

Chapitre 12

Espaces vectoriels pseudo-euclidiens

12.1 Formes quadratiques sur un espace vectoriel réel, théorème de Sylvester.

Soit Q une forme quadratique sur un espace vectoriel réel de dimension finie V . Le théorème 10.4.2 nous dit qu'il existe une base $\{v_1, \dots, v_n\}$ de V qui est orthogonale pour Q . Dans cette base on a

$$x = \sum_{i=1}^n x_i v_i \quad \Rightarrow \quad Q(x) = \sum_{i=1}^n \alpha_i x_i^2.$$

Il est habituel de noter p le nombre de coefficients α_i qui sont positifs et q le nombre de coefficients α_i qui sont négatifs.

Définition 12.1.1. (1) Le couple (p, q) s'appelle la *signature*¹ de Q .

(2) La somme $r = p + q$ s'appelle le *rang* de Q .

(3) La forme quadratique Q est *non dégénérée* si $r = \dim(V)$.

(4) Q est *positive* si $q = 0$ et *négative* si $p = 0$.

(5) La forme quadratique Q est *définie positive* si elle est positive et non dégénérée.

(6) Q est *définie négative* si elle est négative et non dégénérée.

Remarques 1. Une notation utile est la suivante : On dit que $Q > 0$ sur le sous-espace vectoriel $W \subset V$ si la restriction de Q à W est définie positive, c'est-à-dire $Q(x) > 0$ pour tout $x \in W \setminus \{0\}$. On dira dans ce cas que le sous-espace vectoriel W est défini positif pour Q . De même on dit $Q < 0$ si $(-Q) > 0$ sur W .

2. On dit aussi que Q est *semi-définie positive* (resp. semi-définie négative) si $Q(x) \geq 0$ pour tout $x \in V$ (respectivement $Q(x) \leq 0$ pour tout $x \in V$). Il est clair qu'une forme quadratique de signature (p, q) est semi-définie positive si et seulement si $q = 0$ et semi-définie négative si et seulement si $p = 0$.

1. ne pas confondre avec la signature d'une permutation

Le résultat suivant justifie ces définitions.

Théorème 12.1.2 (Théorème d'inertie de Sylvester). *La signature (p, q) ne dépend que de la forme quadratique Q et non de la base choisie. Plus précisément, nous avons la caractérisation suivante de la signature :*

- (i) p est la dimension maximale d'un sous-espace vectoriel de V sur lequel Q est définie positive.
- (ii) q est la dimension maximale d'un sous-espace vectoriel de V sur lequel Q est définie négative.

Preuve. Nous allons démontrer que le coefficient p est la dimension maximale d'un sous-espace vectoriel sur lequel Q est définie positive. Observons tout d'abord que, quitte à permuter les vecteurs de la base orthogonale, on peut supposer que

$$\begin{cases} \alpha_i > 0 & \text{pour } 1 \leq i \leq p, \\ \alpha_i < 0 & \text{pour } p < i \leq r = p + q, \\ \alpha_i = 0 & \text{pour } i > r. \end{cases}$$

Dans cette base on peut écrire

$$Q(x) = \sum_{i=1}^r \alpha_i x_i^2 = \sum_{i=1}^p \alpha_i x_i^2 - \sum_{j=p+1}^r |\alpha_j| x_j^2.$$

Soit maintenant $W \subset V$ un sous-espace vectoriel de dimension maximale tel que la forme quadratique Q restreinte à W est définie positive. Il est clair que $\dim(W) \geq p$ car la restriction de Q au sous-espace $\text{Vec}\{v_1, \dots, v_p\}$ est définie positive.

Nous allons prouver par que $\dim(W) = p$. Supposons par l'absurde que $\dim(W) > p$ et notons $U = \text{Vec}\{v_{p+1}, \dots, v_n\}$. Alors $\dim(U) = n - p$ et donc $\dim(W \cap U) > 0$ car

$$\dim(W \cap U) = \dim(W) + \dim(U) - \dim(W + U) > p + (n - p) - n = 0$$

(on utilise que $\dim(W + U) \leq n$ puisque $W + U \subset V$). Il existe donc un vecteur non nul $x \in W \cap U$, mais ceci est impossible car

$$x \in W \setminus \{0\} \Rightarrow Q(x) > 0 \quad \text{et} \quad x \in U \Rightarrow Q(x) \leq 0.$$

(rappelons que par hypothèse Q est définie positive sur W). Cette contradiction montre que

$$p = \max\{\dim(W) \mid W \subset V \text{ est un sous-espace vectoriel tel que } Q > 0 \text{ sur } W\}.$$

Un argument similaire montre que q est la dimension maximale d'un sous-espace vectoriel sur lequel $Q < 0$. On a ainsi obtenu une caractérisation de la signature (p, q) d'une forme quadratique qui ne dépend pas du choix d'une base orthogonale, ce qui prouve le théorème de Sylvester. \square

Le concept de signature peut aussi se définir pour les formes bilinéaires symétriques :

Définition 12.1.3. La *signature* (p, q) d'une forme bilinéaire symétrique $g : V \times V \rightarrow \mathbb{R}$ sur un espace vectoriel réel de dimension finie V est la signature de la forme quadratique associée $Q(x) = g(x, x)$. On définit de même les notions de *rang*, de *forme bilinéaire symétrique (non) dégénérée* et *définie positive (négative)*.

Remarque. Observons que la forme quadratique Q est définie positive (i.e. de signature $(n, 0)$) si et seulement si la forme bilinéaire associée est un produit scalaire. La norme associée à ce produit scalaire est alors $\|x\| = \sqrt{Q(x)}$.

Du point de vue matriciel, le théorème de Sylvester s'énonce ainsi :

Corollaire 12.1.4. Toute matrice symétrique $A \in M_n(\mathbb{R})$ est congruente à une matrice diagonale de type

$$H_{p,q} = \begin{pmatrix} I_p & 0 & 0 \\ 0 & -I_q & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Les nombres (p, q) ne dépendent que de A et non de la base orthogonale choisie.

Définition. Le couple (p, q) s'appelle alors la *signature* (p, q) de la matrice symétrique $A \in M_n(\mathbb{R})$. On dit que la matrice A est *définie positive* si $(p, q) = (n, 0)$.

Preuve. Notons β la forme bilinéaire sur \mathbb{R}^n dont la matrice de Gram est A . On peut choisir une base $\{v_1, \dots, v_n\}$ de \mathbb{R}^n telle que $\beta(v_i, v_j) = 0$ si $i \neq j$ et

$$\beta(v_i, v_i) > 0, \quad \beta(v_j, v_j) < 0, \quad \beta(v_k, v_k) = 0,$$

pour $1 \leq i \leq p < j \leq (p+q) < k \leq n$. On définit alors une nouvelle base $\{w_1, \dots, w_n\}$ par

$$w_i = \frac{v_i}{\sqrt{\beta(v_i, v_i)}}, \quad w_j = \frac{v_j}{\sqrt{-\beta(v_j, v_j)}}, \quad w_k = v_k,$$

(avec $1 \leq i \leq p < j \leq (p+q) < k \leq n$). Il est facile de vérifier que la matrice de β dans cette base est la matrice $H_{p,q}$. □

Définition 12.1.5. Soit g une forme bilinéaire symétrique sur un espace vectoriel réel V de dimension n . On dit qu'une base $\{e_1, \dots, e_n\}$ de V est une *base de Sylvester*, ou une *base orthonormale généralisée* si

$$g(e_i, e_j) = \begin{cases} +1, & \text{si } 1 \leq i = j \leq p \\ -1, & \text{si } p+1 \leq i = j \leq p+q \\ 0, & \text{sinon,} \end{cases}$$

où (p, q) est la signature de g .

Le corollaire 12.1.4 nous garantit l'existence d'une base de Sylvester (non unique) pour toute forme bilinéaire symétrique sur un espace vectoriel réel V de dimension finie.

Proposition 12.1.6. (i) Deux matrices symétriques réelles $A, B \in M_n(\mathbb{R})$ sont congruentes si et seulement si elles ont la même signature.

(ii) Si $A \in M_n(\mathbb{R})$ est une matrice symétrique de signature (p, q) , alors p est le nombre de valeurs propres strictement positives de A comptées avec multiplicités et q est le nombre de valeurs propres strictement négatives comptées avec multiplicités.

Preuve. (i) L'affirmation découle du fait que toute matrice symétrique $A \in M_n(\mathbb{R})$ est congruente à sa forme de Sylvester et que la congruence est une relation d'équivalence.

(ii) Par le théorème spectral on sait que toute matrice symétrique $A \in M_n(\mathbb{R})$ est orthogonalement diagonalisable, donc à la fois semblable et congruente à une matrice diagonale D qui contient donc les valeurs propres de A , répétées autant de fois que leur multiplicité. □

Rappelons que pour une matrice réelle symétrique, la multiplicité géométrique de chaque valeur propre est égale à sa multiplicité algébrique (c'est une application du théorème spectral, qui nous dit en particulier qu'une telle matrice est diagonalisable).

Voici un exemple pour le point (ii) de cette proposition. Considérons la matrice symétrique

$$A = \begin{pmatrix} 3 & -2 & 4 \\ -2 & 6 & 2 \\ 4 & 2 & 3 \end{pmatrix}.$$

Son polynôme caractéristique est $\chi_A(t) = (t-7)^2(t+2)$, en particulier $\det(A) \neq 0$ et donc A est de rang 3. Les valeurs propres sont $+7$ avec multiplicité 2 et -2 avec multiplicité 1. La signature de A est donc $(p, q) = (2, 1)$. En particulier A n'est pas définie positive.

12.2 Espaces pseudo-euclidiens

Définitions 1. On appelle *espace vectoriel pseudo-euclidien* un espace vectoriel V sur le corps \mathbb{R} de dimension finie muni d'une forme quadratique Q non dégénérée. L'espace (V, Q) est dit *euclidien* si Q est définie positive (i.e. $Q(x) > 0$ pour tout $x \in V$ non nul).

2. Une application affine $f : V_1 \rightarrow V_2$ entre deux espaces pseudo-euclidiens (V_1, Q_1) et (V_2, Q_2) est une *isométrie* si pour tous $x, y \in V_1$ on a

$$Q_2(f(y) - f(x)) = Q_1(y - x).$$

Lorsque f est linéaire, cette condition peut s'écrire $Q_1 = Q_2 \circ f$, c'est-à-dire $Q_2(f(x)) = Q_1(x)$ pour tout $x \in V_1$.

