

Chapitre E-VIII

Thermodynamique statistique.

Joël SORNETTE met ce cours à votre disposition selon les termes de la licence Creative Commons :

- Pas d'utilisation commerciale.
- Pas de modification, pas de coupure, pas d'intégration à un autre travail.
- Pas de communication à autrui sans citer son nom, ni en suggérant son autorisation.

Retrouvez l'intégralité du cours sur le site joelsornette.fr

RÉSUMÉ :

Ce chapitre commence par exposer quelques rudiments de statistiques ; on a cherché un exposé bref mais convainquant. L'objectif est de montrer que, pour un système de la taille d'une mole, la moyenne statistique devient une certitude physique.

Le modèle micro-canonique étudie un système isolé, qui peut se réaliser sous un grand nombre de micro-états. L'hypothèse micro-canonique les considère équiprobables à l'équilibre. On donne une définition statistique de l'entropie, générale puis adaptée à l'état d'équilibre. On définit température et pression statistique à partir de l'équilibre entre parties d'un système isolé. On s'intéresse enfin aux fluctuations d'une variable interne, en particulier de l'énergie d'une partie du système.

On s'intéresse ensuite à la délicate notion d'irréversibilité, que l'on éclaire par le modèle d'Ehrenfest. Le principe du bilan détaillé relatif aux probabilités de transition entre micro-états permet de montrer que l'entropie ne peut qu'augmenter lorsque l'on part d'un état hors d'équilibre.

Le modèle canonique étudie un système de volume constant qui échange de l'énergie avec un thermostat. La probabilité d'un micro-état vérifie la statistique de Maxwell-Boltzmann. On définit la fonction de partition et l'on met en place le formalisme qui permet le calcul des valeurs moyennes des fonctions d'état, voire de leurs fluctuations. On montre la propriété de factorisation de la fonction de partition qui permet de séparer les contributions énergétiques des différents termes.

On met en œuvre à titre d'exemple ce formalisme dans un cas discret (le paramagnétisme de Langevin) et dans un cas continu (le gaz parfait monoatomique), ce dernier exemple servant à montrer l'identité des résultats entre thermodynamique classique et statistique.

Le paradoxe de Gibbs permet d'introduire la notion de particules indiscernables et d'annoncer la thermodynamique statistique quantique.

Enfin, l'on adapte au contexte canonique le principe du bilan détaillé et l'on montre que dans ce contexte, une évolution spontanée se traduit par une diminution de l'énergie libre.

Table des matières

E-VIII Thermodynamique statistique.	1
1 Théorie des grands nombres.	4
1.a Moyenne, variance, écart-type.	4
1.b Inégalité de Bienaymé-Tchebichev.	6
1.c Loi des grands nombres.	7
2 Modèle micro-canonique.	11
2.a Postulat fondamental.	11
2.b Température et pression thermodynamique.	13
2.c Fluctuations d'une variable interne.	15
2.d Irréversibilité.	17
2.e Evolution d'un système dans le modèle micro-canonique	18
3 Modèle canonique.	21
3.a Loi de Maxwell-Boltzmann.	21
3.b Moyennes statistiques des fonctions d'état.	23
3.c Factorisation de la fonction de partition.	26
3.d Un exemple discret : le paramagnétisme de Langevin.	28
3.e Un exemple continu : le gaz parfait monoatomique.	29
3.f Paradoxe de Gibbs.	33
3.g Evolution d'un système dans le modèle canonique	36

1 Théorie des grands nombres.

Il ne s'agit pas ici de faire un cours mathématique de probabilités mais de donner des notions raisonnablement solides pour comprendre l'essence de la thermodynamique statistique. Les résultats ne seront donc démontrés que dans les cas les plus simples.

1.a Moyenne, variance, écart-type.

• Variables aléatoires discrètes.

Une variable aléatoire discrète est une grandeur X réelle ou complexe, scalaire ou vectorielle, pouvant prendre un nombre fini ou dénombrable de valeurs notés x_i distinctes avec $i \in [1, N]$ dans le cas d'un nombre fini de valeurs ou $i \in \mathbb{N}$ dans le cas d'une infinité dénombrable de valeurs. Une loi de probabilité est la donnée de grandeurs réelles positives p_i , appelées probabilités que X ait la valeur x_i , ce que l'on traduit par l'écriture $p_i = \mathcal{P}(X = x_i)$, et telles que $\sum_i p_i = 1$.

Par définition, la probabilité que X ait une valeur contenue dans un sous ensemble \mathcal{A} de valeurs de x_i est

$$\mathcal{P}(X \in \mathcal{A}) = \sum_{i|x_i \in \mathcal{A}} p_i$$

• Variables aléatoires continues.

Une variable aléatoire continue est une grandeur X réelle, scalaire ou vectorielle, pouvant prendre un nombre continu de valeurs de valeurs notées x appartenant à \mathbb{R} ou \mathbb{R}^n (on prendra \mathbb{R}^2 pour les exemples).

La probabilité que X scalaire ait une valeur comprise entre x et $x+dx$ (dans un passage à la limite où dx tend vers 0) est notée $p(x) dx$ où $p(x)$ à valeurs réelles positives vérifie $\int_{\mathbb{R}} p(x) dx = 1$. La fonction $p(x)$ s'appelle *densité de probabilité*.

La probabilité que X vectorielle ait une abscisse comprise entre x et $x + dx$ et une ordonnée comprise entre y et $y + dy$ est notée $p(x, y) dx dy$ où $p(x, y)$ à valeurs réelles positives vérifie $\iint_{\mathbb{R}^2} p(x, y) dx dy = 1$.

• Moyenne.

Une fonction f d'une variable aléatoire est appelée fonction aléatoire. Dans le cas discret, sa moyenne, notée $\langle f \rangle$ est définie par

$$\langle f \rangle = \sum_i p_i f(x_i)$$

et dans le cas continu (à une dimension par exemple) par

$$\langle f \rangle = \int_{\mathbb{R}} p(x) f(x) dx$$

Désormais, on se placera, sauf exception, dans le cas discret ; pour le cas continu, il suffit systématiquement de remplacer la somme par une intégrale.

Remarque 1 : La fonction identité permet de définir la moyenne de la variable aléatoire ; par exemple, dans le cas discret $\langle x \rangle = \sum_i p_i x_i$.

Remarque 2 : si f est une fonction constante de valeur f_0 , sa moyenne est f_0 , en effet :

$$\langle f \rangle = \sum_i p_i f(x_i) = \sum_i p_i f_0 = f_0 \sum_i p_i = f_0 \times 1 = f_0$$

Remarque 3 : la moyenne est une opération linéaire, en effet si λ et μ sont des constantes, on a dans le cas d'un nombre fini de valeurs, grâce aux propriétés de commutativité et de distributivité d'une somme :

$$\langle \lambda f + \mu g \rangle = \sum_i p_i [\lambda f(x_i) + \mu g(x_i)] = \lambda \sum_i p_i f(x_i) + \mu \sum_i p_i g(x_i) = \lambda \langle f \rangle + \mu \langle g \rangle$$

et dans le cas dénombrable, il s'agit d'une propriété élémentaire des séries convergentes.

• Variance et écart-type.

L'idée est de définir une moyenne de l'écart entre les valeurs d'une fonction aléatoire f et sa moyenne ; plus cet écart moyen est petit plus la distribution des valeurs est resserrée autour de la moyenne. L'idée de calculer la moyenne de $f - \langle f \rangle$ s'avère plutôt stupide car, grâce aux remarques précédentes ($\langle f \rangle$ se comporte comme une fonction constante) :

$$\langle f - \langle f \rangle \rangle = \langle f \rangle - \langle \langle f \rangle \rangle = \langle f \rangle - \langle f \rangle = 0$$

simplement parce que les valeurs inférieures à $\langle f \rangle$ compensent celles qui sont supérieures.

L'idée plus raisonnable, car on se sert d'une distance au sens topologique, est de calculer la moyenne de $|f - \langle f \rangle|$ mais les valeurs absolues ne se prêtent pas à des calculs simples.

La bonne idée est de passer au carré¹ de l'écart $(f - \langle f \rangle)^2$ puis de prendre la racine carrée de la moyenne pour obtenir un résultat homogène à $\langle f \rangle$ (c'est-à-dire de même unité). On définit donc la *variance*, notée σ^2 et l'*écart-type*, noté σ par :

$$\sigma^2 = \langle (f - \langle f \rangle)^2 \rangle \quad \text{et} \quad \sigma = \sqrt{\sigma^2}$$

1. carré sur \mathbb{R} et carré scalaire sur \mathbb{R}^2 voire carré hermitien sur \mathbb{C}^2 .

Remarque : Si l'on développe la définition de la variance, on obtient une expression plus aisée à calculer. En effet, grâce à la linéarité et le fait que $\langle f \rangle$ est une constante vis-à-vis des calculs de moyenne, on a :

$$\begin{aligned}\sigma^2 &= \langle f^2 - 2 \langle f \rangle f + \langle f \rangle^2 \rangle = \langle f^2 \rangle - 2 \langle f \rangle \langle f \rangle + \langle \langle f \rangle^2 \rangle \\ \sigma^2 &= \langle f^2 \rangle - 2 \langle f \rangle^2 + \langle f \rangle^2 = \langle f^2 \rangle - \langle f \rangle^2\end{aligned}$$

Donc en pratique, on calcule toujours (ou presque) la variance par $\sigma^2 = \langle f^2 \rangle - \langle f \rangle^2$.

1.b Inégalité de Bienaymé-Tchebichev.

En gros, l'inégalité de BIENAYMÉ-TCHEBICHEV² donne un majorant de la probabilité que la valeur d'une fonction aléatoire s'éloigne à partir sa moyenne de plus d'une quantité donnée, l'écart-type étant pris comme unité d'éloignement.

• Inégalité de Markov.

L'inégalité de MARKOV nous ici sert de lemme. Soit f une fonction aléatoire à *valeurs positives ou nulles* ; on cherche la probabilité pour qu'elle soit supérieure à un certain nombre de fois sa moyenne, disons k fois. On cherche donc $P_k = \mathcal{P}(\{x_i | f(x_i) > k \langle f \rangle\})$ dont la valeur est, par définition (voir au tout début du paragraphe 1.a p. 4) :

$$P_k = \sum_{i|f(x_i) > k \langle f \rangle} p_i$$

Partons de la définition de $\langle f \rangle$:

$$\langle f \rangle = \sum_i p_i f(x_i)$$

C'est une somme de termes tous positifs (f est à valeurs positives ou nulles, par hypothèse), si l'on enlève quelques termes, on obtient quelque chose de plus petit ou à la rigueur égal ; ne conservons donc que les termes pour lesquels $f(x_i) > k \langle f \rangle$, on a alors :

$$\langle f \rangle = \sum_i p_i f(x_i) \geq \sum_{i|f(x_i) > k \langle f \rangle} p_i f(x_i)$$

Pour les termes conservés, les $f(x_i)$ sont minorés par $k \langle f \rangle$ et leurs coefficients sont des p_i positifs par essence, donc :

$$\langle f \rangle = \sum_i p_i f(x_i) \geq \sum_{i|f(x_i) > k \langle f \rangle} p_i f(x_i) > \sum_{i|f(x_i) > k \langle f \rangle} p_i k \langle f \rangle$$

2. comprendre de BIENAYMÉ et de TCHEBYCHEV. On trouve aussi les graphies Tschebyscheff, Chebyshev ou Chebyshev selon le système de translittération du cyrillique.

On factorise la constante (linéarité de la moyenne) et l'on reconnaît la définition de la probabilité recherchée :

$$\langle f \rangle = \sum_i p_i f(x_i) \geq \sum_{i|f(x_i) > k \langle f \rangle} p_i f(x_i) > \sum_{i|f(x_i) > k \langle f \rangle} p_i k \langle f \rangle = k \langle f \rangle \sum_{i|f(x_i) > k \langle f \rangle} p_i = k \langle f \rangle P_k$$

Après simplification par $\langle f \rangle$, on en déduit que $P_k < \frac{1}{k}$.

• **Application à la variance.**

Pour une fonction f à valeurs négatives, positives ou nulles, appliquons ce résultat à la fonction positive $(f - \langle f \rangle)^2$ de moyenne σ^2 et pour une constante notée k^2 . La probabilité pour que $(f - \langle f \rangle)^2 > k^2 \sigma^2$ est inférieure à $\frac{1}{k^2}$, ce que l'on peut exprimer en disant que c'est la probabilité pour que $|f - \langle f \rangle| > k \sigma$, c'est-à-dire pour que l'écart à la moyenne soit supérieur à k fois l'écart-type, où encore pour que $f(x)$ soit en dehors de l'intervalle de confiance $[\langle f \rangle - k \sigma, \langle f \rangle + k \sigma]$.

D'où l'inégalité de BIENAYMÉ-TCHEBICHEV :

$$\mathcal{P}(f(x_i) \notin [\langle f \rangle - k \sigma, \langle f \rangle + k \sigma]) < \frac{1}{k^2}$$

Par exemple la probabilité que $f(x)$ s'éloigne de la moyenne de plus de dix fois l'écart-type est inférieure à 1%, pour plus de 32 fois l'écart-type elle tombe à moins de 1‰.

Si le lecteur a quelques connaissances en probabilité et statistique, ce résultat peut l'étonner : pour des lois de probabilité classiques (loi normale par exemple), on a des majorations bien plus efficaces. L'inégalité de BIENAYMÉ-TCHEBICHEV donne en effet des majorations très grossières; l'avantage est qu'elle ne nécessite aucune hypothèse sur la loi de probabilité qui peut être exotique. La grossièreté de la majoration ne posera pas de problème en thermodynamique où le nombre de particules étudiées est pharaonique; le paragraphe qui suit nous le prouvera.

