes, et dans le cas continu sous la forme

$$(\Delta A)^2 = \int_{-\infty}^{+\infty} w(x)[A(x) - \overline{A}]^2 \, dx = \int_{-\infty}^{+\infty} W(A)[A - \overline{A}]^2 \, dA \, . \tag{30.b}$$

On déduit facilement de la définition l'expression suivante[12] :

$$\boxed{(\Delta A)^2 = \overline{A^2} - \overline{A}^2} \, . \tag{31}$$

Les formules (30) montrent que $(\Delta A)^2$ est une somme de termes qui sont tous positifs ou nuls. Par conséquent, $\Delta A = 0$ implique que chacun de ces termes soit nul : ou bien $P(A_i)$ (ou $W(A)$) est nulle, ou bien A_i (ou A) coïncide avec $\overline{A}$. Donc, *si l'écart quadratique moyen est nul, la grandeur A peut prendre une seule valeur,* qui est alors *certaine* (et avec laquelle s'identifie bien sûr la valeur moyenne $\overline{A}$).

---

12. On démontre la formule (31) en manipulant formellement le symbole de valeur moyenne, selon la règle de linéarité (20). On développe d'abord le carré au second membre de la définition (29) :

$$(\Delta A)^2 = \overline{A^2 - 2\overline{A}A + \overline{A}^2} \, . \tag{a}$$

Par linéarité, on obtient successivement

$$(\Delta A)^2 = \overline{A^2} - 2\overline{\overline{A}A} + \overline{\overline{A}^2} = \overline{A^2} - 2\overline{A}\,\overline{A} + \overline{A}^2 \, . \tag{b}$$

En effet, dans $2\overline{A}A$, $2\overline{A}$ est un nombre qu'on peut sortir de la valeur moyenne; de même, dans le troisième terme $\overline{A}^2$ sort de la valeur moyenne, et il reste à prendre la valeur moyenne de 1, qui vaut 1 à cause de la condition de normalisation (attention à la place de chacun des exposants par rapport au symbole de valeur moyenne). Les deux derniers termes se combinent finalement pour donner la formule (31) cherchée.

De façon générale, l'écart quadratique moyen ΔA caractérise l'ordre de grandeur des écarts par rapport à la moyenne $\overline{A}$ que l'on peut s'attendre à trouver[13] lors de mesures de la grandeur A. Soulignons cependant que la connaissance de la valeur moyenne et de l'écart quadratique moyen n'est pas en général suffisante pour déterminer la distribution statistique[14].

II. DEUX EXEMPLES DE DISTRIBUTIONS STATISTIQUES

Dans cette deuxième partie, nous étudions deux distributions particulières qui interviennent très souvent dans les applications.

1. LA DISTRIBUTION BINOMIALE

a. Énoncé du problème

Reprenons d'abord le jeu de pile ou face. Notons (+) et (−) les événements « pile » et « face », respectivement, et P_+ et P_- les probabilités correspondantes. Ces

13. Quelle que soit la distribution statistique (pourvu quelle admette une valeur moyenne et un écart quadratique moyen), la probabilité pour observer un écart $(A - \overline{A})$ supérieur, en valeur absolue, à λ fois ΔA (où $\lambda > 1$) est inférieure à $1/\lambda^2$ *(inégalité de Tchébycheff) :*

$$P(|A - \overline{A}| \geqslant \lambda \Delta A) \leqslant \frac{1}{\lambda^2} . \qquad \text{(a)}$$

En effet, raisonnons par exemple sur une distribution continue de probabilité (le cas discret se traite de façon analogue). Dans l'expression (30.b) de $(\Delta A)^2$, la fonction à intégrer est constamment positive, de sorte que

$$(\Delta A)^2 \geqslant \int_{|A-\overline{A}| \geqslant \lambda \Delta A} W(A)\left[A - \overline{A}\right]^2 \mathrm{d}A \geqslant \lambda^2 (\Delta A)^2 \int_{|A-\overline{A}| \geqslant \lambda \Delta A} W(A)\,\mathrm{d}A . \qquad \text{(b)}$$

L'intégrale qui subsiste au second membre est par définition la probabilité $P(|A - \overline{A}| \geqslant \lambda \Delta A)$; l'inégalité de Tchébycheff en découle aussitôt.

14. On appelle *moment d'ordre p* de la distribution statistique d'une grandeur A la valeur moyenne de A^p (si elle existe) avec *p entier positif.*

Si l'ensemble des valeurs possibles de A est discret et fini :

$$\{A_j ;\ j = 1, 2, \ldots, J\} \qquad \text{(a)}$$

leur distribution statistique $\{P(A_j)\}$ est complètement déterminée par la donnée des $(J - 1)$ premiers moments. En effet, l'ensemble des J égalités (cf. formule (28.a))

$$\sum_{j=1}^{J} P(A_j) = 1 ,$$

$$\sum_{j=1}^{J} A_j^p P(A_j) = \overline{A^p} , \quad p = 1, 2, \ldots, J-1 , \qquad \text{(b)}$$

peut être traité, si les moments $\overline{A^p}$ sont donnés, comme un système de J équations linéaires à J inconnues, les $P(A_j)$.

Lorsque l'ensemble des valeurs possibles est infini (discret ou continu), on peut montrer que, si certaines conditions de convergence sont vérifiées, leur distribution statistique $\{P(A_i)\}$ ou $W(A)$ est déterminée par la donnée de l'ensemble des moments $\overline{A^p}$, pour p prenant toutes les valeurs entières positives.

probabilités ne seront pas supposées nécessairement égales, mais elles vérifient de toute façon la condition de normalisation :

$$P_+ + P_- = 1 \,. \tag{32}$$

Supposons alors qu'on lance N fois la pièce. Quelle est la probabilité $P(N, n)$ pour obtenir n fois pile au cours de ces N essais?

De façon plus générale, considérons un problème quelconque de probabilités. Choisissons un des événements possibles, que nous noterons (+), sa probabilité étant P_+. Si nous désignons par (−) l'événement contraire, nous sommes ramenés à la situation de pile ou face, et nous posons la même question : *quelle est la probabilité* $P(N, n)$ (où $0 \leqslant n \leqslant N$) *pour que, au cours de N expériences, l'événement* (+) *choisi se réalise n fois?*

b. Solution

L'événement (+) se sera produit n fois, et l'événement (−) $(N-n)$ fois, au cours des N épreuves si, par exemple, on observe la succession d'événements suivante :

$$\underbrace{(+), (+), \ldots, (+)}_{n}, \underbrace{(-), (-), \ldots, (-)}_{N-n} \,. \tag{33}$$

La probabilité de cette séquence particulière s_1 d'événements indépendants s'obtient par la « loi de multiplication » (15) :

$$P(s_1) = P_+^n P_-^{N-n} \,. \tag{34}$$

Mais il existe évidemment d'autres séquences d'événements (+) et (−) dans lesquelles (+) figure n fois et (−) $(N-n)$ fois. La probabilité d'une séquence quelconque s_i comportant n événements (+) et $(N-n)$ événements (−) vaut, quel que soit l'ordre dans lequel ces événements se succèdent,

$$P(s_i) = P_+^n P_-^{N-n} \,. \tag{35}$$

Deux séquences différentes s'excluant nécessairement, la « loi d'addition » (10) fournit la probabilité $P(N, n)$ cherchée comme somme des probabilités $P(s_i)$ de toutes les séquences réalisant n événements (+) parmi les N. Comme ces séquences ont la même probabilité (35), il suffit de déterminer leur nombre. C'est le nombre de manières distinctes de placer les n événements (+) parmi les N événements que comporte chaque séquence. Ceci est un problème classique d'analyse combinatoire : le nombre cherché est celui des combinaisons de N éléments pris n à n, c'est-à-dire

$$C_N^n = \frac{N!}{n!(N-n)!} \,. \tag{36}$$

La probabilité $P(N, n)$ est donc finalement

$$P(N, n) = \sum_i P(s_i) = C_N^n P(s_i) \,, \tag{37}$$

soit, d'après (35),

$$P(N, n) = C_N^n P_+^n P_-^{N-n} = \frac{N!}{n!(N-n)!} P_+^n P_-^{N-n} \,. \tag{38}$$

La distribution de probabilité caractérisée par cette expression est appelée

« binomiale » car les nombres $P(N, n)$ coïncident avec les différents termes du développement de $(P_+ + P_-)^N$ par la formule du binôme :

$$(P_+ + P_-)^N = \sum_{n=0}^{N} C_N^n P_+^n P_-^{N-n} . \tag{39}$$

Cette formule assure la normalisation des probabilités $P(N, n)$ pour N fixé :

$$\sum_{n=0}^{N} P(N, n) = 1 . \tag{40}$$

En effet, le premier membre de (39) est égal à 1 par normalisation des probabilités élémentaires P_+ et P_-.

c. Caractéristiques de la distribution binomiale

Nous allons calculer la valeur moyenne et l'écart quadratique moyen du nombre n d'événements (+) pour *N fixé*. Nous discuterons ensuite les résultats obtenus et leur dépendance en N.

α. Valeur moyenne

La valeur moyenne $\bar{n}$ du nombre n s'écrit

$$\bar{n} = \sum_{n=0}^{N} nP(N, n) = \sum_{n=0}^{N} nC_N^n P_+^n P_-^{N-n} . \tag{41}$$

Cette somme se calcule facilement, si on remarque que

$$\frac{\partial}{\partial P_+}\left[\sum_{n=0}^{N} C_N^n P_+^n P_-^{N-n}\right] = \sum_{n=0}^{N} C_N^n n P_+^{n-1} P_-^{N-n} . \tag{42}$$

Une multiplication par P_+ rétablit ensuite la puissance initiale, de sorte que

$$\bar{n} = P_+ \frac{\partial}{\partial P_+}\left[\sum_{n=0}^{N} C_N^n P_+^n P_-^{N-n}\right] = P_+ \frac{\partial}{\partial P_+}(P_+ + P_-)^N . \tag{43}$$

D'où, grâce à la normalisation des probabilités P_+ et P_-[15],

$$\bar{n} = P_+ N(P_+ + P_-)^{N-1} = NP_+ . \tag{44}$$

Ce résultat n'est pas surprenant : la valeur moyenne du nombre d'événements (+) est le produit, par le nombre total N d'épreuves, de la probabilité P_+ de l'événement (+) dans chaque épreuve.

β. Écart quadratique moyen

Pour obtenir l'écart quadratique moyen Δn, il suffit de calculer la valeur moyenne de n^2 (voir formule (31)) :

$$\overline{n^2} = \sum_{n=0}^{N} n^2 C_N^n P_+^n P_-^{N-n} . \tag{45}$$

15. Dans ce calcul, l'astuce consiste à traiter, à partir de (42), les probabilités P_+ et P_- comme des variables indépendantes. C'est ce qui permet de « faire descendre » juste le facteur n désiré, en dérivant par rapport à P_+. Ce n'est qu'à la fin du calcul qu'on « se rappelle » que P_+ et P_- ont en fait pour somme 1.

Répétons le procédé utilisé au paragraphe α[16] :

$$\overline{n^2} = \left(P_+ \frac{\partial}{\partial P_+}\right)^2 \left[\sum_{n=0}^{N} C_N^n P_+^n P_-^{N-n}\right] = \left(P_+ \frac{\partial}{\partial P_+}\right)^2 \left[(P_+ + P_-)^N\right] . \tag{46}$$

On trouve ainsi aisément

$$\overline{n^2} = NP_+ + N(N-1)P_+^2 , \tag{47}$$

d'où finalement

$$(\Delta n)^2 = \overline{n^2} - \bar{n}^2 = NP_+(1-P_+) . \tag{48}$$

γ. Interprétation des valeurs obtenues

Il est important de bien comprendre la signification des résultats précédents.

Puisqu'on parle de probabilités et de valeurs moyennes pour N fixé, c'est que l'on envisage de répéter un grand nombre de fois $\mathcal{N}$ la série de N expériences. Pour pile ou face, les N expériences peuvent être réalisées simultanément, si l'on dispose de N pièces identiques; on lancera donc $\mathcal{N}$ fois l'ensemble des N pièces, N pouvant être quelconque mais $\mathcal{N}$ étant très grand dans tous les cas; dans chacune de ces $\mathcal{N}$ expériences composées, on comptera le nombre n de pièces, parmi les N, qui montrent pile.

Ce nombre n varie aléatoirement d'une expérience à l'autre; il est tantôt supérieur à $\bar{n}$, tantôt inférieur, et certaines fois égal à $\bar{n}$. L'écart par rapport à la moyenne $|n-\bar{n}|$ peut être quelconque (entre 0 et une valeur maximale, qui est suivant les cas NP_+ ou NP_-); mais, si $\mathcal{N}$ est suffisamment grand, cet écart est la plupart du temps inférieur à Δn (ou à $2\Delta n$) :

$$|n-\bar{n}| \leqslant \Delta n \tag{49}$$

(pour une analyse plus précise, voir complément I.C, § III).

Précisons ce qui se passe lorsque N *croît*. D'après la formule (44), $\bar{n}$ *croît proportionnellement à* N, ce qui est parfaitement compréhensible. Mais la formule (48) indique que Δn *croît lui aussi;* en conséquence, *les écarts par rapport à la moyenne croissent,* en valeur absolue : il ne faut pas croire que, si on lance $N = 1\,000$ pièces à la fois, on a plus de chances d'obtenir un nombre n de piles voisin de $N/2$ (on suppose $P_+ = 1/2$) que si on en lance seulement 100; c'est le contraire qui est vrai. *Cependant,* on voit sur (48) que Δn *croît seulement comme* $\sqrt{N}$, de telle sorte que l'écart par rapport à la moyenne *décroît en valeur relative :*

$$\frac{\Delta n}{\bar{n}} \sim \frac{\sqrt{N}}{N} . \tag{50}$$

C'est ainsi que, dans la limite où N devient très grand, la proportion n/N de pièces donnant pile s'approche de la probabilité P_+.

Ces remarques, faciles à formuler de façon concrète dans le cas du jeu de pile ou face, sont en fait générales, puisque la distribution binomiale intervient dans tous les problèmes de probabilités (voir paragraphe a).

---

16. Rappelons que le carré d'un opérateur est défini comme sa répétition deux fois de suite :

$$\left(P_+ \frac{\partial}{\partial P_+}\right)^2 \widehat{=} \left(P_+ \frac{\partial}{\partial P_+}\right)\left(P_+ \frac{\partial}{\partial P_+}\right) .$$

2. LA DISTRIBUTION GAUSSIENNE

a. Définition

La distribution gaussienne, ou normale, est une distribution continue dont la densité de probabilité s'écrit (fig. 1) :

$$w_G(x) = \frac{1}{\sqrt{2\pi\sigma^2}}\, e^{-(x-x_0)^2/2\sigma^2} . \tag{51}$$

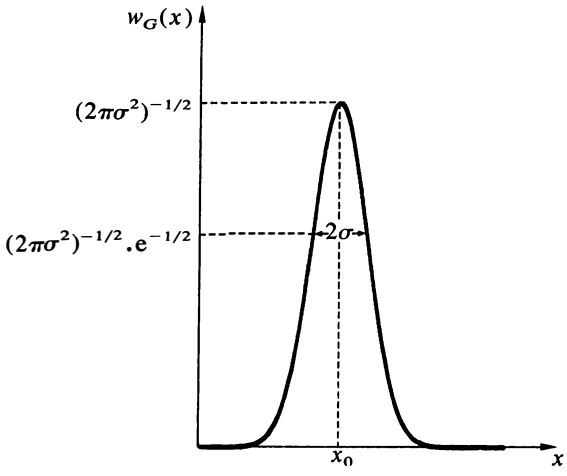


FIGURE 1
La « courbe de Gauss » représentant la densité de probabilité (51). La signification des paramètres x_0 et σ y est illustrée graphiquement.

Elle dépend de *deux paramètres :* x_0 caractérise sa *position* sur l'axe des x (on dit aussi son *« centre »*), et σ sa *largeur* (on prendra $\sigma > 0$). Le facteur $1/\sqrt{2\pi\sigma^2}$ assure la normalisation de la distribution[17] :

$$\int_{-\infty}^{+\infty} w_G(x)\,dx = 1 . \tag{52}$$

17. On sait (voir formulaire) que

$$\int_{-\infty}^{+\infty} e^{-\alpha^2 y^2}\,dy = \frac{1}{\alpha}\sqrt{\pi} \quad (\alpha > 0) . \tag{a}$$

Pour calculer l'intégrale

$$\int_{-\infty}^{+\infty} w_G(x)\,dx = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{+\infty} e^{-(x-x_0)^2/2\sigma^2}\,dx , \tag{b}$$

il suffit donc d'effectuer le changement de variable

$$y = x - x_0 . \tag{c}$$

On vérifie ainsi facilement la condition de normalisation (52).

b. Caractéristiques

On voit immédiatement, en dérivant la formule de définition (51), que la *valeur la plus probable* x_m de la variable est

$$x_m = x_0 \,. \tag{53}$$

Pour calculer la *valeur moyenne* $\bar{x}$ de la variable x, il suffit d'appliquer la définition (19.b), qui s'écrit ici

$$\bar{x} = \int_{-\infty}^{+\infty} x\, w_G(x)\, \mathrm{d}x \,. \tag{54}$$

On trouve facilement[18]

$$\bar{x} = x_0 \,. \tag{55}$$

Dans toute distribution gaussienne, valeur moyenne et valeur la plus probable coïncident[19].

Quant à l'*écart quadratique moyen,* il vaut ici tout simplement[20]

$$\Delta x = \sigma \,. \tag{56}$$

Le paramètre de largeur σ de la distribution gaussienne est donc précisément égal à son écart quadratique moyen.

c. La distribution gaussienne comme limite

Nous avons discuté plus haut (§ 1.c.γ) le comportement de la valeur moyenne $\bar{n}$ et de l'écart quadratique moyen Δn du nombre n d'états (+) pour une distribution binomiale dans laquelle N devient grand. On peut être plus précis, et étudier dans cette situation les probabilités $P(N, n)$ elles-mêmes. On trouve alors le résultat suivant (complément I.C) : *pour N suffisamment grand, la distribution binomiale se confond pratiquement avec une distribution gaussienne,* dont le centre et la largeur coïncident respectivement avec la valeur la plus probable et l'écart quadratique moyen de la distribution binomiale.

---

18. Pour calculer

$$\bar{x} = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{+\infty} x\, \mathrm{e}^{-(x-x_0)^2/2\sigma^2}\, \mathrm{d}x \,, \tag{a}$$

on peut l'écrire sous la forme

$$\bar{x} = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{+\infty} (x - x_0) \mathrm{e}^{-(x-x_0)^2/2\sigma^2}\, \mathrm{d}x + \frac{x_0}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{+\infty} \mathrm{e}^{-(x-x_0)^2/2\sigma^2}\, \mathrm{d}x \,. \tag{b}$$

La première intégrale est alors nulle car la fonction à intégrer est impaire dans la variable $y = x - x_0$; la seconde se calcule aussitôt grâce à la condition de normalisation, et donne

$$\bar{x} = x_0 \,. \tag{c}$$

19. Pour toute distribution continue ayant un maximum unique et symétrique par rapport à ce maximum, la valeur moyenne est égale à la valeur la plus probable.

20. La distribution étant symétrique autour de la valeur moyenne $\bar{x} = x_0$, l'écart quadratique moyen est facile à évaluer directement à partir de sa définition :

$$(\Delta x)^2 = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{+\infty} (x - x_0)^2\, \mathrm{e}^{-(x-x_0)^2/2\sigma^2}\, \mathrm{d}x = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{+\infty} y^2\, \mathrm{e}^{-y^2/2\sigma^2}\, \mathrm{d}y = \sigma^2$$

(voir formulaire).

Ce résultat n'est en réalité qu'un cas particulier d'une propriété générale, qui s'exprime mathématiquement par le *théorème de la limite centrale* (exercice I.12) : si l'on considère la somme de N variables aléatoires indépendantes, sa distribution statistique devient gaussienne lorsque N devient grand, quelle que soit la distribution statistique de chacune des N variables.

La distribution gaussienne intervient de façon quasi-universelle en mécanique statistique, pour les propriétés des *systèmes macroscopiques.* Ceux-ci sont en effet constitués le plus souvent d'un *grand nombre d'éléments faiblement couplés entre eux.* Les grandeurs physiques qui sont libres de fluctuer (c'est-à-dire celles qui ne sont pas fixées de l'extérieur à une valeur certaine) ont alors une *distribution statistique pratiquement gaussienne* (voir notamment II, § II.A.2 et III, § I.c).