Remarque. Lorsque (V, Q) est euclidien, on a les notions de norme d'un vecteur $\|x\| = \sqrt{Q(x)}$ et de distance $d(x, y) = \|y - x\|$ entre deux points. Une isométrie entre deux espaces euclidiens est une bijection qui respecte les distances. Dans le cas général $Q(x)$ n'est pas forcément positif, toutefois même lorsqu'il n'y a pas de norme ou de distance associée à une forme quadratique, celle-ci peut représenter des quantités géométriques intéressantes. Lorsque deux espaces pseudo-euclidiens sont isométriques, on considère que leurs géométries sont équivalentes (par exemple tous les espaces euclidiens de même dimension sont isométriques, leur géométrie sont donc équivalentes à celle de l'espace \mathbb{R}^n muni de son produit scalaire standard).

Les développements du chapitre 10 nous ont appris qu'il y a équivalence entre les trois points de vues suivants :

- (1) La théorie des espaces pseudo-euclidiens.
- (2) La théorie des formes bilinéaires symétriques sur un espace vectoriel réel de dimension finie.
- (3) L'étude des matrices carrées symétriques à coefficients réels de déterminant non nul.

Rappelons ces équivalences :

- Si Q est une forme quadratique sur V , alors il existe une unique forme bilinéaire symétrique, que l'on notera par $g : V \times V \rightarrow \mathbb{R}$, telle que $Q(x) = g(x, x)$ pour tout $x \in V$. Cette forme bilinéaire est donnée par les formules de polarisation, par exemple

$$g(x, y) = \frac{1}{4} (Q(x + y) - Q(x - y)).$$

- Si $\{v_1, \dots, v_n\}$ est une base de V , alors la *matrice de Gram* de Q (ou de g) dans cette base est la matrice $G \in M_n(\mathbb{R})$ définie par

$$G = (g_{ij}), \quad \text{avec } g_{ij} = g(v_i, v_j).$$

Cette matrice est clairement symétrique, i.e. $G^\top = G$ car $g_{ji} = g(v_j, v_i) = g(v_i, v_j) = g_{ij}$. La matrice de Gram est donc déterminée par g .

- Inversement on peut calculer $g(x, y)$ à partir de la formule

$$g(x, y) = \sum_{i,j=1}^n g_{ij} x_i y_j = X^\top G Y$$

où (x_1, \dots, x_n) sont les composantes de x dans la base $\{v_1, \dots, v_n\}$ (c'est-à-dire $x = \sum_{i=1}^n x_i v_i$) et $X \in \mathbb{R}^n$ est le vecteur-colonne de \mathbb{R}^n associé, (de même pour (y_1, \dots, y_n) et Y).

- La condition de non dégénérescence de Q (ou de g) signifie que pour tout $x \in V \setminus \{0\}$ on peut trouver $y \in V$ tel que $g(x, y) \neq 0$.
- Il est facile de vérifier que g est non dégénéré si et seulement $\det(G) \neq 0$, i.e. la matrice de Gram G est inversible.

Nous avons alors le résultat suivant :

Proposition 12.2.1. *Soient (V_1, Q_1) et (V_2, Q_2) deux espaces pseudo-euclidiens de dimension n et $f : V_1 \rightarrow V_2$ un isomorphisme linéaire. Alors les conditions suivantes sont équivalentes :*

- (1) f est une isométrie, i.e. $Q_2 \circ f = Q_1$.
- (2) $g_2(f(x), f(y)) = g_1(x, y)$ pour tous $x, y \in V_1$, où g_i est la forme bilinéaire associée à Q_i (pour $i = 1, 2$).
- (3) Les matrices de Gram de g_1 et g_2 dans des bases $\mathcal{B}_1 \subset V_1$ et $\mathcal{B}_2 \subset V_2$ sont reliées par

$$G_1 = A^\top G_2 A,$$

où $A = M_{\mathcal{B}_2, \mathcal{B}_1}(f)$ est la matrice de f dans ces bases. En particulier les matrices G_1 et G_2 sont congruentes et la congruence est réalisée par la matrice de l'endomorphisme f .

La preuve est un simple exercice.

Remarque. Lorsque $V_1 = V_2 = \mathbb{R}^n$ et $\mathcal{B}_1 = \mathcal{B}_2$ = la base canonique, la formule de congruence est évidente car

$$g_2(f(X), f(Y)) = (AX)^\top G_2 AX = X^\top (A^\top G_2 A) Y = X^\top G_1 Y.$$

Si (V, Q) est un espace pseudo-euclidien, l'ensemble des isométries linéaires de V dans lui même forme un groupe, que l'on appelle le *groupe orthogonal de Q* . On le note

$$O(Q) = \{f \in GL(V) \mid Q \circ f = Q\}.$$

Dans le cas où $V = \mathbb{R}^n$ on peut identifier tout endomorphisme $f \in \mathcal{L}(\mathbb{R}^n)$ avec sa matrice A dans la base canonique. On peut donc écrire

$$O(Q) = \{A \in GL_n(\mathbb{R}) \mid A^\top GA = G\}.$$

Le *groupe spécial orthogonal de Q* est le sous-groupe

$$SO(Q) = O(Q) \cap SL_n(\mathbb{R}) = \{A \in GL_n(\mathbb{R}) \mid A^\top GA = G \text{ et } \det(A) = 1\}.$$

12.3 Base de Sylvester et espaces pseudo-euclidiens modèles.

Rappelons qu'une base $\{v_1, \dots, v_n\}$ d'un espace vectoriel pseudo-euclidien (V, Q) est une *base de Sylvester* (ou une base *orthonormée généralisée*) si $Q(v_i) = \pm 1$ pour tout i et $v_i \perp_Q v_j$ si $i \neq j$. Le théorème de Sylvester nous dit que tout espace vectoriel pseudo-euclidien admet des bases de Sylvester. De plus le nombre p d'éléments de la base tels que $Q(v_i) = +1$ et le nombre q d'éléments tels que $Q(v_j) = -1$ ne dépendent pas de la base choisie. Le couple (p, q) est la *signature* de la forme quadratique Q et nous avons $p + q = n$ car Q est supposée non dégénérée. Il suit du théorème de Sylvester que tout espace pseudo-euclidien est isométrique à l'espace vectoriel \mathbb{R}^n muni de la forme quadratique standard de signature $(p, q) = (p, n - p)$:

$$Q(x) = \sum_{i=1}^p x_i^2 - \sum_{j=p+1}^n x_j^2.$$

On note $\mathbb{E}^{p,q}$ cet espace et on considère que c'est l'*espace pseudo-euclidien modèle* (ou standard) de signature (p, q) . La forme bilinéaire symétrique associée est

$$g(x, y) = \sum_{i=1}^p x_i y_i - \sum_{j=p+1}^n x_j y_j,$$

et la matrice de Gram dans la base canonique est la matrice

$$H_{p,q} = I_p \oplus (-I_q) = \begin{pmatrix} I_p & 0 \\ 0 & -I_q \end{pmatrix} = \begin{pmatrix} 1 & & & & & 0 \\ & \ddots & & & & \\ & & 1 & & & \\ & & & -1 & & \\ & & & & \ddots & \\ 0 & & & & & -1 \end{pmatrix} \quad (12.1)$$

c'est-à-dire

$$H_{p,q} = (\eta_{ij}), \quad \text{avec} \quad \eta_{ij} = \begin{cases} 1 & \text{si } i = j \leq p, \\ -1 & \text{si } i = j > p, \\ 0 & \text{si } i \neq j. \end{cases} \quad (12.2)$$

(*symbole de Kronecker généralisé* de signature (p, q)). Remarquons que $\mathbb{E}^{n,0}$ est l'espace euclidien \mathbb{R}^n muni de son produit scalaire standard, on le note simplement \mathbb{E}^n .

Le groupe des isométries linéaires de $\mathbb{E}^{p,q}$ se note $O(p, q)$:

$$O(p, q) = \{A \in GL_n(\mathbb{R}) \mid A^\top H_{p,q} A = H_{p,q}\}.$$

Le sous-groupe des isométries de déterminant 1 est

$$SO(p, q) = O(p, q) \cap SL_n(\mathbb{R}) = \{A \in O(p, q) \mid \det A = +1\}.$$

12.4 Indicatrices et cône isotrope

Définition. Soit (V, Q) un espace pseudo-euclidien.

(1) On appelle *cône isotrope* de (V, Q) l'ensemble des vecteurs isotropes. On le note

$$S_0(V, Q) = \{x \in V \mid Q(x) = 0\}.$$

(2) On appelle *indatrice positive* de (V, Q) l'ensemble

$$S_+(V, Q) = \{x \in V \mid Q(x) = 1\}.$$

(3) L' *indatrice négative* de (V, Q) est l'ensemble

$$S_-(V, Q) = \{x \in V \mid Q(x) = -1\}.$$

Remarques.

- (i) Le cône isotrope et les indicatrices ne sont pas des sous-espaces vectoriels de V .
- (ii) Si $x \in S_0(V, Q)$, alors $\lambda x \in S_0(V, Q)$ pour tous $\lambda \in \mathbb{R}$.
- (iii) Si $x \in S_+(V, Q)$, alors $\lambda x \in S_+(V, Q)$ si et seulement si $\lambda = \pm 1$. La même propriété est vraie pour $S_-(V, Q)$.
- (iv) Les ensembles $S_0(V, Q)$, $S_+(V, Q)$ et $S_-(V, Q)$ déterminent complètement la forme quadratique Q , i.e. si Q_1 et Q_2 sont deux formes quadratiques telles que

$$S_0(V, Q_1) = S_0(V, Q_2), \quad S_+(V, Q_1) = S_+(V, Q_2), \quad S_-(V, Q_1) = S_-(V, Q_2),$$

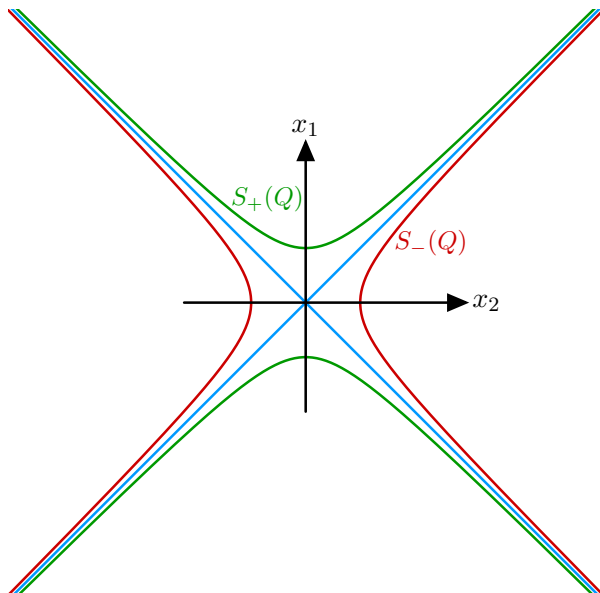
alors $Q_1 = Q_2$.