1.c Loi des grands nombres.

• **Somme et variance de fonctions aléatoires indépendantes.**

Soient deux variables aléatoires, X prenant les valeurs x_i réelles (pour alléger l'exposé) avec les probabilités p_i et Y prenant les valeurs y_j réelles (même remarque) avec les probabilités q_j et deux fonctions aléatoires, l'une f fonction de X et l'autre g fonction de Y , toutes deux à valeurs réelles (même remarque).

Construisons la fonction h de \mathbb{R}^2 sur \mathbb{R} définie par $h(X, Y) = f(X) + g(Y)$ prenant les valeurs $f(x_i) + g(y_j)$. Pour en calculer moyenne et écart-type, il faut connaître la probabilité

pour que $(X, Y) = (x_i, y_j)$, c'est-à-dire $X = x_i$ et en même temps $Y = y_j$. Rien ne permet *a priori* de le savoir.

Par définition deux variables aléatoires sont indépendantes si

$$\mathcal{P}(X = x_i \text{ et } Y = y_j) = \mathcal{P}(X = x_i) \times \mathcal{P}(Y = y_j) = p_i q_j$$

Les résultats expérimentaux en physique sont très souvent conformes à ce modèle et correspondent alors à des phénomènes X et Y sans interaction.

Dans ce cas, on a pour la moyenne (on raisonne dans le cas fini pour alléger) et en détaillant les factorisations successives qui sont plus subtiles qu'il n'y paraît :

$$\begin{aligned} \langle h \rangle &= \sum_{(i,j)} [f(x_i) + g(y_j)] p_i q_j = \sum_{(i,j)} f(x_i) p_i q_j + \sum_{(i,j)} g(y_j) p_i q_j \\ \langle h \rangle &= \sum_j \left(\sum_i f(x_i) p_i q_j \right) + \sum_i \left(\sum_j g(y_j) p_i q_j \right) \\ \langle h \rangle &= \sum_j \left(\sum_i f(x_i) p_i \right) q_j + \sum_i \left(\sum_j g(y_j) q_j \right) p_i \\ \langle h \rangle &= \left(\sum_i f(x_i) p_i \right) \left(\sum_j q_j \right) + \left(\sum_j g(y_j) q_j \right) \left(\sum_i p_i \right) \\ \langle h \rangle &= \langle f \rangle \times 1 + \langle g \rangle \times 1 = \langle f \rangle + \langle g \rangle \end{aligned}$$

La moyenne d'une somme de deux fonctions aléatoires indépendantes est la somme de leurs moyennes.

De la même façon, on peut calculer la moyenne de $h^2 = f^2 + 2fg + g^2$; on laisse le soin³ au lecteur de détailler le calcul sur le modèle du précédent et l'on arrive à :

$$\langle h^2 \rangle = \langle f^2 \rangle + 2 \langle f \rangle \langle g \rangle + \langle g^2 \rangle$$

et la variance de h notée ici σ_h^2 pour éviter les ambiguïtés est

$$\begin{aligned} \sigma_h^2 &= \langle h^2 \rangle - \langle h \rangle^2 = (\langle f^2 \rangle + 2 \langle f \rangle \langle g \rangle + \langle g^2 \rangle) - (\langle f \rangle + \langle g \rangle)^2 \\ \sigma_h^2 &= \langle f^2 \rangle - \langle f \rangle^2 + \langle g^2 \rangle - \langle g \rangle^2 = \sigma_f^2 + \sigma_g^2 \end{aligned}$$

La variance d'une somme de deux fonctions aléatoires indépendantes est la somme de leurs variances.

Ces résultats se généralisent par récurrence à une somme d'un nombre quelconques de variables indépendantes.

3. pour les raisons habituelles : c'est la seule façon que j'aie trouvée pour qu'il ne survole pas les passages délicats.

- **Somme et variance de fonctions aléatoires indépendantes de même loi.**

Les deux variables X et Y du paragraphe précédent peuvent prendre les mêmes valeurs avec les mêmes probabilités. Ça ne veut pas dire pour autant qu'elles soient égales⁴, il est possible d'avoir $(X, Y) = (x_1, x_2)$ avec la probabilité $p_1 p_2$, il est aussi possible d'avoir $(X, Y) = (x_1, x_1)$ avec la probabilité p_1^2 . Dans ce cas si on utilise la même fonction f , les moyennes et variances de $f(X)$ et $f(Y)$ sont bien évidemment égales.

Si l'on considère N variables aléatoires X_n indépendantes et de même loi et une fonction f si l'on étudie la fonction aléatoire définie par

$$F(X_1, X_2, \dots, X_N) = f(X_1) + f(X_2) + \dots + f(X_N)$$

en vertu de ce qui précède la moyenne, la variance et l'écart-type de cette fonction sont respectivement :

$$\langle F \rangle = N \langle f \rangle$$

$$\sigma_F^2 = N \sigma_f^2$$

$$\sigma_F = \sqrt{N} \sigma_f$$

ce qui montre que l'écart-type croît beaucoup moins vite que la moyenne (sauf si $\langle f \rangle$ est nul). Si $\langle f \rangle \neq 0$, il est pertinent de définir l'écart-type relatif comme le rapport de l'écart-type à la moyenne. Dans ce cas, on a :

$$\frac{\sigma_F}{\langle F \rangle} = \frac{1}{\sqrt{N}} \frac{\sigma_f}{\langle f \rangle}$$

ce qui met en évidence une décroissance de l'écart-type relatif en \sqrt{N} , l'étalement relatif des valeurs autour de la moyenne se resserre avec le nombre de variables. C'est la loi des grands nombres, dans son interprétation pour la thermodynamique.

- **Un exemple en guise de conclusion.**

Considérons un atome unique dans un récipient symétrique ; par symétrie la probabilité qu'il soit dans l'une ou l'autre de deux moitiés symétriques (disons gauche et droite) est la même ; la somme des probabilités étant l'unité, chacune vaut $1/2$. Le nombre, noté f , de particules à gauche est donc 1 avec une probabilité $1/2$ et 0 avec une probabilité $1/2$. On a :

$$\langle f \rangle = \frac{1}{2} 1 + \frac{1}{2} 0 = \frac{1}{2}$$

$$\langle f^2 \rangle = \frac{1}{2} 1^2 + \frac{1}{2} 0^2 = \frac{1}{2}$$

4. Si elles l'étaient, la probabilité d'avoir $(X, Y) = (x_1, x_2)$ avec $x_1 \neq x_2$ serait nulle et celle d'avoir $(X, Y) = (x_1, x_1)$ serait p_1 et non p_1^2 .

$$\sigma_f^2 = \langle f^2 \rangle - \langle f \rangle^2 = \frac{1}{4}$$

$$\sigma_f = \sqrt{\frac{1}{4}} = \frac{1}{2}$$

$$\frac{\sigma_f}{\langle f \rangle} = 1$$

Mettons un soixantième de mole (disons $N = 10^{22}$ atomes) sous forme gazeuse dans le même récipient (à température et pression ordinaires, il contiendra un tiers de litre). Dans un gaz parfait les différents atomes sont sans interaction, on peut les considérer comme indépendants au sens statistique. Le nombre F d'atomes dans la partie gauche, en vertu de ce qui précède, sera tel que :

$$\langle F \rangle = N \langle f \rangle = 0,5 \cdot 10^{22}$$

$$\sigma_F^2 = N \sigma_f^2 = 0,25 \cdot 10^{22}$$

$$\sigma_F = \sqrt{N} \sigma_f = 0,5 \cdot 10^{11}$$

$$\frac{\sigma_F}{\langle F \rangle} = \frac{1}{\sqrt{N}} \frac{\sigma_f}{\langle f \rangle} = 10^{-11}$$

Appliquons l'inégalité de BIENAYMÉ-TCHEBICHEV avec $k = 1000$. Il y a moins d'une chance sur un million pour que la valeur de F soit en dehors de l'intervalle dont les bornes sont $\langle F \rangle \pm k \sigma_F = \langle F \rangle \left(1 \pm k \frac{\sigma_F}{\langle F \rangle}\right) = 0,5 \cdot 10^{22} (1 - 10^{-8})$ c'est-à-dire en dehors de l'intervalle noté $[0,5 \cdot 10^{22} (1 - 10^{-8}) \text{ et } 0,5 \cdot 10^{22} \cdot (1 + 10^{-8})]$. F est connu à dix milliardièmes près à mieux que 99,9999 chances sur 100.

Si vous n'êtes toujours pas convaincu que la thermodynamique se prête à une approche statistique, ne lisez pas la suite! Si oui, mon objectif est atteint : l'avoir fait avec des démonstrations accessibles.

Tout ce qui suit résulte des idées de Ludwig BOLTZMANN (physicien autrichien né à Vienne en 1844 et mort près de Trieste en 1906) qui a été le pionnier de cette approche statistique développée pour l'essentiel entre 1866 et 1884. Quand on se souvient que l'essentiel de la thermodynamique classique a été formalisé entre 1824 et 1865, on voit que cette science a progressé à une vitesse record. De son vivant, BOLTZMANN a pourtant été vivement critiqué, surtout pour sa conviction envers l'hypothèse atomique encore minoritaire à l'époque et il n'est pas impossible que ce soit la cause de son suicide.

2 Modèle micro-canonique.

2.a Postulat fondamental.

- **Etats macroscopiques et micro-états.**

Considérons un système *isolé*, ne recevant ni travail (parce que son volume V constant) ni chaleur (parce que ses parois sont isolantes), son énergie interne U est donc constante. Il est caractérisé par la donnée de V et de U qui définissent un état au sens macroscopique.

Supposons que chacun de ses éléments constitutifs ait une énergie quantifiée, l'élément repéré par l'indice n ayant une énergie e_n appartenant à un ensemble de valeurs fini ou dénombrable de valeurs notées $e_{n,i}$. On doit avoir $\sum_n e_{n,i} = U$; il y a un grand nombre fini⁵ de combinaisons de $e_{n,i}$ vérifiant cette contrainte, chacune est appelée micro-état accessible et leur nombre est traditionnellement noté $W(U, V)$ (notation de BOLTZMANN) ou $\Omega(U, V)$ (pour éviter la confusion avec un travail, lui aussi noté W). La réalité est bien évidemment plus compliquée mais l'idée-force est qu'il y a un nombre fini $\Omega(U, V)$ d'états quantiques vérifiant les contraintes que sont la donnée de l'énergie interne U et celle du volume V d'état macroscopique.

Même s'il n'y a pas de quantification apparente comme dans le cas de l'énergie cinétique d'un atome d'un gaz, les lois d'incertitude de HEISENBERG en introduisent une de fait. Nous aurons l'occasion d'y revenir plus en détail au paragraphe 3.e p. 29.

- **Entropie statistique.**

L'agitation thermique est telle que pour un état macroscopique donné, le système passe constamment d'un micro-état à un autre et chacun d'eux est observé avec une certaine probabilité⁶. Nous noterons p_ℓ la probabilité d'un micro-état indicé par la lettre ℓ .

On définit l'entropie statistique S de l'état macroscopique par la moyenne des $\ln(p_\ell)$ multipliée par une constante négative (on a car $p_\ell < 1$ parce que c'est une probabilité donc $\ln(p_\ell) < 0$) notée $-k_B$ ou $-k$, appelée constante de BOLTZMANN et dont la valeur dépend du choix du système d'unités. On a donc par définition :

$$S = -k \langle \ln(p_\ell) \rangle = -k \sum_{\ell} p_\ell \ln(p_\ell)$$

Bien sûr cette définition n'est pas le résultat d'une révélation mais le terme d'un long

5. compte tenu des règles de quantification usuelles.

6. L'identité entre la probabilité à un instant donné pour N systèmes identiques et la probabilité pour un seul système passant par N états successifs ne coule pas de source, c'est l'*hypothèse ergodique* dont la théorie mathématique est beaucoup trop délicate pour être exposée ici. Il s'agit, du reste, de mathématiques pures.

cheminement. L'identification avec l'entropie classique apparaîtra plus tard dans ce chapitre.

• **Hypothèse micro-canonique.**

Les probabilités p_ℓ ne sont bien évidemment pas accessibles et l'on procède, comme toujours en physique, en cherchant un principe (qui joue le même rôle en physique dans sa phase déductive que l'axiome en mathématique) qui rende compte de la réalité expérimentale préexistante. L'hypothèse micro-canonique est en ce sens un diamant de la plus belle eau. D'une simplicité extrême, elle permet de retrouver toutes les lois de la thermodynamique classique. La voici :

Si un système isolé est à l'équilibre, tous les micro-états accessibles sont équiprobables.

Si $\Omega(U, V)$ (Ω pour alléger, s'il n'y a pas d'ambiguïté) est le nombre de micro-états accessibles, la valeur commune p de tous les p_ℓ vérifie, en se souvenant que la somme de toutes les probabilités est égale à l'unité :

$$1 = \sum_1^\Omega p_\ell = \sum_1^\Omega p = p \sum_1^\Omega 1 = p \Omega$$

d'où : $\forall \ell \quad p_\ell = \frac{1}{\Omega}$.

• **Entropie statistique d'un système isolé à l'équilibre.**

Reportons $p_\ell = \frac{1}{\Omega}$ dans $S = -k \sum_\ell p_\ell \ln(p_\ell)$, on en déduit une expression qui est somme de Ω termes identiques et égale au produit du nombre de termes, c'est-à-dire Ω , par la valeur commune de ces termes, soit :

$$S = -k \sum_1^\Omega p_\ell \ln(p_\ell) = -k \sum_1^\Omega \frac{1}{\Omega} \ln\left(\frac{1}{\Omega}\right) = k \sum_1^\Omega \frac{1}{\Omega} \ln \Omega = k \Omega \left(\frac{1}{\Omega} \ln \Omega\right) = k \ln \Omega$$

Retenons :

Si un système isolé est à l'équilibre, son entropie est $S = k \ln \Omega$.

