
PHYSIQUE STATISTIQUE ET GAZ PARFAITS

Patrice BACHER

Table des matières

1	La thermodynamique statistique	1
1.1	Thermodynamique des équilibres	1
1.1.1	Postulats, principes et équations de base	1
1.1.2	Les potentiels thermodynamiques	5
1.1.3	Relations entre variables d'état	10
1.2	Probabilités et statistiques	12
1.2.1	Probabilités, moyennes, distributions	12
1.2.2	La distribution binomiale	13
1.2.3	La distribution multinomiale	20
1.3	Théorie statistique de Gibbs	21
1.3.1	Du macroscopique au microscopique	21
1.3.2	Ensemble statistique, énoncé des postulats	23
1.3.3	Probabilités et valeurs moyennes	24
1.3.4	Confrontation avec la thermodynamique	27
1.4	Les ensembles statistiques	31
1.4.1	Introduction et conventions d'écriture	31
1.4.2	L'ensemble grand canonique	32
1.4.3	L'ensemble T-p	33
1.4.4	L'ensemble canonique	35
1.4.5	L'ensemble microcanonique	36
1.4.6	Les autres ensembles en variance 3	39
1.4.7	Synthèse des résultats	40
1.4.8	Ensembles en variance supérieure à 3	41
1.5	Le deuxième principe et le troisième principe	43
1.5.1	Équivalence entre ensembles statistiques	43
1.5.2	Le deuxième principe	45
1.5.3	Équivalence entre S maxi et E, F, G, H mini	46
1.5.4	Notion de fluctuations aléatoires	47
1.5.5	Le problème ergodique	49
1.5.6	Le troisième principe	52
2	Statistiques quantiques	53
2.1	Les statistiques de réseaux	53
2.1.1	Réseaux de sites tous identiques	53
2.1.2	Réseaux à différents types de sites	55
2.1.3	Recherche de la fonction adaptée au problème	56

2.1.4	Emploi d'autres fonctions de partition	57
2.2	Particules identiques et indépendantes	59
2.2.1	Si les particules quantiques étaient discernables	59
2.2.2	... mais les particules quantiques sont indiscernables	62
2.3	Systèmes de particules indiscernables	66
2.3.1	Limite des hautes températures	66
2.3.2	Fonction de distribution à température quelconque	67
2.3.3	Distributions de Fermi-Dirac et de Bose-Einstein	69
2.4	Le corps noir et le gaz de photons	70
2.4.1	Théorie électromagnétique	70
2.4.2	Les photons	72
2.4.3	Distribution de Planck, loi de Planck	72
2.4.4	Thermodynamique du rayonnement thermique	75
2.5	L'oscillateur harmonique et le gaz de phonons	77
2.5.1	L'oscillateur harmonique linéaire quantique	77
2.5.2	Vibrations de réseau : modèle d'Einstein	79
2.5.3	Vibrations de réseau : modèle de Debye	81
3	Gaz parfait de masses ponctuelles	91
3.1	Quelques ordres de grandeur	91
3.1.1	Les grandeurs énergétiques	91
3.1.2	Densités et pressions	92
3.2	Particule libre dans une boîte cubique	93
3.2.1	Résultats de la mécanique quantique	93
3.2.2	L'approximation du continuum et la densité d'états	93
3.2.3	Fonctions thermodynamiques	95
3.3	Gaz parfait classique de masses ponctuelles	95
3.3.1	Limite hautes températures de la théorie quantique	95
3.3.2	Fonction de partition canonique, paradoxe de Gibbs	97
3.4	Éléments de théorie cinétique des gaz parfaits	99
3.4.1	Distribution de Maxwell-Boltzmann (des énergies)	99
3.4.2	Distribution de Maxwell (des vitesses)	100
3.4.3	Calcul de la pression cinétique	102
3.5	Le mouvement brownien	103
3.5.1	Historique et hypothèses	103
3.5.2	Diffusion sur un réseau unidimensionnel	105
3.5.3	Diffusion sur un réseau tridimensionnel	108
3.5.4	Mouvement de particules dans un fluide	113
3.6	Éléments de mécanique statistique classique	116
3.6.1	Systèmes à une particule	116
3.6.2	Systèmes à N particules	119
3.6.3	Distributions de Maxwell et de Maxwell-Boltzmann	121
4	Gaz parfait d'atomes ou molécules	123
4.1	Les gaz parfaits classiques d'atomes	123
4.1.1	Degrés de liberté électroniques et nucléaires	123
4.1.2	Conséquences sur les propriétés thermodynamiques	124

4.1.3	Gaz adsorbé sur une surface	125
4.2	Les degrés de liberté d'une molécule	129
4.2.1	Indépendance des degrés de liberté de translation	129
4.2.2	Séparation des variables par approximations	130
4.3	Les gaz parfaits de molécules diatomiques	132
4.3.1	Réduction du problème à deux corps	132
4.3.2	Mouvements de vibration	133
4.3.3	Mouvements de rotation	135
4.3.4	Propriétés thermodynamiques	136
4.3.5	Molécules homonucléaires : effets quantiques	137
4.4	Les gaz parfaits de molécules polyatomiques	140
4.4.1	Identification des degrés de liberté	140
4.4.2	Degrés de liberté de rotation	141
4.4.3	Degrés de liberté de vibration	141
4.4.4	Propriétés thermodynamiques	142
4.5	Équilibres chimiques entre gaz parfaits	143
4.5.1	Thermodynamique des équilibres chimiques	143
4.5.2	Statistique des équilibres chimiques entre gaz parfaits	145
4.5.3	Exemples d'équilibres chimiques	146
4.6	Réaction chimique dans un gaz parfait	149
4.6.1	Taux d'avancement d'une réaction chimique	149
4.6.2	Surfaces équipotentielles	151
4.6.3	Théorie d'Eyring	152
4.6.4	Application à la diffusion de surface	156
4.7	Gaz parfait diélectrique sous champ électrique	157
4.7.1	Conditions d'équilibre sous champ de force extérieur	157
4.7.2	Thermodynamique des milieux diélectriques	159
4.7.3	Molécules dipolaires dans un champ électrique	161
5	Les gaz parfaits quantiques	165
5.1	Les différents cas pratiques	165
5.2	L'hélium à haute température	166
5.2.1	Calcul du potentiel chimique	166
5.2.2	Les autres grandeurs thermodynamiques	167
5.3	L'hélium 3 à basse température et l'électron	169
5.3.1	L'état fondamental	169
5.3.2	Le calcul de Sommerfeld	169
5.3.3	Conditions de validité	172
5.3.4	L'effet Pomerantchuk	173
5.3.5	Le réfrigérateur à dilution d'hélium 3 dans l'hélium 4	176
5.4	L'hélium 4 à basse température	177
5.4.1	Température de Bose, volume atomique de Bose	177
5.4.2	Températures légèrement supérieures à T_B	178
5.4.3	Températures inférieures à T_B	179
5.4.4	La courbe $C_V(T)$ et les diagrammes $p-v$ et $p-T$	182
5.4.5	L'hélium 4 et la superfluidité	184

Chapitre 1

La thermodynamique statistique

1.1 Thermodynamique des équilibres

1.1.1 Postulats, principes et équations de base

La thermodynamique est la science des bilans énergétiques (premier principe) et de la prévision des évolutions des systèmes lors de la modification de leurs paramètres externes ou internes (deuxième principe). Elle présente un caractère universel et s'applique à tous les domaines de la physico-chimie.

Ce résumé concerne les *états d'équilibre*. Les flux envisagés sont des flux virtuels, infinitésimaux autour de l'état d'équilibre. On les envisage *par la pensée* dans le seul but d'étudier la stabilité de cet équilibre. Aucun transport n'est envisagé.

Le système étudié est vu à *l'échelle macroscopique*. Le deuxième principe ne rend pas compte des fluctuations locales ou temporelles à l'échelle des particules. On ne peut pas l'appliquer à un système où le nombre de particules est trop faible pour lui appliquer un traitement statistique.

Un système thermodynamique est une partie purement arbitraire de l'univers, isolée par la pensée du *milieu extérieur*. Son **contour** peut être physique (parois d'un récipient), ou tout à fait immatériel (élément de volume isolé par la pensée du reste du récipient), ou même impossible à matérialiser (ensemble de molécules de même nature disséminées dans un mélange).

Postulat des états d'équilibre. On admet qu'il existe un état d'équilibre macroscopique, indépendant de la préparation du système, et défini par la connaissance d'un nombre v , appelé *variance*, de données macroscopiques. Ce sont les *variables d'état* du système. On peut toujours trouver v variables d'état *extensives*, c'est-à-dire *additives* et à valeurs *positives*. Une grandeur est additive si sa valeur dans le système $\Sigma' \cup \Sigma''$ est égale à la somme de ses valeurs dans les systèmes Σ' et Σ'' , à condition que Σ' et Σ'' soient *disjoints* et *initialement en équilibre mutuel*.

Échanges avec le milieu extérieur. Les variables extensives peuvent traverser le contour du système dans un sens ou dans l'autre. C'est ce qu'on appelle un échange. Tout flux est compté positivement du milieu extérieur vers le système, et négativement dans l'autre sens. Pour l'étude d'un équilibre, on n'envisage que des *échanges virtuels*, compatibles avec la nature du contour. C'est uniquement par la pensée, dans le seul but d'étudier la stabilité de l'équilibre.

Échanges internes, contraintes internes. On peut envisager la circulation des grandeurs extensives entre deux parties Σ' et Σ'' d'un système Σ . Si une "cloison" s'oppose à cette circulation, on dit qu'il y a *contrainte interne*. L'équilibre observé est alors un *équilibre sous contraintes internes*. Ce n'est pas réellement l'équilibre puisque la suppression de la cloison, ou *inhibition de la contrainte*, déplacera l'équilibre. L'équilibre proprement dit correspond à l'inhibition de toutes les contraintes internes. Lorsqu'on étudie un équilibre sous contraintes, les flux envisagés sont virtuels et compatibles avec les cloisons ou le contour.

Un équilibre (sous contraintes internes ou non) est un état où toutes les variables extensives sont constantes et où tous les flux internes ou externes sont nuls.

L'énergie. On ne peut avoir qu'une compréhension intuitive de la variable *Énergie* : son existence est un axiome. C'est une grandeur extensive, c'est-à-dire additive et toujours positive. Elle n'est connue qu'à une constante près, d'où la nécessité de choisir à notre convenance une *énergie de point zéro*.

Le premier principe postule que *l'énergie se conserve globalement* passant d'une forme à une autre sans jamais se créer ni disparaître. Si le système gagne de l'énergie, le milieu extérieur en perd : il y a flux d'énergie à travers le contour.

L'énergie interne est par définition l'énergie totale du système. On la note en général E ou U . Dans ce document nous utiliserons la lettre E comme en mécanique.

L'entropie. Comme pour l'énergie, l'existence de la variable *Entropie* S est un axiome. C'est une grandeur extensive définie exclusivement sur les états d'équilibre, indéterminée lors de l'évolution du système.

Le deuxième principe postule que :

- S est *additive*, fonction *positive* et *sans singularités* des variables extensives ;
- S est une fonction *croissante* de E ;
- S *est maximale vis-à-vis des échanges internes, virtuels, infinitésimaux, compatibles avec les cloisons, de toutes les variables extensives y compris E* ; autrement dit, *dans un système isolé*, la suppression d'une contrainte interne entraîne un déplacement *spontané* de l'équilibre, obligatoirement accompagné d'un *accroissement d'entropie*.

Le troisième principe, appelé aussi principe de Nernst, fixe l'origine de l'échelle entropique et permet d'établir le lien avec la thermodynamique statistique :

$$S \longrightarrow 0 \quad \text{quand} \quad \frac{\partial E}{\partial S} \longrightarrow 0$$

Les grandeurs extensives de base pour exprimer l'entropie sont :

- l'énergie interne E ;
- le volume V ;
- le nombre n_i de moles ou N_i de molécules de chacun des p constituants ;
- éventuellement d'autres grandeurs additives X_j ($j = 1$ à q), à condition que leur contribution à l'énergie interne soit déterminée sans ambiguïté par la donnée de l'ensemble des variables.

Terminologie. Un système est dit :

- *isolé* lorsque son contour est imperméable à toutes les variables extensives ;
- *fermé (ouvert)* si son contour est étanche (perméable) à la matière ;
- *adiabatique (diathermique)* si son contour est étanche (perméable) à la chaleur ;
- *isochore (expansible)* si l'on interdit (permet) les variations de volume.