- (v) Les ensembles $S_0(V, Q)$, $S_+(V, Q)$ et $S_-(V, Q)$ sont invariants par l'action du groupe $O(Q)$, c'est-à-dire que si $f \in O(Q)$, alors $x \in S_+(V, Q)$ si et seulement si $f(x) \in S_+(V, Q)$.

Exemples.

- (1) Pour le plan euclidien, on a $S_-(\mathbb{E}^2) = \emptyset$, $S_0(\mathbb{E}^2) = \{0\}$ et $S_+(\mathbb{E}^2)$ est le cercle unité.

- (2) Plus généralement l'indicatrice S_+ d'un espace euclidien \mathbb{E}^n est la *sphère unité* ; c'est l'ensemble des points de \mathbb{E}^n dont la distance à l'origine 0 est égale à 1. L'indicatrice négative est l'ensemble vide et $S_0(\mathbb{E}^n) = \{0\}$.
- (3) Pour $\mathbb{E}^{1,1}$, $S_0(\mathbb{E}^{1,1})$ est la réunion des deux droites $\{x_2 = \pm 1\}$ et $S_{\pm}(\mathbb{E}^{1,1})$ sont deux hyperboles dont les asymptotes sont les droites du cône isotrope.
- (4) Pour $\mathbb{E}^{1,2}$, $S_0(\mathbb{E}^{1,2})$ est le cône circulaire droit $\{x_1^2 = x_2^2 + x_3^2\}$, l'indicatrice négative $S_-(\mathbb{E}^{1,2})$ est une hyperboloïde de révolution à deux nappes et l'indicatrice positive $S_+(\mathbb{E}^{1,2})$ est une hyperboloïde de révolution à une nappe.



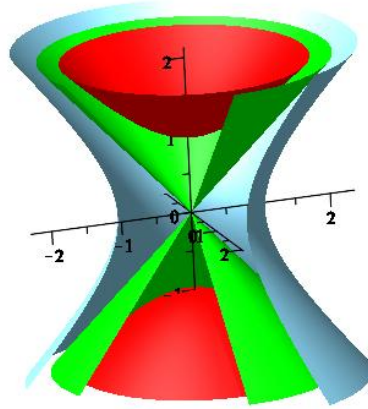


FIGURE 12.1 – Cône isotrope et indicatrices de $\mathbb{E}^{1,2}$.

12.5 L'espace-temps de Lorentz-Minkowski $\mathbb{E}^{1,d}$

Définition. On appelle *espace-temps de Lorentz-Minkowski* (ou simplement espace de Minkowski) de dimension $d + 1$ tout espace pseudo-euclidien de signature $(1, d)$. Dans une base adaptée, on peut donc écrire la forme quadratique

$$Q(x) = c^2 t^2 - x_1^2 - \dots - x_d^2.$$

Un élément $x = (t, x_1, \dots, x_d) \in \mathbb{E}^{1,d}$ s'appelle un *événement*.

L'interprétation physique est la suivante : Les coordonnées x_1, \dots, x_d représentent des coordonnées de l'espace (on supposera en général que $d \leq 3$) et t représente une coordonnée temporelle. Le paramètre c est la vitesse maximale de propagation d'un signal dans l'espace temps. L'expérience nous apprend que c est la vitesse de la lumière.

Il est commode de noter $x_0 = ct$, alors la forme quadratique fondamentale s'écrit

$$Q(x) = x_0^2 - x_1^2 - \dots - x_d^2,$$

cela revient essentiellement à choisir des unités telles que $c = 1$.

Le groupe des isométries linéaires de l'espace de Minkowski est le groupe $O(1, d)$. On l'appelle le *groupe de Lorentz*. Le *principe de relativité d'Einstein* dit que les lois de la physique ne doivent pas dépendre du référentiel choisi. Ce principe se traduit mathématiquement de la manière suivante :

Les notions attachées à l'espace-temps qui ont une signification physique doivent être invariantes sous l'action du groupe de Lorentz.

Les définitions suivantes modélisent les notions liées à la causalité dans l'espace-temps de Lorentz-Minkowski.

Définitions.

- (1) On dit que deux événements $x, y \in \mathbb{E}^{1,d}$ sont *en relation de causalité* si un signal émis depuis l'un des événements peut atteindre l'autre événement. La condition s'écrit mathématiquement par

$$Q(y - x) \geq 0.$$

- (2) On considère qu'un événement ne peut pas être cause d'un événement passé mais seulement d'un événement futur. On dira en conséquence qu'un événement y est dans le *futur causal* de x si

$$Q(y - x) \geq 0 \quad \text{et} \quad y_0 \geq x_0.$$

L'ensemble des événements dans le futur causal de x s'appelle le *cône futur* de x et se note

$$\mathcal{C}_x = \{y \in \mathbb{E}^{1,d} \mid Q(y - x) \geq 0 \text{ et } y_0 \geq x_0\}.$$

- (3) Le *cône isotrope* (ou cône de lumière) issu de x est l'ensemble des y tels que $Q(y - x) = 0$.
(4) Si y est dans le futur de x , on appelle *temps propre* ou *intervalle spatio-temporel* la quantité

$$\tau(x, y) = \sqrt{Q(y - x)} = \sqrt{(y_0 - x_0)^2 - (y_1 - x_1)^2 - \dots - (y_d - x_d)^2}.$$

On utilise aussi la terminologie suivante :

- (i) Un vecteur x est de *type temps* si $Q(x) > 0$.
- (ii) Un vecteur x est de *type lumière*, ou *isotrope* si $Q(x) = 0$.
- (iii) Un vecteur x est de *type espace* si $Q(x) < 0$.

Ces notions ont un sens physique car elles sont invariantes sous l'action du groupe de Lorentz. Par exemple si $f \in O(1, d)$, alors x est de type temps si et seulement si $f(x)$ est de type temps.

La notion de ligne d'univers

Considérons une particule, ou un objet quelconque, qui se déplace au cours du temps. Sa trajectoire dans l'espace est représentée dans un certain référentiel par la fonction $t \mapsto x(t) = (x_1(t), \dots, x_d(t)) \in \mathbb{R}^d$. On appelle *ligne d'univers* de cette particule la fonction

$$t \mapsto x(t) = (t, x_1(t), \dots, x_d(t)) \in \mathbb{E}^{1,d}$$

Proposition 12.5.1. *La ligne d'univers de toute trajectoire physiquement réalisable vérifie la condition suivante :*

$$\text{pour tout } t_2 > t_1, \text{ on a } x(t_2) \in \mathcal{C}_{x(t_1)}.$$

On dit qu'une ligne d'univers est *inertielle*, si elle représente une droite de $\mathbb{E}^{1,d}$ cette droite doit être contenue dans le cône de lumière de chacun de ses points.

12.6 L'inégalité de Cauchy-Schwarz inversée et quelques conséquences

Théorème 12.6.1 (Inégalité de Cauchy-Schwarz inversée). *Si $x, y \in \mathbb{E}^{1,d}$ sont isotropes ou de type temps, alors*

$$|g(x, y)| \geq \sqrt{Q(x)}\sqrt{Q(y)}, \quad (12.3)$$

où g est la forme bilinéaire associée à Q . On a égalité si et seulement si x et y sont colinéaires.

Preuve. Observons tout d'abord que si x ou y est isotrope, alors $Q(x)Q(y) = 0$ et l'inégalité est triviale. Il est par ailleurs facile de vérifier que si x et y sont colinéaires alors $Q(x)Q(y) = g(x, y)^2$. On suppose donc que x et y sont de type temps et linéairement indépendants et nous donnons deux démonstrations de l'inégalité stricte (12.3).

Première preuve : On suppose donc x et y linéairement indépendants et on note $W = \text{Vec}x, y \subset \mathbb{E}^{1,d}$ le sous-espace vectoriel engendré par x et y . La restriction de g à W peut à priori être une forme bilinéaire symétrique de signature $(p, q) = (0, 2)$ ou $(p, q) = (1, 1)$. Or la signature $(0, 2)$ est exclue car nous avons supposé que x et y sont de type temps, i.e. $Q(x) > 0$ et $Q(y) > 0$. Donc g est de signature $(1, 1)$ sur W . Par conséquent le déterminant de la matrice de Gram associée à la base $\{x, y\}$ de W est négatif, on a donc prouvé que

$$Q(x)Q(y) - g(x, y)^2 = \det \begin{pmatrix} g(x, x) & g(x, y) \\ g(x, y) & g(y, y) \end{pmatrix} < 0,$$

ce qui est équivalent à l'inégalité (12.3).