Références

1. J. Bass, *Éléments de calcul des probabilités,* Masson, Paris (1974).
2. V.S. Pugachev, *Théorie des probabilités et statistique mathématique,* Éditions de Moscou (1982).
3. A. Feller, *An Introduction to Probability Theory and Its Applications,* John Wiley, New York (1950).

APPENDICE II

Description quantique d'un système physique

I. Résumé des postulats de la mécanique quantique
   1. Description de l'état d'un système physique
   2. Description des grandeurs physiques
   3. Mesure d'une grandeur physique
      a. Résultats possibles de la mesure
      b. Principe de décomposition spectrale
      c. Réduction du paquet d'ondes
   4. Évolution des systèmes dans le temps

II. Cas particulier important : les systèmes conservatifs
   1. Définition
   2. Solution de l'équation de Schrödinger
      a. Niveaux d'énergie et états stationnaires
      b. État quelconque
   3. Exemple : l'oscillateur harmonique à une dimension
   4. Exemple : la particule libre enfermée dans une « boîte »
      a. Équation aux valeurs propres de l'hamiltonien
      b. Solution pour des conditions aux limites strictes
      c. Solution pour des conditions aux limites périodiques

III. Construction d'espaces d'états par produit tensoriel
   1. Définition du produit tensoriel
   2. Exemples importants
      a. États d'un système de plusieurs particules
      b. États d'une particule de spin 1/2
      c. États d'un système physique divisé en deux sous-systèmes indépendants

IV. Systèmes de particules identiques
   1. Le problème des particules identiques en mécanique quantique
   2. Le postulat de symétrisation
   3. États stationnaires d'un système de N particules identiques et indépendantes

A l'échelle microscopique, les propriétés et l'évolution des systèmes physiques sont régies par la mécanique quantique. C'est donc à partir de leur description quantique que la mécanique statistique détermine les caractéristiques macroscopiques des systèmes constitués d'un grand nombre de particules.

La présentation de la mécanique statistique que nous développons dans cet ouvrage ne nécessite qu'une connaissance très superficielle de la mécanique quantique; nous nous contenterons ici de donner les quelques résultats nécessaires à cette étude élémentaire. Seuls certains compléments et certaines notes font appel à des notions plus élaborées; nous indiquons alors une référence précise dans C. Cohen-Tannoudji, B. Diu et F. Laloë, *Mécanique Quantique,* Hermann, Paris (1985), désigné par le sigle C.D.L.

I. RÉSUMÉ DES POSTULATS DE LA MÉCANIQUE QUANTIQUE[1]

1. DESCRIPTION DE L'ÉTAT D'UN SYSTÈME PHYSIQUE

Premier postulat : l'état d'un système physique à un instant t_0 fixé est défini par la donnée d'un *vecteur d'état,* ou *ket,* $|\psi(t_0)\rangle$ appartenant à l'espace des états $\mathcal{E}$ de ce système[2].

1. Pour plus de détails, voir C.D.L., chapitre III.

2. En réalité, deux kets $|\psi\rangle$ et $|\psi'\rangle$ colinéaires représentent le même état physique (C.D.L., p. 219).

Pour un système constitué d'une seule particule sans spin, on peut de façon équivalente[3] caractériser son état à l'instant t_0 par une *fonction d'onde* $\psi(\vec{r}, t_0)$ à valeurs complexes, de carré sommable. Rappelons l'*interprétation probabiliste* de cette fonction d'onde : si on la normalise de telle sorte que

$$\int d^3r \, |\psi(\vec{r}, t_0)|^2 = 1 , \tag{1}$$

alors $|\psi(\vec{r}, t_0)|^2$ est la densité de probabilité pour que la particule considérée se trouve au point $\vec{r}$ à l'instant t_0.

De façon générale, l'espace des états est un *espace vectoriel* (sur le corps des complexes), muni d'un *produit scalaire.* Dans la *notation de Dirac,* le produit scalaire de deux kets $|\varphi\rangle$ et $|\psi\rangle$ s'écrit

$$\langle \varphi \mid \psi \rangle , \tag{2}$$

où $\langle \varphi|$ est le *«bra conjugué»* du ket $|\varphi\rangle$ (C.D.L., p. 110). Ce produit scalaire définit une *norme :*

$$\langle \psi | \psi \rangle \geqslant 0 \quad \text{pour tout } |\psi\rangle, \tag{3.a}$$

$$\langle \psi | \psi \rangle = 0 \implies |\psi\rangle = 0 . \tag{3.b}$$

Dans le cas d'une seule particule sans spin, le produit scalaire de deux fonctions d'onde $\varphi(\vec{r})$ et $\psi(\vec{r})$ est défini par

$$\langle \varphi | \psi \rangle \widehat{=} \int d^3r \, \varphi^*(\vec{r}) \psi(\vec{r}) \tag{4}$$

(où l'étoile note la conjugaison complexe).

2. DESCRIPTION DES GRANDEURS PHYSIQUES

Deuxième postulat : toute grandeur physique mesurable $\mathcal{A}$ est décrite par un *opérateur linéaire A* agissant dans l'espace des états; cet opérateur est une *observable*[4].

Citons comme exemples[5] les observables $\vec{R}$ (X, Y, Z) décrivant la *position* d'une particule et les observables $\vec{P}$ (P_x, P_y, P_z) décrivant son *impulsion :* dans la représentation des fonctions d'onde, ces opérateurs sont définis par

$$X\psi(\vec{r}) \equiv X\psi(x, y, z) \widehat{=} x\psi(x, y, z) , \tag{5.a}$$

$$P_x\psi(\vec{r}) \widehat{=} \frac{\hbar}{i} \frac{\partial}{\partial x} \psi(\vec{r}) , \tag{5.b}$$

et des égalités analogues pour les autres composantes (C.D.L., p. 149).

---

3. Le formalisme des fonctions d'onde («mécanique ondulatoire») apparaît comme un cas particulier du formalisme plus général des kets et des bras (C.D.L., p. 144).

4. Une *observable* est par définition un opérateur linéaire *hermitique* (C.D.L., p. 120) dont les vecteurs propres constituent une *base de l'espace des états* (C.D.L., p. 137).

5. Certaines grandeurs (spin, par exemple) n'existent que dans le cadre quantique; elles sont alors directement définies par la donnée des observables correspondantes.

Pour les grandeurs qui existent déjà en mécanique classique, l'observable associée découle de *«règles de quantification»* que l'on peut résumer comme suit (C.D.L., p. 222). En mécanique classique (non relativiste), toutes les grandeurs physiques relatives à une particule sont des fonctions de sa position $\vec{r}$ et de son impulsion $\vec{p}$ (appendice III) :

$$\mathcal{A}(\vec{r}, \vec{p}) . \tag{a}$$

Il suffit donc, pour passer à la description quantique, de se donner les observables de position $\vec{R}$ et d'impulsion $\vec{P}$: l'observable A associée à la grandeur $\mathcal{A}$ s'otient alors en écrivant

$$A = \mathcal{A}(\vec{R}, \vec{P}) . \tag{b}$$

3. MESURE D'UNE GRANDEUR PHYSIQUE

a. Résultats possibles de la mesure

Troisième postulat : la mesure d'une grandeur physique $\mathcal{A}$ ne peut donner comme résultat qu'*une des valeurs propres*[6] de l'observable A correspondante.

b. Principe de décomposition spectrale

Quatrième postulat : lorsqu'on mesure la grandeur physique $\mathcal{A}$ sur un système dans l'état $|\psi\rangle$ (normé), la probabilité $P(a_n)$ d'obtenir comme résultat la valeur propre a_n de l'observable A correspondante vaut

$$P(a_n)=\sum_{i=1}^{g_n} |\langle u_n^i|\psi\rangle|^2 , \tag{6}$$

où g_n est le degré de dégénérescence[6] de a_n, et $\{|u_n^i\rangle;\ i=1, 2, \ldots, g_n\}$ un système orthonormé de vecteurs formant une base dans le sous-espace propre $\mathcal{E}_n$ associé à la valeur propre a_n de A.

Lorsque le système est dans l'état $|\psi\rangle$ et qu'on envisage de mesurer la grandeur $\mathcal{A}$, la *valeur moyenne des résultats possibles,* notée $\langle A\rangle_\psi$ est donnée par (C.D.L., p. 227)

$$\langle A\rangle_\psi \widehat{=} \sum_n a_n P(a_n)=\langle\psi|A|\psi\rangle . \tag{7}$$

c. Réduction du paquet d'ondes

Cinquième postulat : si la mesure de la grandeur physique $\mathcal{A}$ sur le système dans l'état $|\psi\rangle$ donne le résultat a_n, l'état du système immédiatement après la mesure est la projection (normée) du ket $|\psi\rangle$ sur le sous-espace propre $\mathcal{E}_n$ associé à a_n.

4. ÉVOLUTION DES SYSTÈMES DANS LE TEMPS

Sixième postulat : l'évolution dans le temps du vecteur d'état $|\psi(t)\rangle$ est régie par l'*équation de Schrödinger :*

$$i\hbar \frac{\mathrm{d}}{\mathrm{d}t} |\psi(t)\rangle = H|\psi(t)\rangle , \tag{8}$$

où H, appelé *«hamiltonien»,* est l'observable associée à l'*énergie totale* du système.

Pour une particule sans spin, non relativiste, de masse m, soumise à des forces dérivant de l'énergie potentielle $U(\vec{r})$, l'équation de Schrödinger donnant l'évolution dans le temps de sa fonction d'onde $\psi(\vec{r}, t)$, s'écrit

$$i\hbar \frac{\partial}{\partial t} \psi(\vec{r}, t)= -\frac{\hbar^2}{2m} \Delta\psi(\vec{r}, t)+ U(\vec{r})\psi(\vec{r}, t) , \tag{9}$$

où Δ est l'opérateur laplacien.

II. CAS PARTICULIER IMPORTANT : LES SYSTÈMES CONSERVATIFS

1. DÉFINITION

Un système est dit «conservatif» lorsque *son hamiltonien H,* observable associée à l'énergie totale, *est indépendant du temps.*

6. Pour une définition des valeurs propres d'un opérateur et de leur degré de dégénérescence, voir C.D.L., p. 132.

2. SOLUTION DE L'ÉQUATION DE SCHRÖDINGER

a. Niveaux d'énergie et états stationnaires

L'étude quantique d'un système conservatif se ramène à la *détermination des valeurs propres E_ℓ de l'hamiltonien H et des vecteurs propres $|\varphi_\ell^i\rangle$ correspondants :*

$$H|\varphi_\ell^i\rangle = E_\ell|\varphi_\ell^i\rangle \; ; \quad \begin{matrix} i = 1, 2, \ldots, g_\ell \\ \ell = 1, 2 \ldots \end{matrix} \tag{10}$$

Les *valeurs propres E_ℓ* sont les *énergies possibles* pour le système (cf. § I.3.a). A un niveau d'énergie E_ℓ non dégénéré (c'est-à-dire pour lequel $g_\ell = 1$) est associé un seul vecteur propre $|\varphi_\ell\rangle$ (à un facteur près); si E_ℓ est dégénérée ($g_\ell > 1$), on peut lui associer une infinité d'états propres distincts, décrits par toutes les combinaisons linéaires des g_ℓ kets orthogonaux $|\varphi_\ell^i\rangle$ ($i = 1, 2, \ldots, g_\ell$).

Si, à l'instant $t = 0$, le système se trouve dans un état décrit par l'un des *vecteurs propres de H :*

$$|\psi(t=0)\rangle = \sum_{i=1}^{g_{\ell_0}} c_i |\varphi_{\ell_0}^i\rangle \; , \tag{11}$$

le vecteur d'état à l'instant t s'écrit[7]

$$|\psi(t)\rangle = \mathrm{e}^{-iE_{\ell_0}t/\hbar}|\psi(t=0)\rangle \; . \tag{12}$$

Ce ket reste constamment colinéaire au ket initial (11); ceci implique que l'état physique correspondant reste inchangé (note 2). Par conséquent, *les propriétés physiques du système sont dans ce cas indépendantes du temps :* on dit que les états propres de l'hamiltonien sont des *états stationnaires*.

Pour une particule sans spin non relativiste d'énergie potentielle $U(\vec{r})$, l'équation (10), parfois appelée «équation de Schrödinger indépendante du temps», prend la forme d'une équation aux dérivées partielles :

$$-\frac{\hbar^2}{2m}\Delta\varphi_\ell^i(\vec{r}) + U(\vec{r})\varphi_\ell^i(\vec{r}) = E_\ell\varphi_\ell^i(\vec{r}) \; . \tag{13}$$

b. État quelconque

Si, à $t = 0$, l'état du système fait intervenir des vecteurs propres de H associés à *plusieurs énergies E_ℓ différentes :*

$$|\psi(0)\rangle = \sum_\ell \sum_{i=1}^{g_\ell} c_\ell^i(0)|\varphi_\ell^i\rangle \; , \tag{14}$$

son état à l'instant t est donné par[8]

$$|\psi(t)\rangle = \sum_\ell \mathrm{e}^{-iE_\ell t/\hbar} \sum_{i=1}^{g_\ell} c_\ell^i(0)|\varphi_\ell^i\rangle \; . \tag{15}$$

Dans ce cas, *les propriétés du système évoluent au cours du temps.*

---

7. On vérifie sans peine que ce ket est bien solution de l'équation de Schrödinger (8) :

$$\frac{\mathrm{d}}{\mathrm{d}t}|\psi(t)\rangle = -i\frac{E_{\ell_0}}{\hbar}\mathrm{e}^{-iE_{\ell_0}t/\hbar}|\psi(t=0)\rangle \; ,$$

$$H|\psi(t)\rangle = \mathrm{e}^{-iE_{\ell_0}t/\hbar}H|\psi(t=0)\rangle = E_{\ell_0}\mathrm{e}^{-iE_{\ell_0}t/\hbar}|\psi(t=0)\rangle \; .$$

8. Il est d'abord évident sur (15) que

$$|\psi(t=0)\rangle = |\psi(0)\rangle \; . \tag{a}$$

De plus, on a d'une part

$$i\hbar\frac{\mathrm{d}}{\mathrm{d}t}|\psi(t)\rangle = i\hbar\sum_\ell\left(-i\frac{E_\ell}{\hbar}\right)\mathrm{e}^{-iE_{\ell_0}t/\hbar}\sum_{i=1}^{g_\ell} c_\ell^i(0)|\varphi_\ell^i\rangle \; , \tag{b}$$

3. EXEMPLE : L'OSCILLATEUR HARMONIQUE A UNE DIMENSION

Une particule de masse m, assujettie à se déplacer le long d'un axe Ox, et soumise à une force de rappel proportionnelle à sa distance à un centre attracteur (le point O), constitue le prototype d'un oscillateur harmonique. L'énergie potentielle est alors de la forme

$$U(x)=\frac{1}{2}kx^2 \tag{16}$$

(où k est une constante positive), de sorte que l'opérateur hamiltonien s'écrit

$$H=\frac{P^2}{2m}+\frac{1}{2}m\omega^2X^2 . \tag{17}$$

Dans cette expression, ω est par définition la *pulsation de l'oscillateur :*

$$\omega \widehat{=} \sqrt{\frac{k}{m}} ; \tag{18}$$

X est l'observable de position de la particule sur l'axe Ox et P son impulsion conjuguée; le commutateur de ces deux observables vaut[9]

$$[X, P]=i\hbar . \tag{19}$$

De façon plus générale, tout système dont l'hamiltonien peut être mis sous la forme

$$H=aP^2+bQ^2 \tag{20}$$

(où a et b sont des constantes positives), en fonction d'observables Q et P vérifiant

$$[Q, P]=i\hbar , \tag{21}$$

constitue un oscillateur harmonique à une dimension.

On montre (C.D.L., p. 486) que les valeurs propres de l'hamiltonien (17) s'écrivent

$$E_\ell=\left(\ell+\frac{1}{2}\right)\hbar\omega ; \quad \ell=0, 1, 2, \ldots . \tag{22}$$

Le spectre des énergies possibles est donc *entièrement discret;* on montre de plus qu'il est *non dégénéré,* c'est-à-dire qu'à chaque niveau d'énergie E_ℓ correspond un état propre $|\varphi_\ell\rangle$ unique (à un facteur près)[10].

---

et d'autre part

$$H|\psi(t)\rangle=\sum_\ell e^{-iE_\ell t/\hbar}\sum_{i=1}^{g_\ell} c^i_\ell(0)\,H|\varphi^i_\ell\rangle=\sum_\ell e^{-iE_\ell t/\hbar}\,E_\ell\sum_{i=1}^{g_\ell} c^i_\ell(0)\,|\varphi^i_\ell\rangle . \tag{c}$$

Les seconds membres des égalités précédentes sont bien égaux entre eux; $|\psi(t)\rangle$ est bien solution de l'équation de Schrödinger (8).

9. Ce commutateur se calcule aisément dans la représentation des fonctions d'onde, à partir des formules (5).

10. L'étude des oscillateurs harmoniques à deux et trois dimensions se ramène à celle-ci.

Par exemple, un oscillateur harmonique à trois dimensions est caractérisé par un hamiltonien de la forme

$$H=\frac{1}{2m}[P_x^2+P_y^2+P_z^2]+\frac{1}{2}m(\omega_x^2X^2+\omega_y^2Y^2+\omega_z^2Z^2) , \tag{a}$$

.../...

4. EXEMPLE : LA PARTICULE LIBRE ENFERMÉE DANS UNE « BOITE »

Considérons maintenant une particule libre (c'est-à-dire qui n'est soumise à aucune force) astreinte à rester à l'intérieur d'une région finie de l'espace, que nous supposons pour simplifier parallélépipédique (« boîte »). Nous choisissons les axes de coordonnées Ox, Oy, Oz parallèles aux arêtes de la boîte, dont nous notons L_x, L_y, L_z les longueurs respectives.

a. Équation aux valeurs propres de l'hamiltonien

La particule étant *libre,* son hamiltonien H ne comporte que le terme d'énergie cinétique :

$$H = \vec{P}^2/2m \,, \tag{23}$$

où m est la masse de la particule et $\vec{P}$ l'opérateur décrivant son impulsion. Si cette particule n'a pas de spin, son état quantique peut être caractérisé par une *fonction d'onde* $\varphi(\vec{r})$. Les états stationnaires (ℓ) et les énergies E_ℓ associées sont alors donnés par l'équation aux valeurs propres

$$-\frac{\hbar^2}{2m}\,\Delta\varphi_\ell(\vec{r}) = E_\ell\,\varphi_\ell(\vec{r}) \,, \tag{24}$$

$\varphi_\ell(\vec{r})$ étant la fonction d'onde caractérisant l'état (ℓ).

On peut montrer facilement (C.D.L., p. 61) qu'on ne restreint pas la généralité du problème en cherchant les fonctions propres φ_ℓ sous la forme factorisée

$$\varphi_\ell(\vec{r}) = u(x).v(y).w(z) \,. \tag{25}$$

On a alors à résoudre trois équations à une dimension découplées,

$$-\frac{\hbar^2}{2m}\frac{\mathrm{d}^2u}{\mathrm{d}x^2} = E_x u(x) \,, \tag{26.a}$$

$$-\frac{\hbar^2}{2m}\frac{\mathrm{d}^2v}{\mathrm{d}y^2} = E_y v(y) \,, \tag{26.b}$$

$$-\frac{\hbar^2}{2m}\frac{\mathrm{d}^2w}{\mathrm{d}z^2} = E_z w(z) \,, \tag{26.c}$$

et les valeurs possibles de l'énergie E_ℓ sont données par

$$E_\ell = E_x + E_y + E_z \,, \tag{27}$$

où E_x, E_y, E_z sont les constantes apparaissant dans les équations (26).