Équation fondamentale. L'état thermodynamique du système est défini par v variables extensives indépendantes. Avec l'entropie, on en a maintenant $v + 1$. Elles sont donc liées par une relation appelée *équation fondamentale* du système^[1] :

$$S = S(E, V, X, N_1, N_2) \quad \text{ou} \quad E = E(S, V, X, N_1, N_2)$$

L'équation fondamentale contient toutes les informations nécessaires à la description thermodynamique du système. Toutes les grandeurs qui y figurent sont additives. Elles varient toutes proportionnellement à la "quantité" de système. C'est ce qui caractérise une *forme homogène de degré 1* ^[2] :

$$\begin{aligned} S(\lambda E, \lambda V, \lambda X, \lambda N_1, \lambda N_2) &= \lambda S(E, V, X, N_1, N_2) \\ E(\lambda S, \lambda V, \lambda X, \lambda N_1, \lambda N_2) &= \lambda E(S, V, X, N_1, N_2) \end{aligned}$$

Variables intensives : S et E étant des formes homogènes de degré 1, leurs dérivées partielles sont des formes homogènes de degré 0, c'est-à-dire indépendantes de la "quantité" de système et, à l'équilibre, *homogènes dans tout le système*. On les appelle *variables intensives* du système :

$$\begin{aligned} dS &= \frac{\partial S}{\partial E} dE + \frac{\partial S}{\partial V} dV + \frac{\partial S}{\partial X} dX + \frac{\partial S}{\partial N_1} dN_1 + \frac{\partial S}{\partial N_2} dN_2 \\ dE &= \frac{\partial E}{\partial S} dS + \frac{\partial E}{\partial V} dV + \frac{\partial E}{\partial X} dX + \frac{\partial E}{\partial N_1} dN_1 + \frac{\partial E}{\partial N_2} dN_2 \end{aligned}$$

$$\text{avec par exemple : } \frac{\partial S}{\partial E}(\lambda E, \lambda V, \lambda X, \lambda N_1, \lambda N_2) = \frac{\partial S}{\partial E}(E, V, X, N_1, N_2)$$

Les dérivées de S sont les *variables intensives entropiques*, celles de E sont les *variables intensives énergétiques*. Ces dernières portent un nom :

- la *température* T est la variable intensive énergétique associée à S ;
- la *pression* p est au signe près la variable intensive énergétique associée à V ;
- le *potentiel chimique* μ_i du constituant i est la variable intensive énergétique associée à N_i ^[3] ;
- la variable intensive énergétique associée à X sera notée P .

$$\text{Autrement dit : } T = \left. \frac{\partial E}{\partial S} \right]_{V, X, N_1, N_2} \quad p = - \left. \frac{\partial E}{\partial V} \right]_{S, X, N_1, N_2}$$

$$P = \left. \frac{\partial E}{\partial X} \right]_{S, V, N_1, N_2} \quad \mu_1 = \left. \frac{\partial E}{\partial N_1} \right]_{S, V, X, N_2} \quad \mu_2 = \left. \frac{\partial E}{\partial N_2} \right]_{S, V, X, N_1}$$

L'équation de Gibbs est l'écriture de dE utilisant ces définitions :

$$dE = T dS - p dV + P dX + \mu_1 dN_1 + \mu_2 dN_2 \quad (1.1)$$

¹ pour alléger les écritures, on travaillera avec $p = 2$ (systèmes biconstituants) et $q = 1$ (une seule variable extensive supplémentaire X). Si $q = 0$, on ignorera les termes contenant P et X .

² une forme homogène de degré n est telle que $f(\lambda x, \lambda y, \lambda z, \dots) = \lambda^n f(x, y, z, \dots)$

³ en thermodynamique statistique, il est préférable de travailler avec les nombres de molécules.

La variation d'énergie interne peut être un apport de *chaleur* ($T dS$) ou de *travail* ($-p dV + P dX$), ou une *modification de l'équilibre chimique* ($\mu_1 dN_1 + \mu_2 dN_2$).

Forces motrices et flux. La variable intensive énergétique P_i associée à une variable extensive X_i est une force motrice pour le flux virtuel δX_i . Ainsi, un écart de température de part et d'autre d'une paroi adiabatique provoque un flux d'entropie (ou de chaleur), un écart de pression de part et d'autre d'une paroi mobile provoque le déplacement de celle-ci, un écart de potentiel chimique de part et d'autre d'une paroi perméable provoque un flux de matière, etc.

À l'équilibre, on peut montrer que le *flux virtuel* de X_i (à travers une paroi ou le contour), s'il est seul à être envisagé, se fait *dans le sens où sa variable associée P_i décroît*. En particulier, le flux de chaleur se fait dans le sens des températures décroissantes, le flux de volume dans le sens des pressions croissantes, et le flux de matière dans le sens des potentiels chimiques décroissants.

Théorème de l'énergie interne minimale. L'équation de Gibbs s'écrit aussi :

$$dS = \frac{1}{T} dE + \frac{p}{T} dV - \frac{P}{T} dX - \frac{\mu_1}{T} dN_1 - \frac{\mu_2}{T} dN_2 \quad (1.2)$$

Les variables intensives entropiques associées à E , V , X , N_1 et N_2 sont donc $1/T$, p/T , $-P/T$, $-\mu_1/T$ et $-\mu_2/T$. Le fait que la variable intensive entropique $1/T$ associée à E soit l'inverse de la variable intensive énergétique T associée à S permet de remplacer le postulat de l'entropie maximale par le *théorème de l'énergie interne minimale* : **à l'équilibre, l'énergie interne est minimale vis-à-vis des échanges internes, virtuels, infinitésimaux, compatibles avec les contraintes internes, de toutes les variables extensives, y compris S** :

$$\text{Deuxième principe : } (dS = 0 \text{ et } d^2S < 0) \quad \text{ou} \quad (dE = 0 \text{ et } d^2E > 0)$$

Autrement dit, *dans un système isolé*, la suppression d'une contrainte interne entraîne un déplacement *spontané* de l'équilibre, avec *diminution* obligatoire de l'énergie interne.

Théorème de Gibbs-Duhem. Comme $E(S, V, X, N_1, N_2)$ est une forme homogène de degré 1, on peut lui appliquer l'équation d'Euler :

$$E \equiv S \frac{\partial E}{\partial S} + V \frac{\partial E}{\partial V} + X \frac{\partial E}{\partial X} + N_1 \frac{\partial E}{\partial N_1} + N_2 \frac{\partial E}{\partial N_2} \equiv TS - pV + PX + \mu_1 N_1 + \mu_2 N_2$$

Cette équation se différencie comme suit :

$$dE = TdS - pdV + PdX + \mu_1 dN_1 + \mu_2 dN_2 + SdT - Vdp + XdP + N_1 d\mu_1 + N_2 d\mu_2$$

On en déduit, en utilisant l'équation de Gibbs, le théorème de Gibbs-Duhem :

$$-S dT + V dp - X dP - N_1 d\mu_1 - N_2 d\mu_2 = 0 \quad (1.3)$$

Équations d'état. On a maintenant $2v + 1$ variables, dont v sont indépendantes. On peut choisir ces dernières arbitrairement et exprimer les $v + 1$ autres par $v + 1$ équations appelées équations d'état. La connaissance de $v + 1$ équations d'état indépendantes décrit entièrement le système. Elle remplace sans équivoque la donnée de l'équation fondamentale. La plus connue est l'équation des gaz parfaits monoconstituants $pV = nRT$ ou $pV = NkT$.

1.1.2 Les potentiels thermodynamiques

Soit $f : x \mapsto f(x)$ une fonction continue et dérivable en tout point. Son graphe peut être considéré comme l'enveloppe de ses tangentes :

$$f(x) = p(x)x + y_0(x) \quad \text{avec} \quad p(x) = \frac{df}{dx}$$

La pente p est fonction de x . Si cette fonction $p = p(x)$ est inversible, c'est-à-dire si on peut l'écrire sous la forme $x = x(p)$, c'est-à-dire encore si la courbe $y = f(x)$ est concave ou convexe (pas de point d'inflexion), on peut exprimer y_0 (c'est l'ordonnée à l'origine de la tangente en x) en fonction de p .

La transformée de Legendre de f est la fonction g qui à p associe $-y_0(p)$:

$$g : p \mapsto g(p) = px - f(x) \quad \text{avec} \quad p = \frac{df}{dx} \quad (x \text{ et } f(x) \text{ exprimés en fonction de } p)$$

Deux propriétés importantes :

- La transformée de Legendre d'une fonction convexe est une fonction convexe ;
- La transformée de Legendre est involutive : si on transforme g , on retrouve f :

$$f : x \mapsto f(x) = xp - g(p) \quad \text{avec} \quad x = \frac{dg}{dp} \quad (p \text{ et } g(p) \text{ exprimés en fonction de } x)$$

Attention : la transformée de Legendre n'est pas une transformation linéaire :

$$\text{TdL}(f_1 + f_2) \neq \text{TdL}(f_1) + \text{TdL}(f_2) \quad \text{TdL}(\lambda f) \neq \lambda \text{TdL}(f)$$

Si f n'est ni concave ni convexe (s'il y a des points d'inflexion), on définit une transformée de Legendre de f sur chacun des intervalles où f est concave ou convexe. Moyennant quoi, les seules fonctions ne possédant pas de transformée de Legendre sont les fonctions linéaires $x \mapsto ax + b$.

Pour une fonction à plusieurs variables, la transformée de Legendre peut s'appliquer à un nombre quelconque de variables. Par exemple la transformée de Legendre de $f(x_1, x_2, x_3, x_4)$ pour x_1 et x_2 est :

$$g_{12}(p_1, p_2, x_3, x_4) = p_1x_1 + p_2x_2 - f(x_1, x_2, x_3, x_4) \quad \text{avec} \quad p_1 = \frac{\partial f}{\partial x_1} \quad \text{et} \quad p_2 = \frac{\partial f}{\partial x_2}$$

$$\text{Pour les variables non transformées, on a : } \frac{\partial g}{\partial x_3} = -\frac{\partial f}{\partial x_3} \quad \text{et} \quad \frac{\partial g}{\partial x_4} = -\frac{\partial f}{\partial x_4}.$$

Attention aux changements de signes lors de la composition de deux transformées de Legendre. Soit par exemple une fonction $f(x_1, x_2)$, et soient T_1 , T_2 et T_{12} les transformations donnant respectivement $g_1(p_1, x_2)$, $g_2(x_1, p_2)$ et $g_{12}(p_1, p_2)$. On a :

$$T_{12}(f) = T_1(-T_2(f)) = T_2(-T_1(f)) \quad (1.4)$$

Les potentiels thermodynamiques sont au signe près les transformées de Legendre de E : avec la fonction de départ, il y en a 2^v . La transformation de Legendre étant involutive, n'importe quel potentiel thermodynamique contient autant d'informations que l'équation fondamentale $E = E(S, V, N)$. Les écritures qui suivent concernent les systèmes monoconstituants et ne faisant pas intervenir de variable X . F est l'énergie libre, H l'enthalpie, Ψ le grand potentiel, G l'enthalpie libre, et Θ le potentiel nul.

Équations initiales :

$$E(S, V, N) \equiv TS - pV + \mu N \quad dE = T dS - p dV + \mu dN$$

$$T = \left. \frac{\partial E}{\partial S} \right]_{V, N} \quad p = - \left. \frac{\partial E}{\partial V} \right]_{S, N} \quad \mu = \left. \frac{\partial E}{\partial N} \right]_{S, V} \quad (1.5)$$

$(-F)$ est la transformée de Legendre de E pour la variable S :

$$F(T, V, N) = E - TS \quad dF = -S dT - p dV + \mu dN$$

$$S = - \left. \frac{\partial F}{\partial T} \right]_{V, N} \quad p = - \left. \frac{\partial F}{\partial V} \right]_{T, N} \quad \mu = \left. \frac{\partial F}{\partial N} \right]_{T, V} \quad (1.6)$$

$(-H)$ est la transformée de Legendre de E pour la variable V :

$$H(S, p, N) = E + pV \quad dH = T dS + V dp + \mu dN$$

$$T = \left. \frac{\partial H}{\partial S} \right]_{p, N} \quad V = \left. \frac{\partial H}{\partial p} \right]_{S, N} \quad \mu = \left. \frac{\partial H}{\partial N} \right]_{S, p} \quad (1.7)$$

On appellera $(-\Phi)$ la transformée de Legendre de E pour la variable N :

$$\Phi(S, V, \mu) = E - \mu N \quad d\Phi = T dS - p dV - N d\mu$$

$$T = \left. \frac{\partial \Phi}{\partial S} \right]_{V, \mu} \quad p = - \left. \frac{\partial \Phi}{\partial V} \right]_{S, \mu} \quad N = - \left. \frac{\partial \Phi}{\partial \mu} \right]_{S, V} \quad (1.8)$$

$(-\Psi)$ est la transformée de Legendre de E pour les variables S et N :

$$\Psi(T, V, \mu) = E - TS - \mu N \quad d\Psi = -S dT - p dV - N d\mu$$

$$S = - \left. \frac{\partial \Psi}{\partial T} \right]_{V, \mu} \quad p = - \left. \frac{\partial \Psi}{\partial V} \right]_{T, \mu} \quad N = - \left. \frac{\partial \Psi}{\partial \mu} \right]_{T, V} \quad (1.9)$$

$(-G)$ est la transformée de Legendre de E pour les variables S et V :

$$G(T, p, N) = E - TS + pV \quad dG = -S dT + V dp + \mu dN$$

$$S = - \left. \frac{\partial G}{\partial T} \right]_{p, N} \quad V = \left. \frac{\partial G}{\partial p} \right]_{T, N} \quad \mu = \left. \frac{\partial G}{\partial N} \right]_{T, p} \quad (1.10)$$

On appellera $(-\Phi')$ la transformée de Legendre de E pour les variables V et N :

$$\begin{aligned}\Phi'(S, p, \mu) &= E + pV - \mu N & d\Phi' &= T dS + V dp - N d\mu \\ T &= \left. \frac{\partial \Phi'}{\partial S} \right]_{p, \mu} & V &= \left. \frac{\partial \Phi'}{\partial p} \right]_{S, \mu} & N &= - \left. \frac{\partial \Phi'}{\partial \mu} \right]_{S, p}\end{aligned}\quad (1.11)$$

$(-\Theta)$ est la transformée de Legendre de E pour toutes les variables :^[4]

$$\begin{aligned}\Theta(M, T, p, \mu) &= E - TS + pV - \mu N & d\Theta &= -S dT + V dp - N d\mu \\ S &= - \left. \frac{\partial \Theta}{\partial T} \right]_{p, \mu} & V &= \left. \frac{\partial \Theta}{\partial p} \right]_{T, \mu} & N &= - \left. \frac{\partial \Theta}{\partial \mu} \right]_{T, p}\end{aligned}\quad (1.12)$$

Attention au potentiel nul Θ : il est nul à l'équilibre (équation d'Euler), sa différentielle également (théorème de Gibbs-Duhem), mais c'est une fonction de (T, p, μ) non identiquement nulle au voisinage de l'équilibre.