Remarque : Cette dernière définition de l'entropie, valable seulement à l'équilibre, est apparue la première (avec la notation $S = k \ln W$ gravée sur la tombe de BOLTZMANN,



FIGURE 1 – Tombeau de Boltzmann.

voir figure 1 p. 13); par la suite, on a donné la définition formulée précédemment, plus générale, valable aussi hors d'équilibre.

Remarque : c'était la première fois que l'on donnait une *vraie* définition à l'entropie, entité mystérieuse en thermodynamique classique.

2.b Température et pression thermodynamique.

Supposons que le système d'énergie U et de volume V soit composé de deux parties faiblement couplées⁷ l'une d'énergie u_1 et de volume v_1 et l'autre d'énergie $u_2 = U - u_1$ et de volume $v_2 = V - v_1$. Le système a beau être isolé, ses parties ne le sont pas et peuvent échanger de l'énergie et de la place (du volume); u_1 et v_1 peuvent donc fluctuer librement. Nous allons calculer la probabilité pour que u_1 et v_1 aient chacun une valeur fixée, respectivement u et v .

Le nombre de micro-états accessibles de la première partie est notée $\omega_1(u, v)$ et celui de la seconde $\omega_2(U - u, V - v)$. Pour le système entier, un micro-état accessible correspondant à cette partition de l'énergie et du volume, et formé d'un micro-état quelconque parmi les ω_1 de la première partie et d'un micro-état quelconque parmi les ω_2 de la seconde partie. Pour le système entier, le nombre de micro-états accessibles correspondant à cette partition de l'énergie et du volume est donc le produit des deux soit, en inventant la notation :

$$\Omega(\{u, U - u\}, \{v, V - v\}) = \omega_1(u, v) \omega_2(U - u, V - v)$$

7. On rappelle que l'énergie n'est pas a priori une grandeur additive (voir le chapitre E-III sur le premier principe) car l'énergie du tout est la somme des énergies des parties *et de l'énergie d'interaction entre les parties*. Toutefois, si les interactions sont dues à des forces à court rayon d'action (forces de VAN DER WAALS par exemple) et n'intéressent donc qu'une mince pellicule de part et d'autre de la surface de séparation entre les parties, alors l'énergie d'interaction est négligeable.

L'hypothèse d'équiprobabilité, donne à chacun de ces micro-états la même probabilité $\frac{1}{\Omega(U,V)}$ et donc la probabilité de la partition étudiée est

$$p(\{u, U - u\}, \{v, V - v\}) = \frac{\omega_1(u, v) \omega_2(U - u, V - v)}{\Omega(U, V)}$$

La loi des grands nombre explique qu'en pratique, on observe la partition la plus probable, soit, à U et V fixés, celle pour laquelle le produit $\omega_1 \omega_2$ est le plus grand, ou encore le logarithme de ce produit le plus grand. En multipliant par la constante de BOLTZMANN, on fait apparaître des grandeurs que l'on interprète comme l'entropie statistique des deux parties. On cherche donc le maximum de :

$$k \ln(\omega_1 \omega_2) = k \ln \omega_1 + k \ln \omega_2 = S_1(u, v) + S_2(U - u, V - v)$$

Le maximum correspond à la nullité de la différentielle⁸ soit en notant $\frac{\partial S}{\partial U}$ et $\frac{\partial S}{\partial V}$ les dérivées partielles⁹ et sans oublier que U et V sont constants :

$$\begin{aligned} \frac{\partial S_1}{\partial U} du + \frac{\partial S_1}{\partial V} dv + \frac{\partial S_2}{\partial U} d(U - u) + \frac{\partial S_2}{\partial V} d(V - v) &= 0 \\ \frac{\partial S_1}{\partial U} du + \frac{\partial S_1}{\partial V} dv - \frac{\partial S_2}{\partial U} du - \frac{\partial S_2}{\partial V} dv &= 0 \\ \left(\frac{\partial S_1}{\partial U} - \frac{\partial S_2}{\partial U} \right) du + \left(\frac{\partial S_1}{\partial V} - \frac{\partial S_2}{\partial V} \right) dv &= 0 \end{aligned}$$

et comme la nullité d'une différentielle signifie la nullité de ses coefficients :

$$\frac{\partial S_1}{\partial U} = \frac{\partial S_2}{\partial U} \quad \text{et} \quad \frac{\partial S_1}{\partial V} = \frac{\partial S_2}{\partial V}$$

La partition la plus probable, donc celle effectivement observée, correspond à l'égalité des dérivées partielles de l'entropie. Or en thermodynamique classique :

$$dU = \delta W_{\text{rév.}} + \delta Q_{\text{rév.}} = -p dV + T dS$$

d'où

$$dS = \frac{1}{T} dU + \frac{p}{T} dV$$

On définit donc par analogie la température micro-canonique T puis la pression micro-canonique p à partir de l'entropie statistique, par :

$$\frac{1}{T} = \frac{\partial S}{\partial U} \quad \text{et} \quad \frac{p}{T} = \frac{\partial S}{\partial V}$$

8. ce qui équivaut, de façon plus lisible, à la nullité des deux dérivées partielles par rapport à u et v .

9. U et V désignant formellement les deux arguments des fonctions entropies, indépendamment de la valeur effective qu'ils prennent u ou $U - u$, v ou $V - v$.

Les relations $\frac{\partial S_1}{\partial U} = \frac{\partial S_2}{\partial U}$ et $\frac{\partial S_1}{\partial V} = \frac{\partial S_2}{\partial V}$ définissant la partition la plus probable s'interprètent donc ainsi :

$$\frac{1}{T_1} = \frac{1}{T_2} \quad \text{et} \quad \frac{p_1}{T_1} = \frac{p_2}{T_2}$$

soit l'égalité des températures et des pressions, ce qui est bien le résultat classique de l'uniformité des pressions et des températures à l'équilibre.

Remarque : la définition de la température et de la pression micro-canoniques est sous-tendue par leur définition classique ; la thermodynamique statistique n'a pu naître qu'après l'achèvement de la thermodynamique classique et naturellement on l'enseigne dans l'ordre historique. De même les mathématiques modernes n'ont pu naître qu'après les mathématiques classiques et la grammaire générative qu'après la grammaire classique mais pourquoi diable des experts autoproclamés en didactique ont-ils pu croire pouvoir faire enseigner directement les mathématiques modernes et la grammaire générative sans passer par les cases mathématiques et grammaire classiques ?

2.c Fluctuations d'une variable interne.

• Cas général.

Une variable interne X est un paramètre n'est pas fixée par les contraintes extérieures (qui sont ici la donnée de U et V), ce pourrait être, par exemple le nombre de particules, dans la moitié gauche du volume ou bien d'autres choses.

Appelons $\Omega(U, V, x)$ le nombre de micro-états accessibles pour lesquels X a la valeur x ; comme U et V sont supposés fixes dans le modèle micro-canonique, nous allégerons l'écriture en $\Omega(x)$. On définira par analogie une *entropie partielle* par $S(U, V, x) = k \ln \Omega(U, V, x)$ ou, de façon allégée $S(x) = k \ln \Omega(x)$.

En raisonnant comme dans le paragraphe précédent, la valeur la plus probable de x , dont celle qui sera observée, correspond au maximum de $\omega(x)$ et de $S(x)$; notons x_0 cette valeur.

Pour x voisin de x_0 , on peut remplacer $S(x)$ par un développement de TAYLOR à l'ordre deux en bonne approximation, soit :

$$S(x) \approx S(x_0) + (x - x_0) S'(x_0) + \frac{(x - x_0)^2}{2} S''(x_0)$$

où, puisque x_0 correspond à un maximum, $S'(x_0) = 0$ et $S''(x_0) < 0$ et nous noterons $S''(x_0) = -\frac{k}{a^2}$ d'où :

$$S(x) = k \ln \Omega(x) \approx S(x_0) - \frac{k(x - x_0)^2}{2a^2}$$

d'où, puisque $S(x_0) = k \ln \Omega(x_0)$ et en passant à l'exponentielle :

$$\Omega(x) \approx \Omega(x_0) \exp \left[-\frac{(x - x_0)^2}{2a^2} \right]$$

Quiconque a étudié un peu de statistiques reconnaît¹⁰ une loi de distribution gaussienne de moyenne x_0 , de variance a^2 et d'écart-type a . Là où l'approximation n'est plus valable, la valeur de $\Omega(x)$, comparée à $\Omega(x_0)$ est négligeable et que le négligeable ne soit plus gaussien, peut nous en chaut¹¹.

Sans préciser la nature du paramètre X , on ne peut pas pousser plus loin ; toutefois, on peut subodorer que, comme en thermodynamique classique, S donc S'' est proportionnel au nombre N de particules du système (comme tout paramètre extensif) donc que l'écart-type est en \sqrt{N} et l'on retrouve un résultat établi dans le prologue mathématique.

• **Cas particulier de l'énergie d'une partie.**

A titre d'exemple, appliquons ce qui précède au cas où X désigne l'énergie u d'une partie du système pour prolonger l'avant-dernier paragraphe. Si l'on ne s'intéresse pas à son volume v , le $S(x)$ du dernier paragraphe, rebaptisé $S(u)$, va s'identifier, à la suppression de la référence à v près, avec $S_1(u) + S_2(U - u)$ du paragraphe 2.b p. 13. U étant constant dans le modèle micro-canonique, on a :

$$\begin{aligned} S(u) &= S_1(u) + S_2(U - u) \\ S'(u) &= S'_1(u) + S'_2(U - u)(-1) \\ S''(u) &= S''_1(u) + S''_2(U - u)(-1)^2 \end{aligned}$$

où les (-1) proviennent de la dérivation de $(U - u)$ par rapport à u .

Or, on a vu que $S'(u) = \frac{\partial S}{\partial u}$ est l'inverse de la température du système, d'où :

$$\begin{aligned} S'(u) &= \frac{1}{T_1(u)} - \frac{1}{T_2(U - u)} \\ S''(u) &= \frac{\partial}{\partial u} \left(\frac{1}{T_1(u)} \right) + (-1)^2 \frac{\partial}{\partial u} \left(\frac{1}{T_2(u)} \right) = -\frac{1}{T_1^2} \frac{\partial T_1}{\partial u} - \frac{1}{T_2^2} \frac{\partial T_2}{\partial u} \end{aligned}$$

Et si $u = u_0$, valeur à l'équilibre, pour lequel $T_1 = T_2$ (voir là encore le paragraphe 2.b p. 13) dont la valeur commune sera notée T ,

$$S''(u_0) = -\frac{1}{T^2} \left(\frac{\partial T_1}{\partial u} + \frac{\partial T_2}{\partial u} \right)$$

10. Sinon, avec un changement d'origine pour alléger et une densité de probabilité $p(x) = A \exp \left[-\frac{x^2}{2a^2} \right]$ où A se tire de $\int_{-\infty}^{\infty} p(x) dx$, la variance est $\sigma^2 = \int_{-\infty}^{\infty} x^2 p(x) dx$ que l'on intègre par parties après report de l'expression de $p(x)$, etc. et l'on montre que l'écart-type est a . Je ne détaille pas, vous saurez le faire et ce n'est pas de la physique.

11. du verbe chaloir, participe présent chalant comme dans nonchalant.

En anticipant sur l'identification des concepts statistiques avec leurs homologues classiques, on appellera capacité thermique à volume constant, noté K_v ou ici K tout court la dérivée $\frac{\partial u}{\partial T}|_v$ et si $\frac{\partial u}{\partial T} = K$, alors $\frac{\partial T}{\partial u} = \frac{1}{K}$ (simple adaptation du théorème sur les dérivées des fonctions réciproques) d'où :

$$S''(u_0) = -\frac{1}{T^2} \left(\frac{1}{K_1} + \frac{1}{K_2} \right)$$

Pour alléger la suite des calculs supposons que la partie 1 soit très petite par rapport à la partie 2 de sorte que $K_1 \ll K_2$. Dans ces conditions, l'écart-type a de la valeur de u , calculé comme dans le précédent sous-paragraphe est donné par :

$$S''(u_0) = -\frac{k}{a^2} \approx -\frac{1}{T^2} \frac{1}{K_1}$$

$$a \approx T \sqrt{k K_1}$$

Toujours en anticipant sur l'identification avec les concepts classiques, pour la plupart des systèmes comportant N molécules, les capacités thermiques varient peu et sont de l'ordre de Nk où (par exemple $(5/2)Nk$ pour un gaz diatomique), d'où $u_0 \sim NkT$ et

$$a \approx T \sqrt{k K_1} \sim T \sqrt{Nk^2} = \sqrt{N} k T$$

où l'on retrouve, comme dans le prologue mathématique un écart-type en \sqrt{N} et un écart-type relatif $\frac{a}{u_0}$ en $1/\sqrt{N}$.

2.d Irréversibilité.