---

avec $[X, P_x] = [Y, P_y] = [Z, P_z] = i\hbar$ (b)

(les observables X, Y, Z commutant entre elles, ainsi que P_x, P_y et P_z). L'expression de H est simplement la somme de trois oscillateurs harmoniques à une dimension :

$$H = H_x + H_y + H_z \,, \tag{c}$$

où $$H_x = \frac{1}{2m}P_x^2 + \frac{1}{2}\,m\omega_x^2 X^2 \tag{d}$$

et des formules analogues pour H_y et H_z. On en déduit les niveaux d'énergie de l'oscillateur harmonique à trois dimensions :

$$E_{\ell_x,\ell_y,\ell_z} = \left(\ell_x + \frac{1}{2}\right)\hbar\omega_x + \left(\ell_y + \frac{1}{2}\right)\hbar\omega_y + \left(\ell_z + \frac{1}{2}\right)\hbar\omega_z \;;\; \ell_x,\ \ell_y,\ \ell_z = 0,\ 1,\ 2,\ \ldots \,. \tag{e}$$

Si l'oscillateur à trois dimensions est *isotrope,* c'est-à-dire si

$$\omega_x = \omega_y = \omega_z = \omega \,, \tag{f}$$

.../...

b. Solution pour des conditions aux limites strictes

A l'intérieur de la boîte, c'est-à-dire pour

$$\begin{cases} 0 \leqslant x \leqslant L_x \\ 0 \leqslant y \leqslant L_y \\ 0 \leqslant z \leqslant L_z \end{cases}, \tag{28}$$

la particule est libre (cf. § a); dire qu'elle est enfermée dans la boîte signifie que les parois de celle-ci l'empêchent de sortir, et donc exercent sur elle une force aussi grande que nécessaire, mais localisée sur la frontière du domaine (28). On traite ce problème commodément en considérant que l'énergie potentielle de la particule, nulle à l'intérieur de la boîte, devient infinie sur les parois. Ceci implique (C.D.L., p. 72) que *la fonction d'onde* décrivant l'état de la particule *doit s'annuler sur les parois.* C'est ce qu'on appelle les *« conditions aux limites strictes »,* qui s'écrivent, en termes des fonctions u, v et w,

$$u(0) = u(L_x) = 0 , \tag{29.a}$$

$$v(0) = v(L_y) = 0 , \tag{29.b}$$

$$w(0) = w(L_z) = 0 . \tag{29.c}$$

Les équations (26) ne diffèrent l'une de l'autre que par le nom de la fonction et de la variable, et l'indice affecté à la constante; il en est de même des conditions (29). On a donc à résoudre trois fois le même problème, une fois sur chaque axe de coordonnées. Il s'agit d'un problème bien connu en mécanique quantique *(« puits infini à une dimension »),* dont la solution est très simple : par exemple, l'équation (26.a) et les conditions aux limites (29.a) impliquent[11]

$$u(x) = \sqrt{\frac{2}{L_x}} \sin\left(\ell_x \pi \frac{x}{L_x}\right) , \tag{30.a}$$

$$E_x = \ell_x^2 \frac{\pi^2 \hbar^2}{2mL_x^2} , \tag{30.b}$$

---

ses niveaux d'énergie sont repérés par un seul entier :

$$E_\ell = \left(\ell + \frac{3}{2}\right)\hbar\omega , \quad \ell = 0, 1, 2, \ldots . \tag{g}$$

Mais ils sont *dégénérés :* pour ℓ fixé, il existe plusieurs choix possibles de ℓ_x, ℓ_y, ℓ_z tels que

$$\ell_x + \ell_y + \ell_z = \ell ; \tag{h}$$

on trouve facilement (C.D.L., p. 550) le degré de dégénérescence g_ℓ correspondant :

$$g_\ell = \frac{1}{2}(\ell + 1)(\ell + 2) \tag{i}$$

(seul le niveau fondamental $E_0 = 3\hbar\omega/2$ est donc non dégénéré).

11. Voir C.D.L., p. 79.

Dans la formule (30.a), le facteur $\sqrt{2/L_x}$ assure simplement que la fonction est normalisée à 1 :

$$\int_0^{L_x} |u(x)|^2 \, dx = 1 .$$

Le paramètre ℓ_x, qui doit être entier pour satisfaire (29.a), est restreint aux valeurs strictement positives : changer ℓ_x de signe a pour seul effet de multiplier $u(x)$ par -1, et deux fonctions d'onde proportionnelles sont physiquement équivalentes (cf. note 2); $\ell_x = 0$ donne une fonction identiquement nulle.

où ℓ_x *est un entier positif* quelconque :

$$\ell_x = 1,\ 2,\ 3,\ \ldots \tag{31}$$

Si nous revenons maintenant aux fonctions $\varphi_\ell(\vec{r})$, nous voyons qu'il faut en réalité *trois indices* pour repérer une de ces fonctions propres, c'est-à-dire l'état (ℓ) correspondant :

$$(\ell) \equiv (\ell_x,\ \ell_y,\ \ell_z)\ . \tag{32}$$

Les fonctions d'onde stationnaires s'écrivent donc

$$\varphi_{\ell_x,\ \ell_y,\ \ell_z}(\vec{r}) = \frac{2\sqrt{2}}{\sqrt{V}} \sin\left(\ell_x \frac{\pi x}{L_x}\right).\sin\left(\ell_y \frac{\pi y}{L_y}\right).\sin\left(\ell_z \frac{\pi z}{L_z}\right) \tag{33}$$

(où $V = L_x L_y L_z$ est le volume de la boîte), l'énergie associée étant

$$E_{\ell_x,\ \ell_y,\ \ell_z} = \frac{\pi^2\hbar^2}{2m}\left[\frac{\ell_x^2}{L_x^2} + \frac{\ell_y^2}{L_y^2} + \frac{\ell_z^2}{L_z^2}\right]. \tag{34}$$

Les valeurs permises, pour chacun des paramètres ℓ_x, ℓ_y, ℓ_z sont les *nombres entiers positifs*.

c. Solution pour des conditions aux limites périodiques

α. Les conditions aux limites périodiques

Les boîtes que nous avons à considérer dans cet ouvrage sont de *taille macroscopique :* nous voulons appliquer les résultats résumés ici à des molécules de gaz contenues dans un récipient, à des électrons enfermés dans un cristal... Par suite, les conditions aux limites strictes que nous avons utilisées au paragraphe b (formules (29)) ne représentent la réalité que de façon très grossière : le comportement des fonctions d'onde au voisinage des parois de la boîte dépend de la forme exacte de ces parois à l'échelle microscopique (elles sont loin d'être planes à cette échelle, si bien polies soient-elles), de leur position précise (la valeur de L_x, L_y ou L_z est entachée d'une incertitude non négligeable au niveau microscopique), et même de leur nature (la particule considérée interagit avec les atomes constituant les parois). Donc, bien que ces conditions aux limites strictes aient paru naturelles quand nous les avons posées, il faut admettre qu'elles sont en fait peu réalistes.

Il est impossible, on le comprend bien, d'avoir accès aux véritables conditions aux limites sur les parois de la boîte. Mais ce n'est pas indispensable : comme toujours en mécanique statistique (voir I, § I.B.2), le comportement microscopique du système nous est inaccessible, mais sa connaissance détaillée n'est pas nécessaire pour comprendre ses propriétés macroscopiques (elle serait même inexploitable étant donné la quantité faramineuse d'information qu'elle apporterait). Ici, c'est principalement la *densité d'états* microscopiques (I, § III.B) qui nous intéresse. Son calcul exige qu'on impose des conditions aux limites, même grossières, puisque ce sont elles qui rendent discret le spectre d'énergie et permettent par suite de compter les états stationnaires. Mais le résultat ne dépend pas de la nature précise de ces conditions (I, § III.B.2.a). Il ne dépend d'ailleurs pas plus de la forme géométrique de la boîte, pourvu que ses dimensions soient toutes très grandes devant la longueur d'onde de de Broglie[12] associée au mouvement de la particule étudiée.

---

12. Rappelons (C.D.L., p. 43) que la longueur d'onde de de Broglie

$$\lambda_B \widehat{=} \frac{h}{p}\ ,$$

où p est l'impulsion de la particule, donne l'ordre de grandeur des distances sur lesquelles se font sentir habituellement les propriétés ondulatoires (quantiques) de la particule.

Les formules du paragraphe b, obtenues avec des conditions aux limites strictes, présentent l'inconvénient de faire intervenir des fonctions sinusoïdales, moins commodes à manipuler que les exponentielles imaginaires que l'on utilise habituellement pour une particule libre car ce sont des *fonctions d'onde d'impulsion bien déterminée.* La liberté de choix concernant les conditions aux limites permet de remédier à cet inconvénient et de retrouver les ondes planes habituelles; il suffit de remplacer les conditions strictes du paragraphe b par des *conditions aux limites périodiques :* on impose aux fonctions d'onde stationnaires de reprendre pour $x = L_x$ la même valeur que pour $x = 0$, et des conditions analogues sur les deux autres axes. Pour les fonctions u, v, w de la formule (25), on posera donc

$$u(L_x) = u(0) , \tag{35.a}$$

$$v(L_y) = v(0) , \tag{35.b}$$

$$w(L_z) = w(0) . \tag{35.c}$$

β. Fonctions d'onde stationnaires

Les solutions de l'équation (26.a), par exemple, peuvent maintenant être prises sous la forme[13]

$$u(x) = \frac{1}{\sqrt{L_x}} \exp(iK_x x) , \tag{36.a}$$

la constante E_x associée ayant pour expression

$$E_x = \frac{\hbar^2 K_x^2}{2m} . \tag{36.b}$$

La condition aux limites (35.a) impose ici

$$K_x = \ell'_x \cdot \frac{2\pi}{L_x} , \tag{37}$$

avec ℓ'_x *entier positif, négatif ou nul*[14] :

$$\ell'_x = \ldots, -1, 0, 1, 2, \ldots . \tag{38}$$

Les nouvelles fonctions d'onde stationnaires s'écrivent donc

$$\varphi'_{K_x, K_y, K_z} = \frac{1}{\sqrt{V}} e^{iK_x x} e^{iK_y y} e^{iK_z z} = \frac{1}{\sqrt{V}} e^{i\vec{K}\cdot\vec{r}} , \tag{39}$$

et les énergies associées

$$E'_{K_x, K_y, K_z} = \frac{\hbar^2}{2m} (K_x^2 + K_y^2 + K_z^2) = \frac{\hbar^2 \vec{K}^2}{2m} , \tag{40}$$

le vecteur $\vec{K}$ étant défini par

$$\vec{K} \widehat{=} \ell'_x \cdot \frac{2\pi}{L_x} \vec{e}_x + \ell'_y \cdot \frac{2\pi}{L_y} \vec{e}_y + \ell'_z \frac{2\pi}{L_z} \vec{e}_z . \tag{41}$$

Les paramètres ℓ'_x, ℓ'_y, ℓ'_z prennent ici toutes les valeurs entières relatives.

La forme (39) des fonctions d'onde stationnaires permet maintenant une interprétation physique simple du vecteur $\vec{K}$: $\hbar\vec{K}$ *est l'impulsion de la particule.*

---

13. Le facteur $1/\sqrt{L_x}$ assure la normalisation de $u(x)$ sur l'intervalle $[0, L_x]$, comme dans la note 11.

14. $\ell'_x = 0$ donne ici une fonction d'onde acceptable (puisque non identiquement nulle), et des valeurs opposées de ℓ'_x caractérisent deux fonctions linéairement indépendantes.

III. CONSTRUCTION D'ESPACES D'ÉTATS PAR PRODUIT TENSORIEL

1. DÉFINITION DU PRODUIT TENSORIEL

Considérons deux systèmes physiques distincts $\mathcal{S}_1$ et $\mathcal{S}_2$, et soit $\mathcal{E}_1$ et $\mathcal{E}_2$ les espaces d'états correspondants; de façon générale, nous affecterons le numéro 1 à tout ce qui a trait au système $\mathcal{S}_1$ et le numéro 2 au système $\mathcal{S}_2$. Si l'on réunit $\mathcal{S}_1$ et $\mathcal{S}_2$ en un système physique unique $\mathcal{S}$, l'espace des états $\mathcal{E}$ de $\mathcal{S}$ est le *produit tensoriel des espaces* $\mathcal{E}_1$ *et* $\mathcal{E}_2$, ce que l'on note

$$\mathcal{E} = \mathcal{E}_1 \otimes \mathcal{E}_2 \tag{42}$$

et que l'on définit comme suit.

On commence par définir le *produit tensoriel de deux vecteurs.* Étant donné $|\varphi(1)\rangle \in \mathcal{E}_1$ et $|\chi(2)\rangle \in \mathcal{E}_2$, on leur fait correspondre un vecteur de $\mathcal{E}$, que nous noterons simplement

$$|\varphi(1)\rangle|\chi(2)\rangle \,. \tag{43}$$

Cette correspondance possède par définition les propriétés suivantes :

(i) Elle est *linéaire* par rapport à la multiplication par les nombres complexes :

$$[\lambda|\varphi(1)\rangle][|\chi(2)\rangle] = \lambda[|\varphi(1)\rangle|\chi(2)\rangle] \,, \tag{44.a}$$

$$[|\varphi(1)\rangle][\mu|\chi(2)\rangle] = \mu[|\varphi(1)\rangle|\chi(2)\rangle] \,. \tag{44.b}$$

(ii) Elle est *distributive* par rapport à l'addition vectorielle :

$$|\varphi(1)\rangle[|\chi(2)\rangle + |\chi'(2)\rangle] = |\varphi(1)\rangle|\chi(2)\rangle + |\varphi(1)\rangle|\chi'(2)\rangle \,, \tag{45.a}$$

$$[|\varphi(1)\rangle + |\varphi'(1)\rangle]|\chi(2)\rangle = |\varphi(1)\rangle|\chi(2)\rangle + |\varphi'(1)\rangle|\chi(2)\rangle \,. \tag{45.b}$$

(iii) Si $\{|u_i(1)\rangle\}$ est une base dans $\mathcal{E}_1$ et $\{|v_j(2)\rangle\}$ une base dans $\mathcal{E}_2$, *l'ensemble des vecteurs* $|u_i(1)\rangle|v_j(2)\rangle$ *constitue une base dans* $\mathcal{E}$ (on voit en particulier que, si $\mathcal{E}_1$ et $\mathcal{E}_2$ sont de dimensions finies n_1 et n_2, la dimension de $\mathcal{E}$ est $n = n_1 n_2$).

$\mathcal{E}$ est bien ainsi un espace vectoriel[15]. Pour qu'il devienne un espace d'états, on le munit d'un *produit scalaire* déduit de ceux que l'on connaît dans $\mathcal{E}_1$ et $\mathcal{E}_2$: en premier lieu, le produit scalaire de deux kets qui sont des produits tensoriels est défini comme le produit des deux nombres complexes obtenus par produit scalaire des deux kets de $\mathcal{E}_1$ d'une part, et des deux kets de $\mathcal{E}_2$ d'autre part :

$$(\langle\varphi'(1)|\langle\chi'(2)|)(|\varphi(1)\rangle|\chi(2)\rangle) \widehat{=} \langle\varphi'(1)|\varphi(1)\rangle \cdot \langle\chi'(2)|\chi(2)\rangle \,; \tag{46}$$

cette définition est ensuite étendue à deux kets quelconques de $\mathcal{E}$ en imposant les propriétés caractéristiques d'un produit scalaire (C.D.L., p. 95).

On généralise sans peine les définitions précédentes à un nombre quelconque (fini) d'espaces d'états $\mathcal{E}_i$.

---

15. Il existe dans $\mathcal{E}$ des vecteurs qui ne sont pas des produits tensoriels d'un vecteur de $\mathcal{E}_1$ par un vecteur de $\mathcal{E}_2$. En effet, d'après la propriété (iii), le vecteur le plus général de $\mathcal{E}$ s'écrit

$$|\psi\rangle = \sum_{i,j} c_{i,j} |u_i(1)\rangle|v_j(2)\rangle \,, \tag{a}$$

et on ne peut pas toujours le mettre sous forme d'un produit tensoriel :

$$|\psi\rangle \neq \Big[\sum_i a_i|u_i(1)\rangle\Big]\Big[\sum_j b_j|v_j(2)\rangle\Big] \,. \tag{b}$$

Par contre, tout vecteur de $\mathcal{E}$ peut être décomposé en une somme de vecteurs produits tensoriels (l'expression (a) est une telle somme).

2. EXEMPLES IMPORTANTS

a. États d'un système de plusieurs particules

L'espace des états d'un système constitué de *deux particules sans spin* s'obtient en multipliant tensoriellement par lui-même l'espace des états d'une particule sans spin. En termes de fonctions d'onde, ceci signifie qu'un état du système des deux particules peut être caractérisé par une fonction de six variables $\psi(\vec{r}_1, \vec{r}_2)$, avec l'interprétation probabiliste habituelle :

$$|\psi(\vec{r}_1, \vec{r}_2)|^2 \, \mathrm{d}^3 r_1 \, \mathrm{d}^3 r_2 \tag{47}$$

est la probabilité pour que la particule numérotée (1) se trouve dans le volume élémentaire $\mathrm{d}^3 r_1$ autour du point $\vec{r}_1$ et la particule (2) dans $\mathrm{d}^3 r_2$ autour de $\vec{r}_2$ (à condition que ψ soit normée :

$$\int |\psi(\vec{r}_1, \vec{r}_2)|^2 \, \mathrm{d}^3 r_1 \, \mathrm{d}^3 r_2 = 1) \, . \tag{48}$$

Un état produit tensoriel a pour fonction d'onde le produit (ordinaire) des fonctions d'onde des deux particules :

$$\varphi(\vec{r}_1) \chi(\vec{r}_2) \, . \tag{49}$$

La fonction d'onde la plus générale $\psi(\vec{r}_1, \vec{r}_2)$ est une combinaison linéaire de tels produits. En particulier, si $\{u_i(\vec{r})\}$ est une base dans l'espace des fonctions d'onde d'une particule, on peut toujours écrire $\psi(\vec{r}_1, \vec{r}_2)$ sous la forme

$$\psi(\vec{r}_1, \vec{r}_2) = \sum_{i,\, j} c_{ij} \, u_i(\vec{r}_1) \, u_j(\vec{r}_2) \, . \tag{50}$$

Ces résultats se généralisent facilement à des systèmes constitués d'un nombre N quelconque de particules. Ils ne sont toutefois valables sous la forme précédente que si les particules sont *de nature différente* (voir au paragraphe IV le traitement particulier que nécessitent les systèmes de particules identiques).

b. États d'une particule de spin 1/2

L'espace des états d'une particule de spin 1/2 s'obtient par produit tensoriel de l'espace des états d'une particule sans spin (isomorphe à l'espace L^2 des fonctions $\psi(\vec{r})$ de carré sommable) et de l'espace des états de spin, qui est de dimension 2[16].

Il faut ainsi *deux fonctions d'onde* de carré sommable pour caractériser l'état d'une particule de spin 1/2; on les écrit le plus souvent en un « spineur à deux composantes » :

$$\begin{pmatrix} \varphi(\vec{r}) \\ \chi(\vec{r}) \end{pmatrix} . \tag{51}$$

Si ces deux fonctions sont proportionnelles, le spineur prend la forme

$$\varphi(\vec{r}) \begin{pmatrix} \alpha \\ \beta \end{pmatrix} , \tag{52}$$

où α et β sont des constantes complexes; l'état de la particule correspond alors au produit tensoriel de la fonction d'onde de position $\varphi(\vec{r})$, normalisée :

$$\int \mathrm{d}^3 r \, |\varphi(\vec{r})|^2 = 1 \, , \tag{53}$$

16. L'observable S_z associée à la composante du spin le long de l'axe Oz n'a que deux valeurs propres, $+\hbar/2$ et $-\hbar/2$, qui ne sont pas dégénérées si la position ou l'impulsion de la particule est fixée.

et de l'état de spin caractérisé par le spineur

$$\begin{pmatrix}\alpha\\ \beta\end{pmatrix}, \tag{54}$$

lui aussi normalisé :

$$|\alpha|^2+|\beta|^2=1\ . \tag{55}$$

c. États d'un système physique divisé en deux sous-systèmes indépendants

Reprenons l'exemple physique général qui nous a servi au paragraphe 1 à introduire le produit tensoriel : un système $\mathcal{S}$ divisé en deux sous-systèmes $\mathcal{S}_1$ et $\mathcal{S}_2$. Si l'état du système global $\mathcal{S}$ est décrit par un ket de $\mathcal{E}$ qui est un produit tensoriel :

$$|\varphi(1)\rangle\,|\chi(2)\rangle\ , \tag{56}$$

alors les propriétés physiques du sous-système $\mathcal{S}_1$ sont déterminées par $|\varphi(1)\rangle$ et ne dépendent pas de $|\chi(2)\rangle$ (c'est-à-dire de l'état du sous-système $\mathcal{S}_2$), et vice versa. Au contraire, si l'état de $\mathcal{S}$ est décrit par une combinaison linéaire de vecteurs produits tensoriels, par exemple

$$a\,|\varphi(1)\rangle\,|\chi(2)\rangle+a'\,|\varphi'(1)\rangle\,|\chi'(2)\rangle\ , \tag{57}$$

les propriétés physiques des deux sous-systèmes sont corrélées[17].