Les transformées de Legendre de l'entropie sont très importantes en physique statistique. Comme précédemment, les écritures qui suivent concernent les systèmes monoconstituants et ne faisant pas intervenir de variable X . On exprimera les transformées de Legendre de S/k (k = constante de Boltzmann $\simeq 1,38 \cdot 10^{-23}$ J/K), en utilisant les notations suivantes :

$$\beta = \frac{1}{kT} \quad \xi = \frac{p}{kT} \quad \gamma = -\frac{\mu}{kT}$$

Avec ces notations, l'équation d'Euler s'écrit :

$$\begin{aligned}S/k(E, V, N) &= \beta E + \xi V + \gamma N & dS/k &= \beta dE + \xi dV + \gamma dN \\ \beta &= \left. \frac{\partial S/k}{\partial E} \right]_{V, N} & \xi &= \left. \frac{\partial S/k}{\partial V} \right]_{E, N} & \gamma &= \left. \frac{\partial S/k}{\partial N} \right]_{E, V}\end{aligned}\quad (1.13)$$

$(+\beta F)$ est la Transformée de Legendre de S/k pour la variable E :

$$\begin{aligned}\beta F(\beta, V, N) &= \beta E - S/k & d\beta F &= E d\beta - \xi dV - \gamma dN \\ E &= \left. \frac{\partial \beta F}{\partial \beta} \right]_{V, N} & \xi &= - \left. \frac{\partial \beta F}{\partial V} \right]_{\beta, N} & \gamma &= - \left. \frac{\partial \beta F}{\partial N} \right]_{\beta, V}\end{aligned}\quad (1.14)$$

$(-\beta \Phi)$ est la Transformée de Legendre de S/k pour la variable V :

$$\begin{aligned}\beta \Phi(E, \xi, N) &= S/k - \xi V & d\beta \Phi &= \beta dE - V d\xi + \gamma dN \\ \beta &= \left. \frac{\partial \beta \Phi}{\partial E} \right]_{\xi, N} & V &= - \left. \frac{\partial \beta \Phi}{\partial \xi} \right]_{E, N} & \gamma &= \left. \frac{\partial \beta \Phi}{\partial N} \right]_{E, \xi}\end{aligned}\quad (1.15)$$

⁴ La seule connaissance des variables intensives ne donne aucune information sur la taille du système ; le potentiel nul doit donc pour être utilisable dépendre d'un certain facteur d'échelle M (exemple : statistiques de réseaux, paragraphe 2.1) ; nous en reparlerons au paragraphe 1.3.4.

$(-\beta H)$ est la Transformée de Legendre de S/k pour la variable N :

$$\begin{aligned} \beta H(E, V, \gamma) &= S/k - \gamma N & d\beta H &= \beta dE + \xi dV - N d\gamma \\ \beta &= \left. \frac{\partial \beta H}{\partial E} \right]_{V, \gamma} & \xi &= \left. \frac{\partial \beta H}{\partial V} \right]_{E, \gamma} & N &= - \left. \frac{\partial \beta H}{\partial \gamma} \right]_{E, V} \end{aligned} \quad (1.16)$$

$(+\beta \Psi)$ est la Transformée de Legendre de S/k pour les variables E et N :

$$\begin{aligned} \beta \Psi(\beta, V, \gamma) &= \beta E + \gamma N - S/k & d\beta \Psi &= E d\beta - \xi dV + N d\gamma \\ E &= \left. \frac{\partial \beta \Psi}{\partial \beta} \right]_{V, \gamma} & \xi &= - \left. \frac{\partial \beta \Psi}{\partial V} \right]_{\beta, \gamma} & N &= \left. \frac{\partial \beta \Psi}{\partial \gamma} \right]_{\beta, V} \end{aligned} \quad (1.17)$$

$(+\beta G)$ est la Transformée de Legendre de S/k pour les variables E et V :

$$\begin{aligned} \beta G(\beta, \xi, N) &= \beta E + \xi V - S/k & d\beta G &= E d\beta + V d\xi - \gamma dN \\ E &= \left. \frac{\partial \beta G}{\partial \beta} \right]_{\xi, N} & V &= \left. \frac{\partial \beta G}{\partial \xi} \right]_{\beta, N} & \gamma &= - \left. \frac{\partial \beta G}{\partial N} \right]_{\beta, \xi} \end{aligned} \quad (1.18)$$

$(-\beta E)$ est la Transformée de Legendre de S/k pour les variables V et N :

$$\begin{aligned} \beta E(E, \xi, \gamma) &= S/k - \xi V - \gamma N & d\beta E &= \beta dE - V d\xi - N d\gamma \\ \beta &= \left. \frac{\partial \beta E}{\partial E} \right]_{\xi, \gamma} & V &= - \left. \frac{\partial \beta E}{\partial \xi} \right]_{E, \gamma} & N &= - \left. \frac{\partial \beta E}{\partial \gamma} \right]_{E, \xi} \end{aligned} \quad (1.19)$$

$(+\beta \Theta)$ est la Transformée de Legendre de S/k pour toutes les variables :^[5]

$$\begin{aligned} \beta \Theta(M, \beta, \xi, \gamma) &= \beta E + \xi V + \gamma N - S/k & d\beta \Theta &= E d\beta + V d\xi + N d\gamma \\ E &= \left. \frac{\partial \beta \Theta}{\partial \beta} \right]_{\xi, \gamma} & V &= \left. \frac{\partial \beta \Theta}{\partial \xi} \right]_{\beta, \gamma} & N &= \left. \frac{\partial \beta \Theta}{\partial \gamma} \right]_{\beta, \xi} \end{aligned} \quad (1.20)$$

On notera la présence d'un signe $(-)$ lorsque la variable E n'est pas transformée $(-\beta \Phi, -\beta H, -\beta E)$, et un signe $(+)$ lorsque la variable E est transformée $(+\beta F, +\beta \Psi, +\beta G, +\beta \Theta)$.

$F = E - TS$ n'est pas une simple opération arithmétique. Elle sous-entend un changement de variable $S \rightarrow T$. On passe d'une différentielle totale exacte dE à une autre différentielle totale exacte dF . Si on écrit F uniquement en fonction de ses variables naturelles (T, V, N) , les autres variables (S, p, μ) sont accessibles par simple calcul de dérivées partielles. Ce sont ces équations d'état $S(T, V, N)$, $p(T, V, N)$, $\mu(T, V, N)$ qui sont exploitables expérimentalement. En clair, tout potentiel thermodynamique, exprimé en fonction de ses variables naturelles contient

⁵ Voir note page 7

autant d'information que l'équation fondamentale $E = E(S, V, N)$. Toute transformée de Legendre de l'entropie, exprimée en fonction de ses variables naturelles contient autant d'information que l'équation fondamentale $S = S(E, V, N)$. Mathématiquement, la raison en est que la transformation de Legendre est involutive : la transformée de Legendre inverse redonne sans ambiguïté la fonction de départ.

L'intérêt des potentiels thermodynamiques est de pouvoir exprimer de façon simple le deuxième principe dans des conditions particulières. On a vu qu'on pouvait, pour tout système isolé, remplacer le principe de l'entropie maximale par le théorème de l'énergie interne minimale. Lorsqu'il y a échange avec l'extérieur, l'utilisation de ce théorème est délicate car il ne peut s'appliquer qu'à l'univers entier. C'est donc l'énergie totale qu'il faut connaître, c'est-à-dire l'énergie interne du système, mais aussi celle du reste de l'univers. Les théorèmes qui suivent utilisent des grandeurs calculables sur le système. Ils concernent les systèmes sans contraintes internes où les seules variables extensives sont S , E , V , N . Ils se démontrent très facilement à partir du deuxième principe et des définitions des potentiels thermodynamiques.

Théorèmes du potentiel thermodynamique minimal :

- Un système adiabatique, isochore, fermé est à l'équilibre lorsque son *énergie interne* E est minimale.
- Un système **isotherme**, isochore, fermé est à l'équilibre lorsque son *énergie libre* F est minimale.
- Un système adiabatique, **isobare**, fermé est à l'équilibre lorsque son *enthalpie* H est minimale.
- Un système adiabatique, isochore, **ouvert**^[6] est à l'équilibre lorsque son potentiel Φ est minimal.
- Un système **isotherme**, isochore, **ouvert** est à l'équilibre lorsque son *grand potentiel* Ψ est minimal.
- Un système **isotherme**, **isobare**, fermé est à l'équilibre lorsque son *enthalpie libre* G est minimale.
- Un système adiabatique, **isobare**, **ouvert** est à l'équilibre lorsque son potentiel Φ' est minimal.
- Un système **isotherme**, **isobare**, **ouvert** est à l'équilibre lorsque son *potentiel nul* Θ est nul.

À noter que l'utilisation d'un potentiel thermodynamique n'est pas restreinte aux conditions particulières pour lesquelles le second principe s'écrit de façon simple. **Tous les potentiels sont équivalents thermodynamiquement.** On peut utiliser n'importe quel potentiel thermodynamique indépendamment de l'environnement du système. Le choix des variables indépendantes est arbitraire, non imposé par l'environnement du système. La raison en est qu'un système, même non isolé, possède à l'équilibre une énergie interne constante, qu'un système, même ouvert, possède à l'équilibre une composition constante, etc.

⁶ par "ouvert" on entend "en contact avec un réservoir de particules qui fixe le potentiel chimique" ; malheureusement, il n'existe pas de terme signifiant "iso-potentiel chimique".

1.1.3 Relations entre variables d'état

Soit le changement de variables $x = x(u, v, w)$, $y = y(u, v, w)$, $z = z(u, v, w)$.

Son **jacobien** est noté : $J = \frac{\partial(x, y, z)}{\partial(u, v, w)} = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} & \frac{\partial x}{\partial w} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} & \frac{\partial y}{\partial w} \\ \frac{\partial z}{\partial u} & \frac{\partial z}{\partial v} & \frac{\partial z}{\partial w} \end{vmatrix}$

Propriétés :

1. Simplification (de deux variables identiques en haut et en bas, à la même place) :

$$\left. \frac{\partial x}{\partial u} \right]_{v,w} = \frac{\partial(x, v, w)}{\partial(u, v, w)} \quad \left. \frac{\partial(x, v, z)}{\partial(u, v, w)} = \frac{\partial(x, z)}{\partial(u, w)} \right]_v \quad \text{etc.}$$

2. permutation de deux variables adjacentes (en haut ou en bas) :

$$\frac{\partial(x, z, y)}{\partial(u, v, w)} = -\frac{\partial(x, y, z)}{\partial(u, v, w)} \quad \frac{\partial(x, y, z)}{\partial(w, v, u)} = -\frac{\partial(x, y, z)}{\partial(u, v, w)} \quad \text{etc.}$$

3. double changement de variables :

$$\frac{\partial(x, y, z)}{\partial(u, v, w)} \times \frac{\partial(u, v, w)}{\partial(r, s, t)} = \frac{\partial(x, y, z)}{\partial(r, s, t)}$$

4. inversion (ci-dessus avec $r, s, t = x, y, z$) :

$$\frac{\partial(x, z, y)}{\partial(u, v, w)} = \frac{1}{\frac{\partial(u, v, w)}{\partial(x, y, z)}}$$

Les relations de Maxwell (-Cauchy) expriment le fait que le résultat du calcul d'une dérivée seconde d'une fonction d'état est indépendant de l'ordre dans lequel s'effectuent les dérivations. Par exemple :

$$\left. \frac{\partial}{\partial T} \left(\left. \frac{\partial F}{\partial V} \right]_{T,N} \right) \right]_{V,N} = \left. \frac{\partial}{\partial V} \left(\left. \frac{\partial F}{\partial T} \right]_{V,N} \right) \right]_{T,N} \Rightarrow \left. \frac{\partial p}{\partial T} \right]_{V,N} = \left. \frac{\partial S}{\partial V} \right]_{T,N}$$

On peut écrire de telles relations pour chacun des potentiels thermodynamiques et pour chaque couple de variables.

Les carrés mnémoniques de Born permettent de retrouver rapidement ces relations. Chaque carré est construit comme suit. On place deux variables extensives et les deux variables intensives associées sur un carré, de sorte que les variables associées soient diagonalement opposées. Chaque flèche relie les variables associées dans le sens extensive $\rightarrow$ intensive. Chaque carré permet d'établir quatre relations :

- à N ou μ constant :

$$\begin{array}{ccc} -p & & T \\ & \swarrow \quad \searrow & \\ S & & V \end{array} \Rightarrow \left\{ \begin{array}{l} + \left. \frac{\partial S}{\partial p} \right]_T = - \left. \frac{\partial V}{\partial T} \right]_p \\ + \left. \frac{\partial T}{\partial V} \right]_S = - \left. \frac{\partial p}{\partial S} \right]_V \end{array} \right. \quad \begin{array}{l} - \left. \frac{\partial V}{\partial S} \right]_p = - \left. \frac{\partial T}{\partial p} \right]_S \\ - \left. \frac{\partial p}{\partial T} \right]_V = - \left. \frac{\partial S}{\partial V} \right]_T \end{array}$$

– à S ou T constant :

$$\begin{array}{ccc} \mu & & -p \\ & \swarrow \quad \searrow & \\ V & & N \end{array} \Rightarrow \left\{ \begin{array}{l} -\left[\frac{\partial V}{\partial \mu}\right]_p = +\left[\frac{\partial N}{\partial p}\right]_\mu \\ -\left[\frac{\partial p}{\partial N}\right]_V = +\left[\frac{\partial \mu}{\partial V}\right]_N \end{array} \right. \quad \begin{array}{l} -\left[\frac{\partial N}{\partial V}\right]_\mu = -\left[\frac{\partial p}{\partial \mu}\right]_V \\ -\left[\frac{\partial \mu}{\partial p}\right]_N = -\left[\frac{\partial V}{\partial N}\right]_p \end{array}$$

– à V ou p constant :

$$\begin{array}{ccc} T & & \mu \\ & \swarrow \quad \searrow & \\ N & & S \end{array} \Rightarrow \left\{ \begin{array}{l} -\left[\frac{\partial N}{\partial T}\right]_\mu = -\left[\frac{\partial S}{\partial \mu}\right]_T \\ +\left[\frac{\partial \mu}{\partial S}\right]_N = +\left[\frac{\partial T}{\partial N}\right]_S \end{array} \right. \quad \begin{array}{l} -\left[\frac{\partial S}{\partial N}\right]_T = +\left[\frac{\partial \mu}{\partial T}\right]_N \\ +\left[\frac{\partial T}{\partial \mu}\right]_S = -\left[\frac{\partial N}{\partial S}\right]_\mu \end{array}$$

Pour la deuxième relation (en haut à droite) du troisième carré par exemple (on le notera $[N, S, \mu, T]_V$ ou $[N, S, \mu, T]_p$), on effectue deux rotations : l'une autour du triangle $S - N - T$, l'autre autour du triangle $\mu - T - N$. On met un signe $(-)$ lorsque la diagonale de retour est parcourue dans le sens contraire de la flèche.