Pour la deuxième preuve, on considère la droite affine qui passe par x et de vecteur directeur y , qui est l'ensemble

$$L = \{x + sy \mid s \in \mathbb{R}\} \subset \mathbb{E}^{1,d}.$$

Si x et y ne sont pas colinéaires, alors L n'est pas contenue dans l'intérieur du cône isotrope de $\mathbb{E}^{1,d}$ et cette droite contient donc des vecteurs de type temps, de type espaces et deux vecteurs isotropes. Par conséquent la fonction

$$f(s) = Q(x + sy) = g(x + sy, x + sy) = Q(x) + 2sg(x, y) + s^2Q(y)$$

est un polynôme du second degré qui s'annule pour exactement deux valeurs de s , ce qui implique que le discriminant de $f(s)$ est strictement positif. On a donc

$$\Delta = g(x, y)^2 - Q(x)Q(y) > 0.$$

□

Lemme 12.6.2. *Si $x, y \in \mathcal{C}_0$ sont deux vecteurs du cône futur de 0, alors $g(x, y) \geq 0$.*

Preuve. L'hypothèse $x, y \in \mathcal{C}_0$ signifie que

$$\sqrt{\sum_{i=1}^d x_i^2} \leq x_0, \quad \text{et} \quad \sqrt{\sum_{i=1}^d y_i^2} \leq y_0.$$

On a donc par l'inégalité de Cauchy-Schwarz classique dans \mathbb{R}^d

$$\sum_{i=1}^d x_i y_i \leq \sqrt{\sum_{i=1}^d x_i^2} \sqrt{\sum_{i=1}^d y_i^2} \leq x_0 y_0,$$

ce qui implique

$$g(x, y) = x_0 y_0 - \sum_{i=1}^d x_i y_i \geq 0.$$

Corollaire 12.6.3. *Soient $x, y \in \mathcal{C}_0$ deux vecteurs du cône futur de 0, alors $Q(x + y) \geq 0$ et*

$$\sqrt{Q(x + y)} \geq \sqrt{Q(x)} + \sqrt{Q(y)}$$

Preuve. Le lemme précédent et l'inégalité de Cauchy-Schwarz inversée impliquent que

$$g(x, y) \geq \sqrt{Q(x)} \sqrt{Q(y)},$$

par conséquent on a

$$\begin{aligned} Q(x + y) &= g(x + y, x + y) \\ &= (g(x, x) + 2g(x, y) + g(y, y)) \\ &\geq \left(Q(x) + 2\sqrt{Q(x)}\sqrt{Q(y)} + Q(y) \right) \\ &= \left(\sqrt{Q(x)} + \sqrt{Q(y)} \right)^2. \end{aligned}$$

□

Le paradoxe des jumeaux

Si A et B sont deux événements de l'espace-temps, on note $A \preceq B$ si B est dans le futur causal de A , i.e. si le vecteur $(B - A) \in \mathcal{C}_0$.

Lemme 12.6.4. *La relation \preceq est une relation d'ordre partiel sur $\mathbb{E}^{1,d}$.*

Preuve. Il faut montrer que si B est dans le futur causal de A et C est dans le futur causal de B alors C est dans le futur causal de A , ce qui est tout-à-fait intuitif. Mathématiquement cela signifie

$$A \preceq B \text{ et } B \preceq C \Rightarrow A \preceq C.$$

Pour le démontrer, on note $x = (B - A)$, $y = (C - B)$ et $z = (C - A)$. Alors $z = x + y$ et par hypothèse on a $x, y \in \mathcal{C}_0$. On doit prouver que $z \in \mathcal{C}_0$.

$$x_0 \leq y_0 \text{ et } y_0 \leq z_0 \Rightarrow x_0 \leq z_0,$$

par conséquent $z_0 > 0$. D'autre part le corollaire précédent implique que $Q(z) \geq 0$, ce qui complète la preuve.

□

Nous pouvons maintenant énoncer le

Paradoxe des jumeaux

Si A , B et C sont trois événements tels que B est dans le futur causal de A et C est dans le futur causal de B , alors C est dans le futur causal de A et les temps propres associés sont reliés par l'*inégalité causale*, aussi appelée l'*inégalité du triangle inversée* :

$$\tau(A, C) \geq \tau(A, B) + \tau(B, C). \quad (12.4)$$

On a égalité si et seulement si l'événement B est situé sur la ligne d'univers inertielle de A vers C .

Rappelons que si A, B, C sont trois points de l'espace Euclidien \mathbb{E}^n , alors leurs distances respectives vérifient l'inégalité du triangle :

$$d(A, C) \leq d(A, B) + d(B, C).$$

d'où le nom de inégalité du triangle inversée pour l'inégalité (12.4)

Preuve. On pose de nouveau $x = (B - A)$, $y = (C - B)$ et $z = (C - A)$, alors on a $z = x + y$. Le corollaire 12.6.3 entraîne alors que

$$\tau(A, C) = \sqrt{Q(z)} = \sqrt{Q(x + y)} \geq \sqrt{Q(x)} + \sqrt{Q(y)} = \tau(A, B) + \tau(B, C).$$

□

Ce résultat a été appelé le “paradoxe des jumeaux” par Paul Langevin², qui l’a formulé de la manière suivante : Si deux frères jumeaux se rencontrent en un lieu précis au même moment, et si l’un des jumeaux s’en va faire un voyage cosmique puis rejoint son frère, alors à son retour l’un des jumeaux (celui qui a voyagé) est plus jeune que l’autre. Ce paradoxe a rendu les commentateurs perplexes, mais il ne signifie pas qu’il y a une contradiction dans la théorie car la situation des deux frères n’est pas symétrique. L’un a une ligne d’univers inertielle et l’autre non.

2. Paul Langevin (1872-1946) est un physicien français qui fut parmi les premiers à admettre et propager la relativité restreinte en France (avec l’exception notable de Henri Poincaré). Il a formulé le paradoxe des jumeaux à ses collègues physiciens et philosophes en 1911. Ce paradoxe a engendré une certaine perplexité et provoqué d’intéressantes discussions sur la relativité einsteinienne et son interprétation.

Chapitre 13

Espaces hermitiens, opérateurs normaux et le théorème spectral

Rappelons qu'un espace vectoriel euclidien est un espace vectoriel de dimension finie sur le corps \mathbb{R} . Dans ce chapitre, nous étudions une notion analogue pour les espaces vectoriels de dimension finie sur le corps \mathbb{C} . Nous étudions en particulier une classe d'opérateurs (i.e. d'endomorphismes) sur de tels espaces, appelés opérateurs *normaux*, pour lesquels nous pouvons généraliser le théorème spectral vu au chapitre 10. Parmi les opérateurs normaux, on peut citer les opérateurs *unitaires* et les *opérateurs auto-adjoints*. Ces opérateurs sont importants notamment en mécanique quantique pour décrire les observables d'un système quantique.

13.1 Formes sesquilineaires et formes hermitiennes sur un espace vectoriel complexe.

Définition 13.1.1. Soient V et W deux espaces vectoriels sur le corps \mathbb{C} des nombres complexes. On dit qu'une application $f : V \rightarrow W$ est *antilinéaire* si elle vérifie

$$f(v + w) = f(v) + f(w) \quad \text{et} \quad f(\lambda v) = \bar{\lambda}f(v).$$

pour tous $v, w \in V$ et tous $\lambda \in \mathbb{C}$.

Remarquons que les applications anti-linéaires sont \mathbb{R} -linéaires, c'est-à-dire qu'elles sont linéaires lorsqu'on regarde V et W comme des espaces vectoriels sur \mathbb{R} .

Exemples 1.) L'application $s : \mathbb{C}^n \rightarrow \mathbb{C}^n$ qui conjugue toutes les coordonnées, i.e. $s(z_1, \dots, z_n) = (\bar{z}_1, \dots, \bar{z}_n)$ est anti-linéaire.

2.) Si $\theta : V \rightarrow \mathbb{C}$ est une forme linéaire, alors l'application $v \mapsto \overline{\theta(v)}$ est anti-linéaire.

3.) On appelle *adjointe* d'une matrice $A \in M_n(\mathbb{C})$ la transposée de la matrice conjuguée, et on note $A^* = \bar{A}^\top$. Il est clair que $A \mapsto A^*$ est une application anti-linéaire de l'espace vectoriel $M_n(\mathbb{C})$ dans lui-même.

Définition 13.1.2. (1) Une application $h : V \times V \rightarrow \mathbb{C}$ est dite *sesquilinéaire* si elle est antilinéaire en la première variable et linéaire en la deuxième variable¹ :

$$h(\lambda v_1 + v_2, w) = \bar{\lambda}h(v_1, w) + h(v_2, w) \quad \text{et} \quad h(v, \mu w_1 + w_2) = \mu h(v, w_1) + h(v, w_2).$$

(2) L'application $h : V \times V \rightarrow \mathbb{C}$ est une *forme hermitienne*² si elle est sesquilinéaire et elle vérifie de plus

$$h(w, v) = \overline{h(v, w)}$$

pour tous $v, w \in V$.

(3) La *forme quadratique* associée à une forme hermitienne h sur V est la fonction $Q : V \rightarrow \mathbb{R}$ définie par

$$Q(w) = h(w, w).$$

Observons que $Q(w)$ est en effet un nombre réel pour tout $w \in V$, car $Q(w) = h(w, w) = \overline{h(w, w)} = \overline{Q(w)}$.

Exemple. Si $\{\varphi_1, \dots, \varphi_n\}$ est une base de l'espace dual V^* , et $a_1, \dots, a_n \in \mathbb{R}$, alors la fonction $h : V \times V \rightarrow \mathbb{C}$ définie par

$$h(x, y) = \sum_{j=1}^n a_j \overline{\varphi_j(x)} \varphi_j(y)$$

est une forme hermitienne sur V . On peut prouver que toutes les formes hermitiennes sur un espace vectoriel de dimension finie sont de ce type.

Lemme 13.1.3. La forme quadratique associée à la forme hermitienne h vérifie les propriétés suivantes :

(i) $Q(\lambda w) = |\lambda|^2 Q(w)$ pour tout $\lambda \in \mathbb{C}$ et tout $w \in V$ (en particulier $Q(\sqrt{-1}w) = Q(w)$).

(ii) On peut retrouver h à partir de Q par la formule de polarisation :

$$h(x, y) = \frac{1}{4} (Q(x + y) - Q(x - y)) - \frac{i}{4} (Q(x + iy) - Q(x - iy)), \quad (13.1)$$

où $i = \sqrt{-1}$.

Preuve. La preuve de la première affirmation est élémentaire :

$$Q(\lambda w) = h(\lambda w, \lambda w) = \bar{\lambda} \lambda h(w, w) = |\lambda|^2 Q(w).$$

Pour prouver la formule de polarisation, on considère séparément les parties réelles et imaginaires. Dans les calculs qui suivent, on utilise les identités

$$h(y, x) = \overline{h(x, y)}, \quad h(x, iy) = ih(x, y), \quad h(iy, x) = -ih(y, x).$$

1. La convention opposée est également utilisée, i.e. certains auteurs demandent que h soit linéaire en la première variable et antilinéaire en la deuxième. Ici nous suivons la convention la plus usuelles parmi les physiciens.