Par conséquent, lorsqu'un système $\mathcal{S}$ isolé est constitué de *deux sous-systèmes $\mathcal{S}_1$ et $\mathcal{S}_2$ isolés l'un de l'autre,* son état quantique est toujours décrit par un produit tensoriel comme (56). Son hamiltonien H est de la forme

$$H=H_1+H_2\ , \tag{58}$$

---

17. Calculons par exemple la valeur moyenne d'une observable $A(1)$ se référant au seul sous-système $\mathcal{S}_1$. Dans l'état produit tensoriel (56), cette valeur moyenne s'écrit (si $|\varphi(1)\rangle$ et $|\chi(2)\rangle$ sont séparément normés)

$$\begin{aligned}\langle A(1)\rangle&=\langle\varphi(1)|A(1)|\varphi(1)\rangle\cdot\langle\chi(2)|\chi(2)\rangle\\&=\langle\varphi(1)|A(1)|\varphi(1)\rangle\ .\end{aligned} \tag{a}$$

Le ket $|\chi(2)\rangle$ disparaît du résultat. Au contraire, dans l'état (57) (nous supposons que $|\varphi(1)\rangle$ et $|\varphi'(1)\rangle$ ne sont pas colinéaires, ni $|\chi(2)\rangle$ et $|\chi'(2)\rangle$; nous supposons de plus que $|\varphi(1)\rangle$, $|\varphi'(1)\rangle$, $|\chi(2)\rangle$, $|\chi'(2)\rangle$ et le ket (57) sont tous normés à 1), on a

$$\begin{aligned}\langle A(1)\rangle=|a|^2\langle\varphi(1)|A(1)|\varphi(1)\rangle+|a'|^2\langle\varphi'(1)|A(1)|\varphi'(1)\rangle\\+2\,\mathrm{Re}\left[a^*a'\langle\varphi(1)|A(1)|\varphi'(1)\rangle\langle\chi(2)|\chi'(2)\rangle\right].\end{aligned} \tag{b}$$

Le résultat dépend cette fois, en général, de $|\chi(2)\rangle$ et $|\chi'(2)\rangle$ (par leur produit scalaire).

Examinons également le cas d'un système de deux particules sans spin. Si sa fonction d'onde s'écrit

$$\psi(\vec{r}_1,\vec{r}_2)=\varphi(\vec{r}_1)\chi(\vec{r}_2) \tag{c}$$

(φ et χ étant séparément normées), la probabilité pour que la particule (1) se trouve en $\vec{r}_1$ à d^3r_1 près et la particule (2) en $\vec{r}_2$ à d^3r_2 près est le produit

$$|\varphi(\vec{r}_1)|^2\,\mathrm{d}^3r_1\times|\chi(\vec{r}_2)|^2\,\mathrm{d}^3r_2\ . \tag{d}$$

Les probabilités de présence des deux particules sont indépendantes l'une de l'autre. Si au contraire $\psi(\vec{r}_1,\vec{r}_2)$ n'est pas de la forme (c), alors les probabilités de présence des deux particules ne sont pas indépendantes :

$$|\psi(\vec{r}_1,\vec{r}_2)|^2\,\mathrm{d}^3r_1\,\mathrm{d}^3r_2\neq\left[\mathrm{d}^3r_1\int\mathrm{d}^3r_2|\psi(\vec{r}_1,\vec{r}_2)|^2\right]\times\left[\mathrm{d}^3r_2\int\mathrm{d}^3r_1|\psi(\vec{r}_1,\vec{r}_2)|^2\right]. \tag{e}$$

où H_1 est l'hamiltonien du sous-système $\mathcal{S}_1$ et H_2 celui de $\mathcal{S}_2$: comme l'énergie d'interaction entre $\mathcal{S}_1$ et $\mathcal{S}_2$ est nulle, *il n'y a pas dans* (58) *de terme d'interaction entre les deux sous-systèmes,* c'est-à-dire pas de terme qui fasse intervenir à la fois des observables associées à $\mathcal{S}_1$ et des observables associées à $\mathcal{S}_2$. Les états stationnaires $|\ell\rangle$ de $\mathcal{S}$ (états propres de H) sont des produits tensoriels d'un état propre $|\ell_1\rangle$ de H_1 et d'un état propre $|\ell_2\rangle$ de H_2 :

$$|\ell\rangle = |\ell_1\rangle\,|\ell_2\rangle\ , \tag{59}$$

avec
$$\begin{cases} H_1\,|\ell_1\rangle = E_{\ell_1}\,|\ell_1\rangle\ , & (60.a) \\ H_2\,|\ell_2\rangle = E_{\ell_2}\,|\ell_2\rangle\ , & (60.b) \end{cases}$$

et donc

$$H\,|\ell\rangle = (E_{\ell_1} + E_{\ell_2})\,|\ell\rangle\ . \tag{61}$$

Pour compter les états accessibles au système $\mathcal{S}$, il suffit alors (II, § I.C.2) de compter les états accessibles à chacun des deux sous-systèmes isolés et de multiplier les nombres ainsi obtenus. Rappelons que ce décompte reste souvent valable, en mécanique statistique, pour des sous-systèmes $\mathcal{S}_1$ et $\mathcal{S}_2$ macroscopiques, même s'ils ne sont pas vraiment isolés l'un de l'autre : si les forces s'exerçant entre les particules microscopiques qui les constituent sont à courte portée, leur couplage est suffisamment faible pour ne pas changer de façon appréciable le résultat (complément II.A, § II).

IV. SYSTÈMES DE PARTICULES IDENTIQUES

1. LE PROBLÈME DES PARTICULES IDENTIQUES EN MÉCANIQUE QUANTIQUE

Deux particules sont dites *identiques* si toutes leurs caractéristiques intrinsèques (masse, spin, charge, ...) sont les mêmes : deux électrons sont identiques, ainsi que deux protons, ou deux atomes de même espèce (deux atomes d'hydrogène, par exemple), ou deux molécules de même espèce (par exemple deux molécules d'eau, ou deux molécules de CO_2, ...).

Les systèmes de particules identiques posent en mécanique quantique un *problème fondamental.* Prenons par exemple deux particules identiques, et supposons qu'elles se situent initialement dans deux régions nettement distinctes de l'espace : leurs fonctions d'onde spatiales ont des supports complètement disjoints. Nous les numérotons alors : (1) et (2). Si leur mouvement les amène par la suite dans une situation où leurs probabilités de présence sont simultanément non nulles dans un même domaine, nous « perdons leur trace » irrémédiablement[18] : même si elles se séparent à nouveau, il est impossible de savoir laquelle avait reçu le numéro 1, laquelle le numéro 2.

---

18. La situation est radicalement différente en mécanique classique, où un système de particules identiques se traite selon les mêmes règles que s'il s'agissait de particules de nature différente. Supposons en effet que, à l'instant initial, on numérote les particules du système : (1), (2), ..., (N). En les « suivant » le long de leur trajectoire, on peut savoir à chaque instant ultérieur où se trouvent celle qui porte le numéro 1, celle qui porte le numéro 2, ... Ainsi, la numérotation, bien qu'arbitraire au départ, se comporte ensuite comme une propriété intrinsèque et permet de distinguer les unes des autres les N particules du système.

C'est donc l'absence de trajectoires définies qui pose en mécanique quantique un problème spécifique dans le cas de particules identiques. On montre d'ailleurs que l'application du postulat de symétrisation n'est pas nécessaire dans le cas où les fonctions d'onde spatiales des particules identiques étudiées ne se recouvrent jamais : il donne dans ce cas les mêmes résultats que le formalisme applicable aux particules non identiques, la numérotation étant alors non ambiguë (C.D.L., chap. XIV, § D.2.b).

2. LE POSTULAT DE SYMÉTRISATION

Dans le cas de particules identiques, les postulats généraux du paragraphe I doivent être complétés par un postulat spécifique.

Prenons par exemple un système de deux particules identiques. L'espace des états de ce système est *a priori* (§ III) le produit tensoriel de deux espaces d'états individuels (évidemment isomorphes l'un de l'autre). Si $\{|u_i\rangle\}$ est une base dans l'espace des états d'une des particules, l'ensemble des vecteurs

$$|(1) : u_i; (2) : u_j\rangle \widehat{=} |u_i\rangle|u_j\rangle \tag{62}$$

est une base dans l'espace produit tensoriel (§ III.1, point (iii)). Du point de vue mathématique, il est clair que[19]

$$|(1) : u_i; (2) : u_j\rangle \neq |(1) : u_j; (2) : u_i\rangle \qquad \text{si } i \neq j . \tag{63}$$

Du point de vue physique, par contre, ces deux vecteurs ne décrivent des états distincts que si les deux particules sont de nature différente; si elles sont identiques, *les deux vecteurs* (63), *ainsi d'ailleurs que toutes leurs combinaisons linéaires, sont associés au même état physique.* Pour un état physique donné, nous avons donc *a priori* le choix entre une infinité de kets mathématiques qui sont susceptibles de le décrire. Cette indétermination est connue sous le nom de «dégénérescence d'échange». On montre facilement (C.D.L., chap. XIV, § A.3.b) que les divers kets possibles conduisent à des prédictions physiques différentes; il faut par conséquent *choisir l'un d'eux comme seul ket «physique»,* c'est-à-dire lever la dégénérescence d'échange.

Postulat de symétrisation : lorsqu'un système comprend plusieurs particules identiques, seuls certains kets de son espace des états peuvent décrire ses états physiques : les kets physiques sont, suivant la nature des particules identiques, soit complètement symétriques, soit complètement antisymétriques par rapport aux permutations de ces particules. On appelle *bosons* les particules pour lesquelles les kets physiques sont symétriques, *fermions* celles pour lesquelles ils sont antisymétriques.

L'expérience montre que les particules de *spin entier* ($s = 0$, ou 1, ...) sont des *bosons,* celles de *spin demi-entier* ($s = 1/2$, ou $3/2$, ...) des *fermions*[20].

L'*espace des états physiques* d'un système de deux particules est donc un sous-espace de l'espace vectoriel engendré par les vecteurs de base (62)[21]. Pour des *bosons,* les kets physiques sont les combinaisons linéaires des vecteurs

---

19. Ces deux vecteurs sont mêmes *orthogonaux* si la base de départ $\{|u_i\rangle\}$ est orthonormée : d'après la définition (46) du produit scalaire dans l'espace produit tensoriel, on a

$$\langle(1) : u_i; (2) : u_j|(1) : u_j; (2) : u_i\rangle = \langle u_i|u_j\rangle . \langle u_j|u_i\rangle = 0 \quad \text{si} \quad i \neq j .$$

20. Les électrons, les protons et les neutrons, qui ont un spin 1/2, sont des fermions; les photons, dont le spin vaut 1, sont des bosons. Les états liés de plusieurs particules sont des bosons ou des fermions suivant que, parmi les particules constituantes, les fermions sont en nombre pair ou impair. Ainsi, les atomes dont le noyau comporte un nombre pair de neutrons sont des bosons, ceux pour lesquels le nombre de neutrons est impair des fermions (dans un atome neutre, le nombre de protons et le nombre d'électrons sont toujours égaux) : par exemple, l'isotope ^{3}He de l'hélium est un fermion, l'isotope ^{4}He un boson.

21. *Attention.* L'espace des états individuels, et donc les vecteurs de base $|u_i\rangle$, sont différents suivant le spin des particules considérées : pour $s = 0$, cet espace des états est isomorphe à l'ensemble L^2 des fonctions de carré sommable; pour $s = 1/2$, c'est le produit tensoriel de L^2 et d'un espace à 2 dimensions (§ III.2.b); etc. Les kets physiques (64) et (65) n'appartiennent donc pas au même espace, contrairement à ce que pourraient faire croire les notations.

$$\begin{cases} |(1) : u_i ; (2) : u_i\rangle \\ \dfrac{1}{\sqrt{2}}\left[|(1) : u_i ; (2) : u_j\rangle + |(1) : u_j ; (2) : u_i\rangle\right] \end{cases} \quad \text{pour } i \neq j . \tag{64}$$

Pour des *fermions,* ce sont les combinaisons linéaires de

$$\frac{1}{\sqrt{2}}\left[|(1) : u_i ; (2) : u_j\rangle - |(1) : u_j ; (2) : u_i\rangle\right] . \tag{65}$$

Notons que *les fermions obéissent au «principe d'exclusion de Pauli»,* selon lequel deux fermions identiques ne peuvent pas se trouver dans le même état individuel. Cette restriction est une conséquence immédiate du postulat de symétrisation : la combinaison antisymétrique (65) est évidemment nulle lorsque $|u_i\rangle$ et $|u_j\rangle$ coïncident.

3. ÉTATS STATIONNAIRES D'UN SYSTÈME DE *N* PARTICULES IDENTIQUES ET INDÉPENDANTES

Les N particules constituant un système sont *indépendantes* si l'hamiltonien H du système est de la forme

$$H = H_1 + H_2 + H_3 + \ldots + H_N , \tag{66}$$

où H_i ne fait intervenir que les observables (position, impulsion, spin, ...) associées à la particule numérotée (i). Les états stationnaires (ℓ) du système, états propres de H, peuvent être caractérisés par la donnée des N états stationnaires individuels :

$$(\ell) \equiv (\lambda_1, \lambda_2, \ldots, \lambda_N) , \tag{67}$$

où (λ_i) est état propre de H_i avec la valeur propre ε_{λ_i}. L'énergie E_ℓ associée à l'état (ℓ) est simplement la somme des énergies individuelles :

$$E_\ell = \varepsilon_{\lambda_1} + \varepsilon_{\lambda_2} + \ldots + \varepsilon_{\lambda_N} . \tag{68}$$

Si maintenant les N particules considérées sont *identiques,* les N hamiltoniens individuels H_i ont nécessairement même forme. Il existe alors un seul et même ensemble d'états individuels (λ), parmi lesquels sont choisis les N états (λ_i), et un seul et même spectre d'énergies correspondantes $\{\varepsilon_\lambda\}$.

Supposons alors donnée une liste de N états individuels. Toute liste obtenue par permutation de l'ordre de ces N états individuels est, par suite de l'identité des particules, physiquement équivalente à la précédente (dégénérescence d'échange). Le postulat de symétrisation implique qu'*à l'ensemble de toutes ces listes* déduites les unes des autres par permutation *est associé un ket physique unique* (éventuellement nul dans le cas de fermions, s'il viole le principe de Pauli). *Un état stationnaire physique du système*[22] *est donc caractérisé par la donnée d'une liste non ordonnée de N états individuels* (λ_i)[23]; pour des *bosons,* il n'y a *pas de restriction* sur la composition de cette liste; pour des *fermions* au contraire, elle ne doit compter que *des états individuels tous différents* (principe de Pauli).

---

22. Si les N particules identiques ne sont pas indépendantes (c'est-à-dire s'il existe dans l'hamiltonien H des termes d'interaction entre les particules), les états que nous construisons dans ce paragraphe ne sont pas les états stationnaires du système. Ils n'en constituent pas moins une *base dans l'espace des états physiques du système :* tout ket physique (c'est-à-dire compatible avec le postulat de symétrisation) peut être écrit comme combinaison linéaire des kets précédents, pourvu que l'ensemble des kets individuels (λ) utilisé forme une base dans l'espace des états d'une particule.

23. Il existe cependant des situations particulières dans lesquelles les particules, bien qu'identiques, sont «discernables» (III, § IV.A.2). Dans de telles situations, l'application du postulat de symétrisation ne restreint pas le nombre d'états microscopiques possibles pour le système, par rapport au cas où les particules seraient de nature différente (chapitre III, note 63).

On peut, de façon équivalente, utiliser la représentation des « nombres d'occupation » : étant donné une liste non ordonnée de N états individuels, on appelle *nombre d'occupation* (ou *population*) N_λ de l'état individuel (λ) le nombre de fois que (λ) apparaît dans cette liste. Ainsi, chaque état microscopique stationnaire (ℓ) d'un système de particules identiques indépendantes[22] peut être caractérisé par une *configuration* $\{N_\lambda\}$, c'est-à-dire par un ensemble de nombres d'occupation, un pour chaque état individuel (λ) possible[24] :

$$(\ell) \iff \{N_\lambda\}, \tag{69}$$

avec

$$\sum_{(\lambda)} N_\lambda = N . \tag{70}$$

Dans cette représentation, l'énergie E_ℓ de l'état stationnaire (ℓ) s'écrit simplement

$$E_\ell = \sum_{(\lambda)} N_\lambda \varepsilon_\lambda . \tag{71}$$

Dans le cas des *bosons*, le nombre d'occupation N_λ d'un état (λ) donné peut prendre *n'importe quelle valeur entière positive ou nulle* (mais inférieure à N d'après (70)). Dans le cas des *fermions* au contraire, *les seules valeurs possibles pour* N_λ *sont* 0 *et* 1, les valeurs supérieures étant interdites par le principe de Pauli.

24. Toute configuration comprend un nombre d'occupation N_λ pour *chacun des états individuels* (λ) *a priori possibles*, ce nombre étant simplement nul pour les états individuels non occupés dans l'état (ℓ) considéré.

APPENDICE III

Description classique d'un système physique

I. Le formalisme lagrangien
   1. Coordonnées et vitesses généralisées
   2. Fonction et équations de Lagrange

II. Le formalisme hamiltonien
   1. Moments conjugués des coordonnées
   2. Fonction de Hamilton et équations canoniques
   3. L'espace des phases
      a. Définition
      b. Théorème de Liouville

De façon générale, c'est la mécanique quantique qu'il convient d'utiliser pour décrire les systèmes physiques au niveau microscopique. Pourtant, lorsque la constante fondamentale $\hbar$ qui caractérise les effets quantiques reste négligeable devant toutes les grandeurs de même dimension que l'on peut construire à partir des caractéristiques du système étudié, il est plus simple (et plus intuitif) de le décrire par la mécanique classique.

La *mécanique classique,* ou *mécanique newtonienne,* est fondée sur les lois de Newton (1687). Son approfondissement et son développement ont conduit à des formulations plus générales et plus abstraites, dont les deux principales portent les noms de Lagrange (1788) et de Hamilton (1830).

I. LE FORMALISME LAGRANGIEN

1. COORDONNÉES ET VITESSES GÉNÉRALISÉES

En mécanique classique, l'étude d'un système commence par le choix de n *coordonnées généralisées* $q_i (i = 1, 2, \ldots, n)$, dont les dérivées $\dot{q}_i$ par rapport au temps sont appelées *vitesses généralisées.* Le choix des coordonnées généralisées doit être tel que la donnée des $2n$ paramètres

$$\{q_i(t_0),\ \dot{q}_i(t_0);\quad i = 1, 2, \ldots, n\} \tag{1}$$

est suffisante (et juste suffisante) pour déterminer l'état du système, c'est-à-dire la position et la vitesse de chacun de ses points, à l'instant t_0. Ainsi, il faut $n = 6$ coordonnées généralisées pour décrire un solide indéformable[1]; l'état (microscopique) instantané d'un ensemble de N particules ponctuelles pourra être défini par les $n = 3N$ coordonnées de ces N particules et les $3N$ vitesses associées[2] :

$$\{\vec{r}_\alpha(t_0),\ \vec{v}_\alpha(t_0);\quad \alpha = 1, 2, \ldots, N\}\ . \tag{2}$$

---

1. On peut prendre les trois coordonnées cartésiennes de son centre d'inertie et les trois angles d'Euler qui repèrent son orientation par rapport au système d'axes choisi.

2. Pour un système donné, le choix des coordonnées généralisées qui permettent de décrire sa position et son mouvement comporte un certain arbitraire. Dans un système de N particules par exemple, on peut évidemment, suivant la commodité, repérer chacune d'elles par ses coordonnées cartésiennes, ou par ses coordonnées sphériques... Ou bien, dans un système de deux particules

.../...

2. FONCTION ET ÉQUATIONS DE LAGRANGE

Les *équations du mouvement,* c'est-à-dire les équations différentielles qui permettent en principe de calculer les coordonnées généralisées q_i en fonction du temps t, découlent du « principe fondamental de la dynamique » ou « deuxième loi de Newton ».

