Physique Statistique PHYS 338
2024-2025

Série 1: Laplace et Legendre

PIERRE-SIMON LAPLACE (1749-1827), mathématicien et astronome français, est célèbre pour ses travaux en mécanique céleste et en théorie des probabilités. Il fut brièvement ministre de l'Intérieur sous Napoléon, puis fait marquis sous la Restauration, illustrant sa capacité à naviguer les changements politiques de son époque. ADRIEN-MARIE LEGENDRE (1752-1833), également mathématicien français, est connu pour ses contributions en analyse et en théorie des nombres. Lorsqu'il découvrit que Gauss travaillait sur des théories similaires, il publia précipitamment ses résultats pour établir sa priorité, montrant que la compétition scientifique existait déjà à cette époque! Leurs travaux se révèlent fondamentaux en physique statistique.

* Exercise 1 Formule de Stirling

Avant de parler de Laplace, nous allons nous intéresser à JAMES STIRLING (1692-1770), un mathématicien écossais connu pour ses contributions significatives à l'analyse et à la théorie des nombres. On veut démontrer la célèbre formule de Stirling donnant le comportement asymptotique de la fonction factoriel dont on rappelle la définition:

$$N! := N \cdot (N-1) \cdot \dots \cdot 2 \cdot 1 = \prod_{i=1}^{N} i \,. \tag{1}$$

cette formule s'écrit

$$\underline{\text{formule de Stirling:}} \qquad \boxed{N! \underset{N \gg 1}{\approx} N^N e^{-N} \sqrt{2\pi N}}\,, \tag{2}$$

qu'il faut entendre comme la formulation informelle de la limite suivante:

$$\lim_{N \to \infty} \frac{N!}{N^N e^{-N} \sqrt{2\pi N}} = 1 \,. \tag{3}$$

Pour la démontrer, on va utiliser la représentation intégrale de la fonction Gamma, qui est la continuation analytique de la fonction factorielle:

$$N! = \Gamma(N+1) = \int_0^\infty e^{-t}\, t^N \, \mathrm{d}t \,. \tag{4}$$

Q1. En utilisant un changement de variable, montrer que l'on peut écrire $N!$ sous la forme:

$$N! = e^{N \ln N} N \int_0^\infty e^{N f(t)} \, \mathrm{d}t \tag{5}$$

où $f(t)$ est une fonction que l'on précisera.

Q2. Comment se comporte la fonction $e^{Nf(t)}$ pour $N \gg 1$? Justifier que l'on peut approximer $f(t)$ par un développement de Taylor au second ordre autour de son maximum.

Q3. On rappelle la formule de l'intégrale Gaussienne

$$I = \int_{-\infty}^{\infty} e^{-\frac{x^2}{2\Delta}}\, \mathrm{d}x = \sqrt{2\pi\Delta} \tag{6}$$

Montrez alors la formule de Stirling pour l'approximation (2) de $N!$. Quelle est son erreur absolue et relative (numérique) pour $N = 10, 50$? Et avez vous un idée pour l'évaluer numériquement pour $N = 10^{23}$?

Solution of Exercise 1

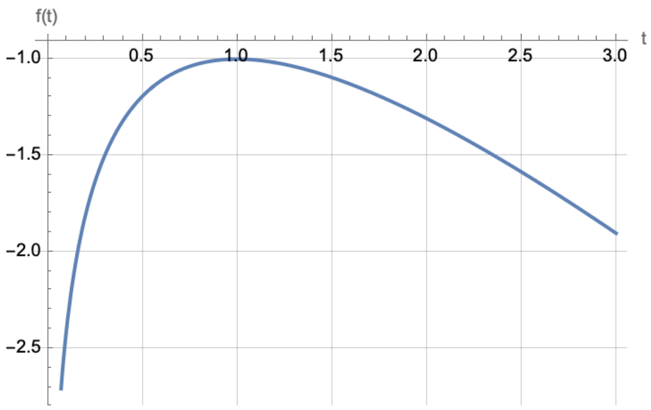
La formule de Stirling Eq. (2) que l'on souhaite prouver a une interprétation simple: si l'on se réfère à la définition de $N!$, le terme en N^N correspond à remplacer dans le produit de l'équation (1), chacun des termes $(N-k)$ avec $k > 1$, par N. Chacun de ces termes est plus petit que N est donc on s'attend à ce qu'il faut "contre-balancer" le terme N^N par des termes décroissant avec N pour espérer avoir une bonne estimation de $N!$. La démonstration qui suit montre que ces termes décroissants sont donnés comme le produit de 2 termes: un décroissance exponentiel e^{-N} et un terme sous-dominant en $O(\sqrt{N})$. Pour illustrer ceci, prenons par exemple $N = 10$, la valeur (exacte) de la factorielle donne $10! = 3628800$ dont l'ordre de grandeur est 10^6 et l'approximation crue $N^N = 10^{10}$ est donc assez loin de la véritable valeur. Si l'on utilise maintenant l'approximation de Stirling (terme de droite dans Eq. (2)), on tombe sur $3,59 \cdot 10^6$, ce qui est beaucoup plus proche de la valeur réelle et pourtant N n'est pas très grand !