2. En référence au mathématicien français Charles Hermite (1822-1901).

On a d'une part

$$\begin{aligned}
Q(x+y) - Q(x-y) &= h(x+y, x+y) - h(x-y, x-y) \\
&= (h(x, x) + h(x, y) + h(y, x) + h(y, y)) - (h(x, x) - h(x, y) - h(y, x) + h(y, y)) \\
&= 2(h(x, y) + h(y, x)) \\
&= 2(h(x, y) + \overline{h(x, y)}) \\
&= 4\operatorname{Ré}(h(x, y)),
\end{aligned}$$

et d'autre part

$$\begin{aligned}
Q(x-iy) - Q(x+iy) &= h(x-iy, x-iy) - h(x+iy, x+iy) \\
&= (h(x, x) - ih(x, y) + ih(y, x) + h(y, y)) - (h(x, x) + ih(x, y) - ih(y, x) + h(y, y)) \\
&= -2i(h(x, y) - h(y, x)) \\
&= -2i(h(x, y) - \overline{h(x, y)}) \\
&= 4\operatorname{Im}(h(x, y)).
\end{aligned}$$

Donc

$$\begin{aligned}
h(x, y) &= \operatorname{Ré}(h(x, y)) + i\operatorname{Im}(h(x, y)) \\
&= \frac{1}{4}(Q(x+y) - Q(x-y)) + \frac{i}{4}(Q(x-iy) - Q(x+iy)).
\end{aligned}$$

□

13.2 Espaces vectoriels hermitiens

Définitions. Soit V un espace vectoriel complexe. On appelle *produit scalaire hermitien* sur V la donnée d'une forme hermitienne $h : V \times V \rightarrow \mathbb{C}$ qui est définie positive. Cette condition signifie que $Q(v) = h(v, v) > 0$ pour tout vecteur non nul de V .

Un *espace vectoriel hermitien* est un espace vectoriel complexe de dimension finie muni d'un produit scalaire hermitien.

On peut résumer la définition de produit scalaire hermitien dans les quatre propriétés suivantes :

- (i) $h : V \times V \rightarrow \mathbb{C}$ est \mathbb{R} -bilinéaire.
- (ii) $h(x, iy) = ih(x, y) = -h(ix, y)$ (où $i = \sqrt{-1}$).
- (iii) $h(y, x) = \overline{h(x, y)}$.
- (iv) $h(x, x) > 0 \quad \forall x \in V \setminus \{0\}$,

pour tous $x, y \in V$.

Exemples 1. Le produit scalaire hermitien standard sur \mathbb{C}^n est défini par

$$\langle z, w \rangle = \bar{z}_1 w_1 + \dots + \bar{z}_n w_n,$$

la forme quadratique associée est

$$Q(w) = \langle w, w \rangle = \sum_{i=1}^n \bar{w}_i w_i = \sum_{i=1}^n |w_i|^2.$$

2. Le produit scalaire L^2 sur l'espace vectoriel \mathcal{F} des fonctions continues $f : [a, b] \rightarrow \mathbb{C}$ est défini par

$$(f | g) = \int_a^b \overline{f(x)} g(x) dx,$$

la forme quadratique associée est

$$Q(f) = \int_a^b |f(x)|^2 dx.$$

3. L'espace des suites complexes de carré intégrable est l'espace vectoriel

$$\ell_2(\mathbb{C}) = \{\zeta = (z_i)_{i \in \mathbb{N}} \mid z_i \in \mathbb{C}, \sum_{i=1}^{\infty} |z_i|^2 < \infty\}.$$

On munit cet espace vectoriel du produit scalaire hermitien : $\langle \zeta, \xi \rangle = \sum_{i=1}^{\infty} \bar{z}_i x_i$. La forme quadratique associée est $Q(\zeta) = \sum_{i=1}^{\infty} |z_i|^2$.

4. Un produit scalaire hermitien est défini sur $M_n(\mathbb{C})$ par

$$\langle A, B \rangle = \text{Trace}(A^* B),$$

où $A^* = \bar{A}^T$ est la *matrice adjointe* de A .

Définition. Lorsque h est un produit scalaire hermitien sur V , on définit la *norme* d'un vecteur $w \in V$ par

$$\|w\| = \sqrt{Q(w)} = \sqrt{h(w, w)}.$$

Nous démontrons maintenant que l'inégalité de Cauchy-Schwarz est encore valable pour un produit scalaire hermitien.

Proposition 13.2.1 (Inégalité de Cauchy-Schwarz hermitienne). *Si V est un espace vectoriel complexe muni d'un produit scalaire hermitien $\langle \cdot, \cdot \rangle$, alors on a pour tous $x, y \in V$*

$$|\langle x, y \rangle| \leq \|x\| \|y\|.$$

De plus on a égalité si et seulement si x et y sont colinéaires.

Preuve. la preuve donnée dans le cas réel (voir Proposition 11.1.2) ne marche pas et doit être légèrement modifiée. On note

$$a = \langle x, y \rangle = \overline{\langle y, x \rangle},$$

et on veut montrer que $|a| \leq \|x\| \|y\|$. Si $a = 0$ il n'y a rien à montrer, sinon on pose $p(t) = \|ax + y\|^2$ où t est un paramètre réel. En utilisant les propriétés du produit scalaire hermitien, on calcule

$$p(t) = \|tax + y\|^2 = \langle tax + y, tax + y \rangle = t^2 \bar{a} a \langle x, x \rangle + t \bar{a} \langle x, y \rangle + t a \langle y, x \rangle + \langle y, y \rangle.$$

En particulier $p(t)$ est un polynôme à coefficients réel de degré 2, qui s'écrit

$$p(t) = \|x\|^2 |a|^2 t^2 + 2|a|^2 t + \|y\|^2,$$

dont le discriminant $\Delta = 4|a|^2 (|a|^2 - \|x\|^2 \|y\|^2)$ doit être négatif, c'est-à-dire $|\langle x, y \rangle| = |a| \leq \|x\| \|y\|$. De plus on a égalité si et seulement s'il existe $t \in \mathbb{R}$ tel que $y = -tax$. \square

On a alors les propriétés suivantes, analogues au cas du produit scalaire sur les espaces vectoriels réels.

Proposition 13.2.2. *La norme associée à un produit scalaire hermitien sur un espace vectoriel complexe V vérifie*

(i) $\|w\| \geq 0$ pour tout $w \in V$ et $\|w\| = 0$ si et seulement si $w = 0$.

(ii) $\|\lambda w\| = |\lambda| \|w\|$ pour tous $w \in V$ et $\lambda \in \mathbb{C}$.

(iii) $\|v + w\| \leq \|v\| + \|w\|$.

On a aussi une version du théorème de Pythagore : si $v \perp w$, i.e. si $\langle v, w \rangle = 0$, alors

$$\|v + w\|^2 = \|v\|^2 + \|w\|^2.$$

La notion d'espace vectoriel hermitien représente donc le pendant complexe de celle d'espace vectoriel euclidien. En particulier nous avons les propositions suivantes :

Proposition 13.2.3. *Soit W un sous-espace vectoriel d'un espace vectoriel hermitien (V, h) . Alors son complément orthogonal, défini par*

$$W^\perp = \{x \in V \mid h(x, w) = 0 \ \forall w \in W\},$$

est un sous-espace vectoriel complexe de V . De plus, on a $V = W \oplus W^\perp$.

Ce résultat se démontre comme dans le cas d'un espace vectoriel euclidien (voir le point (vi) du théorème 11.2.3). \square

Proposition 13.2.4. *Sur tout espace vectoriel hermitien (V, h) , on peut construire des bases orthonormales dans V , i.e. des bases $\{u_1, \dots, u_n\}$ telles que*

$$h(u_i, u_j) = \delta_{ij}.$$

Une telle base s'appelle aussi une base unitaire de (V, h) .

Le procédé de Gram-Schmidt peut être étendu aux espaces hermitiens. Rappelons l'algorithme : Étant donné une base $\{v_1, v_2, \dots, v_m\}$ de l'espace hermitien V , on construit une base unitaire $\{u_1, u_2, \dots, u_n\} \subset V$ en suivant les étapes suivantes :

1. On pose $u_1 = \frac{v_1}{\|v_1\|}$.
2. Pour $k \geq 2$, on suppose que les vecteurs $\{u_1, u_2, \dots, u_{k-1}\}$ ont été construits et on considère le vecteur \hat{u}_k obtenu en soustrayant de v_k les projections des vecteurs précédents :

$$\hat{u}_k = v_k - \sum_{j=1}^{k-1} \langle v_k, u_j \rangle u_j.$$

3. On normalise le vecteur \hat{u}_k pour obtenir u_k :

$$u_k = \frac{\hat{u}_k}{\|\hat{u}_k\|}.$$

13.3 Opérateurs dans les espaces hermitiens

Un endomorphisme \mathbb{C} -linéaire $T : V \rightarrow V$ d'un espace vectoriel hermitien V s'appelle aussi un *opérateur* de V . Il est utile d'observer que la matrice $A = (a_{ij})$ d'un opérateur T de V dans une base unitaire $\{e_1, \dots, e_n\}$ est donnée par

$$a_{ij} = \langle e_i, Te_j \rangle. \quad (13.2)$$

En effet, on a par définition $Te_j = \sum_{k=1}^n a_{kj}e_k$, par conséquent

$$\langle e_i, Te_j \rangle = \langle e_i, \sum_{k=1}^n a_{kj}e_k \rangle = \sum_{k=1}^n a_{kj} \underbrace{\langle e_i, e_k \rangle}_{=\delta_{ik}} = a_{ij}.$$

Il faut être attentif à la place de l'opérateur dans l'équation (13.2), on a en effet $\langle Te_i, e_j \rangle = \overline{a_{ji}}$.

A tout opérateur T de V , on peut associer une application $\phi_T : V \rightarrow \mathbb{C}$ définie par

$$\phi_T(x) = \langle x, Tx \rangle.$$

L'application ϕ_T vérifie $\phi_T(\lambda x) = |\lambda|^2 \phi_T(x)$ pour tout $x \in V$, ça n'est donc pas une forme quadratique au sens classique³ à valeur dans le corps \mathbb{C} .