Considérons le cas où toutes les forces dérivent d'une *énergie potentielle*

$$U(q_1, q_2, \ldots, q_n)\,. \tag{3}$$

Si $$\mathcal{T}(q_1, q_2, \ldots, q_n; \dot{q}_1, \dot{q}_2, \ldots, \dot{q}_n\} \tag{4}$$

est l'*énergie cinétique* totale du système, on montre que les équations du mouvement s'écrivent[3, 4]

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \mathcal{T}}{\partial \dot{q}_i} - \frac{\partial \mathcal{T}}{\partial q_i} = -\frac{\partial U}{\partial q_i}\,; \quad i = 1, 2, \ldots, n\,. \tag{5}$$

Si l'on introduit la *fonction de Lagrange*

$$\mathcal{L}(q_1, q_2, \ldots, q_n; \dot{q}_1, \dot{q}_2, \ldots, \dot{q}_n) \widehat{=} \mathcal{T}(q_1, q_2, \ldots, q_n; \dot{q}_1, \dot{q}_2, \ldots, \dot{q}_n) - U(q_1, q_2, \ldots, q_n)\,, \tag{6}$$

les équations (5) prennent la forme

$$\boxed{\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \mathcal{L}}{\partial \dot{q}_i} - \frac{\partial \mathcal{L}}{\partial q_i} = 0}\ ;\ i = 1, 2, \ldots, n\,. \tag{7}$$

(de masses m_1 et m_2) en interaction, il peut être plus commode de choisir, au lieu des coordonnées $\vec{r}_1$ et $\vec{r}_2$ de chaque particule, celles de leur centre de masse :

$$\vec{r}_G \widehat{=} \frac{m_1\vec{r}_1 + m_2\vec{r}_2}{m_1 + m_2}\,, \tag{a}$$

et les coordonnées relatives :

$$\vec{r} \widehat{=} \vec{r}_1 - \vec{r}_2\,. \tag{b}$$

3. Pour une particule unique de masse m, par exemple, on a

$$\mathcal{T} = \frac{1}{2}\,m\vec{v}^{\,2}\,, \tag{a}$$

de sorte que les équations (5) donnent

$$\frac{\mathrm{d}}{\mathrm{d}t}(m\vec{v}) = -\overrightarrow{\mathrm{grad}}\ U\,, \tag{b}$$

où l'on reconnaît la relation fondamentale de la dynamique.

4. La *dérivée totale par rapport au temps,* notée $\mathrm{d}/\mathrm{d}t$, d'une fonction $f(q_1, q_2, \ldots, q_n; \dot{q}_1, \dot{q}_2, \ldots, \dot{q}_n; t)$ est définie par

$$\frac{\mathrm{d}}{\mathrm{d}t} f(q, \dot{q}; t) \widehat{=} \frac{\partial f}{\partial t} + \sum_{i=1}^{n} \frac{\partial f}{\partial q_i}\frac{\mathrm{d}q_i}{\mathrm{d}t} + \sum_{i=1}^{n} \frac{\partial f}{\partial \dot{q}_i}\frac{\mathrm{d}\dot{q}_i}{\mathrm{d}t}\,. \tag{a}$$

Autrement dit, la dérivée totale $\mathrm{d}/\mathrm{d}t$ est calculée en tenant compte du fait que les q_i et les $\dot{q}_i$ sont des fonctions du temps : c'est la dérivée de

$$f[q(t), \dot{q}(t); t]\,, \tag{b}$$

fonction cette fois du temps seul.

Les *équations de Lagrange* (7) sont tout à fait générales : dans des cas plus complexes, on cherchera une fonction de Lagrange (qui peut dépendre explicitement du temps) :

$$\mathcal{L}(q_1, q_2, \ldots, q_n; \dot{q}_1, \dot{q}_2, \ldots, \dot{q}_n; t) \tag{8}$$

telle que les équations du mouvement prennent la forme (7)[5]. La connaissance de la fonction de Lagrange et de l'état à un instant donné suffit donc, pour un système quelconque, à sa description dynamique.

II. LE FORMALISME HAMILTONIEN

1. MOMENTS CONJUGUÉS DES COORDONNÉES

Le *moment conjugué* p_i de la coordonnée généralisée q_i est défini à partir de la fonction de Lagrange $\mathcal{L}(q, \dot{q}; t)$ par

$$p_i \widehat{=} \frac{\partial \mathcal{L}}{\partial \dot{q}_i} ; \quad i = 1, 2, \ldots, n . \tag{9}$$

On l'appelle aussi *« impulsion généralisée »* conjuguée de q_i.

Pour une particule unique ($n = 3$) soumise à des forces dérivant d'un potentiel, on trouve

$$\vec{p} = m\vec{v} . \tag{10}$$

L'impulsion coïncide dans ce cas avec la « quantité de mouvement »[6] $m\vec{v}$.

2. FONCTION DE HAMILTON ET ÉQUATIONS CANONIQUES

A partir des moments conjugués p_i et de la fonction de Lagrange, on définit la *fonction de Hamilton* $\mathcal{H}$ par

$$\mathcal{H} \widehat{=} \sum_{i=1}^{n} p_i \dot{q}_i - \mathcal{L} . \tag{11}$$

---

5. Prenons par exemple le cas d'une particule, de masse m et de charge q, plongée dans un champ électromagnétique $\{\vec{E}, \vec{B}\}$ décrit par les potentiels scalaire $\phi(\vec{r}, t)$ et vecteur $\vec{A}(\vec{r}, t)$. La relation fondamentale de la dynamique s'écrit dans ce cas

$$m \frac{d\vec{v}}{dt} = q\left[\vec{E} + \vec{v} \wedge \vec{B}\right] = q\left[-\overrightarrow{\text{grad}}\, \phi - \frac{\partial \vec{A}}{\partial t} + \vec{v} \wedge \overrightarrow{\text{rot}}\, \vec{A}\right] . \tag{a}$$

On montre facilement (C.D.L., appendice III, § 4.b) que cette équation peut être déduite, par les formules (7), de la fonction de Lagrange

$$\mathcal{L}(\vec{r}, \vec{v}; t) = \frac{1}{2} m\vec{v}^{\,2} + q\vec{v} \,.\, \vec{A}(\vec{r}, t) - q\phi(\vec{r}, t) . \tag{b}$$

6. Ce n'est plus le cas si la particule, de masse m et de charge q, est plongée dans un champ électromagnétique décrit par les potentiels $\phi(\vec{r}, t)$ et $\vec{A}(\vec{r}, t)$. A partir de la fonction de Lagrange correspondante (note 5), on trouve

$$\vec{p} = m\vec{v} + q\vec{A}(\vec{r}, t) .$$

$\mathcal{H}$ est donc une *transformée de Legendre* de $\mathcal{L}$. *Ses variables naturelles sont les coordonnées généralisées* q_i *et leurs moments conjugués* p_i (outre, éventuellement, le temps t)[7] :

$$\mathcal{H} = \mathcal{H}(q, p; t) . \tag{12}$$

Dans le formalisme hamiltonien, par conséquent, l'*état du système* à un instant t est caractérisé, non plus par l'ensemble des q_i et $\dot{q}_i$, mais par les $2n$ *« variables canoniques »*

$$\{q_i, p_i; \quad i = 1, 2, \ldots, n\} . \tag{13}$$

Chaque grandeur physique associée au système (énergie, moment cinétique, ...) est décrite par une fonction des variables canoniques (et éventuellement du temps) $f(q, p; t)$; la valeur prise à l'instant t par la grandeur est obtenue en remplaçant dans cette fonction les variables q_i et p_i par leur valeur à l'instant t.

Les équations du mouvement s'écrivent ici[8]

$$\begin{cases} \dfrac{\mathrm{d}q_i}{\mathrm{d}t} = \dfrac{\partial \mathcal{H}}{\partial p_i} , & \text{(14.a)} \\ \dfrac{\mathrm{d}p_i}{\mathrm{d}t} = -\dfrac{\partial \mathcal{H}}{\partial q_i} ; \quad i = 1, 2, \ldots, n . & \text{(14.b)} \end{cases}$$

Les équations de Lagrange (7) constituaient un système de n équations différentielles du second ordre par rapport au temps. On l'a remplacé ici par un système équivalent de $2n$ *équations différentielles du premier ordre,* qui sont connues sous le nom d'*équations canoniques,* ou *équations de Hamilton-Jacobi.*

Prenons l'exemple simple d'une particule unique (masse m), d'énergie potentielle $U(\vec{r})$. A partir de la fonction de Lagrange (formule (6))

$$\mathcal{L}(\vec{r}, \vec{v}) = \frac{1}{2} m\vec{v}^{\,2} - U(\vec{r}) , \tag{15}$$

---

7. Calculons la différentielle de la fonction $\mathcal{H}$:

$$\mathrm{d}\mathcal{H} = \sum_{i=1}^{n} (p_i \, \mathrm{d}\dot{q}_i + \dot{q}_i \, \mathrm{d}p_i) - \sum_{i=1}^{n} \left(\frac{\partial \mathcal{L}}{\partial q_i} \mathrm{d}q_i + \frac{\partial \mathcal{L}}{\partial \dot{q}_i} \mathrm{d}\dot{q}_i \right) - \frac{\partial \mathcal{L}}{\partial t} \mathrm{d}t . \tag{a}$$

D'après (9), la différentielle $\mathrm{d}\dot{q}_i$ disparaît de cette expression :

$$\mathrm{d}\mathcal{H} = \sum_{i=1}^{n} \left[\dot{q}_i \, \mathrm{d}p_i - \frac{\partial \mathcal{L}}{\partial q_i} \mathrm{d}q_i \right] - \frac{\partial \mathcal{L}}{\partial t} \mathrm{d}t , \tag{b}$$

qui fournit alors les dérivées de $\mathcal{H}$ par rapport à ses variables « naturelles » :

$$\frac{\partial \mathcal{H}}{\partial q_i} = -\frac{\partial \mathcal{L}}{\partial q_i} , \tag{c}$$

$$\frac{\partial \mathcal{H}}{\partial p_i} = \dot{q}_i , \tag{d}$$

$$\frac{\partial \mathcal{H}}{\partial t} = -\frac{\partial \mathcal{L}}{\partial t} . \tag{e}$$

8. La première de ces équations découle de la définition même de $\mathcal{H}$ (formule (d) de la note 7). La seconde n'est autre que l'équation de Lagrange (7), dans laquelle on a remplacé $\partial \mathcal{L}/\partial \dot{q}_i$ par p_i (définition (9)) et $\partial \mathcal{L}/\partial q_i$ par $-\partial \mathcal{H}/\partial q_i$ (formule (c) de la note 7).

et de l'expression (10) de l'impulsion qui en découle, on trouve aussitôt[9]

$$\mathcal{H}(\vec{r}, \vec{p}) = \frac{\vec{p}^{\,2}}{2m} + U(\vec{r}) . \tag{16}$$

Les équations canoniques (14) s'écrivent dans ce cas

$$\begin{cases} \dfrac{d\vec{r}}{dt} = \dfrac{\vec{p}}{m} , & (17.a) \\ \dfrac{d\vec{p}}{dt} = -\overrightarrow{\text{grad}}\ U . & (17.b) \end{cases}$$

On retrouve effectivement la relation fondamentale de la dynamique.

De façon plus générale, lorsque les forces dérivent d'une énergie potentielle $U(q)$, la fonction de Hamilton est la somme de cette énergie potentielle et de l'énergie cinétique totale[10] :

$$\mathcal{H}(q, p) = \mathcal{T}(q, p) + U(q) . \tag{18}$$

$\mathcal{H}$ n'est donc rien d'autre que l'*énergie totale* du système, exprimée en fonction des variables canoniques q_i et p_i[11] .

3. L'ESPACE DES PHASES

a. Définition

Pour pouvoir raisonner géométriquement, il est commode d'associer à chaque état instantané du système, caractérisé par les variables canoniques (13), un point dans un espace cartésien à $2n$ dimensions appelé *« espace des phases » :* les coordonnées de ce point représentatif ont pour valeur celle des paramètres q_i $(i = 1, 2, \ldots, n)$ et p_i $(i = 1, 2, \ldots, n)$ à l'instant considéré.

Au cours de l'évolution du système dans le temps, le point qui représente son état se déplace dans l'espace des phases; sa « trajectoire » dans l'espace des phases, ainsi que la loi horaire suivant laquelle elle est décrite, sont déterminées par les équations du mouvement (14).

b. Théorème de Liouville

α. Conservation du volume dans l'espace des phases au cours du mouvement

Considérons un domaine infinitésimal $\mathcal{D}$ de l'espace des phases, de volume

$$d\tau = \prod_{i=1}^{n} dq_i . dp_i . \tag{19}$$

Supposons que, à l'instant t, le système étudié se trouve dans un état dont le point représentatif, de coordonnées

$$(q_1, q_2, \ldots, q_n ; p_1, p_2, \ldots, p_n) , \tag{20}$$

---

9. Si la particule, de charge q, est plongée dans un champ électromagnétique de potentiels $\phi(\vec{r}, t)$ et $\vec{A}(\vec{r}, t)$ (cf. note 5), le résultat est

$$\mathcal{H}(\vec{r}, \vec{p} ; t) = \frac{1}{2m}\left[\vec{p} - q\vec{A}(\vec{r}, t)\right]^2 + q\phi(\vec{r}, t) .$$

10. Ceci découle du fait que, même lorsqu'elle dépend des q_i, l'énergie cinétique $\mathcal{T}$ est toujours une *fonction homogène du deuxième degré* des vitesses généralisées $\dot{q}_i$.

11. *Attention :* dans la formule (18) comme dans (16), l'énergie cinétique doit être exprimée en fonction des impulsions p_i, et non plus des vitesses généralisées $\dot{q}_i$, puisque le formalisme hamiltonien se fonde sur l'utilisation des variables canoniques q_i et p_i (cf. formule (12)).

est situé à l'intérieur de ce domaine élémentaire $\mathcal{D}$. A un instant ultérieur $t+\delta t$, le point représentatif se sera déplacé et sera situé en

$$(q'_1, q'_2, \ldots, q'_n; p'_1, p'_2, \ldots, p'_n) . \tag{21}$$

Considérons le domaine $\mathcal{D}'$ où se retrouvent ainsi, à l'instant $t+\delta t$, l'ensemble des points situés dans $\mathcal{D}$ à l'instant t. Le théorème de Liouville affirme[12] que le volume

$$d\tau' = \prod_{i=1}^{n} dq'_i . dp'_i \tag{22}$$

de ce domaine $\mathcal{D}'$ est égal au volume $d\tau$ du domaine originel $\mathcal{D}$:

$$\boxed{d\tau' = d\tau} . \tag{23}$$

En d'autres termes, *le volume d'un domaine de l'espace des phases se conserve au cours du temps* lorsqu'on suit le mouvement des divers points de ce domaine[13].

---

12. Les coordonnées q'_i et p'_i s'écrivent, au premier ordre en δt,

$$q'_i = q_i + \dot{q}_i \delta t , \tag{a}$$

$$p'_i = p_i + \dot{p}_i \delta t , \tag{b}$$

où $\dot{q}_i$ et $\dot{p}_i$ sont donnés par les équations du mouvement (14). Le calcul de $d\tau'$ à partir de $d\tau$ se ramène donc à un problème classique de *changement de variables multiples,* selon les formules ci-dessus : il nous faut déterminer le *jacobien* de ce changement de variables.

Le jacobien est la valeur absolue du déterminant $2n \times 2n$ dont les éléments $a_{k\ell}(k, \ell = 1, 2, \ldots, 2n)$ s'obtiennent comme suit :

$$a_{ij} = \frac{\partial q'_i}{\partial q_j} , \quad a_{i,j+n} = \frac{\partial q'_i}{\partial p_j} ,$$

$$a_{i+n,j} = \frac{\partial p'_i}{\partial q_j} , \quad a_{i+n,j+n} = \frac{\partial p'_i}{\partial p_j} \tag{c}$$

(dans ces expressions, i et j sont limités aux valeurs $i, j = 1, 2, \ldots, n$). On constate sans peine que tous les éléments non diagonaux de ce déterminant sont proportionnels à δt, et que ses éléments diagonaux valent d'après (14)

$$a_{ii} = 1 + \frac{\partial \dot{q}_i}{\partial q_i} \delta t = 1 + \frac{\partial^2 \mathcal{H}}{\partial q_i \partial p_i} \delta t , \tag{d}$$

$$a_{i+n,i+n} = 1 + \frac{\partial \dot{p}_i}{\partial p_i} \delta t = 1 - \frac{\partial^2 \mathcal{H}}{\partial p_i \partial q_i} \delta t . \tag{e}$$

Or la validité du calcul est limitée dès le départ au *premier ordre en* δt (le résultat se généralisant ensuite par intégration sur t). La valeur correspondante du déterminant s'obtient très simplement : seul le produit des éléments diagonaux contribue à cet ordre, et sa contribution s'écrit

$$\det |a_{k\ell}| = 1 + \delta t \sum_{i=1}^{n} \left[\frac{\partial \dot{q}_i}{\partial q_i} + \frac{\partial \dot{p}_i}{\partial p_i} \right] = 1 + \delta t \sum_{i=1}^{n} \left[\frac{\partial^2 \mathcal{H}}{\partial q_i \partial p_i} - \frac{\partial^2 \mathcal{H}}{\partial p_i \partial q_i} \right]$$

$$= 1 \quad \text{au premier ordre en } \delta t . \tag{f}$$

Le jacobien de la transformation (a)-(b) vaut donc 1, ce qui signifie que

$$dq'_1 \, dq'_2 \ldots dq'_n \, dp'_1 \, dp'_2 \ldots dp'_n = dq_1 \, dq_2 \ldots dq_n \, dp_1 \, dp_2 \ldots dp_n . \tag{g}$$

13. Cette propriété ne serait pas valable pour un espace défini à partir des q_i et $\dot{q}_i$: l'élément

$$\prod_{i=1}^{n} dq_i . d\dot{q}_i$$

ne se conserve pas au cours du temps comme l'élément de volume de l'espace des phases.

β. Propriété de la densité de probabilité dans l'espace des phases

Lorsqu'un système peut être décrit par la mécanique classique au niveau microscopique, son état macroscopique est caractérisé, en mécanique statistique, par une densité de probabilité $w(q, p; t)$ dans l'espace des phases qui lui est associé (I, § III.C.2). Cette densité de probabilité est normalisée :

$$\int w(q, p; t)\, \mathrm{d}q\,.\mathrm{d}p = 1\ , \tag{24}$$

où nous utilisons la notation simplifiée

$$\mathrm{d}q\,.\mathrm{d}p \mathrel{\widehat{=}} \prod_{i=1}^{n} \mathrm{d}q_i\,.\mathrm{d}p_i\ . \tag{25}$$

Contrairement à ce qui se passe en mécanique quantique pour les états stationnaires (ℓ) du système, l'état microscopique classique $\{q, p\}$ évolue au cours du temps. Il en résulte, pour la densité de probabilité w, une *condition liant sa dépendance par rapport au point* (q, p) *de l'espace des phases et sa dépendance explicite par rapport au temps* t. Cette condition s'écrit[14]

$$\sum_{i=1}^{n} \left(\frac{\partial w}{\partial q_i}\, \dot{q}_i + \frac{\partial w}{\partial p_i}\, \dot{p}_i \right) + \frac{\partial w}{\partial t} = 0\ , \tag{26}$$

ou de façon plus compacte (cf. note 4)

$$\frac{\mathrm{d}w}{\mathrm{d}t} = 0\ . \tag{27}$$

Elle est connue sous le nom d'*équation de Liouville.*

---

14. Concrétisons la notion de probabilité dans l'espace des phases en supposant que nous disposons d'un très grand nombre $\mathcal{N}$ de systèmes tous identiques et préparés de la même façon (voir I, § IV.B.2). Le nombre de ces systèmes dont le point représentatif se trouve, à l'instant t, dans le domaine élémentaire $\mathrm{d}q\,.\mathrm{d}p$ de l'espace des phases, autour du point (q, p), est alors

$$\mathcal{N} w(q, p; t)\mathrm{d}q\,.\mathrm{d}p \tag{a}$$

(nous oublions pour simplifier les fluctuations; cf. appendice I, § II.1.c.γ). A l'instant $t + \delta t$, l'évolution dans le temps aura amené ces *mêmes systèmes* dans des états microscopiques dont les points représentatifs seront situés dans le domaine $\mathrm{d}q'.\mathrm{d}p'$ autour du point (q', p'), où q'_i et p'_i sont donnés par les formules (a) et (b) de la note 12.