On peut aussi construire des carrés pour les transformées de Legendre de l'entropie :

– à N ou γ constant :

$$\begin{array}{ccc} \xi & & \beta \\ & \swarrow \quad \searrow & \\ E & & V \end{array} \Rightarrow \left\{ \begin{array}{l} -\left[\frac{\partial E}{\partial \xi}\right]_\beta = -\left[\frac{\partial V}{\partial \beta}\right]_\xi \\ +\left[\frac{\partial \beta}{\partial V}\right]_E = +\left[\frac{\partial \xi}{\partial E}\right]_V \end{array} \right. \quad \begin{array}{l} -\left[\frac{\partial V}{\partial E}\right]_\xi = +\left[\frac{\partial \beta}{\partial \xi}\right]_E \\ +\left[\frac{\partial \xi}{\partial \beta}\right]_V = -\left[\frac{\partial E}{\partial V}\right]_\beta \end{array}$$

– à E ou β constant :

$$\begin{array}{ccc} \gamma & & \xi \\ & \swarrow \quad \searrow & \\ V & & N \end{array} \Rightarrow \left\{ \begin{array}{l} -\left[\frac{\partial V}{\partial \gamma}\right]_\xi = -\left[\frac{\partial N}{\partial \xi}\right]_\gamma \\ +\left[\frac{\partial \xi}{\partial N}\right]_V = +\left[\frac{\partial \gamma}{\partial V}\right]_N \end{array} \right. \quad \begin{array}{l} -\left[\frac{\partial N}{\partial V}\right]_\gamma = +\left[\frac{\partial \xi}{\partial \gamma}\right]_V \\ +\left[\frac{\partial \gamma}{\partial \xi}\right]_N = -\left[\frac{\partial V}{\partial N}\right]_\xi \end{array}$$

– à V ou ξ constant :

$$\begin{array}{ccc} \beta & & \gamma \\ & \swarrow \quad \searrow & \\ N & & E \end{array} \Rightarrow \left\{ \begin{array}{l} -\left[\frac{\partial N}{\partial \beta}\right]_\gamma = -\left[\frac{\partial E}{\partial \gamma}\right]_\beta \\ +\left[\frac{\partial \gamma}{\partial E}\right]_N = +\left[\frac{\partial \beta}{\partial N}\right]_E \end{array} \right. \quad \begin{array}{l} -\left[\frac{\partial E}{\partial N}\right]_\beta = +\left[\frac{\partial \gamma}{\partial \beta}\right]_N \\ +\left[\frac{\partial \beta}{\partial \gamma}\right]_E = -\left[\frac{\partial N}{\partial E}\right]_\gamma \end{array}$$

Exemple d'utilisation des jacobiens et des carrés de Born.

On désire calculer en fonction de grandeurs mesurables l'abaissement de température lors d'une désaimantation adiabatique d'un matériau paramagnétique.

Autrement dit, que vaut $\left[\frac{\partial T}{\partial B}\right]_{S,p,N}$, sachant $dE = T dS - p dV + B dM + \mu dN$?

On utilise le formalisme de Jacobi :

$$\begin{aligned}
 \left. \frac{\partial T}{\partial B} \right]_{S,p,N} &\stackrel{(1)}{=} \frac{\partial(T, S, p, N)}{\partial(B, S, p, N)} \stackrel{(3)}{=} \frac{\partial(T, S, p, N)}{\partial(B, T, p, N)} \times \frac{\partial(B, T, p, N)}{\partial(B, S, p, N)} \\
 &\stackrel{(2)}{=} -\frac{\partial(S, T, p, N)}{\partial(B, T, p, N)} \times \frac{\partial(B, T, p, N)}{\partial(B, S, p, N)} \stackrel{(4)}{=} -\frac{\frac{\partial(S, T, p, N)}{\partial(B, T, p, N)}}{\frac{\partial(B, T, p, N)}{\partial(B, S, p, N)}} \\
 &\stackrel{(2)}{=} -\frac{\frac{\partial(S, T, p, N)}{\partial(B, T, p, N)}}{\frac{\partial(S, B, p, N)}{\partial(T, B, p, N)}} \stackrel{(1)}{=} -\frac{\left. \frac{\partial S}{\partial B} \right]_{T,p,N}}{\left. \frac{\partial S}{\partial T} \right]_{B,p,N}} = -\frac{T}{C_{p,B}} \times \left. \frac{\partial S}{\partial B} \right]_{T,p,N}
 \end{aligned}$$

Le carré de Born $[M, S, B, T]_{p,N}$ permet d'écrire :

$$\left. \frac{\partial S}{\partial B} \right]_{T,p,N} = \left. \frac{\partial M}{\partial T} \right]_{B,p,N} \quad \text{et donc} \quad \left. \frac{\partial T}{\partial B} \right]_{S,p,N} = -\frac{T}{C_{p,B}} \times \left. \frac{\partial M}{\partial T} \right]_{B,p,N}$$

La grandeur recherchée est maintenant exprimée en fonction de grandeurs mesurables au cours de la même expérience : à B, p, N constants.

1.2 Probabilités et statistiques

1.2.1 Probabilités, moyennes, distributions

On lance un dé à 6 faces 6000 fois, le résultat de chaque lancer étant indépendant des résultats précédents. Si la nature était parfaite, on observerait exactement 1000 fois "1", 1000 fois "2", 1000 fois "3", 1000 fois "4", 1000 fois "5" et 1000 fois "6". Mais en pratique, ce résultat est fort peu probable. Ce qu'on peut dire intuitivement, c'est que plus le nombre N de lancers est grand plus le résultat tendra vers la répartition $N/6 - N/6 - N/6 - N/6 - N/6 - N/6$.

D'où la définition suivante de la probabilité. Soit une expérience dont le résultat fait partie d'un ensemble fini de p événements possibles (on lance par exemple un dé à p faces). Si cette expérience est répétée N fois, et si n_m est le nombre de fois où on observe le résultat m , la probabilité d'apparition du résultat m est donnée par :

$$P_m = \lim_{N \rightarrow \infty} \frac{n_m}{N} \quad \text{avec} \quad 0 \leq P_m \leq 1 \quad \text{et} \quad \sum_{m=1}^p P_m = 1$$

Les deux opérations de base :

- Soient deux événements e_1 et e_2 s'excluant l'un l'autre. Lors d'une même expérience, la probabilité d'observer " e_1 ou e_2 " est donnée par :

$$\textbf{la loi d'addition} : P(e_1 \text{ ou } e_2) = P(e_1) + P(e_2)$$

- Soient deux expériences à résultats indépendants. La probabilité d'observer e_1 à la première expérience et e_2 à la deuxième est donnée par :

$$\textbf{la loi de multiplication} : P(e_1 \text{ et } e_2) = P(e_1) \times P(e_2)$$

Considérons un dé à p faces numérotées $m = 1$ à p , et X une grandeur calculée sur la valeur de m . On notera P_m la probabilité d'observer la face m , et X_m la valeur de X pour la face m .

La valeur moyenne de X est donnée par : $\bar{X} = \sum_{m=1}^p P_m X_m$

La valeur quadratique moyenne de X est donnée par : $\overline{X^2} = \sum_{m=1}^p P_m X_m^2$

La variance ou écart quadratique moyen est : $\Delta X^2 = \overline{(X - \bar{X})^2} = \overline{X^2} - \bar{X}^2$

L'écart type est par définition : $\Delta X = \sqrt{\overline{X^2} - \bar{X}^2}$

Il est fréquent que plusieurs valeurs de m conduisent à une même valeur $X(i)$. La liste des valeurs possibles $X(i)$ est alors plus courte que la liste des X_m et il est commode de sommer sur les valeurs de i plutôt que sur les valeurs de m . La **distribution statistique** de X est l'ensemble des probabilités $P(i)$ d'observer $X(i)$ pour toutes les valeurs de i . Soit $M(i)$ l'ensemble des valeurs de m donnant la même valeur $X(i)$. $\bar{X}$ peut se calculer à l'aide d'une somme sur m ou d'une somme sur i :

$$X_m = X(i) \quad \forall m \in M(i) \quad \Rightarrow \quad P(i) = \sum_{m \in M(i)} P_m$$

$$\bar{X} = \sum_m X_m P_m = \sum_i \sum_{m \in M(i)} X_m P_m = \sum_i X(i) \sum_{m \in M(i)} P_m = \sum_i X(i) P(i)$$

1.2.2 La distribution binomiale

On joue à pile ou face

C'est une expérience à deux événements possibles. Soient x_1 et x_2 les deux probabilités correspondantes. On répète N fois l'expérience et on compte les nombres n_1 et n_2 de fois où l'événement "1" et l'événement "2" sont observés. La **distribution binomiale** est par définition la distribution de probabilités $P(n_1, n_2)$ d'observer n_1 événements "1" et n_2 événements "2".

Pour calculer la probabilité $P(n_1, n_2)$, on calcule la probabilité d'observer une *configuration* $(m_1, \dots, m_N)$ (m_1 = résultat de l'expérience n° 1, m_2 = résultat de l'expérience n° 2, etc). Si cette configuration appartient à la distribution (n_1, n_2) , autrement dit si dans cette configuration on compte n_1 événements "1" et n_2 événements "2", on a :

$$p(m_1, m_2, \dots, m_N) = x_1^{n_1} x_2^{n_2}$$

On multiplie ensuite par le nombre de configurations faisant apparaître n_1 événements "1" et n_2 événements "2". C'est le nombre de façons de répartir N objets dans deux boîtes, à raison de n_1 objets dans la boîte "1", n_2 dans la boîte "2". A partir d'une configuration donnée, c'est le nombre de permutations possibles entre deux objets, l'un situé dans la boîte "1", l'autre situé dans la boîte "2".

$$\text{C'est-à-dire : } \frac{(n_1 + n_2)!}{n_1! n_2!}$$

On en déduit la loi binomiale : $P(n_1, n_2) = \frac{(n_1 + n_2)!}{n_1! n_2!} x_1^{n_1} x_2^{n_2}$

Elle tire son nom du fait que $P(n_1, n_2)$ est l'un des termes du développement du binôme $(x_1 + x_2)$ à la puissance N :

$$(x_1 + x_2)^N = \sum_{n_1, n_2}^{n_1 + n_2 = N} \frac{(n_1 + n_2)!}{n_1! n_2!} x_1^{n_1} x_2^{n_2} = \sum_{n_1, n_2}^{n_1 + n_2 = N} P(n_1, n_2)$$

Avec $x_1 + x_2 = 1$, on a bien : $\sum_{n_1, n_2}^{n_1 + n_2 = N} P(n_1, n_2) = 1$.

Pour calculer les **valeurs moyennes**, on part de :

$$(x_1 + x_2)^N = \sum_{n_1, n_2}^{n_1 + n_2 = N} \frac{(n_1 + n_2)!}{n_1! n_2!} x_1^{n_1} x_2^{n_2}$$

On lui applique l'opérateur $x_1 \frac{\partial}{\partial x_1}$:

$$N x_1 (x_1 + x_2)^{N-1} = \sum_{n_1, n_2}^{n_1 + n_2 = N} n_1 P(n_1, n_2) = \bar{n}_1$$

D'une part on applique à nouveau $x_1 \frac{\partial}{\partial x_1}$, d'autre part on applique $x_2 \frac{\partial}{\partial x_2}$:

$$N x_1 (x_1 + x_2)^{N-1} + N(N-1) x_1^2 (x_1 + x_2)^{N-2} = \sum_{n_1, n_2}^{n_1 + n_2 = N} n_1^2 P(n_1, n_2) = \overline{n_1^2}$$

$$N(N-1) x_1 x_2 (x_1 + x_2)^{N-2} = \sum_{n_1, n_2}^{n_1 + n_2 = N} n_1 n_2 P(n_1, n_2) = \overline{n_1 n_2}$$

On écrit maintenant la condition $x_1 + x_2 = 1$ et on obtient, avec un traitement identique pour la variable x_2 :

$$\begin{aligned} \bar{n}_1 &= N x_1 & \overline{n_1^2} &= N x_1 (1 - x_1 + N x_1) & \Delta n_1^2 &= N x_1 (1 - x_1) \\ \bar{n}_2 &= N x_2 & \overline{n_2^2} &= N x_2 (1 - x_2 + N x_2) & \Delta n_2^2 &= N x_2 (1 - x_2) \\ \overline{n_1 n_2} &= N(N-1) x_1 x_2 \end{aligned} \quad (1.21)$$

Le nombre N est en général un paramètre fixe. Un léger changement de notation permet d'écrire les résultats en fonction d'une seule variable. En posant $x_1 = x$, x_2 devient $1 - x$. De même, en posant $n_1 = n$, n_2 devient $N - n$. Les écritures deviennent :

$$P(n) = \frac{N!}{n!(N-n)!} x^n (1-x)^{N-n} \quad \bar{n} = N x \quad \Delta n^2 = N x (1-x)$$

On joue à pile ou face en jetant simultanément N pièces de monnaie. On répète l'expérience un très grand nombre de fois ($\mathcal{N} \gg N$) et on compte, pour chaque

valeur de n , le nombre de fois où n pièces présentent le côté "pile". Pour x donné (la probabilité pour une pièce de présenter le côté "pile"), on obtient une courbe en cloche (figure 1.1) plus ou moins centrée sur la valeur moyenne $\bar{n} = Nx$, et de largeur à mi-hauteur environ $2\Delta n$. On constate que n varie aléatoirement entre 0 et N , avec une forte probabilité que $|n - \bar{n}|$ soit inférieur à deux ou trois Δn .