Lemme 13.3.1. *L'application ϕ_T détermine l'opérateur T , i.e. si les opérateurs $T_1, T_2 \in \mathcal{L}(V)$ vérifient $\langle x, T_1 x \rangle = \langle x, T_2 x \rangle$ pour tout $x \in V$, alors $T_1 = T_2$.*

Preuve. Notons $T = (T_1 - T_2)$. Nous devons montrer que si $\langle w, Tw \rangle = 0$ pour tout $w \in V$, alors $T = 0$. Fixons $\alpha \in \mathbb{C}$ et calculons

$$\langle (\alpha x + y), T(\alpha x + y) \rangle = |\alpha|^2 \langle x, Tx \rangle + \bar{\alpha} \langle x, Ty \rangle + \alpha \langle y, Tx \rangle + \langle y, Ty \rangle.$$

Par hypothèse, on a

$$\langle x, Tx \rangle = \langle y, Ty \rangle = \langle (\alpha x + y), T(\alpha x + y) \rangle = 0,$$

donc nous avons

$$\bar{\alpha} \langle x, Ty \rangle + \alpha \langle y, Tx \rangle = 0,$$

pour tout $\alpha \in \mathbb{C}$, ce qui n'est possible que si $\langle x, Ty \rangle = \langle y, Tx \rangle = 0$ (prendre par exemple $\alpha = 1$, puis $\alpha = i$ pour le voir).

□

Remarque. L'analogue du lemme correspondant est faux dans le cas des opérateurs \mathbb{R} -linéaires dans un espace vectoriel euclidien. Dans le cas euclidien, la condition $\langle w, Tw \rangle = 0$ pour tout $w \in V$ entraîne que l'opérateur est antisymétrique. La preuve ci-dessus ne fonctionne pas car le corps de base est \mathbb{R} et donc $\bar{\alpha} = \alpha$.

3. On dit que ϕ est une *forme quadratique hermitienne*.

13.3.1 L'adjoint d'un opérateur

Définition. Soit V un espace vectoriel complexe muni d'un produit scalaire hermitien que l'on note $\langle \cdot | \cdot \rangle$ et soit $T : V \rightarrow V$ un opérateur (c'est-à-dire un endomorphisme \mathbb{C} -linéaire de V). On dit que l'opérateur $T^* : V \rightarrow V$ est l'*adjoint* de T si

$$\langle Tx | y \rangle = \langle x | T^*y \rangle$$

pour tous $x, y \in V$. Par le lemme 13.3.1, on sait que l'adjoint d'un opérateur, s'il existe, est unique.

Proposition 13.3.2. *L'adjoint possède les propriétés suivantes :*

- (a) *L'adjoint de l'adjoint de T est l'opérateur T lui-même : $T^{**} = T$.*
- (b) $(TS)^* = S^*T^*$
- (c) $(S + T)^* = S^* + T^*$
- (d) $(\lambda T)^* = \bar{\lambda}T^*$ pour tout $\lambda \in \mathbb{C}$.
- (e) *Si T est inversible, alors l'inverse de l'adjoint de T est égale à l'adjoint de l'inverse de T .*

Preuve. La preuve de ces propriétés est un simple jeu formel :

- (a) On a pour tous $x, y \in V$:

$$\langle x | Ty \rangle = \overline{\langle Ty | x \rangle} = \overline{\langle y | T^*x \rangle} = \langle T^*x | y \rangle = \langle x | T^{**}y \rangle.$$

- (b) On a pour tous $x, y \in V$:

$$\langle x | (TS)^*y \rangle = \langle TSx | y \rangle = \langle Sx | T^*y \rangle = \langle x | S^*T^*y \rangle$$

- (c) La propriété (c) vient de l'additivité du produit scalaire hermitien en chaque variable :

$$\langle (S + T)x | y \rangle = \langle Sx | y \rangle + \langle Tx | y \rangle = \langle x | S^*y \rangle + \langle x | T^*y \rangle = \langle x | (S^* + T^*)y \rangle$$

- (d) L'argument est semblable :

$$\langle (\lambda T)x | y \rangle = \bar{\lambda} \langle Tx | y \rangle = \bar{\lambda} \langle x | T^*y \rangle = \langle x | \bar{\lambda}T^*y \rangle$$

- (e) Notons $S = T^{-1}$ et I l'identité de V , alors $S^* = (T^*)^{-1}$ car

$$I = I^* = (ST)^* = T^*S^*.$$

□

Proposition 13.3.3. *Si $\dim(V) < \infty$, alors tout opérateur $T : V \rightarrow V$ admet un adjoint, qui est unique.*

Remarque. En dimension infinie, il existe des opérateurs qui n'ont pas d'adjoint.

Preuve. Nous avons déjà mentionné que l'unicité suit du lemme 13.3.1. Pour montrer l'existence, on se donne une base orthonormée $\{e_1, \dots, e_n\}$ de V . Alors on a $Te_j = \sum_{i=1}^n a_{ij}e_i$ avec $a_{ij} = \langle e_i, Te_j \rangle$. Par définition de l'adjoint, on a

$$\langle e_j | T^*e_k \rangle = \langle Te_j | e_k \rangle = \left\langle \sum_{i=1}^n a_{ij}e_i | e_k \right\rangle = \sum_{i=1}^n \bar{a}_{ij} \langle e_i | e_k \rangle = \bar{a}_{kj}.$$

Par conséquent, l'opérateur $T^* : V \rightarrow V$ défini par

$$T^*e_k = \sum_{j=1}^n \bar{a}_{kj}e_j, \quad (13.3)$$

est l'adjoint de T . □

Corollaire 13.3.4. Soit T un opérateur d'un espace hermitien V .

- (a) Si $A \in M_n(\mathbb{C})$ est la matrice de l'opérateur T dans une base orthonormée $\{e_1, \dots, e_n\} \subset V$, alors la matrice transposée-conjuguée \bar{A}^\top est la matrice de l'opérateur T^* dans la même base.
- (b) Si λ est valeur propre de T , alors $\bar{\lambda}$ est valeur propre de T^* .

Preuve. La première affirmation découle immédiatement de (13.3). Pour prouver (b), on observe que le point (a) nous apprend que le polynôme caractéristique de l'adjoint T^* est

$$\chi_{T^*}(t) = \chi_{\bar{A}^\top}(t) = \chi_{\bar{A}}(t),$$

c'est donc le polynôme dont les coefficients sont les conjugués complexes du polynôme $\chi_A(t) = \chi_T(t)$. Par conséquent

$$\chi_T(\lambda) = 0 \quad \Leftrightarrow \quad \chi_{T^*}(\bar{\lambda}) = 0.$$

□

Cette proposition justifie la notation suivante pour la matrice conjuguée d'une matrice $A \in M_n(\mathbb{C})$:

$$A^* = \bar{A}^\top,$$

et on dit que A^* est la *matrice adjointe* de A .

13.4 Endomorphismes normaux et le théorème spectral

Définition 13.4.1. Un opérateur T d'un espace hermitien V est dit *normal* s'il commute avec son adjoint : $TT^* = T^*T$.

Proposition 13.4.2. (a) Pour un opérateur T de V , les conditions suivantes sont équivalentes :

- (i) T est normal.
- (ii) $\langle Tx, Ty \rangle = \langle T^*x, T^*y \rangle$ pour tous $x, y \in V$.
- (iii) $\|Tx\| = \|T^*x\|$ pour tout $x \in V$.

- (b) Si T est normal et v est un vecteur propre de T pour la valeur propre $\lambda \in \mathbb{C}$, alors v est aussi un vecteur propre de T^* pour la valeur propre $\bar{\lambda}$.
- (c) Si $v, w \in V$ sont des vecteurs propres d'un opérateur normal T associés à des valeurs propres distinctes $\lambda, \mu \in \mathbb{C}$, alors $v \perp w$.

Preuve. (a) On montre d'abord (i) \Rightarrow (ii). Supposons que T est normal, alors

$$\langle Tx, Ty \rangle = \langle x, T^*Ty \rangle = \langle x, TT^*y \rangle = \langle T^*x, T^*y \rangle.$$

Pour montrer (ii) \Rightarrow (i), on remarque que si $\langle Tx, Ty \rangle = \langle T^*x, T^*y \rangle$ pour tous $x, y \in V$. Alors le calcul ci-dessus montre que $\langle x, T^*Ty \rangle = \langle x, TT^*y \rangle$ pour tous $x, y \in V$. Mais ceci n'est possible que si $TT^* = T^*T$.

L'implication (ii) \Rightarrow (iii) est évidente et l'implication inverse découle de la formule de polarisation (13.1).

(b) Il est facile de vérifier que $(T - \lambda I_V)$ est normal si et seulement si T est normal. En utilisant le point (a) on a donc pour tout vecteur non nul $v \in V$,

$$\begin{aligned} Tv = \lambda v &\Leftrightarrow \|(T - \lambda I_V)v\| = 0 \\ &\Leftrightarrow \|(T - \lambda I_V)^*v\| = 0 \\ &\Leftrightarrow \|(T^* - \bar{\lambda} I_V)v\| = 0 \\ &\Leftrightarrow T^*v = \bar{\lambda}v. \end{aligned}$$

Ce qui prouve que v est un vecteur propre de T pour la valeur propre λ si et seulement si v est aussi un vecteur propre de T^* pour la valeur propre $\bar{\lambda}$.

(c) Supposons que $Tv = \lambda v$ et $Tw = \mu w$ avec $\mu \neq \lambda$. Alors d'après le point précédent on sait que $T^*v = \bar{\lambda}v$. On a donc

$$\mu \langle v | w \rangle = \langle v | Tw \rangle = \langle T^*v | w \rangle = \langle \bar{\lambda}v | w \rangle = \lambda \langle v | w \rangle.$$

Ainsi $(\lambda - \mu) \langle v | w \rangle = 0$, et puisque $\mu \neq \lambda$ on conclut que $\langle v | w \rangle = 0$. □

Théorème 13.4.3 (Théorème spectral). *Un opérateur $T : V \rightarrow V$ d'un espace hermitien V est normal si et seulement s'il est orthogonalement diagonalisable, c'est à dire qu'il existe une base unitaire $\{e_1, \dots, e_n\} \subset V$ et $\lambda_1, \dots, \lambda_n \in \mathbb{C}$ tels que $Te_i = \lambda_i e_i$ pour $i = 1, \dots, n$.*

Preuve. Supposons que T est orthogonalement diagonalisable, alors on sait par la formule (13.3) que son adjoint est défini par $T^*e_i = \bar{\lambda}_i e_i$, pour $i = 1, \dots, n$. On alors

$$T(T^*e_i) = T(\bar{\lambda}_i e_i) = \bar{\lambda}_i T(e_i) = \bar{\lambda}_i \lambda_i e_i = |\lambda_i|^2 e_i,$$

et de même

$$T^*(Te_i) = T^*(\lambda_i e_i) = \lambda_i T^*(e_i) = \lambda_i \bar{\lambda}_i e_i = |\lambda_i|^2 e_i.$$

On a donc $T(T^*e_i) = T^*(Te_i)$ pour tous les vecteurs de la base $\{e_i\}$ de V , ce qui entraîne que $TT^* = T^*T$, i.e. T est un opérateur normal.