Mais le nombre de systèmes dont le point représentatif se situe, à l'instant $t + \delta t$, dans le domaine $\mathrm{d}q'.\mathrm{d}p'$ entourant le point (q', p') peut aussi être calculé à partir de w :

$$\mathcal{N} w(q', p'; t + \delta t)\mathrm{d}q'.dp'\ . \tag{b}$$

On a donc

$$w(q', p'; t + \delta t)\mathrm{d}q'.\mathrm{d}p' = w(q, p; t)\mathrm{d}q\,.\mathrm{d}p\ . \tag{c}$$

D'après le théorème de Liouville (23), et compte tenu de l'expression des q'_i et p'_i (formules (a) et (b) de la note 12), cette condition s'écrit

$$w(q + \dot{q}\delta t, p + \dot{p}\delta t; t + \delta t) = w(q, p; t) \quad \text{au premier ordre en } \delta t, \tag{d}$$

c'est-à-dire finalement, si l'on explicite à nouveau les q_i et p_i,

$$\sum_{i=1}^{n} \left[\dot{q}_i \frac{\partial w}{\partial q_i} + \dot{p}_i \frac{\partial w}{\partial p_i} \right] + \frac{\partial w}{\partial t} = 0\ . \tag{e}$$

On peut aussi l'exprimer en disant que la probabilité « s'écoule » dans l'espace des phases comme le ferait un fluide de densité w obéissant à la *loi de conservation locale*[15]

$$\frac{\partial w}{\partial t} + \operatorname{div}(w\vec{v}) = 0 , \tag{28}$$

où la divergence est prise par rapport aux $2n$ coordonnées $\{q_i, p_i\}$ de l'espace des phases, et $\vec{v}$ est la « vitesse » dans ce même espace, de composantes $\{\dot{q}_i, \dot{p}_i\}$.

En particulier, les distributions de probabilité décrivant des états d'*équilibre macroscopique* du système doivent être *stationnaires,* c'est-à-dire telles que

$$\frac{\partial w}{\partial t} = 0 . \tag{29}$$

Elles doivent donc vérifier la relation suivante (obtenue à partir de (26) en utilisant les équations du mouvement (14)) :

$$\frac{\partial w}{\partial q_i}\frac{\partial \mathcal{H}}{\partial p_i} - \frac{\partial w}{\partial p_i}\frac{\partial \mathcal{H}}{\partial q_i} = 0 . \tag{30}$$

Notons que cette condition est effectivement réalisée, par exemple, si la densité de probabilité w ne dépend des q_i et p_i qu'à travers la fonction de Hamilton elle-même[16], c'est-à-dire si c'est une fonction de la forme[17]

$$w(q, p) = W\left[\mathcal{H}(q, p)\right] , \tag{31}$$

où $W(u)$ est une fonction d'une variable[18].

---

15. La relation (28) s'explicite sous la forme

$$\frac{\partial w}{\partial t} + \sum_{i=1}^{n}\left[\frac{\partial}{\partial q_i}(w\dot{q}_i) + \frac{\partial}{\partial p_i}(w\dot{p}_i)\right] = 0 , \tag{a}$$

ou encore

$$\frac{\partial w}{\partial t} + \sum_{i=1}^{n}\left[\dot{q}_i\frac{\partial w}{\partial q_i} + \dot{p}_i\frac{\partial w}{\partial p_i} + w\left(\frac{\partial \dot{q}_i}{\partial q_i} + \frac{\partial \dot{p}_i}{\partial p_i}\right)\right] = 0 . \tag{b}$$

Or, d'après les équations du mouvement (14), le dernier terme est nul :

$$\frac{\partial \dot{q}_i}{\partial q_i} + \frac{\partial \dot{p}_i}{\partial p_i} = \frac{\partial}{\partial q_i}\left(\frac{\partial \mathcal{H}}{\partial p_i}\right) - \frac{\partial}{\partial p_i}\left(\frac{\partial \mathcal{H}}{\partial q_i}\right) = 0 . \tag{c}$$

On retrouve donc la condition (26).

16. Nous supposons ici que la fonction de Hamilton $\mathcal{H}(q, p)$ ne dépend pas explicitement du temps (système « conservatif »).

17. On a effet dans ce cas

$$\frac{\partial w}{\partial q_i} = W' . \frac{\partial \mathcal{H}}{\partial q_i} , \tag{a}$$

$$\frac{\partial w}{\partial p_i} = W' . \frac{\partial \mathcal{H}}{\partial p_i} \tag{b}$$

(où W' est la dérivée de la fonction W), de sorte que la condition (30) est évidemment satisfaite.

18. On pourrait prendre pour w une *fonction des constantes du mouvement* du système.

Une constante du mouvement est une grandeur physique G qui ne dépend pas explicitement du temps :

$$\frac{\partial G}{\partial t} = 0 \tag{a}$$

et qui ne change pas de valeur lorsqu'on suit l'évolution temporelle du système.

.../...

Références

1. J.B. Marion, *Classical Dynamics of Particles and Systems,* Academic Press, New York (1965).
2. H. Goldstein, *Mécanique Classique,* P.U.F., Paris (1964).
3. L.D. Landau et E.M. Lifshitz, *Mécanique,* Éditions Mir, Moscou (1969).

Une telle grandeur vérifie donc

$$\frac{\partial G}{\partial q_i}\frac{\mathrm{d}q_i}{\mathrm{d}t}+\frac{\partial G}{\partial p_i}\frac{\mathrm{d}p_i}{\mathrm{d}t}=0\,, \tag{b}$$

c'est-à-dire, compte tenu des équations du mouvement (14),

$$\frac{\partial G}{\partial q_i}\frac{\partial \mathcal{H}}{\partial p_i}-\frac{\partial G}{\partial p_i}\frac{\partial \mathcal{H}}{\partial q_i}=0\,. \tag{c}$$

Pour un système conservatif (note 16), la fonction de Hamilton $\mathcal{H}$ est elle-même une constante du mouvement, ce qui traduit la conservation de l'énergie. L'impulsion totale d'un système isolé, ainsi que son moment cinétique total, sont également des constantes du mouvement. Les conditions (29)-(30) font apparaître les distributions d'équilibre possibles comme des constantes du mouvement.

APPENDICE IV

La méthode des multiplicateurs de Lagrange

I. Recherche des points stationnaires d'une fonction de plusieurs variables indépendantes

II. Recherche des points stationnaires d'une fonction de plusieurs variables liées
1. Étude du problème
2. La méthode de Lagrange

I. RECHERCHE DES POINTS STATIONNAIRES D'UNE FONCTION DE PLUSIEURS VARIABLES INDÉPENDANTES

Soit

$$y = f(x_1, x_2, \ldots, x_n) \tag{1}$$

une fonction continue, dérivable, de n variables indépendantes x_i et

$$\mathrm{d}y = \frac{\partial f}{\partial x_1}\,\mathrm{d}x_1 + \frac{\partial f}{\partial x_2}\,\mathrm{d}x_2 + \ldots + \frac{\partial f}{\partial x_n}\,\mathrm{d}x_n \tag{2}$$

sa différentielle. Le « point » (d'un espace à n dimensions)

$$(x_1^{(0)}, x_2^{(0)}, \ldots, x_n^{(0)}) \tag{3}$$

est un *point stationnaire* de la fonction y si les valeurs $x_i^{(0)}$ annulent la différentielle de y.

La recherche des points stationnaires est aisée en principe. Les x_i étant des variables *indépendantes*, la condition

$$\mathrm{d}y = \frac{\partial f}{\partial x_1}\,\mathrm{d}x_1 + \frac{\partial f}{\partial x_2}\,\mathrm{d}x_2 + \ldots + \frac{\partial f}{\partial x_n}\,\mathrm{d}x_n = 0 \tag{4}$$

doit être vérifiée quels que soient les $\mathrm{d}x_i$, c'est-à-dire que chacune des dérivées partielles doit s'annuler. Les n coordonnées $x_i^{(0)}$ $(i = 1, 2, \ldots, n)$ des points stationnaires sont donc données par le système de n équations[1]

$$\frac{\partial f}{\partial x_j} = 0 \; ; \quad j = 1, 2, \ldots, n \; . \tag{5}$$

1. Assez souvent, on désire savoir en plus, pour chacun des points stationnaires trouvés, s'il s'agit d'un minimum, d'un maximum ou d'un col, et, dans le cas d'un extremum, s'il est absolu ou seulement local.

Le point stationnaire est un extremum si la différentielle seconde de y garde un signe constant, quels que soient les $\mathrm{d}x_i$, autour du point considéré. Pour savoir si cet extremum est local ou absolu, il n'y a pas d'autre méthode que de le comparer aux autres extremums de même nature.

II. RECHERCHE DES POINTS STATIONNAIRES D'UNE FONCTION DE PLUSIEURS VARIABLES LIÉES

Nous cherchons à nouveau les points stationnaires de la fonction (1), mais maintenant dans des cas où les variables x_i ne sont pas indépendantes.

1. ÉTUDE DU PROBLÈME

Supposons d'abord que les n variables x_i soient liées par une seule relation, que nous écrivons sous la forme

$$R(x_1, x_2, \ldots, x_n) = 0 . \tag{6}$$

Parmi les n variables figurant dans y, $(n-1)$ seulement sont alors indépendantes; en particulier, les n accroissements $\mathrm{d}x_i$ doivent être tels que

$$\mathrm{d}R = \frac{\partial R}{\partial x_1}\,\mathrm{d}x_1 + \frac{\partial R}{\partial x_2}\,\mathrm{d}x_2 + \ldots + \frac{\partial R}{\partial x_n}\,\mathrm{d}x_n = 0 . \tag{7}$$

Le point stationnaire cherché doit donc remplir à la fois les conditions (4) et (7).

Il est évidemment possible, en principe, de se ramener au cas de variables indépendantes : on peut utiliser la relation (6) pour exprimer l'une des variables, disons x_n, en fonction des $(n-1)$ autres. Pour trouver les points stationnaires, il suffit en fait de tirer $\mathrm{d}x_n$ de (7) et de le reporter dans (4). L'expression ainsi obtenue pour $\mathrm{d}y$ ne comporte plus que les $(n-1)$ différentielles indépendantes $\mathrm{d}x_1, \mathrm{d}x_2, \ldots, \mathrm{d}x_{n-1}$. Le coefficient de chacune d'elles doit être nul aux points stationnaires, ce qui donne $(n-1)$ équations. En leur adjoignant (6), on forme ici aussi un système de n équations pour déterminer les valeurs $x_i^{(0)}$ des n variables.

Mais on peut également procéder comme suit. Combinons linéairement les deux conditions (4) et (7) : pour tout λ, on doit avoir

$$\mathrm{d}y + \lambda\,\mathrm{d}R = \left(\frac{\partial f}{\partial x_1} + \lambda\,\frac{\partial R}{\partial x_1}\right)\mathrm{d}x_1 + \left(\frac{\partial f}{\partial x_2} + \lambda\,\frac{\partial R}{\partial x_2}\right)\mathrm{d}x_2 + \ldots + \left(\frac{\partial f}{\partial x_n} + \lambda\,\frac{\partial R}{\partial x_n}\right)\mathrm{d}x_n = 0 . \tag{8}$$

Seuls les $(n-1)$ premiers accroissements $\mathrm{d}x_i$ sont indépendants. Mais supposons que l'on arrive à choisir le paramètre λ de façon que, pour $x_i = x_i^{(0)}$ $(i = 1, 2, \ldots, n)$, on ait[2]

$$\frac{\partial f}{\partial x_n} + \lambda\,\frac{\partial R}{\partial x_n} = 0 . \tag{9}$$

Le terme en $\mathrm{d}x_n$ disparaît alors de la condition (8). Celle-ci devant être vérifiée quelles que soient les valeurs de $\mathrm{d}x_1, \mathrm{d}x_2, \ldots, \mathrm{d}x_{n-1}$, on en tire les $(n-1)$ équations

$$\begin{cases} \dfrac{\partial f}{\partial x_1} + \lambda\,\dfrac{\partial R}{\partial x_1} = 0 , \\ \ldots\ldots \\ \dfrac{\partial f}{\partial x_{n-1}} + \lambda\,\dfrac{\partial R}{\partial x_{n-1}} = 0 . \end{cases} \tag{10}$$

2. Ceci suppose le problème résolu : il faut connaître les $x_i^{(0)}$ pour calculer λ. Mais on va aboutir à un système d'équations déterminant simultanément les $x_i^{(0)}$ et λ.

Finalement, en associant la relation initiale (6) à (9) et (10), on obtient un système de $(n+1)$ équations :

$$\begin{cases} \dfrac{\partial f}{\partial x_j} + \lambda \dfrac{\partial R}{\partial x_j} = 0 \ ; \quad j = 1, 2, \ldots, n \ , & \text{(11.a)} \\ R(x_1, x_2, \ldots, x_n) = 0 \ , & \text{(11.b)} \end{cases}$$

dont les solutions donnent les n valeurs $x_i^{(0)}$ et celle du paramètre λ supplémentaire.

La seconde façon de procéder, plus simple et plus élégante que la première car elle traite les n variables x_i de façon symétrique, se généralise facilement au cas où ces variables sont liées par plusieurs relations. Elle est à la base de la *méthode de Lagrange.*

2. LA MÉTHODE DE LAGRANGE

On cherche les points stationnaires de la fonction (1) dans le cas où les n variables x_i sont liées par m relations $(m < n)$:

$$R_\alpha(x_1, x_2, \ldots, x_n) = 0 \ ; \quad \alpha = 1, 2, \ldots, m \ . \tag{12}$$

On introduit pour cela m *« multiplicateurs de Lagrange »* λ_α, et l'on construit la fonction

$$\widetilde{y} = f(x_1, x_2, \ldots, x_n) + \lambda_1 R_1(x_1, x_2, \ldots, x_n) + \lambda_2 R_2(x_1, x_2, \ldots, x_n) + \ldots + \lambda_m R_m(x_1, x_2, \ldots, x_n) \ . \tag{13}$$

On cherche ensuite les valeurs $x_i^{(0)}$ des variables x_i qui rendent cette fonction $\widetilde{y}$ stationnaire, *en traitant les n variables comme si elles étaient indépendantes.* Ceci donne les n équations suivantes (analogues à (11.a)) :

$$\frac{\partial f}{\partial x_j} + \lambda_1 \frac{\partial R_1}{\partial x_j} + \lambda_2 \frac{\partial R_2}{\partial x_j} + \ldots + \lambda_m \frac{\partial R_m}{\partial x_j} = 0 \ ; \quad j = 1, 2, \ldots, n \ , \tag{14}$$

dont les solutions dépendent des m paramètres λ_α. On détermine finalement ceux-ci en utilisant les relations (12)[3], comme en (11.b)[4].

---

3. Dans les applications pratiques, cette dernière étape n'est pas toujours indispensable : on se contente souvent de l'expression des n $x_i^{(0)}$ en fonction des m paramètres λ_α.

4. Il n'est pas difficile de vérifier que l'on résout bien ainsi le problème posé initialement, c'est-à-dire qu'on détermine les points stationnaires de la fonction y des $(n-m)$ variables indépendantes. Dans le cas d'une seule relation R, par exemple, l'élimination de dx_n entre (7) et (2) donne

$$dy = \frac{\partial f}{\partial x_1} dx_1 + \frac{\partial f}{\partial x_2} dx_2 + \ldots + \frac{\partial f}{\partial x_{n-1}} dx_{n-1} - \frac{\partial f}{\partial x_n} \times \frac{1}{\partial R / \partial x_n} \left[\frac{\partial R}{\partial x_1} dx_1 + \frac{\partial R}{\partial x_2} dx_2 + \ldots + \frac{\partial R}{\partial x_{n-1}} dx_{n-1} \right] , \tag{a}$$

de sorte que les points stationnaires de y sont déterminés par les équations

$$\frac{\partial f}{\partial x_1} - \frac{\partial f}{\partial x_n} \frac{1}{\partial R / \partial x_n} \frac{\partial R}{\partial x_1} = 0 \ ,$$

.....

$$\frac{\partial f}{\partial x_{n-1}} - \frac{\partial f}{\partial x_n} \frac{1}{\partial R / \partial x_n} \frac{\partial R}{\partial x_{n-1}} = 0 \ . \tag{b}$$

C'est bien ce que l'on obtient en calculant λ à partir de (9) et en reportant son expression dans les équations (10).

Index

Lorsque n, e *ou* f *suit un numéro de page, le renvoi concerne une* note, *un* exercice *ou une* figure.

Absorbant (pouvoir) 912, 916, 917
Absorption 907, 911
Accepteur (état) 845, 849
Accessibles (états) 147, 151
Action de masse (loi d') 516e
Addition (loi d') 950
Additivité
— de l'entropie 152
— de l'entropie statistique **11**, 99, 104, 127
Adiabatique
— désaimantation 189, 317
— paroi 160, 247e
— transformation 215, 219
Adsorption 676, 758e
— avec dissociation 684, 759e
Agitation thermique 26, 572
Aimantation 309, 449
— diamagnétique 319
— ferromagnétique **445**f, 446, **448**f, 456, 466f, 472
— paramagnétique 313
Aléatoire (variable) 949
Alliages 478
Amplitudes de probabilité 596
Analyse dimensionnelle 23, 793n
Anharmonicité des vibrations
— cristallines 398
— moléculaires 348, 508e
Antiferromagnétisme 525e
Anti-particule 937e
Approximation par exponentiation 63, **65**n, 154, 155n, 264
Archimède (poussée d') 761ee
Astrophysique(applicationssimples à l') 936e...
Autocohérence (équation d')
— pour les alliages 489
— pour le ferromagnétisme **454**, 457f, 461f, 462f
— pour un gaz sur réseau 768e
— pour des particules chargées 437, 695, 700n
Avogadro (nombre d') 4

Bande
— interdite, permise 89, 835
— de conduction 92, 835, 845, 848
— d'énergie 777n, 835
— de valence 835, 845, 848
Bernoulli (système de) 566, 569
Bethe-Peierls (approximation de) 488
Big-bang 922
Bilan détaillé **547**, 548, 549, 908, 914n
Binomiale (distribution) 958
— approximation gaussienne de la 67, 69, 70f
— limite des grands nombres pour la 62
Bit 102
Bleu du ciel 764e
Bloch
— loi de 944e
— théorème de 88
Boltzmann
— constante de 9, 81, 510e
— équation de 603, **611**, 636e...
Born-Oppenheimer (approximation de) 334
Born-von Karman (conditions de) 90
Bose
— condensation de **801**, 805, 808, 814, 892n, 926e, 928e, 929e
— facteur de 794, 795f
— température de 799, 874, 928e
Bose-Einstein (distribution de) 244e, 671, 762e
Bosons **289**, 342, 398, 757e..., **978**
— gaz de 794, 798, 925e...
- pour $T > T_B$ 800, 875
- pour $T < T_B$ 801, 808, 892, 895
- pour T voisine de T_B 874
Boulanger (transformation du) 567
Bragg-Williams (approximation de) 483, 528e
Bras de levier (règle du) 430n, 414n
Brillouin
— fonctions de 314, 322
— zone de **88**, 384, 387, 389
Brisure spontanée de symétrie 447n
Brownien (mouvement) 72, 79, 632e...