Q1. Le changement de variable a considéré est $t = Nt'$. on a donc

$$N! = \int_0^{\infty} e^{-t}\, t^N\, dt = \int_0^{\infty} e^{-t}\, e^{N\log t}\, dt = \int_0^{\infty} e^{-Nt'}\, e^{N\log Nt'}\, N dt' = e^{N\log N} N \int_0^{\infty} e^{-Nt'}\, e^{N\log t'}\, dt' \tag{7}$$

et donc $f(t) = \log(t) - t$.

Q2. $e^{Nf(t)}$ prend la valeur maximale $e^{Nf(1)} = e^{-N}$. Comme $f < 0$, pour $N \gg 1$ et n'importe quelle autre valeur $t \neq 1$, $e^{Nf(t)}$ se retrouve être " exponentiellement plus petit" que sa valeur max (en valeur relative). C'est à dire que l'intégrale devient très piquée autour de son maximum. Cela justifie l'approximation de $f(t)$ par un développement de Taylor au second ordre autour de ce maximum.
Le graphique de la fonction est



Le maximum de $f(t) = \ln(t) - t$ se trouve à $t = 1$ ($f'(1) = 0$). Autour de ce point, le développement de Taylor donne :

$$f(t) = f(1) + \frac{1}{2}f''(1)(t-1)^2 + \frac{1}{6}\frac{2}{\xi(t)^3}(t-1)^3 = -1 - \frac{1}{2}(t-1)^2 + \frac{1}{6}\frac{2}{\xi(t)^3}(t-1)^3 \tag{8}$$

pour quelque $\xi(t) \in [1, t]$ Cette approximation capture le comportement essentiel de la fonction près de son pic, ce qui est crucial pour l'intégrale quand N est grand.

Q3. En utilisant le développement de Taylor de $f(t)$ autour de $t = 1$, nous avons :

$$N! = e^{N \ln N} N \int_0^{\infty} e^{N(-1-\frac{1}{2}(t-1)^2)+\frac{1}{6}\frac{2}{\xi^3}(t-1)^3} dt \tag{9}$$

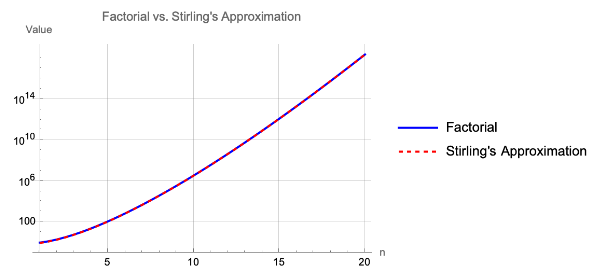
on peut oublier le reste pour nous simplifier le calcul. En effectuant le changement de variable $x = \sqrt{N}(t-1)$, on obtient

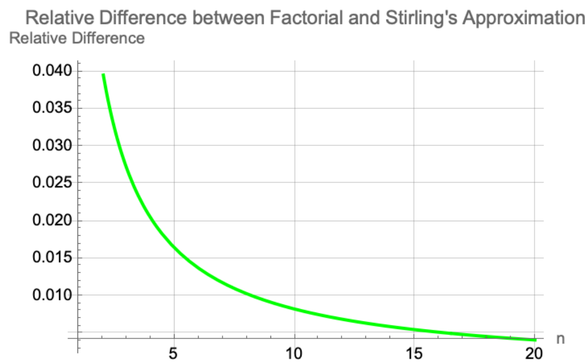
$$N! \approx e^{N \ln N} N e^{-N} \frac{1}{\sqrt{N}} \int_{-\sqrt{N}}^{\infty} e^{-x^2/2}\, \mathrm{d}x \tag{10}$$