Plusieurs approches, dont les deux premières prolongent ce qui précède, sont possibles pour décrire l'irréversibilité dans le modèle microcanonique :

Si à un instant donné une variable interne X a une valeur initiale x_i très éloigné de la valeur d'équilibre, lors de l'évolution du système, X balayera toute sorte de valeurs avec une moyenne égale à la valeur d'équilibre et un écart-type minuscule de sorte que X passera en pratique de x_i à x_0 et il sera hautement improbable que X repasse un jour par la valeur x_i . Comme x_0 correspond (cf supra) à un maximum d'entropie partielle, le passage de x_i à x_0 se traduit par une augmentation d'entropie

Si à un instant donné une contrainte impose une énergie u et un volume v à une partie du système, le nombre de micro-états accessibles est $\omega_1(u, v) \omega_2(U - u, V - v)$ (cf supra). Si on libère la contrainte, le nombre de micro-états passe, dans un contexte où u et v sont formellement quantifiés, pour alléger l'exposé (sinon on remplace la somme double par une intégrale double), à quelque chose comme $\Omega = \sum_u \sum_v \omega_1(u, v) \omega_2(U - u, V - v)$ qui va être beaucoup plus grand ; le nombre d'états augmente, l'entropie aussi et pour les mêmes raisons que dans le cas précédent, le retour à l'état initial, s'il est trop éloigné de l'état d'équilibre est très peu probable.

Néanmoins, l'impossible retour aux conditions initiales a longtemps posé problème; les détracteurs de BOLTZMANN disaient : « Les lois de la physique sont invariantes par retournement du temps, il suffit donc à un instant donné de retourner toutes les vitesses de toutes les particules pour repasser par tous les états antérieurs. », ce à quoi BOLTZMANN, agacé, répondait : « Allez-y, retournez les ! ». Pour expliquer cet apparent paradoxe, les époux EHRENFEST (Paul, physicien et Tatiana, mathématicienne) ont développé le modèle dit des urnes d'EHRENFEST, vite remplacé par un modèle de puces et de chiens.

Dans cette histoire, il y a deux chiens. A un instant initial, N puces sont toutes sur le même chien. A intervalle de temps régulier, disons tous les $\tau = 1$ picoseconde (l'ordre de grandeur du temps séparant deux chocs pour une molécule d'un gaz aux conditions ordinaires), une puce tirée au hasard saute du chien où elle est à l'autre. Des simulations montrent, avec de petites valeurs de N , que rapidement chaque chien a près de $N/2$ puces mais avec des fluctuations importantes et il arrive de temps à temps que les N puces se retrouvent sur le chien où elles étaient au début de cette histoire. On a pu démontrer que le temps moyen T_m séparant deux instants où les N puces se retrouvent toutes sur le chien du départ est $T_m = 2^N \tau$ (théorème de KAC¹² en 1947). En utilisant l'approximation informatique $2^{10} = 10^3$,

- si $N = 30$, $T_m = 10^9 \tau = 1$ seconde
- si $N = 60$, $T_m = 10^{18} \tau = 10^9$ secondes, disons 30 ans
- si $N = 90$, $T_m = 10^{27} \tau = 10^{18}$ secondes, disons 30 milliards d'années, soit environ deux fois l'âge de l'univers
- si $N = 120$, $T_m = 10^{36} \tau = 10^{27}$ secondes, disons 30 milliards de milliards d'années, soit environ deux milliards de fois l'âge de l'univers, c'est à dire jamais.
- si $N = \mathcal{N}_A$, le nombre d'AVOGADRO (environ $6 \cdot 10^{23}$), il faut attendre l'éternité¹³ !

2.e Evolution d'un système dans le modèle micro-canonique

Nous allons ici introduire un peu plus de formalisme et un second postulat pour étudier un peu plus en détail l'évolution d'un système.

• Le contexte de libération d'une contrainte.

Prenons à titre d'exemple un système isolé séparé en deux par une paroi fixe et isolante qui le sépare en deux sous-systèmes isolés d'énergies et de volumes U_1 et V_1 pour l'un et U_2 et V_2 pour l'autre. Seuls les micro-états vérifiant cette contrainte sont accessibles.

Si l'on escamote cette paroi (on *libère* la contrainte), d'autres micro-états deviennent accessibles, correspondant à un autre partage du volume total et de l'énergie totale.

Juste au moment où la contrainte est libérée, la probabilité des états juste libérés est encore nulle, ce qui est contraire au principe d'équipartition. Il va donc exister une phase

12. prononcer "katz", le nom est d'origine polonaise.

13. et l'éternité, c'est long... surtout vers la fin (Woody Allen).

transitoire pendant laquelle les probabilités vont évoluer jusqu'à cette équipartition et c'est cette phase que nous voulons analyser.

• **Probabilités de transition.**

Supposons que notre système (toujours dans l'hypothèse micro-canonique) soit à l'instant t dans un micro-état indicé par la lettre ℓ . L'agitation thermique rend cet état instable et à l'instant $t + dt$ il peut encore être dans l'état ℓ mais peut aussi être dans un tout autre micro-état, par exemple un micro-état particulier indicé par la lettre m . Pour un instant initial t donné, la probabilité que le système initialement dans le micro-état ℓ à t soit dans le micro-état m à $t + dt$ dépend *a priori* de ℓ , m , t et dt . L'invariance des lois de la physique par translation dans le temps indique qu'elle ne doit pas dépendre de t ; notons-la $\delta\mathcal{P}_{\ell,m}(dt)$ et appelons-la *probabilité de transition*.

Cette probabilité est bien évidemment nulle pour dt nul et croît bien évidemment avec dt . On peut donc pour dt suffisamment petit se contenter d'une approximation d'ordre 1 (en fait un développement de TAYLOR). On notera :

$$\delta\mathcal{P}_{\ell,m}(dt) = W_{\ell,m} dt$$

où $W_{\ell,m}$ est appelée *probabilité de transition par unité de temps*.

Revenons à une situation classique où l'on ne sait pas dans quel micro-état est le système. Appelons $p_\ell(t)$ la probabilité qu'il soit dans le micro-état ℓ à un instant initial t . Nous nous proposons de donner l'expression de la probabilité $p_\ell(t + dt)$ qu'il soit dans le même micro-état à l'instant final $t + dt$.

Il faut d'abord déduire les probabilités qu'il soit dans le micro-état ℓ à t (avec la probabilité p_ℓ) et qu'il soit passé dans un micro-état m (avec la probabilité $W_{\ell,m} dt$), donc déduire la somme des produits $p_\ell W_{\ell,m} dt$ (probabilités composées). Mais il faut tenir compte du mécanisme inverse et ajouter les probabilités qu'il soit dans le micro-état m à t (avec la probabilité p_m) et qu'il soit passé dans un micro-état ℓ (avec la probabilité $W_{m,\ell} dt$), donc déduire la somme des produits $p_m W_{m,\ell} dt$. Donc :

$$p_\ell(t + dt) = p_\ell(t) - \sum_m p_\ell W_{\ell,m} dt + \sum_m p_m W_{m,\ell} dt$$

On en déduit par passage à la limite que :

$$\frac{dp_\ell}{dt} = \frac{p_\ell(t + dt) - p_\ell(t)}{dt} = \sum_m (p_m W_{m,\ell} - p_\ell W_{\ell,m})$$

• **Principe du bilan détaillé.**

Cette relation est en particulier valable si le système est à l'équilibre; or dans ce cas l'hypothèse microcanonique affirme que les probabilités des micro-états accessibles sont

égales, donc égales à $1/\Omega$ et donc constantes. La relation précédente où $\frac{dp_\ell}{dt}$ est nulle et où l'on peut simplifier par la valeur commune des probabilités devient :

$$0 = \sum_m (W_{m,\ell} - W_{\ell,m})$$

Une condition suffisante mais non nécessaire pour qu'il en soit ainsi est que pour tout couple (ℓ, m) , on ait $W_{m,\ell} = W_{\ell,m}$. Le principe du bilan détaillé postule que cette condition est réalisée.

$$\boxed{\forall \ell \forall m \quad W_{m,\ell} = W_{\ell,m}}$$

• **Retour à l'équilibre.**

Revenons au cas d'une situation initiale hors d'équilibre. Avec le principe du bilan détaillé, on a :

$$\frac{dp_\ell}{dt} = \sum_m (p_m - p_\ell) W_{\ell,m}$$

On a donné comme définition première de l'entropie, valable même hors d'équilibre, la relation $S = -k \sum_\ell p_\ell \ln(p_\ell)$; par dérivation on en déduit :

$$\begin{aligned} \frac{dS}{dt} &= -k \sum_\ell \frac{dp_\ell}{dt} \ln(p_\ell) - k \sum_\ell p_\ell \frac{1}{p_\ell} \frac{dp_\ell}{dt} = -k \sum_\ell \frac{dp_\ell}{dt} \ln(p_\ell) - k \sum_\ell \frac{dp_\ell}{dt} = \dots \\ &\dots = -k \sum_\ell \frac{dp_\ell}{dt} \ln(p_\ell) - k \frac{d}{dt} \sum_\ell p_\ell = -k \sum_\ell \frac{dp_\ell}{dt} \ln(p_\ell) - k \frac{d}{dt} (1) = \dots \\ &\dots = -k \sum_\ell \frac{dp_\ell}{dt} \ln(p_\ell) \end{aligned}$$

où l'on s'est souvenu que la somme des probabilités de tous les cas possibles est égale à l'unité.

Reportons-y le résultat concernant les dérivées des probabilités élémentaires :

$$\frac{dS}{dt} = -k \sum_\ell \ln(p_\ell) \sum_m (p_m - p_\ell) W_{\ell,m} = -k \sum_\ell \sum_{m \neq \ell} \ln(p_\ell) (p_m - p_\ell) W_{\ell,m}$$

Regroupons-les deux termes relatifs à un même couple d'indices, c'est-à-dire le terme $\ln(p_\ell) (p_m - p_\ell) W_{\ell,m}$ et le terme $\ln(p_m) (p_\ell - p_m) W_{m,\ell}$ et tenons compte du principe du bilan détaillé. On peut alors écrire, en remplaçant $m \neq \ell$ par $m > \ell$ pour ne pas compter deux fois ces termes :

$$\frac{dS}{dt} = -k \sum_\ell \sum_{m > \ell} [\ln(p_\ell) (p_m - p_\ell) W_{\ell,m} + \ln(p_m) (p_\ell - p_m) W_{m,\ell}]$$

$$\frac{dS}{dt} = -k \sum_{\ell} \sum_{m>\ell} [\ln(p_{\ell}) (p_m - p_{\ell}) W_{\ell,m} - \ln(p_m) (p_m - p_{\ell}) W_{\ell,m}]$$

$$\frac{dS}{dt} = -k \sum_{\ell} \sum_{m>\ell} [\ln(p_{\ell}) - \ln(p_m)] (p_m - p_{\ell}) W_{\ell,m}$$

$$\frac{dS}{dt} = k \sum_{\ell} \sum_{m>\ell} [\ln(p_{\ell}) - \ln(p_m)] (p_{\ell} - p_m) W_{\ell,m}$$

La fonction logarithme est croissante, donc $\ln(p_{\ell}) - \ln(p_m)$ est du même signe que $p_{\ell} - p_m$; chacun des termes $[\ln(p_{\ell}) - \ln(p_m)] (p_{\ell} - p_m)$ est donc positif ou nul et par ailleurs $W_{\ell,m}$ est par essence positif (c'est une probabilité) et non nul car l'état m est accessible donc il y a une probabilité non nulle d'y accéder. On retrouve ici un résultat essentiel : l'entropie ne peut diminuer en aucune façon ; toute évolution s'accompagne d'une augmentation d'entropie, tout au moins jusqu'au maximum.

- **Etat d'entropie maximum.**

Si S est maximum, $\frac{dS}{dt}$ est nul ; comme c'est une somme de terme positifs ou nuls, ce n'est possible que si chacun des termes est nul. Comme les $W_{\ell,m}$ sont non nuls par essence, tous les $[\ln(p_{\ell}) - \ln(p_m)] (p_{\ell} - p_m)$ sont nuls et on en déduit aisément que toutes les probabilités de micro-états sont égales.

On vient ici de déduire l'hypothèse micro-canonique du principe du bilan détaillé. Dans un axiomatic optimisée on peut supprimer l'hypothèse micro-canonique qui devient un principe, mais on préférera respecter l'ordre historique d'apparition des concepts.

L'hypothèse micro-canonique seule permet de reconstruire la thermodynamique de l'équilibre. Le principe du bilan détaillé permet d'affiner la compréhension du retour à l'équilibre.

3 Modèle canonique.

3.a Loi de Maxwell-Boltzmann.

Le modèle micro-canonique ne correspond guère à la réalité où le système que l'on étudie est en général en équilibre thermique avec l'air ambiant et non pas isolé.

Soit un système \mathcal{S} , assez petit, en contact avec un extérieur \mathcal{E} , très grand, appelé thermostat, le tout formant un système isolé d'énergie totale constante U_0 . On suppose que le couplage entre système et extérieur est faible (cf supra) de sorte que U_0 soit somme des énergies de \mathcal{S} et de \mathcal{E} en excellente approximation. Par contre, on suppose que le volume V du système \mathcal{S} ne varie pas.