La valeur moyenne $\bar{n}$ varie proportionnellement à N , alors que l'écart-type Δn varie proportionnellement à $\sqrt{N}$. Le critère de convergence est *l'écart relatif moyen à la moyenne* :

$$\frac{\Delta n}{\bar{n}} = \sqrt{\frac{1-x}{Nx}} \propto \frac{1}{\sqrt{N}} \rightarrow 0 \text{ quand } N \rightarrow \infty \quad (1.22)$$

Sur la figure 1.1, on constate que, du point de vue qualitatif, plus N augmente :

- plus le graphe est dense en valeurs de n ,
- plus les valeurs de n de probabilité appréciable se resserrent autour de $\bar{n}$,
- plus le graphe de $P(n)$ est symétrique par rapport à $\bar{n}$.

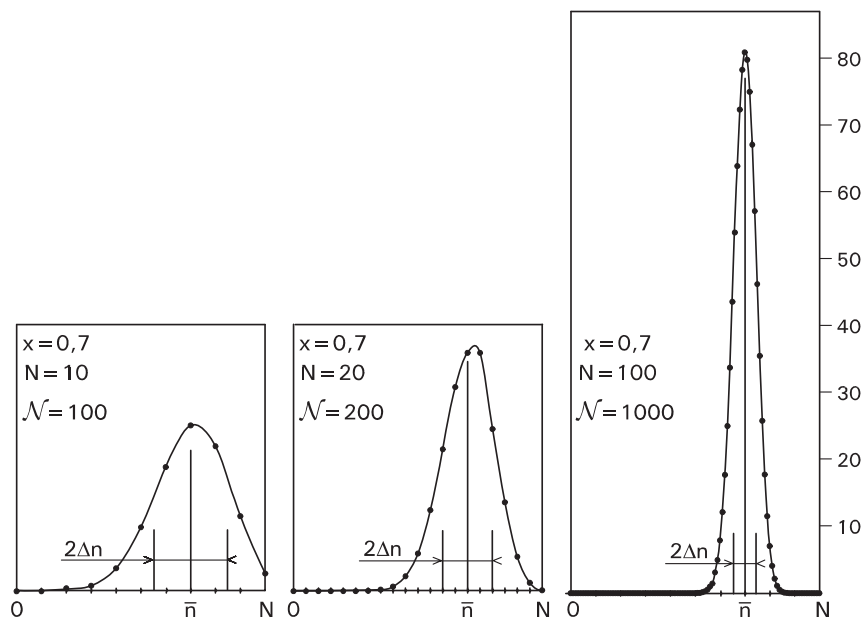


FIG. 1.1 – Distribution binomiale

La limite des grands nombres

C'est maintenant que l'on commence à quitter la rigueur mathématique. Face au nombre d'Avogadro, l'unité est ridiculement petite. Les fractions d'unités encore plus. Lorsqu'on applique la formule de Stirling, on associe à un nombre **entier** $k!$ un nombre **réel** (sauf si dans la formule de Stirling 1.23 on pousse le développement jusqu'à l'infini). A cette *rigueur mathématique* se substitue la *réalité physique* : lorsque, physiquement, aucune expérience réaliste ne permet de distinguer deux valeurs, on dira que ces deux valeurs sont rigoureusement égales à l'échelle macroscopique.

Cette ambiguïté est également présente en thermodynamique : lorsqu'un thermodynamicien écrit dn , où n est un nombre de moles, il oublie qu'une quantité de

matière ne peut varier que de manière discrète. Il quitte alors la rigueur mathématique pour la réalité physique : la variation d'une molécule par rapport au nombre d'Avogadro peut être, *physiquement*, considérée comme infinitésimale.

Ce problème concerne en fait toutes les variables : à l'échelle microscopique, c'est la mécanique quantique qui définit les états possibles pour un système donné. En mécanique quantique, tout est quantifié, tout varie de façon discrète. Mais pour une grandeur X donnée, l'écart entre deux valeurs consécutives est toujours, pour des systèmes macroscopiques, extrêmement petit face à la valeur moyenne de cette grandeur : on peut toujours trouver un intervalle δX à la fois inappréciable à l'échelle macroscopique et suffisamment grand pour contenir un très grand nombre de valeurs à l'échelle microscopique. En un mot pour ce qui nous concerne ici : dans l'intervalle $[0, 10^{24}]$, k devient une variable continue.

La *formule de Stirling* donne un infiniment grand équivalent à $k!$ lorsque $k \rightarrow \infty$:

$$k! \approx \frac{k^k}{e^k} \sqrt{2\pi k} \left(1 + \frac{1}{12k} + \frac{1}{288k^2} - \frac{139}{51840k^3} - \frac{571}{2488320k^4} \cdots \right) \quad (1.23)$$

Les valeurs reportées dans le tableau ci-dessous ont été calculées sans tenir compte de la parenthèse ci-dessus. On s'aperçoit que les résultats sont bons dès les premières valeurs de k , mais que la contribution de $\ln \sqrt{2\pi k}$ est non négligeable. Dans cet gamme, il faut donc utiliser :

$$k! \approx \frac{k^k}{e^k} \sqrt{2\pi k} \quad \ln(k!) \approx k \ln k - k + \ln \sqrt{2\pi k} \quad \text{et} \quad \frac{d \ln(k!)}{dk} \approx \ln k + \frac{1}{2k} \quad (1.24)$$

k	$\ln(k!)$	Stirling	$\ln \sqrt{2\pi k}$
1	0,0000	-0,0811	0,9189
10	15,1044	15,0961	2,0702
100	363,7394	363,7385	3,2215
1000	5912,1282	5912,1281	4,3728
10000	82108,9278	82108,9278	5,5241

Dans la gamme des très grands nombres, $\ln \sqrt{2\pi k}$ est négligeable devant $k \ln k - k$. On peut considérer que la formule de Stirling simplifiée est quasi rigoureuse :

$$k! \approx \frac{k^k}{e^k} \quad \ln(k!) \approx k \ln k - k \quad \text{et} \quad \frac{d \ln(k!)}{dk} \approx \ln k \quad (1.25)$$

k	$k \ln k - k$	$\ln \sqrt{2\pi k}$
10^{10}	$2,203 \times 10^{11}$	12,43
10^{20}	$4,505 \times 10^{21}$	23,94
10^{30}	$6,808 \times 10^{31}$	35,46

Pour $k = 10^{20}$, $k!$ est de l'ordre de $10^{2 \times 10^{21}}$. C'est un nombre à 2×10^{21} chiffres. A raison de 2 chiffres par centimètre, il mesure plus de 1 000 années-lumière.

Lorsqu'on oublie le terme multiplicatif $\sqrt{2\pi k}$, qui vaut tout de même 25 milliards, on commet une erreur négligeable :

$$10^{10} \times 10^{2 \times 10^{21}} = 10^{(2 \times 10^{21} + 10)} = 10^{2\,000\,000\,000\,000\,000\,000\,010}$$

En écrivant $10^{2\,000\,000\,000\,000\,000\,000\,010} \approx 10^{2\,000\,000\,000\,000\,000\,000\,000}$, on oublie une dizaine de chiffres par rapport aux 2×10^{21} déjà présents, et les 1 000 années-lumière sont réduites de ... cinq centimètres.

Revenons à la distribution $P(n)$. Dans la limite des très grands nombres, l'écart entre les valeurs possibles de n (une unité) devient extrêmement faible comparativement à l'intervalle de définition $[0, N]$, à tel point que l'on pourra considérer $P(n)$ exprimé en fonction de $y = n/N$ comme une fonction continue $W(y)$:

$$W(y) = P(n) = P(Ny) = \frac{N!}{(Ny)!(N - Ny)!} x^{Ny} (1 - x)^{N - Ny}$$

Une valeur moyenne se calcule maintenant à l'aide d'une intégrale :

$$\bar{X} = \sum_{n=0}^N P(n)X(n) \approx \int_0^1 P(Ny)X(Ny) d(Ny) = N \int_0^1 W(y)X(Ny) dy$$

Pour que les calculs qui suivent soient valables, il faut que toutes les valeurs de n donnant une valeur non négligeable à $P(n)$ soient de grands nombres. Ces valeurs sont situées dans un intervalle de quelques $\Delta n = \sqrt{Nx(1 - x)}$ autour de $\bar{n} = Nx$.

Seront donc supposés très grands, non seulement N , mais aussi $\bar{n}$ et $N - \bar{n}$. D'où :

Pour que toutes les valeurs de n donnant une probabilité $P(n)$ non négligeable soient de grands nombres il faut, pour x donné, choisir le nombre N pour que :

$$Nx(1 - x) \gg 1$$

À noter que, en physique statistique, cette condition n'est pas très restrictive. Avec $N \sim 10^{20}$, même si x est très petit, disons 10^{-10} , $\bar{n}$ reste encore très grand (10^{10}). Il reste encore assez de place entre 0 et $\bar{n}$ pour un grand nombre de Δn qui ne vaut que 10^5 . Les problèmes n'apparaissent que si x est inférieur à quelques $1/N$.

À la valeur n^* la plus probable de n correspond une valeur y^* telle que $n^* = Ny^*$. Cette valeur y^* rend $\ln W(y)$ maximum^[7]. Elle doit annuler $\frac{d}{dy} \ln W$, avec :

$$\ln W(y) = \ln(N!) - \ln((Ny)!) - \ln((N - Ny)!) + Ny \ln x + (N - Ny) \ln(1 - x)$$

La formule de Stirling simplifiée 1.25 n'est pas satisfaisante ici car elle conduit à $P(n^*) = 1$ et donc à $P(n) = 0$ pour tout $n \neq n^*$. C'est donc la formule de Stirling complète 1.24 qu'il faut utiliser. On trouve :

$$\left. \frac{d \ln W}{dy} \right|_{y=y^*} = 0 \quad \Rightarrow \quad \frac{x}{1 - x} = \frac{y^*}{1 - y^*} \exp \frac{1 - 2y^*}{2Ny^*(1 - y^*)}$$

On résout en exprimant x en fonction de y^* . Par hypothèse, on a $Nx(1 - x) \gg 1$, et par conséquent $Ny^*(1 - y^*) \gg 1$ puisque x et y^* sont très proches. On peut développer l'exponentielle au premier ordre. On trouve :

$$x \approx \left(1 - \frac{1}{N}\right)y^* + \frac{1}{2N} \quad \Rightarrow \quad y^* \approx \left(1 + \frac{1}{N}\right)x - \frac{1}{2N} \quad \Rightarrow \quad n^* \approx \bar{n} + x - 1/2$$

⁷ $W(y)$ varie extrêmement rapidement au voisinage de x . Pour éviter les problèmes lors des dérivations, on travaille avec $\ln W(y)$ qui présente des variations plus douces.

Comme $0 < x < 1$, n^* et $\bar{n}$ diffèrent au maximum d'une demi unité (avec la formule de Stirling simplifiée, on aurait trouvé exactement $n^* = \bar{n}$). Conclusion :

La valeur la plus probable n^ correspond quasi exactement à la valeur moyenne $\bar{n}$.*

On remplace y par $y^* = x$ dans $\ln W(y)$ en utilisant la formule de Stirling complète :

$$\ln W(y^*) = \ln W(x) = \ln P(\bar{n}) = \ln \frac{1}{\sqrt{2\pi N x(1-x)}} = \ln \frac{1}{\sqrt{2\pi \Delta n^2}}$$

Il reste à calculer la valeur de $P(n)$ au voisinage de $\bar{n}$. On assimile y^* à x et on écrit le développement de Taylor de $\ln W(y)$ au voisinage de x :

$$\ln W(y) = \ln \frac{1}{\sqrt{2\pi \Delta n^2}} + (y-x) \left. \frac{d \ln W}{dy} \right|_{y=x} + \sum_{k=2}^{\infty} \frac{(y-x)^k}{k!} \left. \frac{d^k \ln W}{dy^k} \right|_{y=x}$$

Les calculs donnent : $\left. \frac{d \ln W}{dy} \right|_{y=x} = \frac{x - 1/2}{x(1-x)} = N \frac{x - 1/2}{\Delta n^2}$

Et pour $k > 1$:

$$\left. \frac{1}{k!} \frac{d^k \ln W}{dy^k} \right|_{y=x} = \frac{\varphi_k(x)}{x^k(1-x)^k} + N \frac{\psi_k(x)}{x^{k-1}(1-x)^{k-1}} = \left(\frac{\varphi_k(x)}{\Delta n^k} + \frac{\psi_k(x)}{\Delta n^{k-2}} \right) \frac{N^k}{\Delta n^k}$$

Les fonctions φ_k et ψ_k prennent des valeurs de l'ordre de l'unité :

$$\begin{aligned} \varphi_2(x) &= x^2 - x + 1/2 & \psi_2(x) &= -1/2 \\ \varphi_3(x) &= \frac{(x - 1/2)(x^2 - x + 1)}{3} & \psi_3(x) &= \frac{1/2 - x}{3} \end{aligned}$$