Pour N grand, on peut étendre la borne inférieure à $-\infty$ sans introduire d'erreur significative. En utilisant la formule de l'intégrale gaussienne, on obtient :

$$N! \approx e^{N \ln N} N e^{-N} \frac{1}{\sqrt{N}} \sqrt{2\pi} = N^N e^{-N} \sqrt{2\pi N} \tag{11}$$

Ce qui démontre la formule de Stirling.
On montre la difference entre le vrai valuer de $N!$ et l'approximation





* Exercise 2 Méthode de Laplace

La méthode de l'exercice précédent est en fait très générale, et se nomme méthode de Laplace, formellement ette méthode se résume au résultat suivant, si l'on dénote par

$$I_N := \int_a^b e^{Nf(x)}\, \mathrm{d}x \tag{12}$$

alors on a

$$I_N \asymp e^{Nf(x^\star)} \qquad \text{pour } N \gg 1\,, \tag{13}$$

où $x^\star$ est la valeur maximisant la fonction $f(x)$ et $\asymp$ désigne (formellement) l'équivalence "à l'échelle logarithmique" [1] dont on donne la formulation rigoureuse à l'équation (19).

Cette technique (qui permet de remplacer une intégration a priori difficile par une simple maximisation) est tellement fondamentale en physique statistique que nous allons devoir la démontrer proprement: Pour simplifier le problème, on suppose dans la suite du problème l'hypothèse suivante:

Hyp: $f \in C^2([a,b])^2$, $\exists!\, x^* \in]a,b[$ *tel que* $f'(x^*) = 0$ *et de plus on suppose que* $f''(x^*) \neq 0$.

[1] c'est à dire que pour deux séquences $(a_N)_N, (b_N)_N$, on écrit $a_N \asymp b_N$ si $\lim_{N\to\infty} \frac{1}{N}\log(a_N) = \lim_{N\to\infty} \frac{1}{N}\log(b_N)$.

[2] On rappelle que $\mathcal{C}^k(I)$ est le sous-ensemble de l'ensemble des fonctions de I vers $\mathbb{R}$ qui sont k fois dérivables constitué des fonctions dont la k-ième dérivée est continue.

Pour obtenir le résultat voulu, considérons l'intégrale suivante :

$$\mathcal{I}_N := \frac{I_N}{e^{Nf(x^*)}} = \int_a^b e^{N(f(x)-f(x^\star))}\,\mathrm{d}x = \int_{a-x^\star}^{b-x^\star} e^{N(g(x)-g(0))}\,\mathrm{d}x \tag{14}$$

avec $g(x) = f(x+x^\star)$ qui a son maximum en 0.

Q1. Justifiez que l'on peut toujours, pour une certaine valeur de γ, diviser l'intervalle d'intégration en ne conservant qu'un intervalle $[-\delta,\delta]$ autour duquel la fonction $g(x)$ admet une dérivée seconde négative.

$$\exists\gamma,\delta>0 : \int_{-\delta}^{\delta} e^{N(g(x)-g(0))} \le \mathcal{I}_N \le (b-a)e^{-\gamma N} + \int_{-\delta}^{\delta} e^{N(g(x)-g(0))} \tag{15}$$

Q2. Justifier (par exemple graphiquement) que l'on peut toujours écrire, dans l'intervalle $[-\delta,\delta]$, que

$$\frac{x^2}{2}g''(0)(1+\epsilon) \le g(x)-g(0) \le \frac{x^2}{2}g''(0)(1-\epsilon) \tag{16}$$

avec $\epsilon>0$ arbitrairement petit quand l'on rétrécie l'intervalle $[-\delta,\delta]$.

Q3. Montrez que l'on peut par ailleurs écrire, en utilisant la borne de Hoeffding pour l'intégrale Gaussienne[3]

$$\int_{-\delta}^{\delta} e^{-NC\frac{x^2}{2}}\,\mathrm{d}x = \sqrt{\frac{2\pi}{NC}} + O(e^{-N}) \tag{17}$$

Q4. Finalement montrer que cela implique que pour tout ϵ on a:

$$O(e^{-N}) + \sqrt{\frac{2\pi}{-Ng''(0)(1+\epsilon)}} \le \mathcal{I}_N \le O(e^{-N}) + \sqrt{\frac{2\pi}{-Ng''(0)(1-\epsilon)}}$$