Soit un micro-état du système, d'énergie U , l'extérieur a donc l'énergie $U_0 - U$ avec $U \ll U_0$ car le système est réputé très petit. Le nombre de micro-états de l'ensemble système-extérieur correspondant à cette situation est égal au produit du nombre de micro-états du système (égal à 1 par hypothèse) et de celui de l'extérieur pour une énergie $U_0 - U$, soit :

$$\Omega = 1 \times \Omega_{\mathcal{E}}(U_0 - U)$$

L'ensemble système-extérieur est isolé et relève du modèle micro-canonique ; la probabilité de la situation étudiée, somme des probabilités élémentaires, est donc, à l'équilibre, proportionnelle au nombre de micro-états accessibles car ils sont équiprobables, nous noterons donc :

$$p(U) = Cte \cdot \Omega_{\mathcal{E}}(U_0 - U)$$

Introduisons l'entropie micro-canonique partielle de l'extérieur et effectuons-en un développement de TAYLOR à l'ordre 1, ce qui sera, dans le contexte, une excellente approximation :

$$k \ln \Omega_{\mathcal{E}}(U_0 - U) = S_{\mathcal{E}}(U_0 - U) \approx S_{\mathcal{E}}(U_0) - U \frac{dS_{\mathcal{E}}}{dU}(U_0)$$

où $S_{\mathcal{E}}(U_0)$ est une constante dont il importe peu de connaître la valeur et où l'on reconnaît dans $\frac{dS_{\mathcal{E}}}{dU}(U_0)$ la définition de l'inverse la température statistique de l'extérieur, noté ici $T_{\mathcal{E}}$ (puis T tout court). On a donc :

$$k \ln \Omega_{\mathcal{E}}(U_0 - U) = S_{\mathcal{E}}(U_0 - U) \approx Cte - \frac{U}{T_{\mathcal{E}}}$$

$$p(U) = Cte \cdot \Omega_{\mathcal{E}}(U_0 - U) \approx Cte \cdot e^{-\frac{U}{kT}}$$

Remarque 1 : La constante multiplicative se déduit du fait que la somme des probabilités est par essence égale à l'unité, soit :

$$1 = \sum_i Cte \cdot e^{-\frac{U_i}{kT}} = Cte \cdot \sum_i e^{-\frac{U_i}{kT}}$$

On appelle *fonction de partition*, notée traditionnellement Z , la somme

$$Z = \sum_i e^{-\frac{U_i}{kT}}$$

d'où $Cte = 1/Z$ et :

$$p(U) = \frac{1}{Z} e^{-\frac{U}{kT}}$$

Remarque 2 : Vu la démarche, il y a un terme $e^{-\frac{U_i}{kT}}$ par micro-état et non par valeur de U ; si plusieurs micro-états ont la même énergie U_i , il faudra en tenir compte en comptant plusieurs fois le même terme, en pratique en le multipliant par le nombre de micro-états ayant cette énergie.

Remarque 3 : Dans le modèle micro-canonique, l'énergie et le volume sont fixés ; dans le modèle canonique, c'est le volume et la température qui sont fixés et l'énergie fluctue.

3.b Moyennes statistiques des fonctions d'état.

- **Généralités.**

Soit une fonction d'état X ; dans le micro-état d'indice i , on note X_i sa valeur. Sa moyenne est définie (cf supra) par $\langle X \rangle = \sum_i p_i X_i$, soit ici :

$$\langle X \rangle = \sum_i p_i X_i = \sum_i \frac{1}{Z} e^{-\frac{U_i}{kT}} X_i = \frac{1}{Z} \sum_i X_i e^{-\frac{U_i}{kT}}$$

Cette définition permet aussi le calcul de la variance et de l'écart-type en se souvenant que $\sigma^2 = \langle X^2 \rangle - \langle X \rangle^2$ où :

$$\langle X^2 \rangle = \frac{1}{Z} \sum_i X_i^2 e^{-\frac{U_i}{kT}}$$

- **Application à l'énergie.**

La valeur de l'énergie du micro-état d'indice i n'est autre que celle qui intervient dans p_i ; il suffit de remplacer X_i par U_i dans la formule précédente pour calculer la moyenne de l'énergie, soit :

$$\langle U \rangle = \frac{1}{Z} \sum_i U_i e^{-\frac{U_i}{kT}}$$

Pour évaluer cette somme, désolé, il faut sortir quelques outils mathématiques. Posons $\beta = \frac{1}{kT}$ et considérons la fonction de partition Z comme une fonction de β :

$$Z(\beta) = \sum_i e^{-\frac{U_i}{kT}} = \sum_i e^{-U_i \beta}$$

Les règles de quantifications rencontrées en physique valident les théorèmes qui permettent la dérivation terme à terme ; nous l'admettrons car ça nous mènerait trop loin de chercher à le démontrer. Donc on peut affirmer (cf dérivation des fonctions composées) que :

$$\frac{dZ}{d\beta} = - \sum_i U_i e^{-U_i \beta} = - \sum_i U_i e^{-\frac{U_i}{kT}}$$

Considérons maintenant que Z est fonction de T par l'intermédiaire du lien entre T et β , alors, par dérivation d'une fonction composée :

$$\frac{dZ}{dT} = \frac{dZ}{d\beta} \frac{d\beta}{dT} = \frac{dZ}{d\beta} \frac{d}{dT} \left(\frac{1}{kT} \right) = \frac{dZ}{d\beta} \left(-\frac{1}{kT^2} \right)$$

d'où, successivement :

$$\frac{dZ}{d\beta} = -kT^2 \frac{dZ}{dT}$$

$$\sum_i U_i e^{-\frac{U_i}{kT}} = -\frac{dZ}{d\beta} = kT^2 \frac{dZ}{dT}$$

$$\langle U \rangle = \frac{1}{Z} \sum_i U_i e^{-\frac{U_i}{kT}} = \frac{kT^2}{Z} \frac{dZ}{dT} = kT^2 \frac{d}{dT}(\ln Z)$$

Retenons provisoirement :

$$\boxed{\langle U \rangle = kT^2 \frac{d}{dT}(\ln Z)}$$

• **Application aux fluctuations de l'énergie.**

En vu du calcul de la variance puis de l'écart-type de l'énergie, calculons la moyenne de son carré. On reprend le schéma du calcul précédent en l'adaptant ; j'allège donc la rédaction. On a donc :

$$\langle U^2 \rangle = \frac{1}{Z} \sum_i U_i^2 e^{-\frac{U_i}{kT}} = \frac{1}{Z} \sum_i U_i^2 e^{-\frac{U_i}{kT}} = \frac{1}{Z} \frac{d^2 Z}{d\beta^2}$$

Par ailleurs, on avait établi que :

$$\frac{dZ}{d\beta} = -kT^2 \frac{dZ}{dT}$$

d'où :

$$\frac{d^2 Z}{d\beta^2} = \frac{d}{d\beta} \left(\frac{dZ}{d\beta} \right) = \frac{d}{dT} \left(\frac{dZ}{d\beta} \right) \cdot \frac{dT}{d\beta} = \frac{d}{dT} \left(\frac{dZ}{d\beta} \right) \cdot \left(\frac{d\beta}{dT} \right)^{-1}$$

$$\frac{d^2 Z}{d\beta^2} = \frac{d}{dT} \left(-kT^2 \frac{dZ}{dT} \right) \left(-\frac{1}{kT^2} \right)^{-1}$$

$$\frac{d^2 Z}{d\beta^2} = \left(-2kT \frac{dZ}{dT} - kT^2 \frac{d^2 Z}{dT^2} \right) (-kT^2) = 2k^2 T^3 \frac{dZ}{dT} + k^2 T^4 \frac{d^2 Z}{dT^2}$$

$$\langle U^2 \rangle = \frac{1}{Z} \frac{d^2 Z}{d\beta^2} = \frac{2k^2 T^3}{Z} \frac{dZ}{dT} + \frac{k^2 T^4}{Z} \frac{d^2 Z}{dT^2}$$

Or, on a montré plus haut que

$$\langle U \rangle = \frac{kT^2}{Z} \frac{dZ}{dT}$$

d'où :

$$\sigma_U^2 = \langle U^2 \rangle - \langle U \rangle^2 = \frac{1}{Z} \frac{d^2 Z}{d\beta^2} = \frac{2k^2 T^3}{Z} \frac{dZ}{dT} + \frac{k^2 T^4}{Z} \frac{d^2 Z}{dT^2} - \frac{k^2 T^4}{Z^2} \left(\frac{dZ}{dT} \right)^2$$

ce qui est le résultat mais un peu indigeste. On peut le présenter plus simplement en remarquant que si $\langle U \rangle = \frac{kT^2}{Z} \frac{dZ}{dT} = kT^2 \frac{1}{Z} \frac{dZ}{dT}$, alors par dérivation :

$$\frac{d\langle U \rangle}{dT} = 2kT \frac{1}{Z} \frac{dZ}{dT} - kT^2 \frac{\frac{dZ}{dT}}{Z^2} \frac{dZ}{dT} + kT^2 \frac{1}{Z} \frac{d^2Z}{dT^2}$$

et l'on reconnaît

$$\sigma_U^2 = kT^2 \frac{d\langle U \rangle}{dT}$$

où l'on identifiera par analogie avec la thermodynamique classique, $\frac{d\langle U \rangle}{dT}$ à K_v capacité thermique à volume constant. On retrouve ainsi, sous une forme légèrement différente, un résultat du modèle micro-canonique. C'est du reste en recherchant cette similitude que l'on a pensé à calculer $\frac{d\langle U \rangle}{dT}$, ce qui a permis de mettre le résultat sous forme exploitable.

On peut bien sûr trouver un moyen plus rapide d'arriver à ce résultat mais il ne serait pas dans la logique de cet exposé ; l'élégance ponctuelle ne doit pas nuire à la cohérence d'ensemble.

• Application à l'entropie.

Revenons à la définition la plus générale de l'entropie partielle d'un état, à savoir $S = -k \langle \ln(p_i) \rangle$ (cf supra) ; on va donc, dans la formule générale remplacer X_i par :

$$-k \ln(p_i) = -k \ln \left(\frac{1}{Z} e^{-\frac{U_i}{kT}} \right) = k \left(\ln Z + \frac{U_i}{kT} \right)$$

La formule $\langle X \rangle = \frac{1}{Z} \sum_i X_i e^{-\frac{U_i}{kT}}$ s'adapte donc en :

$$\langle S \rangle = \frac{1}{Z} \sum_i k \left(\ln Z + \frac{U_i}{kT} \right) e^{-\frac{U_i}{kT}}$$

soit en scindant en deux sommes :

$$\langle S \rangle = \frac{k \ln Z}{Z} \sum_i e^{-\frac{U_i}{kT}} + \frac{1}{T Z} \sum_i U_i e^{-\frac{U_i}{kT}}$$

où l'on reconnaît $Z = \sum_i e^{-\frac{U_i}{kT}}$ (par définition) et $\sum_i U_i e^{-\frac{U_i}{kT}} = kT^2 \frac{dZ}{dT}$ (résultat partiel antérieur), d'où :

$$\langle S \rangle = \frac{k \ln Z}{Z} Z + \frac{1}{T Z} kT^2 \frac{dZ}{dT} = k \ln Z + \frac{kT}{Z} \frac{dZ}{dT}$$

soit encore en ouvrant bien les yeux

$$\langle S \rangle = k \cdot 1 \cdot \ln Z + k \cdot T \cdot \frac{d \ln Z}{dT} = k \frac{d(T \ln Z)}{dT}$$

Retenons provisoirement :

$$\langle S \rangle = k \frac{d(T \ln Z)}{dT}$$

- **Energie libre.**

Par analogie avec la thermodynamique classique, on peut définir l'énergie libre moyenne par

$$\langle F \rangle = \langle U \rangle - T \langle S \rangle$$

(on ne perd pas de vue que T est imposé par le thermostat et que U et S fluctuent), soit en reportant les résultats précédents :

$$\langle F \rangle = k T^2 \frac{d}{dT}(\ln Z) - k T \frac{d(T \ln Z)}{dT} = k T^2 \frac{d}{dT}(\ln Z) - k T \left(\ln Z + T \frac{d(\ln Z)}{dT} \right)$$

d'où :

$$\langle F \rangle = -k T \ln Z$$

- **Optimisation du formalisme.**

En confrontant les formules $\langle F \rangle = -k T \ln Z$ et $\langle S \rangle = k \frac{d(T \ln Z)}{dT}$, on tire :

$$\langle S \rangle = -\frac{dF}{dT}$$

et par ailleurs $\langle F \rangle = \langle U \rangle - T \langle S \rangle$ peut se réécrire $\langle U \rangle = \langle F \rangle + T \langle S \rangle$. On peut donc pour un système donné calculer, à partir des énergies U_i des micro-états accessibles, la fonction de partition Z puis les fonctions d'état selon l'algorithme suivant :

$$\begin{aligned} Z &= \sum_i e^{-\frac{U_i}{kT}} \\ \langle F \rangle &= -k T \ln Z \\ \langle S \rangle &= -\frac{dF}{dT} \\ \langle U \rangle &= \langle F \rangle + T \langle S \rangle \end{aligned}$$

3.c Factorisation de la fonction de partition.

- **Un exemple simple.**

Supposons que l'énergie U d'un micro-état soit somme de deux termes U' et U'' pouvant prendre, indépendamment l'un de l'autre, les valeurs u'_1 et u'_2 pour U' et les valeurs u''_1 et

u_2'' pour U'' . U a donc quatre valeurs possibles et l'on a successivement :

$$Z = e^{-\frac{u_1'+u_1''}{kT}} + e^{-\frac{u_1'+u_2''}{kT}} + e^{-\frac{u_2'+u_1''}{kT}} + e^{-\frac{u_2'+u_2''}{kT}}$$

$$Z = e^{-\frac{u_1'}{kT}} e^{-\frac{u_1''}{kT}} + e^{-\frac{u_1'}{kT}} e^{-\frac{u_2''}{kT}} + e^{-\frac{u_2'}{kT}} e^{-\frac{u_1''}{kT}} + e^{-\frac{u_2'}{kT}} e^{-\frac{u_2''}{kT}}$$

$$Z = e^{-\frac{u_1'}{kT}} \left(e^{-\frac{u_1''}{kT}} + e^{-\frac{u_2''}{kT}} \right) + e^{-\frac{u_2'}{kT}} \left(e^{-\frac{u_1''}{kT}} + e^{-\frac{u_2''}{kT}} \right)$$

$$Z = \left(e^{-\frac{u_1'}{kT}} + e^{-\frac{u_2'}{kT}} \right) \left(e^{-\frac{u_1''}{kT}} + e^{-\frac{u_2''}{kT}} \right)$$

On voit que Z se factorise en un produit $Z' Z''$ où Z' et Z'' sont les fonctions de partition que l'on obtiendrait si U se résumait au seul terme U' ou U'' .