On pose $u = \frac{n - \bar{n}}{\Delta n} = \frac{N(y-x)}{\Delta n}$ et on range les termes en puissances de $\frac{1}{\Delta n}$:

$$\ln W(y) = \ln \frac{1}{\sqrt{2\pi \Delta n^2}} - \frac{u^2}{2} + \frac{(x - 1/2)(u - u^3/3)}{\Delta n} + \dots$$

Par hypothèse, Δn est très grand ($\Delta n \sim 10^{10}$), et on sait que les valeurs significatives de n sont telles que u vaut quelques unités. Dans la limite des très grands nombres, *la convergence est extrêmement rapide*. Les deux premiers termes suffisent :

$$W(y) = \frac{e^{-u^2/2}}{\sqrt{2\pi \Delta n^2}} \Rightarrow P(n) = \frac{1}{\sqrt{2\pi \Delta n^2}} \exp \left[-\frac{1}{2} \left(\frac{n - \bar{n}}{\Delta n} \right)^2 \right]$$

C'est l'équation d'une *gaussienne* (figure 1.2) caractérisée par :

- sa *position* $\bar{n}$;
- sa *demi-largeur* Δn ;
- son *amplitude* $A = \frac{1}{\sqrt{2\pi \Delta n^2}}$

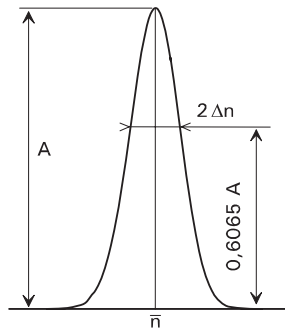


FIG. 1.2 – Distribution gaussienne

On peut vérifier qu'on a bien $S = \sum_{n=0}^N P(n) = 1$ et $\sigma^2 = \sum_{n=0}^N (n - \bar{n})^2 P(n) = \Delta n^2$:

$$S \approx \frac{N}{\sqrt{2\pi\Delta n^2}} \int_0^1 e^{-1/2(N(y-x)/\Delta n)^2} dy \approx \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-u^2/2} du = 1$$

$$\sigma^2 \approx \frac{N^3}{\sqrt{2\pi\Delta n^2}} \int_0^1 (y-x)^2 e^{-1/2(N(y-x)/\Delta n)^2} dy \approx \frac{\Delta n^2}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} u^2 e^{-u^2/2} du = \Delta n^2$$

Pour calculer S et σ^2 , on a effectué le changement de variable $u = N(y-x)/\Delta n$ et assimilé les bornes $-Nx/\Delta n$ à $-\infty$ et $N(1-x)/\Delta n$ à $+\infty$.

Récurrence utile pour ces deux calculs :

$$J_p = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} u^p e^{-u^2/2} du = (p-1)J_{p-2} \quad \text{avec} \quad J_0 = 1 \text{ et } J_1 = 0$$

Physiquement, on retiendra que, pour $N \approx 10^{20}$, la distribution binomiale et la distribution gaussienne sont quasi identiques dans un domaine de plusieurs Δn autour de $\bar{n}$. Au delà, l'écart relatif est important, mais les valeurs de $P(n)$ sont tout à fait négligeables.

Considérons maintenant une grandeur X calculable pour chaque valeur de n . Soient $X(n)$ ces valeurs. La valeur moyenne de X vaut :

$$\overline{X(n)} = \sum_{n=0}^N X(n)P(n) \approx \int_0^1 X(Ny)W(y)N dy$$

$W(y)$ prend des valeurs appréciables dans un intervalle extrêmement petit centré autour de x . On peut donc sans crainte développer $X(Ny)$, que l'on notera $f(y)$, au premier ordre :

$$f(y) \approx f(x) + (y-x) \left. \frac{df}{dy} \right]_{y=x} = f(x) + (y-x)f'(x)$$

$$\Rightarrow \quad \overline{X(n)} \approx f(x) \int_{-\infty}^{+\infty} W(y) N dy + f'(x) \int_{-\infty}^{+\infty} (y-x)W(y) N dy$$

La première intégrale vaut 1 et la deuxième est nulle, d'où :

$$\overline{X(n)} = \sum_{n=0}^N X(n)P(n) \approx X(\bar{n})$$

La valeur moyenne de X est égale à la valeur qu'elle prend pour la valeur de n la plus probable. C'est ***l'approximation du terme maximum*** qui, dans la limite des très grands nombres peut être considérée comme rigoureuse.

$$\text{À retenir donc : } \overline{X(n)} \approx X(\bar{n})$$

1.2.3 La distribution multinomiale

On lance un dé à p faces

On étend le raisonnement utilisé pour la loi binomiale à une expérience à p événements possibles ($m = 1$ à p) que l'on répète N fois. On compte le nombre n_m de fois où chaque événement m est observé.

Pour calculer la probabilité $P(n_1 \dots n_p)$ correspondante, on calcule d'abord la probabilité d'observer une configuration $(m_1, m_2, \dots, m_N)$ faisant apparaître n_1 événements "1", n_2 événements "2", etc (on dira que la configuration $(m_1, m_2, \dots, m_N)$ appartient à la distribution $(n_1 \dots n_p)$) :

$$p(m_1, m_2, \dots, m_N) = x_1^{n_1} \dots x_p^{n_p}$$

On multiplie ensuite par le nombre de configurations faisant partie de la distribution $(n_1 \dots n_p)$:

$$P(n_1 \dots n_p) = \frac{N!}{n_1! \dots n_p!} x_1^{n_1} \dots x_p^{n_p} = N! \prod_{i=1}^p \frac{x_i^{n_i}}{n_i!}$$

Les calculs sont identiques à ceux du paragraphe précédent.

Partant de $\sum_{n_1 \dots n_p}^{n_1 + \dots + n_p = N} P(n_1 \dots n_p) = (x_1 + \dots + x_p)^N$, on trouve :

$$\bar{n}_i = N x_i \quad \overline{n_i^2} = N x_i (1 - x_i + N x_i) \quad \Delta n_i^2 = N x_i (1 - x_i)$$

Limite des grands nombres

Dans la *limite des grands nombres*, on travaille avec :

$$\ln P = \ln(N!) + \sum_{i=1}^p (n_i \ln(x_i) - \ln(n_i!))$$

que l'on réécrit en fonction de $y_1 = \frac{n_1}{N}, \dots, y_p = \frac{n_p}{N}$:

$$\ln W(y_1 \dots y_p) = \ln P(n_1 \dots n_p) = \ln \left[\left(N \sum_{i=1}^p y_i \right)! \right] + \sum_{i=1}^p \left[N y_i \ln(x_i) - \ln((N y_i)!) \right]$$

Soit $[y]^* = \{y_1^* \dots y_p^*\}$ le multiplet le plus probable.

Ce multiplet doit rendre $\ln W$ maximum, sous la contrainte $\sum_{i=1}^p y_i = S$.

On introduit un *multiplieur de Lagrange* α . Les y_j^* sont solutions des équations :

$$\frac{\partial}{\partial y_j} \left(\ln W(y_1, \dots, y_p) - \alpha \sum_{i=1}^p y_i \right) = 0 \quad \forall j = 1 \text{ à } p$$

Tous les n_i étant supposés très grands, on applique la formule de Stirling à chaque $\ln N(y_i)$. On trouve :

$$N \ln N + N \ln(x_j) - N \ln(N y_j^*) - \alpha = 0 \quad \forall j \Rightarrow y_j^* = x_j e^{-\alpha/N}$$

Pour déterminer α , on remplace les y_j^* dans la contrainte :

$$\sum_{i=1}^p y_i^* = S \text{ et } \sum_{i=1}^p x_i = S \Rightarrow e^{-\alpha/N} = 1 \Rightarrow y_j^* = x_j \Rightarrow n_j^* = N x_j = \bar{n}_j$$

Le développement de Taylor au voisinage de $[n]^*$ conduit à un produit de p gaussiennes :

$$P(n_1, n_2, \dots, n_p) = \prod_{j=1}^p \frac{1}{\sqrt{2\pi \Delta n_j^2}} \exp \left[-\frac{1}{2} \left(\frac{n_j - n_j^*}{\Delta n_j} \right)^2 \right]$$

Et *l'approximation du terme maximum* conduit à :

$$\overline{X(n_1, n_2, \dots, n_p)} \approx X(\bar{n}_1, \bar{n}_2, \dots, \bar{n}_p)$$

1.3 Théorie statistique de Gibbs

1.3.1 Du macroscopique au microscopique

La thermodynamique macroscopique permet de trouver des relations entre les différentes propriétés d'un système en équilibre. Elle ne donne aucune information sur l'origine ou sur l'interprétation physique de ces propriétés.

La physique statistique permet de relier les propriétés d'un système macroscopique aux propriétés des particules qui le constituent. Elle a débuté à la fin du XIX^e siècle grâce à Boltzmann, en même temps que la renaissance de l'hypothèse atomique. Cette hypothèse devint rapidement une théorie grâce au développement de la mécanique quantique et à la découverte des rayons X, qui ont permis de confirmer l'existence des atomes et d'en décrire leur structure, puis d'identifier les particules élémentaires de la matière.

Le problème mécanique est impossible à résoudre puisqu'il met en jeu un nombre N de particules de l'ordre de 10^{22} . Sa résolution exacte serait d'ailleurs sans intérêt : il faudrait une dizaine de milliers de milliards de giga-octets pour stocker une configuration instantanée ($3N$ coordonnées de position et $3N$ composantes de vitesse),

et une centaine de millions d'années pour l'écrire sur le disque dur de votre PC. Pour imprimer cette configuration instantanée, il faudrait environ 10^{20} feuilles A4 en recto-verso : la pile sur le coin de votre bureau aurait une hauteur d'environ une année-lumière.

On doit donc se tourner vers des méthodes probabilistes : les mesures ne sont pas sensibles à notre échelle au comportement individuel des particules, mais seulement à des moyennes. Les méthodes probabilistes conduisent habituellement à un résultat plus ou moins approximatif. Ici, du fait de l'énormité des nombres utilisés, les prédictions sont *certaines*, ce qui conduit à une description *exacte* des propriétés macroscopiques du système étudié.

Le passage entre le macroscopique et le microscopique n'est pas toujours trivial. Entre ces deux mondes, le facteur d'échelle est énorme, à tel point qu'il a fallu donner un nom au domaine intermédiaire : le domaine *mésoscopique*, où se situent, pour une grandeur extensive donnée, les valeurs à la fois *suffisamment petites* pour être inappréciables à l'échelle macroscopique et *suffisamment grandes* à l'échelle microscopique pour lui appliquer un traitement statistique. Il faut pour s'en convaincre se rendre compte de l'énormité du nombre d'Avogadro : un millionième de millionième de mole, inappréciable à l'échelle macroscopique, contient environ un million de millions de molécules, suffisamment pour lui appliquer avec une précision plus que raisonnable les résultats de la théorie statistique.

À l'échelle microscopique, c'est la mécanique quantique qui régit le comportement des particules : les états possibles pour un système donné forment un ensemble discret, dénombrable. Les calculs statistiques se feront au moyen de signes sommes. Chaque état possible est appelé *microétat*. Il faut, pour le définir, connaître l'état individuel de chaque particule (sa position, sa vitesse et son état quantique "interne"), c'est-à-dire au moins $6N$ nombres. Pour faire l'analogie avec le lancer de dés, considérons les deux expériences suivantes.

- Expérience 1 : on jette 10 000 dés à 6 faces sur une table. Pour décrire le microétat observé, il faut aligner les 10 000 dés dans un ordre prédéterminé (chaque dé porte une étiquette), et relever la série ordonnée des 10 000 nombres : la face présentée par le dé 1, la face présentée par le dé 2, etc.
- Expérience 2 : on jette 10 000 fois un seul dé à 6 faces. Pour fabriquer le microétat de l'expérience 1, on doit ici photographier le dé après chaque lancé et au final juxtaposer toutes les photos dans l'ordre où elles sont apparues.

L'hypothèse ergodique que nous verrons plus loin consiste à dire ici que les deux expériences sont équivalentes en termes de nombres d'apparitions de chacune des faces.

Pour mesurer la probabilité d'apparition de chacune des faces, il n'est nul besoin de savoir dans quel ordre les différents résultats sont apparus. Les nombres d'apparition suffisent. La donnée de ces six nombres définit le *macroétat* du système constitué des 10 000 dés. La connaissance de ce macroétat suffit pour calculer la valeur moyenne d'une grandeur sur ce système. La précision obtenue dépendra uniquement du nombre de dés ou de lancers.

L'état macroscopique d'un système physique est défini en thermodynamique par la donnée de quelques grandeurs extensives (E , V , N , etc). Du point de vue statistique, un système physique est analogue à un dé possédant un très grand nombre de

faces, chaque face représentant un microétat possible. Au lieu de lancer 10 000 fois le dé (expérience 2), on observe le système de façon aléatoire un très grand nombre de fois au cours d'un certain intervalle de temps Δt . Le *macroétat* du système est défini par la donnée du nombre de fois où chacun des *microétats* est observé.

Cet intervalle de temps Δt nécessaire à la mesure d'une certaine grandeur X , peut être très court à notre échelle, mais sera toujours extrêmement grand à l'échelle microscopique. Pour fixer les idées, considérons un système constitué de 10^{18} particules. On estime qu'une particule donnée change d'état en moyenne 10^{12} fois par seconde. Le système change donc de microétat en moyenne 10^{30} fois par seconde. Même en utilisant un capteur ultra perfectionné capable d'effectuer la mesure en une picoseconde, le système aura changé 10^{18} fois de microétat pendant la mesure.^[8] La valeur mesurée de X correspond à une valeur moyenne, moyenne faite sur les différents microétats observés pendant cette picoseconde.