Q5. Cela nous permet d'écrire plus rigoureusement la signification de la méthode de Laplace, que l'on peut formuler de deux façons : la première, la plus precise

$$\underline{\text{Méthode de Laplace } (2^{nd} \text{ ordre}):} \quad \lim_{N\to\infty} \frac{\sqrt{-Nf''(x^\star)}}{\sqrt{2\pi}} \frac{\int_a^b e^{Nf(x)}\,\mathrm{d}x}{e^{Nf(x^\star)}} = 1 \tag{18}$$

et la seconde, plus utile en physique statistique, étant

$$\underline{\text{Méthode de Laplace } (1^{er} \text{ ordre}):} \quad \boxed{\lim_{N\to\infty}\frac{1}{N}\log\int_a^b e^{Nf(x)}\,\mathrm{d}x = f(x^\star) = \max_x f(x)} \tag{19}$$

Prouvez la version précise et donnez l'intuition pour la version que on va utiliser en physique.

Q6. Argumenter pourquoi si le maximum x^* n'est plus unique mais qu'il existe un nombre fini de MAXIMUM, on peut écrire a la place, en somment sur tout les supremum x_i^*, que

$$I_N = \int_a^b e^{Nf(x)}\,\mathrm{d}x \approx \sum_i \sqrt{\frac{2\pi}{f''(x_i^\star)}} e^{Nf(x_i^\star)} \text{ quand } N\to\infty \tag{20}$$

et que

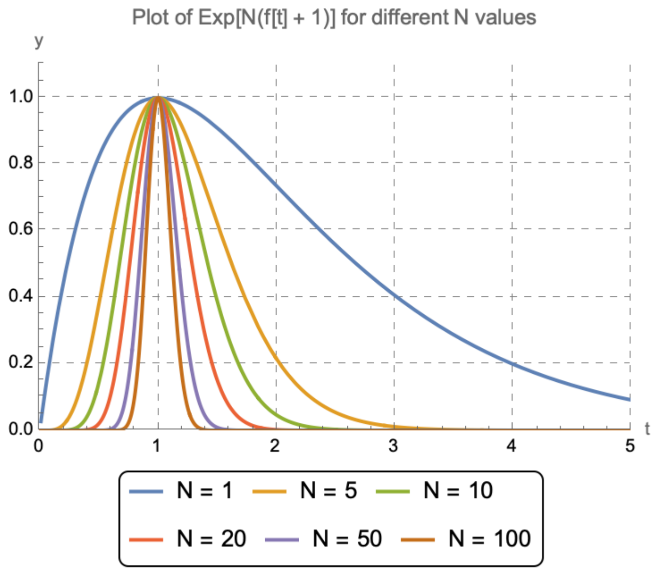
$$\lim_{N\to\infty}\frac{1}{N}\log\int_a^b e^{Nf(x)}\,\mathrm{d}x = \sup_x f(x) \tag{21}$$

[3]Nous verrons en effet dans le prochain cours une formule très utile pour une variable Gaussienne de moyenne nulle on a que pour tout $k\ge 0$:

$$\mathbb{P}(X>k) = \frac{1}{2\pi\Delta}\int_k^\infty e^{-x^2/2\Delta} \le e^{-k^2/2\Delta}$$

Solution of Exercise 2

La méthode de Laplace (et ses variantes!) est un outil fondamental en physique statistique/théorique, elle stipule que "la somme de termes exponentiellement grands est dominée par le plus grand d'entre eux". Pour s'en convaincre reprenons l'exemple de l'exercice précédent, où l'on a tracé la valeur de $\mathrm{e}^{N(f(x)-f(x^*))}$ *pour différentes valeurs de* x*:*



où l'on voit la concentration de l'intégrant autour de x^*. *Avant de passer à la preuve rigoureuse, donnons d'abord une preuve heuristique de ce résultat: comme* $\mathrm{e}^{N(f(x)-f(x*))}$ *est petit, par théorème de Taylor, on a*

$$I_N \approx \int_a^b \exp\Big\{N\big(f(x^*) + f''(x^*)(x-x^*)^2/2\big) + \mathcal{E}\Big\}\,\mathrm{d}x\ , \tag{22}$$

où $\mathcal{E}$ *est le terme d'erreur que l'on va négliger. On a donc*

$$I_N \approx \mathrm{e}^{Nf(x^*)} \cdot \int_a^b \exp\Big\{N\big(f''(x^*)(x-x^*)^2/2\big)\Big\}\,\mathrm{d}x\ , \tag{23}$$