Il résulte du passage au logarithme dans l'algorithme du paragraphe précédent que $\langle F \rangle = \langle F' \rangle + \langle F'' \rangle$ où $\langle F' \rangle = -kT \ln Z'$ et $\langle F'' \rangle = -kT \ln Z''$ et ainsi de suite pour $\langle S \rangle$ et $\langle U \rangle$.

• Généralisation.

Si U est somme de plusieurs termes U_j pouvant prendre les valeurs u_{ji} , on démontre de même avec plein de signes Σ et Π , un produit d'ensembles et de n-uplets éléments de ce produit que $Z = \prod_j Z_j$ où $Z_j = \sum_i e^{-\frac{u_{ji}}{kT}}$ et aussi que $\langle F \rangle = \sum_j \langle F_j \rangle$ où l'on a $\langle F_j \rangle = -kT \ln Z_j$ et ainsi de suite pour $\langle S \rangle$ et $\langle U \rangle$.

Ceci permet, quand une énergie est somme de différents termes indépendants, de les étudier indépendamment les uns des autres et d'en déduire leurs contributions individuelles aux fonctions d'état, ce qui donne beaucoup de souplesse à ce formalisme.

• Cas du système formés d'entités identiques.

Si un système est formés de N entités identiques mais statistiquement indépendantes et dont les énergies peuvent prendre exactement les mêmes valeurs, on aura, d'après ce qui précède $Z = \prod_{j=1}^{j=N} Z_j$ mais avec les Z_j tous égaux, d'où

$$Z = Z_1^N$$

$$\langle F \rangle = N \langle F_1 \rangle$$

$$\langle S \rangle = N \langle S_1 \rangle$$

$$\langle U \rangle = N \langle U_1 \rangle$$

Ce qui, somme toute, ne nous surprend pas car les fonctions d'état énergétiques sont extensives.

Remarque de synthèse : si l'énergie d'une entité est lui-même somme de plusieurs termes indépendants, Z_1 se factorisera et $\langle F_1 \rangle$, $\langle S_1 \rangle$ et $\langle U_1 \rangle$ seront les sommes des contributions des différents termes énergétiques.

3.d Un exemple discret : le paramagnétisme de Langevin.

Dans le cas de matériaux paramagnétiques¹⁴, l'énergie d'un atome contient un terme magnétique lié à l'interaction de son moment dipolaire magnétique m et du champ magnétique B dont l'expression est $U = -mB$ (il s'agit en fait d'un produit scalaire).

Le cas le plus simple est le paramagnétisme à deux niveaux dit paramagnétisme de LANGEVIN. Le moment dipolaire m peut prendre deux valeurs symétriques $\pm\mu$ d'énergies $u_+ = -\mu B$ et $-\mu$ d'énergie $u_- = \mu B$. On a donc successivement et moyennant une petite révision de trigonométrie hyperbolique :

$$Z = e^{-\frac{\mu B}{kT}} + e^{\frac{\mu B}{kT}} = 2 \operatorname{ch} \frac{\mu B}{kT}$$

$$\langle F \rangle = -kT \ln \left(2 \operatorname{ch} \frac{\mu B}{kT} \right)$$

$$\langle S \rangle = -\frac{d\langle F \rangle}{dT} = k \ln \left(2 \operatorname{ch} \frac{\mu B}{kT} \right) + kT \frac{2 \operatorname{sh} \frac{\mu B}{kT} \left(\frac{-\mu B}{kT^2} \right)}{2 \operatorname{ch} \frac{\mu B}{kT}} = k \ln \left(2 \operatorname{ch} \frac{\mu B}{kT} \right) - \frac{\mu B}{T} \operatorname{th} \frac{\mu B}{kT}$$

$$\langle U \rangle = \langle F \rangle + T \langle S \rangle = -\mu B \operatorname{th} \frac{\mu B}{kT}$$

Bien sûr si un système est formés de N atomes paramagnétiques identiques, ces résultats sont à multiplier par N (cf supra).

Ces résultats, obtenu avec une concision extrême sont conformes à la réalité expérimentale et illustrent bien la puissance de la thermodynamique statistique.

Remarque 1 : Tout bon physicien se doit maintenant de commenter ce résultat. Mais ici, il s'agit d'illustrer la méthode ; le commentaire se trouvera en bonne et due forme dans le chapitre C-XII, déjà mentionné en note.

Remarque 2 : Le moment dipolaire moyen se calcule, selon la définition d'une moyenne par :

$$\langle m \rangle = p_+ \mu + p_- (-\mu) = \frac{\mu}{Z} \left(e^{-\frac{u_+}{kT}} - e^{-\frac{u_-}{kT}} \right) = \frac{\mu}{Z} \left(e^{\frac{\mu B}{kT}} - e^{-\frac{\mu B}{kT}} \right)$$

$$\langle m \rangle = \frac{\mu}{2 \operatorname{ch} \frac{\mu B}{kT}} \left(2 \operatorname{sh} \frac{\mu B}{kT} \right) = \mu \operatorname{th} \frac{\mu B}{kT}$$

14. Voir le chapitre C-XII sur les propriétés électromagnétiques de la matière.

résultat qu'on trouvait bien plus vite car on doit avoir $\langle U \rangle = -\langle m \rangle B$; mais je voulais ici illustrer le principe de calcul d'une moyenne.

3.e Un exemple continu : le gaz parfait monoatomique.

- Comment discrétiser le continu.

Un gaz parfait est un gaz sans interaction entre ses molécules, son énergie est donc somme des énergies de ses molécules sans terme d'énergie potentielle d'interaction; on est donc, pour un système à N molécules, dans le cadre d'une factorisation de la fonction de partition en $Z = Z_1^N$ (cf supra).

Si le gaz est monoatomique, les atomes identifiables à des points matériels ont une énergie qui se réduit à l'énergie cinétique $U = \frac{1}{2} m \vec{v}^2$ ou encore, en introduisant la quantité de mouvement $\vec{p} = m \vec{v}$, $U = \frac{\vec{p}^2}{2m}$ où \vec{v} (ou \vec{p}) peut prendre toutes les valeurs possibles de façon continue. Jusqu'ici nous avons travaillé sur un nombre fini ou dénombrable de valeurs de l'énergie. Bien sûr, on conçoit bien que l'on remplacera les sommes finies ou infinies par des intégrales et les probabilités par des densités de probabilité, encore faut-il le faire de façon clairement explicitée et que les résultats théoriques concordent avec la réalité expérimentale, ce qui est le cœur même du raisonnement physique.

Je ne suis pas historien des sciences et ne dispose pas d'un corpus de documents permettant de restituer avec certitude la genèse des idées, mais il est certain que les choses ont dû évoluer par tentatives successives jusqu'à obtenir un résultat satisfaisant. Oserais-je me comparer à Jean-Jacques ROUSSEAU rédigeant le *Discours sur l'origine et les fondements de l'inégalité parmi les hommes* en affirmant que les choses ont dû se passer comme suit ?

La première idée a dû être d'associer un micro-état à une énergie U et à découper l'ensemble des valeurs de U , à savoir \mathbb{R}^+ en tranches égales d'une « largeur » δU arbitraire (que l'on fait tendre ultérieurement vers zéro) de façon à avoir une infinité dénombrable et non plus continue de micro-états.

A cela, deux objections. La première est que l'hypothèse canonique prévoit qu'il puisse y avoir plusieurs états de même énergie, donc ici correspondant à une même tranche d'énergie de largeur δU , et donc que les différentes tranches d'énergie n'ont pas forcément le même nombre de micro-états. La seconde, qui est de taille, est que le volume V n'apparaît pas dans les expressions de l'entropie et de l'énergie libre qui en résultent, au contraire du résultat classique.

Peu importe ensuite le détail du cheminement¹⁵ mais le formalisme auquel on est arrivé. Tout d'abord, le micro-état d'un atome ponctuel est la donnée non seulement de sa vitesse ou de sa quantité de mouvement (le second choix s'est avéré, à l'épreuve de l'expérience, le bon), mais aussi de sa position et l'on doit aussi tenir compte de la structure

15. Que les bouddhistes pour qui « le chemin est le but » me pardonnent cet écart. Ils le feront volontiers car la démarche de la physique, c'est d'abord et avant tout le chemin.

tridimensionnelle de l'espace; le micro-état est donc décrit par un point de l'*espace des phases* à six dimensions et dont les coordonnées sont celles de la position (x, y, z) et les composantes de la quantité de mouvement (p_x, p_y, p_z) . Cet espace est divisés en *cellules* d'hyper-volume arbitraire \mathcal{C} qui fait office de *deus ex machina*¹⁶.

Ultérieurement, cette idée s'est avérée être une fabuleuse prémonition de la mécanique quantique, plus précisément des relations d'incertitude de HEISENBERG qui affirment que, pour chacune des trois dimensions (disons selon x), le produit de l'incertitude Δx sur la position et de l'incertitude Δp_x ne peut descendre en dessous d'un minimum de l'ordre de h , constante de PLANCK, la valeur exacte dépendant de la façon précise dont on définit l'incertitude. En trois dimensions, l'hyper-volume de la cellule est donc $\mathcal{C} \sim h^3$. La valeur précise s'avérant peu importante, nous prendrons par la suite $\mathcal{C} = h^3$.

• Calcul de la fonction de partition.

Considérons un des atomes du gaz parfait enfermé dans un domaine \mathcal{V} de volume V , situé au point $M \in \mathcal{V}$ de coordonnées (x, y, z) et de quantité de mouvement \vec{p} de composantes (p_x, p_y, p_z) et donc d'énergie $U = \frac{\vec{p}^2}{2m} = \frac{p^2}{2m}$ (où $p = \|\vec{p}\|$).

La première étape de notre calcul consiste à calculer combien de cellules d'hyper-volume $\mathcal{C} = h^3$ donc combien de micro-états correspondent à une énergie comprise entre $U = \frac{p^2}{2m}$ et $U + dU = \frac{(p+dp)^2}{2m}$. Notre atome est dans ce cas si M est dans \mathcal{V} de volume V et si dans un espace à trois dimensions le point de coordonnées (p_x, p_y, p_z) se trouve entre les sphères centrées sur l'origine de rayons p et $p + dp$, dans un volume :

$$\frac{4}{3} \pi (p + dp)^3 - \frac{4}{3} \pi p^3 = d \left(\frac{4}{3} \pi p^3 \right) = 4 \pi p^2 dp$$

Dans l'espace des phases le point de coordonnées (x, y, z, p_x, p_y, p_z) correspondant à notre intervalle d'énergie occupe donc un hyper-volume $4 \pi V p^2 dp$ correspond à un nombre de cellules :

$$dn = \frac{4 \pi V p^2 dp}{h^3}$$

Si dU est assez petit devant U , le terme de la fonction de partition est pratiquement $e^{-\frac{U}{kT}}$ pour toutes les énergies entre U et $U + dU$, approximation d'autant meilleure que dU est petit.

Ce dénombrement de micro-états réalisé, pour calculer la fonction de partition Z_1 relative à un seul atome, nous devons, tranche d'énergie par tranche d'énergie, additionner

16. Traduit du grec *Απο μηχανης θεος* : dans le théâtre grec, pour sortir d'une intrigue insoluble, un dieu descendait sur un nuage mû par une machinerie et résolvait le problème grâce à ses prétendus super-pouvoirs. Cette tradition a perduré sous des formes diverses jusque dans le théâtre classique et au delà, où, au dernier acte, les jeunes héros s'avèrent être les enfants de protagonistes plus âgés, jadis enlevés par les Barbaresques et reconnus grâce à signe quelconque (voir les *Fourberies de Scapin* de MOLIÈRE ou le *Mariage de Figaro* de BEAUMARCHAIS).

les produits du nombre dn de micro-états par tranche et du facteur de BOLTZMANN $e^{-\frac{U}{kT}}$, soit encore $e^{-\frac{p^2}{2mkT}}$, d'un état de cette tranche. L'approximation s'améliore quand dU diminue et par passage à la limite, on passe d'une somme discrète à l'intégrale :

$$Z_1 = \int_{p=0}^{p=\infty} \frac{4\pi V p^2 dp}{h^3} e^{-\frac{p^2}{2mkT}} = \frac{4\pi V}{h^3} \int_0^\infty e^{-\frac{p^2}{2mkT}} p^2 dp$$

La suite relève provisoirement du calcul intégral. Introduisons u , réel positif tel que $u^2 = \frac{p^2}{2mkT}$ d'où :

$$u = \frac{p}{\sqrt{2mkT}} \quad \text{et} \quad du = \frac{dp}{\sqrt{2mkT}}$$

et donc, avec une intégration par parties :

$$\begin{aligned} Z_1 &= \frac{4\pi V}{h^3} (2mkT)^{\frac{3}{2}} \int_0^\infty e^{-u^2} u^2 du \\ Z_1 &= \frac{4\pi V}{h^3} (2mkT)^{\frac{3}{2}} \int_0^\infty -2u du e^{-u^2} \left(-\frac{u}{2}\right) = \frac{4\pi V}{h^3} (2mkT)^{\frac{3}{2}} \int_0^\infty \left(-\frac{u}{2}\right) d(e^{-u^2}) \\ Z_1 &= \frac{4\pi V}{h^3} (2mkT)^{\frac{3}{2}} \left[-\frac{u}{2} e^{-u^2}\right]_0^\infty - \frac{4\pi V}{h^3} (2mkT)^{\frac{3}{2}} \int_0^\infty d\left(-\frac{u}{2}\right) (e^{-u^2}) \\ Z_1 &= 0 + \frac{2\pi V}{h^3} (2mkT)^{\frac{3}{2}} \int_0^\infty e^{-u^2} du \end{aligned}$$