Mais que sont ces 10^{18} états par rapport à tous ceux disponibles ? RIEN.

Imaginez que chaque particule est une pièce de monnaie, c'est-à-dire ne disposant que de deux états individuels.

Cela donne pour le système $2^{10^{18}} \approx 10^{301\,029\,995\,663\,981\,000}$ microétats possibles. Le système est un dé à $10^{301\,029\,995\,663\,981\,000}$ faces ! Le nombre d'états observés pendant la picoseconde (10^{18}) est ridiculement petit par rapport à ce nombre.

Cela voudrait-il dire que le résultat de la mesure est complètement aléatoire ? Non ! Parce que parmi tous ces états possibles, seuls quelques uns ont une probabilité non négligeable d'apparaître, ceux compatibles quasi exactement avec l'état thermodynamique du système. C'est du moins ce qui est sous-entendu dans l'énoncé du principe ergodique.

1.3.2 Ensemble statistique, énoncé des postulats

Le macroétat d'un système macroscopique, défini par la donnée de quelques variables thermodynamiques (E , V , N , etc.), correspond à un "état moyen", moyenne faite sur l'ensemble des microétats accessibles, et pondérée par leurs probabilités d'apparition. Cette moyenne devra être faite sur une durée suffisamment longue pour que le résultat ne soit pas affecté par la durée ou les instants de début et de fin du comptage. N étant très grand, cette moyenne temporelle est impossible à calculer sans l'appui d'un certain nombre de postulats : c'est la base de la *méthode des ensembles de Gibbs*.

Soit un système en équilibre thermodynamique avec le milieu extérieur. Cet équilibre n'est pas figé. Les particules sont en perpétuel mouvement, et le macroétat observé correspond à un état stationnaire, à une moyenne temporelle sur des microétats fluctuant autour d'un certain état moyen. Si on photographie le système à un instant donné, on le trouvera dans tel ou tel microétat qui ne correspondra pas forcément à l'état d'équilibre. Un *ensemble statistique* est un rassemblement d'un très grand nombre $\mathcal{N}$ de ces photos prises au hasard dans un certain intervalle de temps Δt . Cet ensemble statistique constitue un super-système que l'on place dans

⁸ C'est pourquoi les puristes font la distinction entre "mesure" et "enregistrement" : un enregistrement est un ensemble de mesures prises simultanément ; une mesure est en toute rigueur impossible à réaliser.

le même environnement que le système d'origine. Si Δt est suffisamment long, on verra passer le système par tous les microétats compatibles avec le macroétat d'équilibre. Le temps de passage dans chaque microétat sera proportionnel à la probabilité d'apparition de ce microétat. De la même façon, si $\mathcal{N}$ est suffisamment grand, tous les microétats seront présents dans l'ensemble statistique, en nombre proportionnel à leur probabilité d'apparition.

De cette analogie découle le *principe ergodique* ^[9] (premier postulat) :

La moyenne dans le temps sur une durée Δt d'une variable mécanique d'un système thermodynamique est égale à la moyenne de cette même variable sur un ensemble statistique construit à partir du système, dans la limite où Δt et le nombre $\mathcal{N}$ de systèmes constituant l'ensemble statistique sont suffisamment grands.

Le deuxième postulat s'appelle le *principe d'équiprobabilité* et s'énonce :

Pour un système isolé à l'équilibre, tous les microétats accessibles sont équiprobables. Si Ω est le nombre de micro-états accessibles, la probabilité pour que le système soit dans un microétat donné vaut $\frac{1}{\Omega}$, quel que soit le microétat considéré.

Autrement dit, à énergie constante, il n'y a aucune raison de favoriser un microétat par rapport aux autres. Ω est appelé *nombre de complexions* du système.

1.3.3 Probabilités et valeurs moyennes

Soit un système macroscopique ayant accès à un très grand nombre de microétats. Ceux-ci, rangés dans un ordre quelconque, sont numérotés (indice j). Pour un microétat j donné, on notera E_j l'énergie du système, V_j son volume, et N_j le nombre de particules qu'il contient. Si parmi les variables extensives décrivant le macroétat du système, il en existe une autre X , on notera X_j sa valeur lorsque le système est dans le microétat j . De même, si le système contient des particules de différentes natures, les nombres de particules seront notées N_{1j} , N_{2j} , etc. Les démonstrations qui suivent concernent les systèmes décrits par les seules variables extensives E , V , N . Elles sont très facilement généralisables aux systèmes poly-constituants ou nécessitant l'emploi d'autres variables extensives. On supposera le problème mécanique résolu : tous les microétats accessibles au système sont répertoriés, et toutes les valeurs E_j , V_j , N_j sont connues.

L'objectif est de calculer les valeurs moyennes $\bar{E}$, $\bar{V}$ et $\bar{N}$. Dans une première étape, après calcul des P_j (P_j est la probabilité qu'un système pris au hasard dans l'ensemble soit dans le microétat j), ces valeurs moyennes sont exprimées en fonction de paramètres indéterminés provenant de la méthode des multiplicateurs de Lagrange. La deuxième étape confronte les résultats aux équations thermodynamiques, ce qui permet de relier les paramètres indéterminés aux variables intensives et de calculer l'entropie statistique en fonction des P_j .

⁹ L'énoncé du principe ergodique semble couler de source. En fait, le *problème ergodique* est un sujet qui est loin d'être clos à l'heure actuelle. Nous en reparlerons au paragraphe 1.5.5.

Le système étudié est le plus général qui soit : toutes les variables extensives sont libres de fluctuer. Il se trouve dans un bain très grand (capable d'accueillir $\mathcal{N}$ systèmes) fixant la température T , la pression p , et le potentiel chimique μ . Le contour du système est parfaitement conducteur de chaleur et de matière et permet les variations de volume.

L'ensemble est constitué de $\mathcal{N}$ systèmes séparés par des parois mobiles parfaitement perméables à la chaleur et à la matière. Tous les systèmes sont macroscopiquement identiques au système de départ : ils ont tous accès aux mêmes microétats. On place l'ensemble dans le bain, et une fois l'équilibre atteint, *on l'isole* et on le retire du bain. On obtient ainsi un *super-système isolé*, caractérisé par une énergie totale E_t , un volume total V_t et un nombre total de particules N_t .

On photographie à un instant donné le super-système. Celui-ci est dans un "super-microétat" $(j_1, \dots, j_{\mathcal{N}})$ donné (système 1 dans le microétat $j_1, \dots$, système $\mathcal{N}$ dans le microétat $j_{\mathcal{N}}$). On cherche la probabilité P_j qu'un système choisi au hasard sur une photo prise au hasard soit dans un microétat j donné.

On calcule dans un premier temps la probabilité d'avoir pris une photo donnée parmi toutes celles qu'il était possible de prendre, c'est-à-dire toutes celles compatibles avec (E_t, V_t, N_t) . Soit Ω_t le nombre total de super-microétats (ou de photos) compatibles avec (E_t, V_t, N_t) .

Le super-système étant isolé, on peut lui appliquer le *principe d'équiprobabilité* : tous les super-microétats compatibles avec (E_t, V_t, N_t) ont la même probabilité d'être observés. Si on numérote les photos possibles $\omega = 1$ à Ω_t , on a :

$$P(\omega) = \frac{1}{\Omega_t} \quad \forall \omega = 1 \text{ à } \Omega_t$$

Dans un deuxième temps, on tire sur une photo ω un système au hasard. Puisque $\mathcal{N}$ est très grand, la probabilité que ce système soit dans un microétat j donné vaut :

$$p_j(\omega) = \frac{n_j(\omega)}{\mathcal{N}} \quad \forall \omega = 1 \text{ à } \Omega_t$$

($n_j(\omega)$ est le nombre de systèmes dans l'état j sur la photo ω)

$$\text{On en déduit : } P_j = \sum_{\omega=1}^{\Omega_t} P(\omega) \times p_j(\omega) = \frac{1}{\mathcal{N}\Omega_t} \sum_{\omega=1}^{\Omega_t} n_j(\omega).$$

On peut réécrire cette somme en utilisant la notion de distribution. On classe les Ω_t photos et on les range dans différentes "boîtes". Chaque boîte contient toutes les photos se déduisant l'une de l'autre par permutation sur les systèmes. En clair, toutes les photos d'une boîte donnée possèdent la même liste de n_j . Cette liste $(n_1, \dots, n_j, \dots)$, appelée *distribution*, sera notée $[n]$. À chaque boîte correspond une distribution, et à chaque distribution correspond une boîte. On suppose qu'il y a suffisamment de photos dans chaque boîte pour y faire des moyennes statistiques. Toutes les distributions $[n] = ((n_1, \dots, n_j, \dots))$ possibles doivent être compatibles avec (E_t, V_t, N_t) :

$$(1) \quad \sum_j n_j = \mathcal{N} \qquad (2) \quad \sum_j E_j n_j = E_t = \mathcal{N} \bar{E}$$

$$(3) \quad \sum_j V_j n_j = V_t = \mathcal{N} \bar{V} \qquad (4) \quad \sum_j N_j n_j = N_t = \mathcal{N} \bar{N}$$

Le nombre de façons d'obtenir la distribution $[n]$ vaut $\Omega_{[n]} = \frac{\mathcal{N}!}{\prod_j n_j!}$.

Le nombre total de photos est la somme de tous les $\Omega_{[n]}$: $\Omega_t = \sum_{[n]}^{1,2,3,4} \Omega_{[n]}$

Le nombre total de systèmes dans le microétat j vaut : $\sum_{\omega=1}^{\Omega_t} n_j(\omega) = \sum_{[n]}^{1,2,3,4} \Omega_{[n]} n_j$

Notation : $\sum_{[n]}^{1,2,3,4}$ = somme sur les distributions vérifiant (1), (2), (3), (4).

Ce qui donne finalement : $P_j = \frac{1}{\mathcal{N}} \frac{\sum_{[n]}^{1,2,3,4} \Omega_{[n]} n_j}{\sum_{[n]}^{1,2,3,4} \Omega_{[n]}}$ avec $\Omega_{[n]} = \frac{\mathcal{N}!}{\prod_j n_j!}$

On utilise maintenant la *méthode du terme maximum* : parmi toutes les distributions $[n]$ possibles, seules quelques unes participent significativement aux sommes en numérateur et dénominateur. Ces distributions sont réparties autour de la distribution la plus probable, notée $[n^*]$. La méthode du terme maximum permet d'écrire :

$$P_j = \frac{1}{\mathcal{N}} \frac{\sum_{[n]}^{1,2,3,4} \Omega_{[n]} n_j}{\sum_{[n]}^{1,2,3,4} \Omega_{[n]}} \approx \frac{1}{\mathcal{N}} \frac{\Omega_{[n^*]} n_j^*}{\Omega_{[n^*]}} = \frac{n_j^*}{\mathcal{N}}$$

Cette distribution est celle qui rend $\ln \Omega_{[n]}$ maximum. C'est un *maximum sous contraintes* : la distribution $[n^*]$ recherchée doit vérifier les conditions (1), (2), (3), (4). On introduit donc quatre multiplicateurs de Lagrange $\alpha, \beta, \xi, \gamma$. Les n_j^* constitutifs de la distribution $[n^*]$ doivent vérifier :

$$\frac{\partial}{\partial n_j} \left(\ln \Omega_{[n]} - \alpha \sum_i n_i - \beta \sum_i E_i n_i - \xi \sum_i V_i n_i - \gamma \sum_i N_i n_i \right) = 0$$

Attention : avant de dériver, il faut écrire $\Omega_{[n]}$ uniquement en fonction des n_i :

$$\Omega_{[n]} = \frac{\mathcal{N}!}{\prod_i n_i!} = \frac{\left(\sum_i n_i \right)!}{\prod_i n_i!}$$

Cette précaution étant prise, le calcul conduit à : $n_j^* = \mathcal{N} e^{-\alpha} e^{-\beta E_j} e^{-\xi V_j} e^{-\gamma N_j}$.

Pour déterminer α, β, ξ et γ , on introduit le résultat obtenu dans les contraintes.

L'équation (1) conduit à : $e^\alpha = \sum_j e^{-\beta E_j} e^{-\xi V_j} e^{-\gamma N_j}$

C'est ce qu'on appelle une *fonction de partition*. On la note $I(\beta, \xi, \gamma)$ et on l'appelle *fonction de partition totale* car il en existe d'autres dont les sommations

portent sur des ensembles d'indices plus restreints :

$$e^\alpha = I(\beta, \xi, \gamma) = \sum_j e^{-\beta E_j} e^{-\xi V_j} e^{-\gamma N_j} \quad \text{et} \quad n_j^* = \mathcal{N} \frac{e^{-\beta E_j} e^{-\xi V_j} e^{-\gamma N_j}}{I(\beta, \xi, \gamma)}$$

Les équations (2), (3) et (4) conduisent à :

$$\bar{E} = \sum_j E_j e^{-\alpha} e^{-\beta E_j} e^{-\xi V_j} e^{-\gamma N_j} = \sum_j E_j \frac{e^{-\beta E_j} e^{-\xi V_j} e^{-\gamma N_j}}{I(\beta, \xi, \gamma)}$$

$$\bar{V} = \sum_j V_j e^{-\alpha} e^{-\beta E_j} e^{-\xi V_j} e^{-\gamma N_j} = \sum_j V_j \frac{e^{-\beta E_j} e^{-\xi V_j} e^{-\gamma N_j}}{I(\beta, \xi, \gamma)}$$

$$\bar{N} = \sum_j N_j e^{-\alpha} e^{-\beta E_j} e^{-\xi V_j} e^{-\gamma N_j} = \sum_j N_j \frac{e^{-\beta E_j} e^{-\xi V_j} e^{-\gamma N_j}}{I(\beta, \xi, \gamma)}$$

Ces trois expressions ainsi que celle de n_j^* conduisent toutes à :

$$P_j(\beta, \xi, \gamma) = e^{-\alpha} e^{-\beta E_j} e^{-\xi V_j} e^{-\gamma N_j} = \frac{e^{-\beta E_j} e^{-\xi V_j} e^{-\gamma N_j}}{I(\beta, \xi, \gamma)}$$

1.3.4 Confrontation avec la thermodynamique

Cette confrontation est nécessaire car l'introduction du résultat dans les contraintes n'a pas permis de déterminer $\alpha, \beta, \xi, \gamma$ en fonction de $\bar{E}, \bar{V}, \bar{N}$: on a juste exprimé $\bar{E}, \bar{V}, \bar{N}$ en fonction de $\alpha, \beta, \xi, \gamma$.