comme $f''(x^*) < 0$*, pour* $N \gg 1$*;* $\exp\Big\{N\big(f''(x^*)(x-x^*)^2/2\big)$ *décroît très fortement (avec* x*) dès que l'on s'éloigne de* x^* *et l'on peut remplacer l'intégration entre* $[a, b]$ *par l'intégration sur toute la droite réelle (cf. inégalité de Hoeffding de l'exercice précédent):*

$$I_N \approx \mathrm{e}^{Nf(x^*)} \cdot \int_{-\infty}^{\infty} \exp\Big\{N\big(f''(x^*)(x-x^*)^2/2\big)\Big\}\,\mathrm{d}x\ , \tag{24}$$

et obtient le résultat voulu par intégration Gaussienne. L'exercise propose de rendre cet argument heuristique rigoureux.

Q1. Par hypothèse, on a $f''(x^\star) < 0$ et donc aussi que $g''(0) < 0$. Comme la fonction exponentielle est toujours positive, on a que $\forall \delta < \delta^\star = \min(x^\star - a, b - x^\star)$ si on réduit la borne d'intégration

$$\int_{a-x^\star}^{b-x^\star} e^{N(g(x)-g(0))}\,\mathrm{d}x \geq \int_{-\delta}^{\delta} e^{N(g(x)-g(0))}\,\mathrm{d}x \geq 0 \tag{25}$$

donc la première partie de l'inégalité est démontrée.

On veut donc fixer le valeur de δ pour avoir que la dérivée seconde de $g(x)$ soit positive in $[-\delta, \delta]$. On sait que la fonction $g''(.)$ est continue sur $[a, b]$ et donc notamment en 0 donc si on considère $g''(0)$ on a par définition de la continuité que

$$\forall \eta > 0 \, \exists \delta > 0 : \forall x \, |x| < \delta \implies \left|g''(x) - g''(0)\right| < \eta \tag{26}$$

vu que on peu choisir n'importe quel $\eta > 0$ on fixe $\tilde{\eta} = -g''(0)$ et donc la définition de continuité implique que il existe $\tilde{\delta} > 0$ tel que

$$\forall x \in [-\tilde{\delta}, \tilde{\delta}] \quad g''(x) < 0 \tag{27}$$

comme la fonction est continue et par unicité du maximum, on a:

$$\forall \delta > 0 : \quad \exists \gamma > 0 \forall x : |x| > \delta \implies g(x) < g(0) - \gamma \tag{28}$$

Pour la borne supérieure on peut dire que $\forall \delta < \delta^\star$ on a que

$$\begin{aligned} \int_{a-x^\star}^{b-x^\star} e^{N(g(x)-g(0))} \, \mathrm{d}x &= \int_{a-x^\star}^{-\delta} e^{N(g(x)-g(0))} \, \mathrm{d}x + \int_{-\delta}^{\delta} e^{N(g(x)-g(0))} \, \mathrm{d}x + \int_{\delta}^{b-x^\star} e^{N(g(x)-g(0))} \, \mathrm{d}x \\ &\leq (b-a)e^{-\gamma N} + \int_{-\delta}^{\delta} e^{N(g(x)-g(0))} \, \mathrm{d}x \end{aligned} \tag{29}$$

Donc on a que $\exists \gamma > 0, \forall \delta < \min(\delta^\star, \tilde{\delta})$

$$\int_{a-x^\star}^{b-x^\star} e^{N(g(x)-g(0))} \, \mathrm{d}x \leq (b-a)e^{-\gamma N} + \int_{-\delta}^{\delta} e^{N(g(x)-g(0))} \, \mathrm{d}x \, . \tag{30}$$

Q2. On commence avec l'intervalle $[-\delta, \delta]$ de la question précédente, en se rappelant que cet intervalle a été choisi pour avoir la seconde dérivée positive dessus. Par théorème de Taylor avec reste de Peano, on a:

$$g(x) - g(0) = \frac{1}{2} g''(0) x^2 + h(x) x^2 \quad \text{ou} \quad \lim_{x \to 0} h(x) = 0 \, , \tag{31}$$

où la fonction $h(x)$ est continue. Par continuité, on a

$$\forall \epsilon > 0, \exists \delta > 0, \forall |x| < \delta, \text{we have } |h(x)| < \epsilon \tag{32}$$

et donc

$$\forall x \in [-\delta, \delta] \, : \frac{x^2}{2} g''(0)(1+\epsilon) \leq g(x) - g(0) \leq \frac{x^2}{2} g''(0)(1-\epsilon) \tag{33}$$