On démontre la réciproque par récurrence sur $n = \dim(V)$, en supposant que $n \geq 2$ car il n'y a rien à prouver si $n = 1$. Tout endomorphisme d'un espace vectoriel de dimension finie possède au moins un vecteur propre, que l'on peut supposer de norme 1. Notons e_1 ce vecteur propre et λ_1 la valeur propre correspondante. On a donc $Te_1 = \lambda_1 e_1$ et on a vu plus haut que $T^*e_1 = \bar{\lambda}_1 e_1$. Notons $W = e_1^\perp = \{x \in V \mid \langle e_1, x \rangle = 0\}$ l'orthogonal de e_1 . On sait que W est un sous-espace vectoriel de dimension $n - 1$ de V , montrons que W est invariant par T : en effet supposons que $x \in W$, alors

$$\langle e_1, Tx \rangle = \langle T^*e_1, x \rangle = \langle \bar{\lambda}_1 e_1, x \rangle = \lambda_1 \langle e_1, x \rangle = 0,$$

ce qui signifie que $Tx \in W$. On a ainsi montré que $T(W) \subset W$ et on note $T_W = T|_W : W \rightarrow W$ l'opérateur obtenu par restriction de T .

La première affirmation de la Proposition 13.4.2 implique que T_W est un opérateur normal de W et par hypothèse de récurrence, il existe donc une base orthonormée $\{e_2, \dots, e_n\}$ de W et $\lambda_2, \dots, \lambda_n \in \mathbb{C}$ tels que $T_W e_j = \lambda_j e_j$ pour $j = 2, \dots, n$.

Il est clair que la famille de vecteurs $\{e_1, e_2, \dots, e_n\}$ est une base unitaire de V et que c'est une base propre pour T . □

Nous pouvons reformuler le théorème spectral sous la forme importante suivante :

Théorème 13.4.4 (Théorème spectral, variante). *Soit V un espace vectoriel hermitien et T un opérateur linéaire de V . Alors T est normal si et seulement s'il peut être écrit sous la forme suivante :*

$$T = \sum_{j=1}^r \lambda_j P_j, \quad (13.4)$$

où $\sigma(T) = \{\lambda_1, \dots, \lambda_r\} \subset \mathbb{C}$ est le spectre de T et $P : V \rightarrow E_j$ est le projecteur orthogonal sur le sous-espace propre $E_j = \text{Ker}(T - \lambda_j I_V)$.

Notons que V se décompose en somme directe orthogonale :

$$V = E_1 \oplus \dots \oplus E_r, \quad (E_i \perp E_j, \text{ si } i \neq j),$$

et que les projecteurs P_j vérifient les propriétés suivantes :

- (1) $P_j^2 = P_j$.
- (2) $P_i \circ P_j = 0$ si $i \neq j$.
- (3) $\text{Im}(P_j) = E_j$ et $\text{ker}(P_j) \perp E_j$.
- (4) $I_V = \sum_{j=1}^r P_j$.

On formule ces propriétés en disant que $\{P_k\}$ est un *système complet de projecteurs orthogonaux*.

Lorsque l'opérateur T possède n valeurs propres distinctes, alors la décomposition spectrale prend la forme simplifiée suivante :

$$T(x) = \sum_{j=1}^n \lambda_j \langle e_j, x \rangle e_j. \quad (13.5)$$

13.5 Opérateurs auto-adjoints et unitaires

Définition 13.5.1. Un opérateur $T : V \rightarrow V$ d'un espace hermitien V est dit :

- (1) Autoadjoint (ou hermitien), si $T = T^*$.
- (2) Anti-autoadjoint, si $T = -T^*$.
- (3) Unitaire, si $TT^* = I_V$.

Ces trois types d'opérateurs sont clairement normaux.

Proposition 13.5.2. *Un opérateur T de V est autoadjoint si et seulement $\langle x, Tx \rangle$ est réel pour tout $x \in V$.*

Preuve. Supposons que $T : V \rightarrow V$ est auto-adjoint, alors

$$\langle x, Tx \rangle = \langle T^*x, x \rangle = \langle Tx, x \rangle = \overline{\langle x, Tx \rangle},$$

ce qui implique que $\langle x, Tx \rangle$ est réel. Supposons inversement que $\langle x, Tx \rangle$ est réel pour tout vecteur $x \in V$, alors

$$\langle x, Tx \rangle = \overline{\langle x, Tx \rangle} = \langle Tx, x \rangle = \langle x, T^*x \rangle,$$

et le lemme 13.3.1 entraîne alors que $T = T^*$. □

Le théorème spectral permet de prouver facilement les caractérisation suivantes :

- Corollaire 13.5.3.** (a) *Un opérateur de V est autoadjoint si et seulement s'il est normal et toutes ses valeurs propres sont réelles.*
- (b) *Un opérateur de V est anti-autoadjoints si et seulement s'il est normal et toutes ses valeurs propres sont imaginaires.*
- (c) *Un opérateur de V est unitaire si et seulement s'il est normal et toutes ses valeurs propres des nombres complexes de modules 1.*

Proposition 13.5.4. (a) *L'ensemble des opérateurs autoadjoints de V forme un sous-espace vectoriel réel de $\mathcal{L}(V)$.*

(b) *L'ensemble des opérateurs unitaires de V forme un sous-groupe de $\text{GL}(V)$, que l'on note $\text{U}(V)$ et qui s'appelle le groupe unitaire de V .*

Le groupe unitaire de \mathbb{C}^n , pour le produit scalaire hermitien standard est noté $\text{U}(n)$.

Noter par contraste que l'ensemble des opérateurs normaux ne forme pas un sous-espace vectoriel. Si S et T sont normaux alors $S + T$ n'est en général pas un opérateur normal.

13.6 Espaces de Hilbert et opérateurs auto-adjoints

Cette section est facultative

Définition 13.6.1. On appelle *espace de Hilbert* un espace vectoriel \mathcal{H} sur le corps \mathbb{R} ou \mathbb{C} muni d'un produit scalaire $\langle \cdot, \cdot \rangle$ (qui est un produit scalaire hermitien dans le cas complexe) et qui est complet pour la norme $\| \cdot \|$ associée au produit scalaire. Cette condition signifie que toute suite de Cauchy de \mathcal{H} doit converger.

Exemples 1. Tout espace vectoriel de dimension finie sur \mathbb{R} ou \mathbb{C} muni d'un produit scalaire (produit scalaire hermitien dans le cas complexe) est un espace de Hilbert car le théorème de Bolzano-Weierstrass entraîne que toute suite de Cauchy de \mathbb{R}^n ou \mathbb{C}^n converge. La notion d'espace de Hilbert généralise donc à la fois les espaces euclidiens et les espaces hermitiens et autorise la dimension infinie.

2. L'espace ℓ_2 est un espace de Hilbert (c'est le prototype d'espace de Hilbert en dimension infinie).

3. L'espace $C^0([a, b], \mathbb{C})$ des fonctions continues sur un intervalle $[a, b]$ n'est pas un espace de Hilbert (car la limite d'une suite de fonctions continues n'est pas toujours continue si la convergence n'est pas uniforme, ce qui implique que l'espace $C^0([a, b], \mathbb{C})$ n'est pas complet pour la norme associée au produit scalaire L^2).

Définition 1.) Soit \mathcal{H} un espace de Hilbert. Un *opérateur* de \mathcal{H} est une application \mathbb{R} -linéaire ou \mathbb{C} -linéaire de \mathcal{H} dans lui-même qui est continue pour la norme associée.

2.) L'opérateur $T : \mathcal{H} \rightarrow \mathcal{H}$ est dit *auto-adjoint* si la condition suivante est vérifiée :

$$\langle Tx, y \rangle = \langle x, Ty \rangle,$$

pour tous $x, y \in \mathcal{H}$.

Exemple. On peut associer à toute fonction continue $\varphi : [a, b] \times [a, b] \rightarrow \mathbb{C}$ un opérateur T sur $C^0([a, b], \mathbb{C})$ par la formule

$$(Tf)(x) = \int_a^b f(y) \varphi(x, y) dy.$$

Cet opérateur est auto-adjoint si et seulement si $\varphi(y, x) = \overline{\varphi(x, y)}$ pour tous $x, y \in [a, b]$.

Définition. L'opérateur T sur l'espace de Hilbert \mathcal{H} admet une *décomposition spectrale finie* si on peut écrire

$$T = \lambda_1 P_1 + \cdots + \lambda_r P_r$$

où P_1, \dots, P_r sont des projecteurs de \mathcal{H} qui sont deux-à-deux orthogonaux.

Le théorème spectral nous dit que tout espace de Hilbert de dimension finie admet une décomposition spectrale. Un opérateur autoadjoint d'un espace de Hilbert de dimension infinie admet aussi une décomposition spectrale qui peut être infinie. Cela donne une écriture du type

$$T = \sum_{i=1}^{\infty} \lambda_i P_i,$$

lorsqu'il y a un nombre dénombrable de valeurs propres. Dans le cas le plus général, le spectre n'est pas forcément un sous-ensemble discret de \mathbb{R} et la décomposition spectrale s'écrit sous la forme d'une intégrale que l'on étend au spectre continu de T :

$$T = \int_{\sigma(T)} \lambda P_\lambda d\lambda$$

Un chapitre important de l'*analyse fonctionnelle* est de donner un sens précis à ce genre de formules et de les démontrer rigoureusement.

13.7 Applications en mécanique quantique.

Cette section est facultative

En mécanique Newtonienne, l'état d'un système évolue selon une loi déterministe et les quantités observables sont des fonctions des variables d'état que l'on peut (en principe) connaître exactement à chaque instant.