Or $\int_0^\infty e^{-u^2} du = \frac{\sqrt{\pi}}{2}$ est un résultat classique¹⁷, d'où :

$$Z_1 = \frac{V}{h^3} (2\pi mkT)^{\frac{3}{2}}$$

Pour alléger la suite de l'exposé nous regrouperons les constantes sous un nom unique $\alpha = \frac{1}{h^3} (2\pi mk)^{\frac{3}{2}}$, d'où $Z_1 = \alpha V T^{\frac{3}{2}}$ puis, grâce à la règle de factorisation, la fonction de partition pour un gaz à N atomes :

$$Z = \left(\alpha V T^{\frac{3}{2}}\right)^N$$

17. Une intégrale ne dépend pas du nom donné à la variable d'intégration donc

$$\int_{-\infty}^\infty e^{-x^2} dx = \int_{-\infty}^\infty e^{-y^2} dy$$

Si l'on appelle I cette intégrale, on a

$$I^2 = \int_{-\infty}^\infty e^{-x^2} dx \int_{-\infty}^\infty e^{-y^2} dy = \iint_{\mathbb{R}^2} e^{-(x^2+y^2)} dx dy$$

On passage en polaire (\mathbb{R}^2 est à la fois un carré de coté infini et un cercle de rayon infini) d'où

$$I^2 = \iint_{\mathbb{R}^2} e^{-r^2} r dr d\theta = 2\pi \int_0^\infty e^{-r^2} r dr = \pi \left[-e^{-r^2}\right]_0^\infty = \pi$$

et par symétrie $\int_0^\infty e^{-u^2} du = \frac{I}{2} = \frac{\sqrt{\pi}}{2}$

• **Bilan et perspectives.**

La suite relève de l'application automatique de l'algorithme mis en place plus haut, qui donne ici les fonctions d'état du gaz parfait monoatomique en moyenne statistique.

$$\langle F \rangle = -k T \ln Z = -N k T \ln \left(\alpha V T^{\frac{3}{2}} \right) = -N k T \left(\ln \alpha + \ln V + \frac{3}{2} \ln T \right)$$

$$\langle S \rangle = -\frac{d\langle F \rangle}{dT} = N k \ln \left(\alpha V T^{\frac{3}{2}} \right) + N k T \frac{3}{2T} = N k \ln \left(\alpha V T^{\frac{3}{2}} \right) + \frac{3}{2} N k$$

$$\langle U \rangle = \langle F \rangle + T \langle S \rangle = \frac{3}{2} N k T$$

On retrouve le résultat classique pour lequel $pV = nRT$ et $U = \frac{3}{2} nRT$ où n est la quantité de matière (nombre de moles) si l'on prend $k = R/\mathcal{N}_A$, quotient de la constante des gaz parfaits par le nombre d'AVOGADRO (car alors, $nR = \frac{N}{\mathcal{N}_A} R = Nk$). Pour l'énergie (interne) l'identité saute aux yeux ; pour l'entropie, en thermodynamique classique, on raisonne ainsi :

$$dU = \delta W_{\text{rév.}} + \delta Q_{\text{rév.}} = -p dV + T dS$$

soit ici

$$\frac{3}{2} n R dT = -\frac{n R T}{V} dV + T dS$$

d'où :

$$dS = \frac{3}{2} n R \frac{dT}{T} + n R T \frac{dV}{V}$$

$$S = \frac{3}{2} n R \ln T + n R T \ln V + Cte = n R \ln \left(V T^{\frac{3}{2}} \right) + Cte = N k \ln \left(V T^{\frac{3}{2}} \right) + Cte$$

C'est bien la même expression mais, cerise sur la gateau, le modèle canonique précise, sans qu'il soit besoin d'évoquer le troisième principe ou principe de NERNST, devenu inutile, la valeur de la constante :

$$Cte = N k \ln \alpha + \frac{3}{2} N k = N k \ln \left(\frac{1}{h^3} (2\pi m k)^{\frac{3}{2}} \right) + \frac{3}{2} N k$$

L'identification de l'énergie et de l'entropie dans les deux théories entraîne celle de l'énergie libre.

Désormais dans la suite de ce chapitre, on écrira U pour $\langle U \rangle$ et ainsi de suite puisque l'identification avec les expressions classiques et la valeur infime des fluctuations statistiques le permettent.

On ne s'intéressera pas ici au gaz parfait polyatomique, ni au gaz non parfait, car ce serait trop long et nuirait à la fluidité d'un chapitre déjà très riche. Le chapitre E-VI est entièrement consacré aux gaz avec une approche classique et statistique.

• **Et la pression, dans tout ça ?**

La démarche du modèle canonique sous-entend que le volume du système est fixé et qu'il n'échange que de l'énergie avec le thermostat. Il n'est pas inenvisageable d'adapter le raisonnement avec un échange possible de volume. En raisonnant un peu comme dans le paragraphe "température et pression thermodynamique" du modèle microcanonique (paragraphe 2.b p. 13), on montrerait que la probabilité d'un micro-état d'énergie U_i et de volume V_j est proportionnelle à $e^{-\frac{U_i+pV_j}{kT}}$ où T et p sont la température et la pression du thermostat, la fonction de partition serait une somme double sur i et sur j se factorisant en un facteur relatif aux U_i , la fonction de partition dont nous avons parlé jusqu'ici, et un facteur :

$$Z^* = \sum_j e^{-\frac{pV_j}{kT}}$$

Le volume moyen défini par :

$$\langle V \rangle = \frac{1}{Z^*} \sum_j V_j e^{-\frac{pV_j}{kT}}$$

où $\sum_j V_j e^{-\frac{pV_j}{kT}}$ s'identifierait à $-\frac{dZ^*}{d\gamma}$ où $\gamma = \frac{p}{kT}$; mais c'est là que les choses deviennent difficiles à gérer. Avec $\beta = \frac{1}{kT}$ donc l'inverse de T à une constante près, on pouvait lier $\frac{dZ}{d\beta}$ à $\frac{dZ}{dT}$, ici γ fait intervenir p et T et les choses deviennent complexes¹⁸.

Personne à ma connaissance ne s'est lancé là-dedans car c'est inutile : le modèle canonique est si puissant qu'il exprime les fonctions d'état du système automatiquement en fonction de son volume V . Même si dans le contexte V est réputé constant (constant mais quelconque), la relation obtenue, liant F (c'est le plus simple) à T et V , est valable quels que soient T et V , donc se présente comme une fonction de T et V . L'identification avec les notions classiques permet alors d'affirmer que $dF = -p dV + T dS$ et donc que $p = -\frac{\partial F}{\partial V}$.

Par exemple, dans le cas du gaz parfait monoatomique, on a :

$$F = -N k T \ln \left(\alpha V T^{\frac{3}{2}} \right) = -N k T \ln \alpha - N k T \ln V - \frac{3}{2} N k T \ln T$$

d'où $p = -\frac{\partial F}{\partial V} = \frac{N k T}{V}$ et $pV = N k T$. On retrouve l'équation d'état du gaz parfait instantanément ; pourquoi donc élaborer une théorie délicate ?

3.f Paradoxe de Gibbs.

L'expression de l'entropie d'un gaz parfait en fonction de la température et du volume pose un problème, en thermodynamique classique comme en thermodynamique statistique

18. Après avoir justifié que $\frac{dZ^*}{d\gamma} = T \left. \frac{\partial Z^*}{\partial p} \right|_T$, les choses s'éclairent ; le chemin est balisé, un prix NOBEL est à prendre mais n'oubliez pas de le partager avec moi !

puisque cette expression y est la même. Ce problème est connu sous le nom de paradoxe de GIBBS. Il n'est évoqué que marginalement en thermodynamique classique car elle n'a pas les moyens de le lever¹⁹.

C'est que cette expression ne fait pas de l'entropie une grandeur extensive additive. En effet considérons un gaz parfait monoatomique comportant $2N$ atomes dans un volume $2V$ à la température T , chaque moitié (numérotées 1 et 2) comportant N atomes dans un volume V à la température T .

Les entropies S_1 et S_2 des deux moitiés, en utilisant la formule ci-dessus sont :

$$S_1 = S_2 = N k \ln \left(\alpha V T^{\frac{3}{2}} \right) + \frac{3}{2} N k$$

et l'entropie S du tout est :

$$S = (2N) k \ln \left[\alpha (2V) T^{\frac{3}{2}} \right] + \frac{3}{2} (2N) k$$

alors que la somme des entropies des moitiés est :

$$S_1 + S_2 = 2 N k \ln \left(\alpha V T^{\frac{3}{2}} \right) + 2 \frac{3}{2} N k$$

S n'est pas égal à $S_1 + S_2$, on a, en effet, puisque le logarithme d'un produit est somme des logarithmes :

$$\Delta S = S - (S_1 + S_2) = 2 N k \ln 2$$

Réfléchissons plus en profondeur : si la moitié 1 est du néon et la moitié 2 de l'argon et le tout un mélange homogène néon-argon, la thermodynamique classique justifie aisément la chose par une expérience de pensée et l'on conçoit aisément qu'un mélange homogène néon-argon ne soit pas la même chose qu'une juxtaposition néon à gauche, argon à droite ; la différence ΔS est la traduction du mélange irréversible des deux gaz et s'appelle donc *entropie de mélange*.

Par contre si les deux gaz sont de même nature, où siège la distinction ? Dans le fait que les atomes qui étaient initialement d'un côté ou de l'autre se retrouvent intimement mélangés de façon irréversible. Certes mais comment le faire apparaître dans les calculs ?

La réponse se trouve dans un nouveau concept, celui d'atomes *indiscernables*. Prenons un exemple : dans un problème de choc élastique entre deux points matériels²⁰ numérotés 1 et 2 de vitesses initiales connues, on est capable de trouver des couples de deux vitesses après le choc, notées ici \vec{v} et \vec{w} . Or ce couple correspond à deux situations, la première est celle où le point 1 a la vitesse \vec{v} et le point 2, \vec{w} , la seconde celle où le point 1 a la vitesse \vec{w} et le point 2, \vec{v} . Autrement dit, après le choc, on ne peut plus reconnaître quel

19. Un paradoxe se lève, un contre-exemple s'exhibe, il y a des associations verbe-nom consacrées par l'usage, entre scientifiques comme pour tout le monde.

20. schématisant le choc entre deux atomes du gaz.

est le point 1 et le point 2 puisqu'ils peuvent échanger leur rôle. C'est cela le sens profond de l'indiscernabilité des atomes d'un gaz ou plus généralement d'un fluide. A contrario, le paramagnétisme de LANGEVIN s'applique à des atomes d'un réseau cristallin immobiles donc identifiables, ils sont *discernables*.

Plaçons-nous dans le cas d'un gaz à N atomes dans des conditions telles que le nombre de cellules de l'espace des phases de probabilité non négligeable soit grand devant N ; sauf de rares cas une cellule ne contient pas d'atome ou n'en contient qu'un. Idéalisons la situation en disant que les N atomes occupent N cellules différentes. Pour un N -uplet donné de cellules, les $N!$ façons de placer les N atomes sur ces N cellules-là ont été ci-dessus comptés comme $N!$ micro-états différents, alors que l'indiscernabilité fait que tous ces micro-états n'en forment qu'un seul. Dans la fonction de partition chaque terme a donc été compté $N!$ fois et la fonction de partition a été surévaluée d'un facteur $N!$.

Notons d'un tilde les fonctions corrigées de ce facteur et sans tilde celle présentées plus haut; on a donc $\tilde{Z} = \frac{Z}{N!}$ et l'on en déduit que :

$$\tilde{F} = -kT \ln(\tilde{Z}) = -kT \ln Z - kT \ln(N!) = F - kT \ln(N!)$$

On connaît la formule de STIRLING²¹ qui donne $\ln N! \approx N \ln N - N$; en thermodynamique $N \sim 10^{24}$ d'où $\ln N \sim 24 \ln 10 \approx 55$ assez grand devant l'unité, ce qui allège encore en $\ln N! \approx N \ln N$, on a donc en bonne approximation :

$$\tilde{F} = F - N k T \ln(N)$$

On poursuit ainsi :

$$\tilde{S} = -\frac{d\tilde{F}}{dT} = -\frac{dF}{dT} - N k \ln(N) = S - N k \ln(N)$$

En reportant $S = N k \ln\left(\alpha V T^{\frac{3}{2}}\right) + \frac{3}{2} N k$, on tire avec les propriétés du logarithme :

$$\tilde{S} = N k \ln\left(\alpha \frac{V}{N} T^{\frac{3}{2}}\right) + \frac{3}{2} N k$$

On laisse au lecteur, qui commence à s'assoupir, de vérifier, dans le contexte précédent, que cette fois, on a bien $\tilde{S} = \tilde{S}_1 + \tilde{S}_2$ pour un système considéré comme juxtaposition de ses moitiés. Le paradoxe est levé.

Dans le cas où le nombre de cellules de la probabilité d'occupation non négligeable n'est plus grand devant N , il faudra trouver une autre approche, connue sous le nom de *modèle grand-canonique* ou encore de *thermodynamique statistique quantique*. C'est suffisamment vaste pour que l'on y consacre le chapitre E-IX.