Q3. On a

$$\int_{-\delta}^{\delta} e^{-NC\frac{x^2}{2}} \, \mathrm{d}x = \int_{-\infty}^{\infty} e^{-NC\frac{x^2}{2}} \, \mathrm{d}x - \int_{-\infty}^{-\delta} e^{-NC\frac{x^2}{2}} \, \mathrm{d}x - \int_{\delta}^{\infty} e^{-NC\frac{x^2}{2}} \, \mathrm{d}x \tag{34}$$

$$= \sqrt{\frac{2\pi}{NC}} - 2 \int_{\delta}^{\infty} e^{-NC\frac{x^2}{2}} \, \mathrm{d}x \tag{35}$$

et avec l'inégalité de Hoeffding on a

$$\mathbb{P}[X \geq k] \leq e^{-\frac{k^2}{2}} \quad X \sim \mathcal{N}(0,1) \tag{36}$$

et donc

$$0 < \int_{\delta}^{\infty} e^{-NC\frac{x^2}{2}} \, \mathrm{d}x \leq e^{-NC\frac{\delta^2}{2}} \, , \tag{37}$$

et donc
$$\left|\int_\delta^\infty e^{-NC\frac{x^2}{2}}\,\mathrm{d}x\right| < Me^{-N} \tag{38}$$
pour une constante $M > 0$, donc
$$\int_{-\delta}^{\delta} e^{-NC\frac{x^2}{2}}\,\mathrm{d}x = \sqrt{\frac{2\pi}{NC}} + O(e^{-N}) \tag{39}$$

Q4. cela découle immédiatement des 2 questions précédentes.

Q5. Avec les bornes précédentes on a dans la limite $N \to \infty$ et pour tout ϵ suffisament petit,
$$\sqrt{\frac{1}{1+\epsilon}} \leq \lim_{N\to\infty} \frac{\sqrt{-Nf''(x^\star)}}{\sqrt{2\pi}} \frac{\int_a^b e^{Nf(x)}\,\mathrm{d}x}{e^{Nf(x^\star)}} \leq \sqrt{\frac{1}{1-\epsilon}} \tag{40}$$
et comme cette relation est vraie pour tout ϵ, cela implique la limite désirée.

Q6. Supposons que la fonction $f(x)$ possède un nombre fini de maxima locales en des points $x_i^* \in]a,b[$. Près de chaque maximum $x_i^\star$, l'intégrale peut être approchée en développant $f(x)$ au second ordre autour de x_i^* dans chaque sub-interval :
$$f(x) \approx f(x_i^*) + \frac{1}{2}f''(x_i^*)(x-x_i^*)^2.$$
En substituant cette approximation dans l'intégrale, on obtient :
$$I_N \approx \sum_i \int_a^b e^{N\left(f(x_i^*)+\frac{1}{2}f''(x_i^*)(x-x_i^*)^2\right)}\,\mathrm{d}x\,.$$
L'intégrale autour de chaque x_i^* est approximativement gaussienne ce qui donne:
$$I_N \approx \sum_i \sqrt{\frac{2\pi}{N|f''(x_i^*)|}}e^{Nf(x_i^*)}.$$
comme $f(x_i)$ es identique pour tout i, en prenant le logarithme et en considérant la limite $N \to \infty$ seulement le maximum globale domine la somme et on obtient :
$$\lim_{N\to\infty} \frac{1}{N}\log I_N = \sup_x f(x).$$

* Exercise 3 De Laplace à Legendre

Appliquons la méthode de Laplace à des fonctions de type $f(x,\lambda) = -e(x) + \lambda x$. On définit la séquence de fonctions
$$s_N(\lambda) = \frac{1}{N}\log\int_a^b e^{N(-e(x)+\lambda x)}\,\mathrm{d}x \;\text{ et }\; s(\lambda) = \lim_{N\to\infty} s_N(\lambda)$$

Q1. Montrer que $s_N(\lambda)$ est une fonction convexe, en montrant que sa dérivée seconde est toujours positive.

Q2. Montrer en utilisant la méthode de Laplace que $s(\lambda)$ est la transforme de Legendre (ou plus précisément de Legendre-Fenchel [4]) de la fonction $e(x)$. Pourquoi est-elle forcement convexe?
$$s(\lambda) = \sup\,[-e(x) + x\lambda] = -e(x^\star) + x^\star\lambda \tag{41}$$

[4] Il faut rendre justice à Fenchel qui a explicitement généralisé les travaux de Legendre qui étaient restreints aux fonctions convexes. C'est Fenchel qui a étendu ces concepts aux fonctions non différentiables et non convexes.