Canonique
— distribution **259**, 299, 533e, 548, 718
— ensemble 260
— entropie 272
— opérateur-densité 732
— situation 255
— température 257
Canoniques (équations) 984
Caoutchouc 371
Capacité calorifique 269, 165n
— électronique 839
— ferromagnétique 460, 472, 473f
— paramagnétique 315, 316f
— d'adsorption 679, 683
— d'un alliage **480**f, 486, 488f, 493, **496**f
— de bosons libres 809, 813f, 874
— de fermions libres 789
— d'un gaz parfait classique **297**, **331**, 336, 342, **348**
— d'un gaz réel 421, 765e
— de l'hélium 4 liquide 885, 892, 899
— à pression constante 752
— des solides 378, 392
Caractéristique (fonction) 59, 132e
Celsius (température) 159n
Centrifuge (force et potentiel) 228

Chaleur 160, 214, 497, **502**
— latente
- dans la condensation de Bose 928e
- de fusion 412n, 882n
- de sublimation 514e
- de vaporisation 141e, 411, **431**, 513e, 769e, 941e
- ordre de grandeur des 45, 141e
— spécifique : voir « capacité calorifique »
— d'adsorption 680, 684
— de réaction 44, 518e
— diffusion de la 890
— sens des échanges de 165
— supraconductivité de la 890, 891nn
Champ extérieur 688, 694, 761e, 767e
Champ moléculaire 454
Champ moyen (approximation de)
— pour les alliages 483, 488
— pour l'antiferromagnétisme 525e
— pour les électrons dans un cristal 88, 837n
— pour le ferromagnétisme **453**, 470, 474
— pour les gaz réels 418
— pour les plasmas ou électrolytes 436, 627
— pour les vibrations cristallines 379
Chandrasekhar (limite de) 872, 939e
Chaos moléculaire **596**, 605, 609, 617, 643e
Chapman-Kolmogorov (équation de) 592
Chargées (particules) 434
Chimisorption 676n, 681n
Cinétique
— pression 355, 356
— théorie des gaz 350, 509e...
Clapeyron (formule de)
— pour la condensation de Bose 928e
— pour la fusion 882n
— pour la sublimation 514e, 760e
— pour la vaporisation 411n, **513**e, 769e, 941e
Classique (mécanique)
— aimantation en 323
— description d'un système en 24, 981
— distribution canonique en 299, 520e...
— distribution microcanonique en 236
— domaine de validité de la 44, 48, 303
— entropie microcanonique en 240
Clausius (inégalité de) 499
Coercitif (champ) 446
Collisions (dans un gaz)
— fréquence des 604, 134e, 140e
— intégrale de 609, 615
— rôle des 604, 636e...
— section efficace de 573, 606
Combinaison linéaire d'orbitales atomiques 94n
Compressibilité électronique 840
Compression (module de) 531e, 841
Concavité (de l'entropie statistique) 106, 128
Concentration 400, 412
— gradient de 79n
Condensat 892, 894, 904
Conditions aux limites
— périodiques 18, 384, 819, 972
— strictes 971
— de Born-von Karman 90
Conducteur (équilibre d'un) 694
Conduction (électrons de) 49, 92, 579
Conductivité 639e, 834
Configuration
— énergie de 480
— intégrale de 401
Conjugué (moment ou impulsion) 969, 983
Conservatif (système) 967, 988n
Conservation (lois de) 146, 215, 583
— locale 612, 628, 988
Constante du mouvement 988n
Contact
— entre deux systèmes 159, 161, 168
— d.d.p. de 699, 701n
Continuité (équation de) : voir « conservation locale »
Cooper (paires de) 808, 880n, 928e
Corde vibrante 511e
Corps blanc, gris 914
Corps noir 818, 913, 919, 930e
— densité spectrale du 825, 827f
— rayonnement du 916, 918
Corrélations 752, 755
Corrélation (fonction de) 361, 591, 634e, **708**
Cosmologique
— modèle standard 922, 924
— rayonnement 919, 939e
Couplage
— thermique 160
— entre systèmes macroscopiques 160, 173
Cristal
— défauts dans un 242ee
— états électroniques dans un 88, 835, 844
— ordres de grandeur pour un 39, 41
Critique(s)
— coordonnées 424-425, 524e, 768e
— exposants 449, 465, 472, 476, 527e
— facteur de compression 425, 524e
— point 272n, 884f
— température
- magnétique **445**, 457, 476, 524e, 525e
- pour un alliage **479**, 486, 492, 529ee
- pour un fluide 424
Cryoscopie 412n
Curie
— loi de 315
— température de **445**, **446**, 458, 472
Curie-Weiss (loi de) 462

Dalton (loi de) 193, 506e
De Broglie (longueur d'onde thermique de) 296, 794n
Debye
— approximation de 390
— fréquence de 41
— longueur de 438, 697n, 934e
— température de **391**, **392**, 840, 897
Debye-Hückel (approximation de) 436
Dégénéré (gaz quantique) 49, 758e, **793**, 839, 865, 926e
Dégénérescence (degré de) 13, 500
Degrés de liberté
— internes 330, 331, 347
— gel des 331, 532e
— voir « électroniques », « rotation », « translation », « vibrations »
Démixtion 529e

Densité d'états 15, 151
— électroniques **91**, 92, 96
— d'un gaz parfait 20, 82, 197
— de N éléments faiblement couplés 23
— d'une particule libre 17, 138e...
— des phonons 397, 897
— des photons 821
Densité de modes normaux (ou propres)
— électromagnétiques d'une enceinte 820
— pour les vibrations cristallines 389, 391
Densité de probabilité 949
— dans l'espace des phases 236, 299, 621
Désaimantation adiabatique 189, 317
Désordonnée (phase) 447n, 478, 494
Désordre maximum (principe du) 150, 554
Déterministe (équation) 558, 568, 596
Diagramme d'état, ou de phase : voir « phase »
Diamagnétisme **309**, **318**, 325, 841n, 932e
Diatherme (cloison) 160
Diatomiques (molécules) 334, 339, 341, 342
Diffusion
— de la chaleur 890
— coefficient de **79**, 81n, 135e, 633e, 635e
— équation de la 635e
Dilatation (coefficient de) 377, 531e
Dipolaire
— interaction 176n, 451
— moment électrique 42, 507e
Discernables (particules) 190, **287**, 300, 380, 883
Dispersion (relation de)
— pour un cristal à une dimension 384, 395
— pour un cristal à trois dimensions 387, 391, 396
— pour les ondes de spin 943e
— pour les quasiparticules dans l'hélium II 895, 896f
Dissociation
— énergie de 38, 390n
— équilibre de 515e
Distribution
— canonique **259**, 299, 533e, 548, 718
— canonique généralisée 722
— grand-canonique 650
— microcanonique **145**, 148, 236, 546
— statistique d'une fonction de variables aléatoires 53
— $T-p$ 368
— voir « Bose-Einstein », « Fermi-Dirac », « Maxwell-Boltzmann », « Planck »
Domaines de Weiss 446
Donneur (état) 845, 848
Dopage 934e
Doppler
— effet 920n, 921n
— élargissement des raies 354, 509e
Double couche (de charge) 524e, 853, 857f
Dulong et Petit (loi de) **378**, 381, 390, 393, 396
Dynamique (pression) 355, 360

Ébullioscopie 412
Écart quadratique moyen 957, 961
Échange(s)
— interaction d' 451, 453n
— sens des 165, 170, 220, 278
Échelle (loi d') 475, 527e
Écrantage 173, 176n, **438**, **696**, 699n, 933e
Effusion 354, 509e
Ehrenfest (modèle d') 640e
Einstein
— coefficients d' 906
— modèle d' 379, 396
Électrolytes 435, 440, 443, 524e
Électroniques
— degrés de liberté 333, 347, 532e
— états dans un cristal : voir « cristal »
Électrons
— libres
- modèle des 92, 837, 838n
- diamagnétisme des 320
— de conduction 49, 92, 579
Émission 911
— induite ou stimulée 907, 909
— spontanée 906, 909
Émissivité 911, 916, 917
Énergie
— « chimique » 214
— interne 222, 229
— minimale (condition d') 250e, 889
— moyenne 268, 656
— distribution statistique de l' 262, 654
— écart quadratique moyen sur l' 268, 657
Énergie libre 268
— équilibre comme minimum de l' 277, 555
Ensemble statistique 30, 641e
— canonique 260
— grand-canonique 651
— microcanonique 149
— $T-p$ 365, 507e
Enthalpie libre 368, 399, 405
Entropie
— canonique 272
— grand-canonique 658
— microcanonique **150**, 240, 242e..., 248e
— microcanonique partielle 154, 163n
— statistique **8**, 99, 103, 110, 124
— d'un gaz parfait **190**, 197, 298, 674
— de mélange 190, 193, 198
— d'un système totalement isolé 230
— additivité de l' 152
— création d' 616
— équilibre comme maximum de l' 158, 553
— évolution de l' 552, 613
— invariance de l' 223
— variations d' 203, 217, 497
Équation d'état
— magnétique 449
— d'un gaz parfait classique 191, **298**, **331**, 359
— d'un gaz parfait de bosons 801, 811
— d'un gaz parfait de fermions 781, 792
— d'un gaz de photons 831n
— d'un gaz réel 364, **421**, 523e, **712**, 714, 768e
— d'un solide 530e
Équation maîtresse
— établissement de l' 539, **543**, 590, 595
— propriétés de l' 582, 630e...
Équilibre
— macroscopique 32, 122
— liquide-vapeur **428**, 512e, 519e, 940e

— solide-vapeur 514e, 759e
— de l'Univers primordial 762e, 936e
— conditions d' : voir ci-après
— distributions d'— d'un gaz dilué 617, 637e
— évolution vers l' 549, 588
— solutions d'— de l'équation maîtresse 546, 587, 762e
Équilibre (condition d')
— «chimique» 515e, 684, 691, 758e..., 889
— dans un champ extérieur 688, 691
— comme minimum de l'énergie 250e, 889
— dans un système isolé 162, **169**, 209, 225, 245e
— à température fixée 277, 512e...
Équipartition de l'énergie
— applications de l' 337, 341, 353, 381, 828
— théorème d' 304, 520e
Ergodique
— principe 32, 149, 578, 583, 643e
— problème 560n, 562, 570
— système 137e, 562
Espace des phases 25, 985
— densité dans l' 236, 299
— évolution dans l' 607, 621
— volume élémentaire de l' 239, 301
État
— macroscopique 13, 25, 114
— microscopique 13, 24
— pur 12, 121
— stationnaire 968, 988
États correspondants (loi des) 425n
Étoile à neutrons 873n, 939e
Événement contraire ou complémentaire 951
Évolution
— macroscopique 27, 30, 123
— spontanée 156, **158, 277, 371,** 554, 556, **662**
— d'une distribution de probabilités 584
— de l'entropie 552, 613
— vers l'équilibre 549, 588
— dans l'espace des phases 607, 621
— des fonctions de répartition 623
— irréversibilité de l' 556, 617, 640e
— postulat d' 539, 590
Excitations collectives 382, 893
Expansion de l'Univers 920, 924
Extensive (grandeur) 156n, **177**, 249e, 863
Extraction (travail d') 702, 931e

Fermi
— énergie (ou niveau) de **783**, 839, 930e
— fonction (ou facteur) de 774, 775f, 776f
— impulsion de 868
— sphère, surface de 783, 783n, 784f
— température de **783**, 793, 839, 868, 882
— vecteur d'onde de 783, 838
— vitesse de 839
Fermi-Dirac (distribution de) 244e, 670, 762e, 774
Fermions **289**, 342, 757e..., **978**
— gaz de 774, 778, 925e...
- à $T=0$ **783**, 842, 848, 868, 925e, 926e
- à basse température 785, 925e
Ferrimagnétisme 444n
Ferroélectrique 45
Ferromagnétisme 45, **444**, 523e, 942e
Fluctuations 7, 26, 29, 156, 510e..., **745**, 764e...
— critiques 272n, 751
— de la densité 750, 764e
— de l'énergie 265, 751
— du nombre de particules 7, 748
— de la température 754
— d'une variable interne 156, 272, 280, 745
— du volume 749
Fluctuation-dissipation (théorème de) 634e
Fluide classique (évolution d'un) 621
Fluide simple 270, 399, 661
— dans un champ extérieur 693
— fluctuations dans un 748
Fonction d'onde 966
Fonction de partition 260, **267**, 300, 732
— généralisée 723, 736
— $T-p$ 368
— factorisation de la 274, 288, 291
Fontaine (effet) 888
Force conjuguée d'un paramètre extérieur 202, 207, 532e
Forces intermoléculaires 50, 174, 416
Frenkel (défauts de) 242e
Friction (coefficient de) 634e
Fugacité 299, 331, 517e, 655
Fusion (courbe de) 884f

Gap 835, 837, 844
Gaussienne
— approximation **67**, 69, 70f, **155**, **264**, 653, 663
— distribution 133e, 636e, 962
— comme limite 963
Gaz d'électrons 49, 92, 320, 579
Gaz parfaits classiques
— constante des 297
— définition des 47, 293, 329
— équation d'état des 191, **298**, **331**, 359
— ordres de grandeur dans les 47, 572
— propriétés des 197, 245e, 247e, **297**, **331**, 506e..., 673, 781n
Gaz parfaits quantiques **773**, 925e..., 936e
— voir «bosons (gaz de)», «fermions (gaz de)»
Gaz réels
— dans le champ de pesanteur 767e
— équation d'état des 415, 523e, 704
— théorie cinétique des 350, 360, 603
Gibbs (paradoxe de) **190**, 195, 197n, 302
Grand-canonique
— distribution 650
— ensemble 651
— opérateur-densité 733
— potentiel chimique 649
— situation 647
— température 649
Grande fonction de partition 655, 733
— factorisation de la 659, 667, 681
Grand potentiel 655
— équilibre comme minimum du 662
Granulation
— de l'espace des phases 566n, 570n
— du temps 596, 600

Gravitationnelle
— interaction 140e, 862
— pression 863
Gyromagnétique (rapport) 180

H (théorème) 552, 554, 613
Hamilton (fonction de) 983
Hamilton-Jacobi (équations de) 984
Hamiltonien
— formalisme 983
— opérateur 967
Harmonique (approximation)
— dans les cristaux 386, 388, 398
— dans les molécules 336, 338, 348
Heisenberg (hamiltonien de) 450, 455
Hélium 878
— hélium 3 878, 880, 940e
- polarisé 883n, 940e
— hélium 4 807, 878, 884
— hélium I 885
— hélium II 807, 885, 891
- densité normale de l' 900, 904
- modèle des deux fluides pour l' 886, 900
- viscosité de l' 887, 894, 904
— liquéfaction et solidification de l' 878
Hertzprung-Russel (diagramme de) 861f
Hétéropolaire (molécule) 341, 508e
Hiérarchie BBGKY 623
Homopolaire (molécule) 342
Hooke (loi de) 376, 531e
Hubble
— constante de 921n
— loi de 920
Hydrogène (ortho et para) 242e, 516e
Hystérésis (cycle d') 446, 445f

Identiques (particules) 773, 977
Impulsion (opérateur) 966
Impuretés (états d') 844, 850
Indépendants (événements) 951
Indépendantes (particules)
20, 190, **286**, 309, 403, 418, 454, **667**, 773, **979**
Indiscernables (particules)
196, 241, 244e, **289**, 302, 621n
Inertie (moments principaux d') 227, 231, 345
Infinitésimal (système macroscopique) 221
Influence électrique 699
Information **8**, **98**, 108, 112
— perte d' 558, 617
Instabilité des équations dynamiques 559, 566, 577
Intensive (grandeur) 156n, 177, 249e
Interaction
— effective **400**, 403, **418**, 453n, 818n
— entre molécules 50, 174, 416
— systèmes de particules en 173, 350, 378, 399, 415, 434, 444, 478, 522e..., 603, 765e...
Ioniques (systèmes) 140e, **176**, **434**, 439, 442, 627, 638e
Ionisation (énergie d') 37, 38
Irréversibilité
— de l'évolution spontanée 556, 617, 640e...
— d'une transformation 217
— aspect stochastique de l' 558

Isentropique (transformation) 188n, **220**, 317, 927e
Ising (modèle d') **466**, 482, 524e, 526e, 527e, 769e
Isolant 834
Isolé (système) **146**, 147n, 232, 539, 572
— totalement 221, 224, 230
Isothermes (courbes)
— d'adsorption 678, 682, 685
— pour un fluide 422, 429f, 815f
— pour un gaz de bosons 814, 815f, 928e

Jonction $p-n$ 852
Joule
— détente de 156, 213n, 502n
— expérience de 503n

K-système 566
Kac (anneau de) 641e
KAM (théorème) 564
Kappler (expérience de) 510e
Kelvin (degré) 159
Kirchhoff (loi de) 916

Lagrange
— équations de 982
— fonction de 982
— multiplicateurs de 10n, 131e, 225, 720, 721, **990**
Lagrangien (formalisme) 981
Lambert (loi de) 916, 917
Landau
— diamagnétisme de 841, 932e
— longueur de 140e, 439
— modèle de — pour l'hélium II 895
— niveaux de 320
Langevin
— équation de 632e...
— fonction de 322, 324, 375
— magnétisme de 323, 327
Langmuir (isothermes de) 679
Laser 909
Le Chatelier (principe de) 680n
Legendre (transformation de) 370n, 665n, 724, 984
Lennard-Jones (potentiel de) 51n, **417**, 420
Liaison (énergie de) 38, 348
Liaisons fortes (modèle des) 94
Libre (états quantiques d'une particule) 970
Libre parcours moyen 573, 922n
Limite centrale (théorème de la) 74, 133e, 964
Lindemann (critère de) 879n
Liouville
— équation de 621, 987
— théorème de 238n, 559n, 566n, 985

Macromolécules (masse des) 761e
Magnétisme 180, 309, 444
Magnéton de Bohr 43, 311, 187n
Magnons 942e
Maille élémentaire 86
Marche au hasard 72, 137e, 635e...
Markovien (processus) 73, 540, **592**, 597
Maser 909, 910n
Masse effective 838n, 844

Matrice-densité 120
Maxwell
— démon de 113
— distribution des vitesses de 350, 618
Maxwell-Boltzmann
— approximation de **196**, 199, **291**, 672
— distribution de 671
Mayer (relation de) 753n
Mécanique (système purement) 206, 210, 251e
Mélange 399, 519e
— idéal 191, 760e, 886
— statistique **13**, 114, 119, 120n
— entropie de 107n, **190**, 193, 198, 506e
— propriété de 565
Mélangeant (système) 562, 565
Membrane semi-perméable 409
Métastable (état) 432, 464, 551
Métaux (propriétés électroniques des) 694, **834**, 837, 838, 930e...
Microcanonique
— distribution **145**, 148, 236, 546
— ensemble 149
— entropie **150**, 240, 242e..., 248e
— opérateur-densité 730
— température **158**, 165, 166, 232, 242e
Micro-ondes 910
Mobilité (d'un corpuscule) 633e, 834
Modes normaux (ou propres)
— électromagnétiques 819
— d'un cristal 382, 385, 386, 388
— d'une molécule 338
Mole 4
Moléculaire (interaction) 50, 173, 416
Molécules linéaires 339, 345, 348
Moment
— cinétique
- conservation du 225, 230
- échanges de 232
— électrique 42, 507e
— magnétique 43, **180**, 311, 326, 518e
— d'une distribution statistique 958n
Monoatomiques (gaz) 190, 293, 415
Mouvement macroscopique 221, 226, 521e
Moyenne
— temporelle 29, 137e, 563
— sur un ensemble 30, 137e
— voir « valeur »
Multiplication (loi de) 951

Naines blanches 860
Néel (température de) 525e
Nernst (principe de) 500, 791
Neutrinos 936e, 939e
Neutrons
— diffusion de 895
— étoiles à 873n, 939e
Niveaux d'énergie 968
Nivellement barométrique 247e, 694, 767e
Nombre
— baryonique, leptonique 937e
— d'occupation **289**, 667, 758e, 980
- moyen 398, 669
— de particules
- moyen 656
- distribution statistique du 651
- écart quadratique moyen sur le 657

Observable 966
Ohm (loi d') 834
Opalescence critique 751
Opérateur-densité 114, 729, 769e...
Ordonnée (phase) 447n, 478, 494
Ordre (paramètre d') **447**n, 481, 483, 491
Ordre-désordre : voir « transition »
Ordres de grandeur
35, 44, 140e..., 244e..., 572, 837, 860, 862, 864
Oscillateur(s) harmonique(s) 134e, 137e, 243ee, 563, **969**
Osmotique (pression) 409, 443, 761e
Overhauser (effet) 631e

Paires (création de) 938e
Palier de liquéfaction 429f, 430
Paramagnétique
— cristal 180, 187n
— gaz 246e, 841, 931e
Paramagnétisme 309, 445
— étude statistique du 311, 841, 931e
— théorie semi-classique du 323
Paramètres extérieurs 145, 161, 230
Pauli
— paramagnétisme de 841, 931e
— principe de **289**, 451, 579, 784, 790, 882, **979**
Périodiques (conditions aux limites) 18, 384, 819, **972**
Permittivité 435, 507e
Perturbations
— extérieures aléatoires 574, 578, 580, 598
— méthode des 474, 738
Pesanteur (champ de) 694, 767e
Peuplement (d'un état) **539**, **546**, 547, 610, 908
Phases
— aléatoires (hypothèse des) 597, 600
— diagramme de 413, 530e, 816, 881f, 884f
— équilibre entre 410, 414n, **426**, 431n, 529e, 817
— voir « désordonnée », « espace », « ordonnée », « transition »
Phonons **397**, 895, 897, 903
Photons **818**, 906, 913n, 914, 929e...
— densité d'états des 821
— fonction de distribution des 823
— potentiel chimique des 824
— propriétés statistiques des 822
— spin des 821
Physisorption 676n, 681n, 758e
Piège 681, 686, 934e
Pile ou face 947, 958, 961
Planck (distribution ou loi de) **825**, 827f, 917, 923, 929e
Planck-Einstein (relations de) 821
Plasma 140e, **434**, 439, 442
— hors d'équilibre 627, 638e
Poincaré (théorème de) 566n, 640e, 641e
Point zéro (énergie de) 390
Poisson
— crochet de 622
— distribution de 62n, 131e, **135**e, 758e
Polaire (fluide) 177n
Polarisation
— dynamique 631e