21. On part de la méthode classique de comparaison d'une somme et d'une intégrale et en très gros

$$\ln(N!) = \sum_{k=1}^{k=N} \ln k \approx \int_1^N \ln x dx = [x \ln x - x]_0^N = N \ln N - N + 1$$

où 1 est négligeable devant N .

3.g Evolution d'un système dans le modèle canonique

On se place, comme dans le modèle micro-canonique, dans le cadre d'une évolution par libération d'une contrainte.

- **Probabilités de transition.**

L'ensemble système-thermostat relève du modèle micro-canonique, on peut donc lui en appliquer les concepts. Un micro-état de l'ensemble est la donnée simultanée d'un micro-état ℓ du système et d'un micro état α du thermostat ; un autre micro-état de l'ensemble est la donnée simultanée d'un micro-état m du système et d'un micro état β du thermostat. A un changement de notation près, la probabilité de transition s'écrit comme plus haut :

$$\frac{dp_{\alpha\ell}}{dt} = \sum_{\beta m} W_{\beta m, \alpha\ell} (p_{\beta m} - p_{\alpha\ell})$$

Notons la probabilité $p_{\alpha\ell}$ (ou $p_{\beta m}$) que l'ensemble soit dans l'état noté $\alpha\ell$ (ou βm) comme le produit de la probabilité p_ℓ (ou p_m) que le système soit dans l'état ℓ (ou m) et d'une probabilité²² p_α (ou p_β) définie par cette notation comme probabilité que le thermostat soit dans l'état α (ou β) compatible avec le micro-état p_ℓ (ou p_m) du système. On peut donc réécrire ainsi ce qui précède :

$$\frac{d(p_\alpha p_\ell)}{dt} = \sum_{\beta} \sum_m W_{\beta m, \alpha\ell} (p_\beta p_m - p_\alpha p_\ell)$$

où il faut lire dans \sum_{β} la somme sur les micro-états β du thermostat compatibles avec l'état m du système.

Si l'on ne s'intéresse plus qu'à la probabilité de transition du système dans le micro-état ℓ , indépendamment du micro-état dans lequel se trouve le thermostat, il suffit de sommer sur les micro-états de celui-ci et compatibles avec le micro-état ℓ du système et de se convaincre que cela se traduit par $p_\ell = \sum_{\alpha} p_\alpha p_\ell$ (et donc $\sum p_\alpha = 1$). D'où :

$$\sum_{\alpha} \frac{d(p_\alpha p_\ell)}{dt} = \frac{d(p_\ell \sum_{\alpha} p_\alpha)}{dt} = \frac{d(p_\ell \cdot 1)}{dt} = \frac{dp_\ell}{dt} = \sum_{\alpha} \sum_{\beta} \sum_m W_{\beta m, \alpha\ell} (p_\beta p_m - p_\alpha p_\ell)$$

En inversant l'ordre des sommations pour regrouper les terme relatifs à un indice m donné, on arrive à :

$$\frac{dp_\ell}{dt} = \sum_m \left[\sum_{\alpha} \sum_{\beta} W_{\beta m, \alpha\ell} p_\beta p_m - \sum_{\alpha} \sum_{\beta} W_{\beta m, \alpha\ell} p_\alpha p_\ell \right]$$

22. La notation standard en statistique serait $p_{\alpha/\ell}$, cf la notion de probabilités composées.

$$\frac{dp_\ell}{dt} = \sum_m \left[\left(\sum_\alpha \sum_\beta W_{\beta m, \alpha \ell} p_\beta \right) p_m - \left(\sum_\alpha \sum_\beta W_{\beta m, \alpha \ell} p_\alpha \right) p_\ell \right]$$

soit en notant $\tilde{W}_{m\ell} = \sum_\alpha \sum_\beta W_{\beta m, \alpha \ell} p_\beta$ et $\tilde{W}_{\ell m} = \sum_\alpha \sum_\beta W_{\beta m, \alpha \ell} p_\alpha$, la formulation condensée suivante :

$$\frac{dp_\ell}{dt} = \sum_m \left(\tilde{W}_{m\ell} p_m - \tilde{W}_{\ell m} p_\ell \right)$$

qui ressemble à une formule analogue du modèle micro-canonique (dont, du reste, on est parti dans cette étude), à la différence colossale près que si²³ les coefficients W sont symétriques dans le modèle micro-canonique, les \tilde{W} , à la simple vue de leur définition, ne le sont pas.

• Application à l'équilibre.

Tous les micro-états du thermostat notés α (ou β) correspondent au micro-état ℓ (ou m) du système d'énergie U_ℓ (ou U_m)

Si le système est suffisamment petit pour ne pas perturber le thermostat, celui-ci sera en équilibre micro-canonique et la probabilité qu'il soit dans un des $\Omega(U_{\text{tot.}} - U_\ell)$ états d'énergie $U_{\text{tot.}} - U_\ell$ est $1/\Omega$, lié à l'entropie partielle par $S(U_{\text{tot.}} - U_\ell) = k \ln \Omega$. Or, un développement limité conduit à :

$$S(U_{\text{tot.}} - U_\ell) = S(U_{\text{tot.}}) - U_\ell \frac{dS}{dU} = S(U_{\text{tot.}}) - \frac{U_\ell}{T}$$

par définition de la température thermodynamique du thermostat. On en déduit que :

$$\Omega(U_{\text{tot.}} - U_\ell) = Cte \cdot e^{-\frac{U_\ell}{kT}}$$

donc que :

$$p_\alpha = \frac{1}{\Omega(U_{\text{tot.}} - U_\ell)} = Cte \cdot e^{\frac{U_\ell}{kT}}$$

et de même

$$\Omega(U_{\text{tot.}} - U_m) = Cte \cdot e^{-\frac{U_m}{kT}}$$

donc que :

$$p_\beta = \frac{1}{\Omega(U_{\text{tot.}} - U_m)} = Cte \cdot e^{\frac{U_m}{kT}}$$

Il en résulte que :

$$\tilde{W}_{m\ell} = \sum_\alpha \sum_\beta W_{\beta m, \alpha \ell} Cte \cdot e^{\frac{U_m}{kT}}$$

23. Attention, c'est un « si concessif ».

et

$$\tilde{W}_{\ell m} = \sum_{\alpha} \sum_{\beta} W_{\beta m, \alpha \ell} Cte \cdot e^{\frac{U_{\ell}}{kT}}$$

Par contre, on a vu plus haut (dans le paragraphe 3.a p. 21) que p_{ℓ} est proportionnel à $\Omega(U_{\text{tot.}} - U_{\ell})$ donc à $e^{-\frac{U_{\ell}}{kT}}$ et p_m à $e^{-\frac{U_m}{kT}}$. On a donc

$$\forall m \quad \tilde{W}_{m\ell} p_m = \tilde{W}_{\ell m} p_{\ell}$$

qui assure que $\frac{dp_{\ell}}{dt}$ soit nulle et p_{ℓ} constant ; il s'agit de la version canonique du bilan détaillé.

• Retour à l'équilibre.

Imaginons maintenant que le système soit hors d'équilibre, on ne peut donc plus affirmer que p_{ℓ} soit proportionnel à $e^{-\frac{U_{\ell}}{kT}}$ ni p_m à $e^{-\frac{U_m}{kT}}$.

L'étude menant à la formule :

$$\frac{dp_{\ell}}{dt} = \sum_m \left(\tilde{W}_{m\ell} p_m - \tilde{W}_{\ell m} p_{\ell} \right)$$

ne nécessite pas que le système soit à l'équilibre ; elle reste valable.

Si le système est suffisamment petit pour ne pas perturber le thermostat, celui-ci sera en équilibre micro-canonique, le calcul des p_{α} et p_{β} reste valable et l'on a donc encore :

$$\tilde{W}_{m\ell} = \sum_{\alpha} \sum_{\beta} W_{\beta m, \alpha \ell} Cte \cdot e^{\frac{U_m}{kT}}$$

$$\tilde{W}_{\ell m} = \sum_{\alpha} \sum_{\beta} W_{\beta m, \alpha \ell} Cte \cdot e^{\frac{U_{\ell}}{kT}}$$

que l'on peut noter $\tilde{W}_{m\ell} = w_{m\ell} e^{\frac{U_m}{kT}}$ et $\tilde{W}_{\ell m} = w_{\ell m} e^{\frac{U_{\ell}}{kT}}$ avec $w_{m\ell} = w_{\ell m}$

On ne peut pas raisonner en thermodynamique statistique sans avoir en tête les résultats de thermodynamique classique. Or dans le contexte du modèle canonique (température imposée par le thermostat et volume donné), l'approche classique démontre que l'énergie libre F ne peut que diminuer. C'est ce que nous nous proposons donc de démontrer.

Reprenons ici les notations de moyennes ; nous avons donc, par définition de $\langle U \rangle$ et $\langle S \rangle$

$$\langle F \rangle = \langle U \rangle - T \langle S \rangle = \sum_{\ell} p_{\ell} U_{\ell} - T \left(-k \sum_{\ell} p_{\ell} \ln p_{\ell} \right) = \sum_{\ell} p_{\ell} (U_{\ell} + kT \ln p_{\ell})$$

d'où, en dérivant par rapport au temps et en s'imprégnant du fait que les U_ℓ sont des constantes, en tant qu'énergies de micro-états dûment identifiés :

$$\frac{d\langle F \rangle}{dt} = \sum_\ell \frac{dp_\ell}{dt} (U_\ell + kT \ln p_\ell) + \sum_\ell p_\ell \left(kT \frac{\frac{dp_\ell}{dt}}{p_\ell} \right) = \sum_\ell \frac{dp_\ell}{dt} (U_\ell + kT \ln p_\ell + kT)$$

reportons-y :

$$\frac{dp_\ell}{dt} = \sum_m \left(\tilde{W}_{m\ell} p_m - \tilde{W}_{\ell m} p_\ell \right) = \sum_m \left(w_{m\ell} e^{\frac{U_m}{kT}} p_m - w_{\ell m} e^{\frac{U_\ell}{kT}} p_\ell \right)$$

on arrive à :

$$\frac{dp_\ell}{dt} = \sum_\ell \sum_m w_{m\ell} p_m e^{\frac{U_m}{kT}} (U_\ell + kT \ln p_\ell + kT) - \sum_\ell \sum_m w_{\ell m} p_\ell e^{\frac{U_\ell}{kT}} (U_\ell + kT \ln p_\ell + kT)$$

Tout le monde suit ? Jusqu'ici, l'écriture est complexe, mais les calculs sont de routine. On s'accroche un peu, ça va devenir subtil !

Posons $q_\ell = p_\ell e^{\frac{U_\ell}{kT}}$ d'où $\ln q_\ell = \ln p_\ell + \frac{U_\ell}{kT}$ et, de la même façon $q_m = p_m e^{\frac{U_m}{kT}}$ d'où $\ln q_m = \ln p_m + \frac{U_m}{kT}$; on peut donc ré-écrire :

$$\frac{dp_\ell}{dt} = \sum_\ell \sum_m w_{m\ell} q_m kT (1 + \ln q_\ell) - \sum_\ell \sum_m w_{\ell m} q_\ell kT (1 + \ln q_\ell)$$

L'expression obtenue est plus lisible. Utilisons une seconde astuce : dans la seconde somme double, permutons la notation de indices. On arrive à :

$$\frac{dp_\ell}{dt} = \sum_\ell \sum_m w_{m\ell} q_m kT (1 + \ln q_\ell) - \sum_\ell \sum_m w_{m\ell} q_m kT (1 + \ln q_m)$$

$$\frac{dp_\ell}{dt} = \sum_\ell \sum_m w_{m\ell} q_m kT (\ln q_\ell - \ln q_m)$$

Utilisons une seconde fois la seconde astuce, en remontant au point de départ de sa première utilisation, mais cette fois en permutant les indices dans la première somme double. On arrive cette fois à :

$$\frac{dp_\ell}{dt} = \sum_\ell \sum_m w_{\ell m} q_\ell kT (\ln q_m - \ln q_\ell)$$

On s'accroche, c'est presque fini !

On utilise enfin, grâce à une troisième astuce, la symétrie $w_{m\ell} = w_{\ell m}$ en effectuant la demi-somme de deux expressions de $\frac{dp_\ell}{dt}$, soit :

$$\frac{dp_\ell}{dt} = \frac{1}{2} \sum_\ell \sum_m w_{m\ell} q_m kT (\ln q_\ell - \ln q_m) + \frac{1}{2} \sum_\ell \sum_m w_{\ell m} q_\ell kT (\ln q_m - \ln q_\ell)$$

$$\frac{dp_\ell}{dt} = -\frac{kT}{2} \sum_\ell \sum_m w_{m\ell} q_m (\ln q_m - \ln q_\ell) + \frac{kT}{2} \sum_\ell \sum_m w_{m\ell} q_\ell (\ln q_m - \ln q_\ell)$$

$$\frac{dp_\ell}{dt} = -\frac{kT}{2} \sum_\ell \sum_m w_{m\ell} (q_m - q_\ell) (\ln q_m - \ln q_\ell)$$

Il ne reste plus qu'à utiliser, comme pour l'évolution dans le modèle micro-canonique, le fait que la fonction logarithme est croissante, donc que $\ln q_m - \ln q_\ell$ est du même signe que $q_m - q_\ell$, pour affirmer que $\langle F \rangle$ ne peut que diminuer.

Ici aussi, le principe du bilan détaillé, adapté au contexte canonique, permet d'affiner la compréhension du retour à l'équilibre.