Q3. Montrer le résultat fondamental des Transformée $s(.)$ de Legendre (Attention! $x^\star$ est une fonction de λ! On devrai plutôt écrire $s(\lambda) = -e(x^\star(\lambda)) + x^\star(\lambda)\lambda$ pour être correct):

$$\frac{d}{d\lambda} s(\lambda) = x^\star \tag{42}$$

Nous allons voir bientôt que le fait que les Transformées de Legendre apparaissent partout en physique statistique (et par conséquent, en thermodynamique) est une conséquence de la formule de Laplace pour les intégrales.

Solution of Exercise 3

Appliquons la méthode de Laplace à des fonctions de type $f(x, \lambda) = -e(x) + \lambda x$. On définit la séquence de fonctions

$$s_N(\lambda) = \frac{1}{N} \log \int_a^b e^{N(-e(x)+\lambda x)}\, \mathrm{d}x \ \text{ et } \mathrm{s}(\lambda) = \lim_{\mathrm{N}\to\infty} \mathrm{s_N}(\lambda)$$

Q1. Maintenant, calculons la deuxième dérivée.

$$\frac{d^2}{d\lambda^2} s_N(\lambda) = \frac{d}{d\lambda}\left(\frac{\int_a^b x \cdot e^{N(-e(x)+\lambda x)}\, \mathrm{d}x}{\int_a^b e^{N(-e(x)+\lambda x)}\, \mathrm{d}x}\right) \tag{43}$$

En utilisant la règle du quotient, nous obtenons :

$$\begin{aligned} \frac{d^2}{d\lambda^2} s_N(\lambda) =& \frac{(\int_a^b e^{N(-e(x)+\lambda x)}\, \mathrm{d}x) \cdot (\int_a^b N x^2 \cdot e^{N(-e(x)+\lambda x)}\, \mathrm{d}x)}{(\int_a^b e^{N(-e(x)+\lambda x)}\, \mathrm{d}x)^2} \\ & - \frac{(\int_a^b x \cdot e^{N(-e(x)+\lambda x)}\, \mathrm{d}x) \cdot (\int_a^b N x \cdot e^{N(-e(x)+\lambda x)}\, \mathrm{d}x)}{(\int_a^b e^{N(-e(x)+\lambda x)}\, \mathrm{d}x)^2} \end{aligned} \tag{44}$$

Pour montrer que cette deuxième dérivée est toujours positive, nous pouvons réécrire l'expression sous une autre forme. Définissons une mesure de probabilité $p(x)$ sur $[a, b]$:

$$p(x) = \frac{e^{N(-e(x)+\lambda x)}}{\int_a^b e^{N(-e(x)+\lambda x)}, \mathrm{d}x} \tag{45}$$

(il est clair que $p > 0$ et $\int p = 1$). La dérivée seconde peut se réécrire comme:

$$\frac{d^2}{d\lambda^2} s_N(\lambda) = N\big(\mathbb{E}[X^2] - (\mathbb{E}[X])^2\big) = N \cdot \mathbb{V}\mathrm{ar}(X) > 0 \tag{46}$$

où $\mathbb{E}$ (resp. $\mathbb{V}$ar) représente l'espérance par rapport à la mesure $p(.)$ (resp. la variance de $X \sim p$). La variance d'une variable aléatoire étant toujours non-négative (inégalité de Jensen), cela conclut la preuve.

Q2. La solution est donne directement par l'application du dernière exercice.

$$s(\lambda) = \sup\,[-e(x) + x\lambda] = -e(x^\star) + x^\star\lambda \tag{47}$$

ou $x^\star = x^\star(\lambda)$ est le point de minumum de $e(x) - x\lambda$ pour toutes valuers de λ.
$s(\lambda)$ est forcément convexe car elle est obtenue comme un supremum de fonctions affines en λ. La propriété que la transformée de Legendre-Fenchel d'une fonction est toujours convexe découle du fait que le supremum d'une famille de fonctions convexes ou affines est toujours convexe.

Q3. On peut directement dire que

$$\frac{d}{d\lambda}s(\lambda) = -e'(x^\star(\lambda))(x^\star)'(\lambda) + (x^\star)'(\lambda)\lambda + x^\star(\lambda) = \left[-e'(x^\star(\lambda)) + \lambda\right](x^\star)'(\lambda) + x^\star(\lambda) = x^\star(\lambda) \tag{48}$$

parce que $x^\star(\lambda)$ est le maximum du $-e(x) + x\lambda$ et donc il est le point qui annule la derivee.