— électrique 507e
— de l'hélium 3 883n, 940e
— d'un mode normal 387
— d'une onde électromagnétique 820
Polyatomiques
— gaz parfaits 329
— mailles cristallines 393
— molécules 337, 345, 346n, 416n
Polymères linéaires 371, 531e
Pomerantchuk (effet) 883
Population (inversion de) 184
Porteur (de charge) 846, 848, 853
Position (opérateur) 966
Positron 937e, 938e
Postulat fondamental **147**, 148n, 234, 237, 554, 725
— d'évolution 539, 590
— généralisé 717
Potentiel chimique
— canonique 273, 652n
— grand-canonique 649, 652n
— interne 226, 228
— local 691, 693
— microcanonique 167, 170
— d'un fluide simple 399
— d'une solution diluée 408
Potentiel électrochimique 700n, 701n
Pression
— canonique 273
— cinétique, dynamique 355
— grand-canonique 659
— microcanonique 167, 170, 207
— osmotique 409, 443, 761e
— partielle 193, 506e
— quantique 784, 864
— de radiation 831, 866n
— de vapeur saturante 430, 769e
Principe(s) de la thermodynamique 201
— premier 215
— second 220, 499
— troisième 331n, **500**, 791
Probabilité(s) 947
— conditionnelles 591, 952n
— applications des 134e...
— définition des 948
— densité de 949
— évolution des 583
— propriétés des 130e...
Propagation (équation de) 387

Quantique(s)
— mécanique 965
— statistiques 670
Quantité de mouvement
— conservation de la 225, 230
— échanges de 232, 235n
Quasiparticules **397**, 837n, 893, 895, 900, 943e
Quasistatique (transformation) 215

Raie (largeur de) 907n, 354, 509e
Raoult (lois de) 412
Rayleigh-Jeans (formule de) 826, 829
Rayon atomique 37, 38
Rayonnement
— cosmique fossile 919, 939e
— thermique **825**, 829, 911, 917
— du corps noir 916, 918
— thermodynamique du 818
Redresseur $p-n$ 858
Réflexion (théorème de) 75
Relativiste(s)
— fermions 866, 926e
— gaz parfait 506e
— particule libre (densité d'états d'une) 139e
— transformation — de la température 223n
Relaxation (temps de) 187, **550**, 589, 638e
Renormalisation (méthode du groupe de) 475, 527e
Répartition (fonction de) **56**, 57, 607, 623
Renversement du temps
— dans les équations cinétiques 627
— dans l'équation maîtresse 540, 549, 763e
— dans les interactions microscopiques 541, 542nn, 835n
— dans des modèles simples 568, 643e
Réseau
— cristallin 85
— de Bravais 85, 86n
— réciproque 87, 387
— gaz sur 767e
Réservoir
— d'énergie 255
— d'énergie et de particules 648
— d'énergie et de volume (ou de longueur) 366, 371
Résonance magnétique 631e
Réversible (transformation) 188n, **216**, 218, 317, 498, **500**, 503
Richardson-Dushman (loi de) 931e
Rotateur rigide (approximation du) 339, 347
Rotation
— degrés de liberté de **339**, 348, 507e, 508e, 520e, 521e, 532e
— énergie de (ordre de grandeur de l') 41, 42
Rotons 897, 898, 903
— loveurs 1066
Rouge (décalage vers le) 920, 921n
Rydberg 37

Sackur-Tétrode (formule de) 197, 298, 674
Saturation (courbe de) 429f
Schottky (défauts de) 242e
Schrödinger (équation de) 542n, 590, 967
Section efficace 573, 606, 922n
Semi-conducteur 512e, **834**, 844, 934e...
— extrinsèque 845, 848, 936ee
— intrinsèque 845, 846, 935e
— non dégénéré 847
— densités de porteurs dans un 846, 848, 850, 853
Shannon (formule de) 102
Smoluchowski (équation de) 636e
Solide indéformable 227
Solutions
— diluées 399, 518e, 760e
— solides 478
Sommerfeld (développement de) 785
Son
— second 890, 904
— vitesse du 384, 387

Sphères dures (modèles des) 419n, 523e, 560n, 566
Spin
— états de 975
— ondes de 942e
— température de 187
Spinodale (courbe) 429f
Stabilité 107n
— de l'équilibre entre systèmes 164, 278, 662
— des naines blanches 870n
— et séparation en deux phases 426
Stationnaire
— état 968, 988
— point 990, 991
Statistique (distribution) 955
Stefan-Boltzmann (loi de) 829, 917, 929e
Stirling (formule de) 64n, 83
Stochastique (processus) 559, 590
Stosszahlansatz 597, 605
Structure
— fine 333, 507e
— hyperfine 333n, 631e
Sublimation 513e, 759e, 884f
Supercritique (fluide) 424
Superfluidité 807, 880n, **886**, **894**, 896
Superfuite 888
Supraconductivité 807, 890, 891nn
Surchauffé (liquide), sursaturée (vapeur) 432
Susceptibilité 310
— diamagnétique 319
— ferromagnétique 464
— paramagnétique 313, 462, 843
Symétrie
— brisure spontanée de 447n
— facteur de 345, 346
Symétrisation (postulat de) 978

Tchebycheff (inégalité de) 958n
Température
— canonique 257
— caractéristique des mouvements moléculaires 336, 338
— grand-canonique 649
— microcanonique **158**, 165, 166, 232, 242e...
— négative 180, 185, 245e, 246e
— dans un changement de référentiel 223
— comme multiplicateur de Lagrange 720
— de spin 187
— échelle de 46, 47, 159n
— obtention de basses 189, 317, 883
— unité de 159
— voir « Bose », « critique », « Curie », « Debye », « Fermi », « Néel »
Temps-énergie (relation d'incertitude) 578
Tension de surface 141e
Tensoriel (produit) 974
Thermique
— contact 159, 161
— équilibre 161
Thermodynamique
— limite 280
- des diverses descriptions 281, 282, 370, 663, 665
- pour un gaz de bosons 797
- pour un gaz de fermions 777
— voir « principes », « transformation »
Thermoionique (effet) 930e
Thermomécanique (effet) 888
Thermostat **255**, 275, 497, 532e, 543, 554
Thomas-Fermi
— longueur de 697n, 934e
— modèle de 699n
Tonométrie 412
Toupies (molécules polyatomiques) 346n
Trace (d'un opérateur) 116n
Transformation thermodynamique 201, 211
— adiabatique 215, 219
— infinitésimale 212, 217
— isentropique 220
— quasistatique 215
— réversible 216, 218
Transition
— entre états microscopiques 537
- probabilités de **538**, 592, 598, 600
— de phase
- ferro-paramagnétique 445
- gaz-liquide 415, 422
- λ 884
- ordre-désordre 478, 528e
- dans un gaz de bosons 812, 816f, 892
- dans l'hélium 878, 884, 892
- du premier ordre 430, 817n
- du second ordre **447n**, 471, 527e, 817n, 885
Translation (degrés de liberté de) 330, 347, 351
Transport (phénomènes de) 619, 639e
Travail **212**, 497, 502, 503n
Triple (point) 884f
Trou 845
— noir 873n

Univers
— primordial 762e, 936e
— expansion de l' 920
Universalité (des exposants critiques) **449**, 465, 475, 886n

Valeur
— moyenne 954, 956
— la plus probable 956
Van der Waals
— équation de 140e, 364, 415, **421**, 711, 766e
— forces de 174n, 879
Van Hove (singularités de) 92n, 96n
Van Leeuwen (théorème de) 327
Van't Hoff (loi de) 410
Vapeur saturante (pression de) 430, 769e
Vaporisation (courbe de) 429f, 513e, 884f
Variable interne 145, 162, 745
— distribution statistique d'une
- dans la situation canonique 262, 271, 512e...
- dans la situation grand-canonique 651, 654, 657
- dans la situation microcanonique 153, 245e..., 250e
— voir « fluctuations »

Variationnelle (méthode) 742
Vibrations
— dans les cristaux 378, 382
— dans les molécules
334, 348, 507e, 508e, 532e
— énergies de
(ordre de grandeur des) 40, 41
Viriel 704
— développement du 712, 766e, 767e
— théorème du 705
Vitesse quadratique moyenne 354, 572, 634e
Vlasov (équation de) 617n, 626
Volume exclu 418, 419

Wien
— loi de 826
— loi du déplacement de 827
Wigner-Seitz (maille de) 87

Young (module d') 140e

Zermélo (paradoxe de) 642e
Zéro absolu 159n

Ensemble	microcanonique	canonique	grand-canonique	$T-p$
Paramètres fixés	$E,\ V,\ N,\ \ldots$	$T,\ V,\ N,\ \ldots\quad (\beta \mathrel{\widehat{=}} 1/kT)$	$T,\ V,\ \mu,\ \ldots\quad (\beta \mathrel{\widehat{=}} 1/kT)$	$T,\ p,\ N,\ \ldots\quad (\beta \mathrel{\widehat{=}} 1/kT)$
Fonctions fondamentales associées	Nombre d'états accessibles : $\Omega(E,\ V,\ N,\ \ldots)$ Entropie : $S(E,\ V,\ N,\ \ldots) = k \ln \Omega$	Fonction de partition : $Z(T,\ V,\ N,\ \ldots) \mathrel{\widehat{=}} \sum_{(\ell)} \mathrm{e}^{-\beta E_\ell}$ Énergie libre : $F(T,\ V,\ N,\ \ldots) \mathrel{\widehat{=}} -kT \ln Z$	Grande fonction de partition : $\Xi(T,\ V,\ \mu,\ \ldots) \mathrel{\widehat{=}} \sum_{(\ell)} \mathrm{e}^{-\beta(E_\ell - \mu N_\ell)}$ Grand potentiel : $J(T,\ V,\ \mu,\ \ldots) \mathrel{\widehat{=}} -kT \ln \Xi$	Fonction de partition $T-p$: $\widetilde{Z}(T,\ p,\ N,\ \ldots) \mathrel{\widehat{=}} \int_0^\infty \mathrm{d}V \sum_{(\ell)} \mathrm{e}^{-\beta(E_\ell + pV)}$ Enthalpie libre : $G(T,\ p,\ N,\ \ldots) \mathrel{\widehat{=}} -kT \ln \widetilde{Z}$
Probabilité d'un état	$P_\ell = \begin{cases} \dfrac{1}{\Omega} & \text{si } E < E_\ell < E + \delta E \\ 0 & \text{sinon} \end{cases}$	$P_\ell = \dfrac{1}{Z} \mathrm{e}^{-\beta E_\ell}$	$P_\ell = \dfrac{1}{\Xi} \mathrm{e}^{-\beta(E_\ell - \mu N_\ell)}$	$\mathrm{d}P_\ell = \dfrac{1}{\widetilde{Z}} \mathrm{e}^{-\beta(E_\ell + pV)}\, \mathrm{d}V$
Relations importantes	$\mathrm{d}S = \dfrac{1}{T}\mathrm{d}E + \dfrac{p}{T}\mathrm{d}V - \dfrac{\mu}{T}\mathrm{d}N + \ldots$	$\overline{E} = -\dfrac{\partial}{\partial \beta} \ln Z$ $\mathrm{d}F = -S\mathrm{d}T - p\mathrm{d}V + \mu \mathrm{d}N + \ldots$ $F = \overline{E} - TS$	$\overline{N} = kT \dfrac{\partial}{\partial \mu} \ln \Xi$ $\overline{E} - \mu\overline{N} = -\dfrac{\partial}{\partial \beta} \ln \Xi$ $\mathrm{d}J = -S\mathrm{d}T - p\mathrm{d}V - \overline{N}\mathrm{d}\mu + \ldots$ $J = \overline{E} - \mu\overline{N} - TS$ (pour un fluide simple, $J = -pV$)	$\overline{V} = -kT \dfrac{\partial}{\partial p} \ln \widetilde{Z}$ $\overline{E} + p\overline{V} = -\dfrac{\partial}{\partial \beta} \ln \widetilde{Z}$ $\mathrm{d}G = -S\mathrm{d}T + \overline{V}\mathrm{d}p + \mu \mathrm{d}N + \ldots$ $G = \overline{E} + p\overline{V} - TS$ (pour un fluide simple, $G = \mu N$)

Définition statistique de l'entropie : $S \mathrel{\widehat{=}} -k \sum_{(\ell)} P_\ell \ln P_\ell$.

Distributions de Fermi-Dirac et de Bose-Einstein : nombre d'occupation moyen d'un état individuel (λ) :

$$\overline{N}_\lambda = \frac{1}{\mathrm{e}^{\beta(\varepsilon_\lambda - \mu)} \pm 1} \qquad \begin{matrix} + \text{ fermions} \\ - \text{ bosons} \end{matrix}$$

Unités et ordres de grandeur

Angström : $1\ \text{Å} = 10^{-10}\ \text{m}$

Électron-Volt : $1\ \text{eV} = 1{,}602\ldots \times 10^{-19}\ \text{J}$

Atmosphère : $1\ \text{atm} = 1{,}013\,25 \times 10^{5}\ \text{Pa}$

$kT = 1\ \text{eV} \iff T \simeq 12000\ \text{K}\ (11\,604{,}5(1)\ \text{K})$

$kT \simeq \dfrac{1}{40}\ \text{eV}$ à température ordinaire ($T \simeq 300\ \text{K}$)

Volume molaire (gaz parfait) :

22,4 ℓ à 0 °C sous 1 atm $(22{,}414\,10(19) \times 10^{-3}\ \text{m}^3)$

22,7 ℓ à 0 °C sous 10^5 Pa $(22{,}711\,08(19) \times 10^{-3}\ \text{m}^3)$

Formulaire

Formule de Stirling :

pour $N \gg 1$, $\quad N! \simeq N^N e^{-N} \sqrt{2\pi N}$

$\ln N! \simeq N \ln N - N$

Intégrales définies comportant une gaussienne :

$$I(n) \widehat{=} \int_0^{+\infty} dx\, x^n\, e^{-ax^2} \quad (a > 0)$$

$$I(0) = \frac{1}{2}\sqrt{\frac{\pi}{a}}, \qquad I(1) = \frac{1}{2a}, \qquad I(2) = \frac{1}{4}\sqrt{\frac{\pi}{a^3}},$$

$$I(3) = \frac{1}{2a^2}, \qquad I(4) = \frac{3}{8}\sqrt{\frac{\pi}{a^5}}, \qquad I(5) = \frac{1}{a^3}.$$

Fonction gamma :

$$\Gamma(z) \widehat{=} \int_0^{+\infty} dx\, x^{z-1}\, e^{-x} \quad (z > 0)$$

$$\Gamma(z+1) = z\,\Gamma(z)$$

$$\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}, \quad \Gamma(n+1) = n! \text{ pour } n \text{ entier positif}$$

Constantes physiques

L'incertitude sur les derniers chiffres significatifs est donnée entre parenthèses.

Nombre d'Avogadro :

$$N_A = 6{,}022\,136\,7\,(36) \times 10^{23}\ \text{mol}^{-1}$$

Constante de Boltzmann :

$$k = 1{,}380\,658\,(12) \times 10^{-23}\ \text{J/K}$$
$$= 8{,}617\,385\,(73) \times 10^{-5}\ \text{eV/K}$$

Constante des gaz parfaits :

$$R = N_A k = 8{,}314\,510\,(70)\ \text{J/K.mol}$$

Constante de Planck :

$$h = 6{,}626\,075\,5\,(40) \times 10^{-34}\ \text{J.s}$$
$$= 4{,}135\,669\,2\,(12) \times 10^{-15}\ \text{eV.s}$$
$$\hbar = \frac{h}{2\pi} = 1{,}054\,572\,66\,(63) \times 10^{-34}\ \text{J.s}$$
$$= 6{,}582\,122\,0\,(20) \times 10^{-16}\ \text{eV.s}$$

Vitesse de la lumière dans le vide :

$$c = 299\,792\,458\ \text{m/s par définition}$$

Perméabilité et permittivité du vide :

$$\mu_0 = 4\pi \times 10^{-7}\quad \text{SI (H/m) par définition}$$
$$\varepsilon_0 = \frac{1}{\mu_0 c^2} = 8{,}854\,187\,817\ldots \times 10^{-12}\quad \text{SI (F/m) par définition}$$
$$\frac{1}{4\pi\varepsilon_0} = \frac{\mu_0 c^2}{4\pi} = 8{,}987\,551\,787\ldots \times 10^{9}\quad \text{SI}$$

Charge élémentaire :

$$e = 1{,}602\,177\,33\,(49) \times 10^{-19}\ \text{C}$$

Masse de l'électron :

$$m_e = 9{,}109\,389\,7\,(54) \times 10^{-31}\ \text{kg}$$

Masse du proton et du neutron :

$$m_p = 1{,}672\,623\,1\,(10) \times 10^{-27}\ \text{kg}$$
$$m_n = 1{,}674\,928\,6\,(10) \times 10^{-27}\ \text{kg}$$

Constante de structure fine :

$$\alpha = \frac{e^2}{4\pi\varepsilon_0 \hbar c} = 1/137{,}035\,989\,5\,(61)$$

Rayon de Bohr :

$$a_0 = \frac{\hbar^2}{m_e e^2 / 4\pi\varepsilon_0} = 0{,}529\,177\,249\,(24)\ \text{Å}$$

Énergie de Rydberg (noyau immobile) :

$$Ry = \frac{m_e}{2\hbar^2}\left(\frac{e^2}{4\pi\varepsilon_0}\right)^2 = 2{,}179\,874\,1\,(13) \times 10^{-18}\ \text{J}$$
$$= 13{,}605\,698\,1\,(40)\ \text{eV}$$

Longueur d'onde de Compton de l'électron :

$$\lambda\!\!\!^{-}_c = \frac{\hbar}{m_e c} = 3{,}861\,593\,23\,(35) \times 10^{-13}\ \text{m}$$

Magnéton de Bohr :

$$\mu_B = \frac{e\hbar}{2m_e} = 9{,}274\,015\,4\,(31) \times 10^{-24}\ \text{J/T}$$
$$= 5{,}788\,382\,63\,(52) \times 10^{-5}\ \text{eV/T}$$

Magnéton nucléaire :

$$\mu_N = \frac{e\hbar}{2m_p} = 5{,}050\,786\,6\,(17) \times 10^{-27}\ \text{J/T}$$
$$= 3{,}152\,451\,66\,(28) \times 10^{-8}\ \text{eV/T}$$

Constante gravitationnelle :

$G = 6{,}672\,59\,(85) \times 10^{-11}$ SI ($\text{Nm}^2.\text{kg}^{-2}$)

Accélération de la pesanteur (niveau de la mer, latitude 45°) :

$g = 9{,}806\,2\ \text{m/s}^2$.

Imprimé en France
Imprimerie Moderne de Bayeux, Bayeux
Dépôt légal : août 2001
numéro d'édition : 6065 d

HERMANN, ÉDITEURS DES SCIENCES ET DES ARTS

Ce livre présente la mécanique statistique et son articulation avec la thermodynamique, au niveau de la maîtrise de physique des Universités. La mécanique statistique est développée de façon progressive à partir du postulat fondamental microanonique ; il est en effet important que les étudiants la perçoivent d'emblée comme une théorie fondamentale qui constitue l'un des piliers de la physique moderne.

Des applications physiques illustrent et concrétisent au fur et à mesure les considérations théoriques (physique des gaz et des solides, transition liquide-vapeur, magnétisme, systèmes de particules chargées, rayonnement fossile à trois kelvins, équation de Bolzmann...).

Un effort pédagogique constant a été réalisé pour rendre ces sujets accessibles aux étudiants du niveau choisi. Aucune connaissance préalable de la thermodynamique n'est indispensable, et les notions nécessaires de mécanique quantique (qui sont rappelées dans un appendice) restent très élémentaires. Certains passages, nettement délimités, se situent plutôt au niveau du troisième cycle, ils sont, cependant, totalement compréhensibles dans le contexte de cet ouvrage.

Le livre réunit deux parties distinctes mais imbriquées : six chapitres et des compléments. Les chapitres constituent un cours structuré, dont la compréhension ne nécessite la connaissance d'aucun des compléments. Les compléments sont de courtes monographies illustrant ou prolongeant le cours, indépendantes les unes des autres. Chaque ensemble chapitre-compléments est suivi d'une série d'exercices groupés par thèmes et classés par ordre de difficulté croissante.

Claudine Guthmann, Bernard Diu et Bernard Roulet sont professeurs à l'Université de Paris VII, Danielle Lederer est maître de conférences à l'Université de Paris VII.

Bernard Diu est co-auteur, avec Claude Cohen-Tannoudji et Franck Laloë, de *Mécanique quantique*, publié chez Hermann.

C'est une somme... une cathédrale magnifique dans son appareil architectural. BUP

Chaque chapitre contient de nombreux compléments très intéressants, qui sont de petites monographies indépendantes. Bulletin critique du livre français

Vient combler un vide certain...très clair et très détaillé...grandes qualités...un ouvrage de référence. La Recherche

ISBN 2 7056 6065 8

460 F 70.13 €

HERMANN ÉDITEURS DES SCIENCES ET DES ARTS