L'exemple le plus simple est la mécanique classique du point matériel. L'état du système au temps t est déterminé par la position $x(t) \in \mathbb{R}^3$ et le moment $p = m\dot{x} \in \mathbb{R}^3$ de la particule. La loi d'évolution est prescrite par l'équation de Newton : $\frac{dp}{dt} = F$.

La formalisation mathématique de la mécanique quantique repose sur un certain nombre de postulats que nous décrivons brièvement ci-dessous sans chercher à être ni rigoureux ni exhaustif :

Premier postulat : L'état d'un système quantique (par exemple une particule) au temps t est représenté par un élément non nul $\psi = \psi(t)$ d'un espace de Hilbert \mathcal{H} , appelé *vecteur d'état* du système (typiquement une *fonction d'onde*).

Deux vecteurs $\psi_1, \psi_2 \in \mathcal{H}$ représentent le même état si l'un est multiple de l'autre. On supposera donc souvent que le vecteur d'état est normalisé $\|\psi\| = 1$.

Deuxième postulat : On associe à chaque observable du système un opérateurs autoadjoints T de l'espace de Hilbert \mathcal{H} .

Par le théorème spectral, Il existe alors une décomposition de \mathcal{H} en somme directe orthogonale de sous-espaces propres pour T . Les vecteurs propres de T s'appellent des *états propres* pour l'opérateur T , et tout état ψ est une "superposition" d'états propres (le mot "superposition" veut dire ici "combinaison linéaire", la décomposition de ψ en combinaison linéaire de vecteurs propres de T s'appelle en mécanique quantique une *superposition d'états propres*).

Troisième postulat : Lors d'une expérience, on ne peut observer que les états propres. En conséquence, ce qui est mesuré est une valeur propre de l'opérateur T . Puisque cet opérateur est autoadjoint, ses valeurs propres sont des nombres réels.

Si on a la décomposition spectrale $T = \sum_i \lambda_i P_i$, alors toute expérimentation visant à mesurer l'observable T a pour effet de projeter un état ψ (c'est-à-dire un vecteur $\psi \in \mathcal{H}$) sur un sous-espace propre de T :

$\psi \mapsto P_j(\psi)$, de plus, la mesure obtenue est la valeur propre λ_i , avec une probabilité

$$p_j = \left| \frac{\langle \psi | P_j(\psi) \rangle}{\|\psi\| \|P_j(\psi)\|} \right|^2.$$

Remarques.

- (a) Si $\psi, \phi \in \mathcal{H}$ sont deux vecteurs non nuls de \mathcal{H} qui représentent deux états possibles d'un système quantique, alors on dit que le nombre complexe

$$\frac{\langle \psi | \phi \rangle}{\|\psi\| \|\phi\|} \in \mathbb{C}$$

représente l'*amplitude de probabilité* que le système préparé dans l'état ψ soit observé dans l'état ϕ . La probabilité elle-même est le carré du module de cette amplitude.

L'importance de cette notion vient du fait qu'en mécanique quantique, les calculs de transitions et de comportement des systèmes se font en manipulant algébriquement les amplitudes de probabilité et non directement les probabilités.

- (b) Si l'opérateur auto-adjoint T possède $n = \dim(\mathcal{H})$ valeurs propres distinctes, alors la décomposition spectrale prend la forme simplifiée dans une base unitaire (voir (13.5)) :

$$T(\psi) = \sum_{j=1}^n \lambda_j \langle e_j, \psi \rangle e_j.$$

Les amplitudes de probabilités sont alors données par $\langle e_j, \psi \rangle$ et les probabilités de transition par $p_j = |\langle e_j, \psi \rangle|^2$ (on suppose $\|\psi\| = 1$), cette quantité représente la probabilité que le système dans l'état ψ soit observé dans l'état propre e_j .

- (c) Si l'espace des états est un espace de Hilbert de dimension infinie, alors le spectre de l'opérateur T (l'ensemble des valeurs propres) peut-être continu ou discret. Lorsque le spectre est discret, cela implique que la variable observée ne peut prendre que des valeurs discrètes (principe de quantification).

Quatrième postulat : Le dernier postulat nous dit que loi d'évolution du système est prescrite par l'équation de Schrödinger :

$$i\hbar \frac{d}{dt} \psi = H(\psi)$$

où $i = \sqrt{-1}$, $\hbar = \frac{ih}{2\pi}$ est la constante de Plank réduite et H est un opérateur autoadjoint de l'espace de Hilbert \mathcal{H} qu'on appelle le *Hamiltonien du système*.

Remarques.

- (1) Ces postulats n'ont clairement rien d'intuitifs, ils ont été développés dans les années 1920-1940 par les Pères fondateurs de la mécanique quantique (Bohr, Heisenberg, Dirac...). L'histoire de la mécanique quantique est l'un des chapitres les plus complexes et passionnants de l'histoire des sciences.
- (2) Suivant les auteurs, l'ordre des postulats, leur nombre et leurs formulations exactes peuvent présenter des variations assez importantes.

- (3) En principe l'espace de Hilbert est de dimension infinie, mais des modèles simplifiés peuvent être développés sur la base d'un espace de Hilbert de dimension finie (en négligeant donc une partie de l'information). Cette approche est semblable à celle qui consiste à réduire le nombre de dimensions (par exemple de 3 à 2 ou à 1) dans un problème de mécanique classique. Ces modèles de dimension finie ont l'avantage de ramener la mécanique quantique à de l'algèbre linéaire. Ils sont très efficaces en chimie pour modéliser les orbitales des électrons dans un modèle quantique d'atome.
- (4) Si l'espace de Hilbert est de dimension finie, l'équation de Schrödinger peut en principe se résoudre directement en calculant l'exponentielle de la matrice du Hamiltonien H :

$$i\hbar \frac{d}{dt} \psi = H(\psi) \quad \Rightarrow \quad \psi(t) = e^{-\frac{i}{\hbar} t H} \cdot \psi_0$$

- (5) Le hamiltonien est lui-même un opérateur autoadjoint. Selon le deuxième postulat il lui correspond donc une quantité observable. La théorie montre que cette observable est l'*énergie totale* du système (qui est donc constante au cours du temps, conformément au principe de conservation de l'énergie).

Sur la fonction d'onde

Dans le modèle de Schrödinger, l'état de la particule-onde est représenté à chaque instant par une *fonction d'onde* ψ_t , qui est une fonction

$$\psi_t : \mathbb{R}^3 \rightarrow \mathbb{C} \quad \text{telle que} \quad \int_{\mathbb{R}^3} |\psi_t|^2 = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |\psi_t(x, y, z)|^2 dx dy dz < \infty.$$

On considère que deux fonctions d'onde ψ et ϕ représentent le même état si $\phi = \alpha\psi$ où α est un nombre complexe non nul. Cela nous permet de normaliser la fonction d'onde, i.e. de supposer que

$$\int_{\mathbb{R}^3} |\psi_t|^2 = 1.$$

On a alors l'interprétation suivante : *la probabilité de présence de la particule dans une région $\Omega \subset \mathbb{R}^3$ est donnée par l'intégrale*

$$\text{Prob}(\psi_t | \Omega) = \int_{\Omega} |\psi_t|^2.$$

Ainsi la quantité $|\psi_t(x)|^2$ représente une *densité de probabilité de présence de la particule au point x* . Par comparaison l'argument complexe de ψ représente une simple phase et n'a pas d'interprétation physique (car les fonctions d'onde ψ et $e^{i\theta}\psi$ représentent le même état de la particule).

L'ensemble des fonctions intégrables⁴ $\phi : \mathbb{R}^3 \rightarrow \mathbb{C}$ vérifiant

$$\int_{\mathbb{R}^3} |\phi|^2 < \infty$$

4. intégrable au sens de Lebesgue... mais ne nous inquiétons pas de cela.

forme un espace de Hilbert, que l'on note $L^2(\mathbb{R}^3, \mathbb{C})$. Le produit scalaire hermitien sur cet espace est défini par

$$\langle \phi, \psi \rangle = \int_{\mathbb{R}^3} \overline{\phi(x)} \psi(x) dx.$$

A priori une fonction $\phi \in L^2(\mathbb{R}^3, \mathbb{C})$ n'est pas forcément continue, elle n'a a fortiori pas de dérivée partielle. Mais on démontre en analyse fonctionnelle que l'espace $L^2(\mathbb{R}^3, \mathbb{C})$ contient un sous-espace vectoriel $\mathcal{H} \subset L^2(\mathbb{R}^3, \mathbb{C})$ qui est dense (c'est-à-dire que tout élément de $L^2(\mathbb{R}^3, \mathbb{C})$ peut s'approximer par une suite d'éléments de \mathcal{H} , de même que tout nombre réel peut s'approximer par une suite de nombres rationnels) et qui a la propriété suivante :

$$\psi \in \mathcal{H} \quad \Rightarrow \quad \Delta \psi \in \mathcal{H},$$

où Δ est l'opérateur de Laplace défini par

$$\Delta \psi = \frac{\partial^2 \psi}{\partial x^2} + \frac{\partial^2 \psi}{\partial y^2} + \frac{\partial^2 \psi}{\partial z^2}.$$

Le hamiltonien est alors donné par

$$H(\psi) = \left(V - \frac{\hbar^2}{2m} \Delta \right) \cdot \psi,$$

où m est la masse de la particule et $V : \mathbb{R}^3 \rightarrow \mathbb{R}$ est une fonction qui représente l'énergie potentielle. Ce hamiltonien représente la version quantique de la quantité "énergie totale = énergie cinétique + énergie potentielle".

Exercice. Prouver que H est autoadjoint, i.e.

$$\int_{\mathbb{R}^3} \overline{\phi} H(\psi) dx = \int_{\mathbb{R}^3} \overline{H(\phi)} \psi dx$$

(on peut supposer pour cet exercice que ϕ et ψ sont des fonctions de classe C^2 qui tendent vers 0, ainsi que leur dérivées, lorsque x tend vers l'infini).

En conclusion, l'état de la particule est représenté par sa fonction d'onde $\psi_t \in \mathcal{H}$ qui est normalisée, i.e. $\int_{\mathbb{R}^3} |\phi|^2 = 1$, et qui vérifie l'équation de Schrödinger :

$$i\hbar \frac{\partial \psi}{\partial t} = \left(V - \frac{\hbar^2}{2m} \Delta \right) \psi.$$