* Exercise 4 Tranformée et double-tranformée de Legendre-Fenchel

On définit la *transformée de Legendre-Fenchel* de la fonction $e(x)$ par

$$s(\lambda) = \overline{e}(\lambda) = \sup_x\left[-e(x) + x\lambda\right] = -e(x^\star) + x^\star\lambda \tag{49}$$

On peut aussi calculer la transformée de la transformée :

$$\overline{\overline{e}}(x) = \overline{s}(\lambda) = \sup_\lambda\left[-s(\lambda) + x\lambda\right] . \tag{50}$$

On va se poser la question de la relation entre la double Transformée $\overline{\overline{e}}(.)$ et la fonction originale $e(.)$.

Q1. Est-il est possible que $\overline{\overline{e}}(x) = e(x)$ pour toute fonction $e(.)$? Donner un contre-exemple évident en utilisant la question $3-1$.

Q2. Montrer que si $e(.)$ est strictement convexe, alors il ne peut y a qu'un seul maximum, et que la relation entre λ et x^* est unique et inversible [5].

Q3. Montrer que pour toute fonction convexe $e(.)$ (et seulement dans ce cas) on a bien $\overline{\overline{e}}(x) = e(x)$.

Q4. Vérifier numériquement que si $e(.)$ n'est pas convexe, la double Transformée $\overline{\overline{e}}(.)$ donne l'enveloppe convexe de la fonction $e(x)$.

Solution of Exercise 4

Q1. On a montré que la transformée de Legendre-Fenchel est toujours une fonction convexe, quelle que soit la fonction d'origine. Donc, si $e(x)$ n'est pas convexe, $\overline{\overline{e}}(x)$ est convexe et, par conséquent, les deux fonctions ne peuvent pas être égales.

Q2. On a que le maximum doit satisfaire

$$\partial_x[-e(x) + x\lambda] = 0 \implies e'(x) = \lambda \tag{51}$$

De plus, si la fonction $e(x)$ est strictement convexe, alors $e'(x)$ est strictement croissante et donc l'équation précédente admet une seule solution.

Q3. On peut commencer par prendre la dérivée de la définition :

$$\partial_\lambda[-s(\lambda) + x\lambda] \implies x = x^\star(\lambda), \tag{52}$$

où le $\arg\sup$ en λ est noté $\lambda^\star$. C'est la valeur pour laquelle on a $x = x^\star(\lambda^\star)$. De plus, cette valeur $\lambda^\star$ existe toujours car $s(\lambda)$ est convexe, et donc $x^\star(\lambda)$ est une fonction croissante de λ. En substituant cela dans la définition :

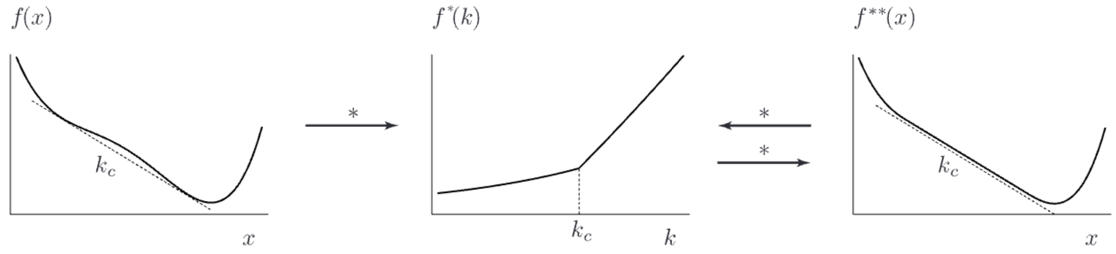
$$\overline{\overline{e}}(x) = -s(\lambda^\star) + x\lambda^\star = e(x^\star(\lambda^\star)) - x^\star(\lambda^\star)\lambda^\star + x\lambda^\star = e(x), \tag{53}$$

[5] On rappelle le théorème suivant qui explicite la dérivée de la réciproque d'une fonction bijective et dérivable en fonction de la dérivée:

$$\left[f^{-1}\right]'(a) = \frac{1}{f'\left(f^{-1}(a)\right)}$$

.

Q4. Dan le case ou $e(.)$ n'est pas convexe on a la chose suivante



out la flèche indique une application de la transforme de Legendre-Fenchel.