Point de départ de la confrontation : dans le but de relier les résultats de l'étude statistique à l'équation de Gibbs, on identifie les valeurs moyennes $\bar{E}, \bar{V}, \bar{N}$ aux variables E, V, N de la thermodynamique.

De l'expression de P_j on tire : $-\ln P_j = \beta E_j + \xi V_j + \gamma N_j + \alpha$

On multiplie les deux membres par P_j et on somme sur j . On obtient :

$$-\alpha - \sum_j P_j \ln P_j = \beta \bar{E} + \xi \bar{V} + \gamma \bar{N}$$

En thermodynamique on a : $\frac{S}{k} = \frac{1}{kT} E + \frac{p}{kT} V - \frac{\mu}{kT} N$. On est tenté d'écrire :

$$\frac{S}{k} = -\alpha - \sum_j P_j \ln P_j \quad \beta = \frac{1}{kT} \quad \xi = \frac{p}{kT} \quad \gamma = -\frac{\mu}{kT}$$

L'entropie trouvée est appelée *entropie statistique*. Les constantes α et k sont purement arbitraires. Leurs valeurs peuvent être choisies à notre convenance. Ce sont des constantes universelles, indépendantes du système étudié. Dans le cadre de la thermodynamique, on choisit les valeurs de k et α de sorte que *l'entropie statistique coïncide avec l'entropie thermodynamique*.

- α est choisie de façon à ce que l'entropie statistique soit additive.
Soit un système Σ d'entropie S résultant de la réunion de deux sous systèmes Σ' et Σ'' , d'entropies S' et S'' . Un micro-état du système Σ est caractérisé par 2 indices j' et j'' . Le premier caractérise le micro-état du système Σ' , et le deuxième celui du système Σ'' . La probabilité de trouver le système dans le micro-état $j'j''$ est égale à la probabilité de trouver le système Σ' dans le micro-état j' et le système Σ'' dans le micro-état j'' : $P_{j'j''} = P_{j'} \times P_{j''}$. Le calcul de S en fonction de S' et S'' conduit à :

$$\begin{aligned} \frac{S}{k} &= -\alpha - \sum_{j'} \sum_{j''} P_{j'j''} \ln P_{j'j''} = -\alpha - \sum_{j'} \sum_{j''} P_{j'} P_{j''} (\ln P_{j'} + \ln P_{j''}) \\ &= -\alpha - \sum_{j'} P_{j'} \ln P_{j'} \sum_{j''} P_{j''} - \sum_{j'} P_{j'} \sum_{j''} P_{j''} \ln P_{j''} \\ &= \left(-\alpha - \sum_{j'} P_{j'} \ln P_{j'} \right) + \left(-\alpha - \sum_{j''} P_{j''} \ln P_{j''} \right) + \alpha = \frac{S'}{k} + \frac{S''}{k} + \alpha \end{aligned}$$

Pour avoir $S = S' + S''$, on doit choisir $\alpha = 0$.

- La valeur de k , toujours dans le cadre de la thermodynamique, est calculée à partir de résultats expérimentaux. Sa valeur est indépendante du système étudié et de la technique expérimentale utilisée (c'est une *constante universelle*). L'expérience la plus simple consiste en des mesures de pressions d'un gaz parfait. Elle donne :

$$k = k_B = (1,380\,658 \pm 0,000\,012) 10^{-23} \text{ J/K} \quad \textbf{Constante de Boltzmann.}^{[10]}$$

Les résultats importants obtenus de cette confrontation :

$$\begin{aligned} \beta &= \frac{1}{kT} & \xi &= \frac{p}{kT} & \gamma &= -\frac{\mu}{kT} & S(\beta, \xi, \gamma) &= -k \sum_j P_j \ln P_j \\ P_j(\beta, \xi, \gamma) &= \frac{e^{-\beta E_j} e^{-\xi V_j} e^{-\gamma N_j}}{I(\beta, \xi, \gamma)} & I(\beta, \xi, \gamma) &= \sum_j e^{-\beta E_j} e^{-\xi V_j} e^{-\gamma N_j} \end{aligned}$$

Attention : $\alpha = 0$ est une valeur d'équilibre. Au voisinage de l'équilibre, c'est une fonction de β, ξ, γ . Si on veut étudier le comportement du système au voisinage de l'équilibre, il est nécessaire de conserver l'écriture $\alpha = \ln I(\beta, \xi, \gamma)$.

On identifie maintenant les valeurs moyennes et les variables thermodynamiques.

Le calcul a donné : $S/k = \beta \bar{E} + \xi \bar{V} + \gamma \bar{N} + \alpha = \beta \bar{E} + \xi \bar{V} + \gamma \bar{N} + \ln I(\beta, \xi, \gamma)$. En thermodynamique on a (voir paragraphe 1.1.2) : $S/k = \beta E + \xi V + \gamma N - \beta \Theta$

¹⁰ La constante de Boltzmann est la valeur de k utilisée en thermodynamique statistique. Son unité, le J/K, est celle de l'entropie thermodynamique. En théorie de l'information (voir paragraphe 1.4.5), l'entropie statistique est une grandeur sans dimension mais possédant une unité (le bit), et la valeur $k = 1/\ln 2$ choisie est telle que, dans le jeu "pile ou face", l'entropie statistique prend la valeur maximale $S = 1$ bit lorsque les deux événements "pile" et "face" sont équiprobables ($S(1/2, 1/2) = -(1/2 \ln 1/2 + 1/2 \ln 1/2)/\ln 2 = 1$), et la valeur $S = 0$ lorsque l'un des événements est certain et l'autre impossible (en posant $0 \ln 0 = 0$, $S(0, 1) = -(0 \ln 0 + 1 \ln 1)/\ln 2 = 0$).

On en déduit : $E = \bar{E} \quad V = \bar{V} \quad N = \bar{N} \Rightarrow -\beta\Theta(\beta, \xi, \gamma) = \ln I(\beta, \xi, \gamma)$

Cette relation fait le lien entre l'étude statistique et la thermodynamique. Elle relie la *fonction de partition* au *potentiel thermodynamique*. Ce potentiel thermodynamique, ici le potentiel nul, est la *fonction caractéristique* de l'ensemble étudié. Pour le distinguer des suivants, l'ensemble étudié ici sera nommé *ensemble T - p - μ* .

Le choix $k = k_B$ a permis d'identifier entropie thermodynamique et entropie statistique. Ce qui précède permet également d'identifier les grandeurs thermodynamiques E, V, N aux moyennes statistiques :

$$\begin{aligned} E(\beta, \xi, \gamma) &= \left. \frac{\partial \beta\Theta}{\partial \beta} \right]_{\xi, \gamma} = \bar{E}(\beta, \xi, \gamma) = - \left. \frac{\partial \ln I}{\partial \beta} \right]_{\xi, \gamma} = \sum_j P_j E_j \\ V(\beta, \xi, \gamma) &= \left. \frac{\partial \beta\Theta}{\partial \xi} \right]_{\gamma, \beta} = \bar{V}(\beta, \xi, \gamma) = - \left. \frac{\partial \ln I}{\partial \xi} \right]_{\gamma, \beta} = \sum_j P_j V_j \\ N(\beta, \xi, \gamma) &= \left. \frac{\partial \beta\Theta}{\partial \gamma} \right]_{\beta, \xi} = \bar{N}(\beta, \xi, \gamma) = - \left. \frac{\partial \ln I}{\partial \gamma} \right]_{\beta, \xi} = \sum_j P_j N_j \end{aligned}$$

Remarque sur les sommes sur j : Dans le calcul de $I(\beta, \xi, \gamma)$, on peut regrouper entre eux tous les micro-états j donnant les mêmes valeurs E, V, N à E_j, V_j, N_j . Si on note Ω_{EVN} le nombre de micro-états j tels que $E_j = E, V_j = V, N_j = N$, on obtient la transformée de Laplace de Ω_{EVN} :

$$I(\beta, \xi, \gamma) = \sum_{EVN} \sum_{j \in J_{EVN}} e^{-\beta E_j} e^{-\xi V_j} e^{-\gamma N_j} = \sum_{EVN} \Omega_{EVN} e^{-\beta E} e^{-\xi V} e^{-\gamma N} \quad (1.26)$$

Notation : J_{EVN} = ensemble des indices j donnant la même valeur E à l'énergie interne, la même valeur V au volume, et la même valeur N au nombre de particules.

Il nous reste à écrire les différentes probabilités :

– probabilité d'observer le micro-état j :

$$P_j(\beta, \xi, \gamma) = \frac{1}{I(\beta, \xi, \gamma)} e^{-\beta E_j} e^{-\xi V_j} e^{-\gamma N_j}$$

– probabilité de mesurer E, V et N simultanément :

$$P_{EVN}(\beta, \xi, \gamma) = \frac{\Omega_{EVN}}{I(\beta, \xi, \gamma)} e^{-\beta E} e^{-\xi V} e^{-\gamma N}$$

– probabilité de mesurer E et V , indépendamment de N :

$$P_{EV}(\beta, \xi, \gamma) = \frac{Q_{EV}(\gamma)}{I(\beta, \xi, \gamma)} e^{-\beta E} e^{-\xi V} \quad \text{avec} \quad Q_{EV}(\gamma) = \sum_N \Omega_{EVN} e^{-\gamma N}$$

– probabilité de mesurer E et N , indépendamment de V :

$$P_{EN}(\beta, \xi, \gamma) = \frac{Q_{EN}(\xi)}{I(\beta, \xi, \gamma)} e^{-\beta E} e^{-\gamma N} \quad \text{avec} \quad Q_{EN}(\xi) = \sum_V \Omega_{EVN} e^{-\xi V}$$

- probabilité de mesurer V et N , indépendamment de E :

$$P_{VN}(\beta, \xi, \gamma) = \frac{Q_{VN}(\beta)}{I(\beta, \xi, \gamma)} e^{-\xi V} e^{-\gamma N} \quad \text{avec} \quad Q_{VN}(\beta) = \sum_E \Omega_{EVN} e^{-\beta E}$$

- probabilité de mesurer E indépendamment de N et V :

$$P_E(\beta, \xi, \gamma) = \frac{Z_E(\xi, \gamma)}{I(\beta, \xi, \gamma)} e^{-\beta E} \quad \text{avec} \quad Z_E(\xi, \gamma) = \sum_{VN} \Omega_{EVN} e^{-\xi V} e^{-\gamma N}$$

- probabilité de mesurer V indépendamment de E et N :

$$P_V(\beta, \xi, \gamma) = \frac{Z_V(\beta, \gamma)}{I(\beta, \xi, \gamma)} e^{-\xi V} \quad \text{avec} \quad Z_V(\beta, \gamma) = \sum_{EN} \Omega_{EVN} e^{-\beta E} e^{-\gamma N}$$

- probabilité de mesurer N indépendamment de E et V :

$$P_N(\beta, \xi, \gamma) = \frac{Z_N(\beta, \xi)}{I(\beta, \xi, \gamma)} e^{-\gamma N} \quad \text{avec} \quad Z_N(\beta, \xi) = \sum_{EV} \Omega_{EVN} e^{-\beta E} e^{-\xi V}$$

Remarque : se fixer β, ξ, γ revient à se fixer T, p, μ . Il est donc possible de travailler avec le potentiel thermodynamique $\Theta(T, p, \mu)$ et la fonction de partition $I(T, p, \mu)$ exprimée en fonction des variables intensives énergétiques habituelles. On a donc le choix de travailler :

- avec les fonctions $\beta\Theta(\beta, \xi, \gamma)$ et $I(\beta, \xi, \gamma)$:

$$I(\beta, \xi, \gamma) = \sum_j e^{-\beta E_j} e^{-\xi V_j} e^{-\gamma N_j} = \sum_{EVN} \Omega_{EVN} e^{-\beta E} e^{-\xi V} e^{-\gamma N}$$

$$-\beta\Theta(\beta, \xi, \gamma) = \ln I(\beta, \xi, \gamma) \quad d(\beta\Theta) = E d\beta + V d\xi + N d\gamma$$

$$E = \left. \frac{\partial \beta\Theta}{\partial \beta} \right]_{\xi, \gamma} \quad V = \left. \frac{\partial \beta\Theta}{\partial \xi} \right]_{\gamma, \beta} \quad N = \left. \frac{\partial \beta\Theta}{\partial \gamma} \right]_{\beta, \xi} \quad \frac{S}{k} = \beta E + \xi V + \gamma N - \beta\Theta$$

- avec les fonctions $\Theta(T, p, \mu)$ et $I(T, p, \mu)$:

$$I(T, p, \mu) = \sum_j e^{-E_j/kT} e^{-pV_j/kT} e^{\mu N_j/kT} = \sum_{EVN} \Omega_{EVN} e^{-E/kT} e^{-pV/kT} e^{\mu N/kT}$$

$$\Theta(T, p, \mu) = -kT \ln I(T, p, \mu) \quad d\Theta = -S dT + V dp - N d\mu$$

$$S = - \left. \frac{\partial \Theta}{\partial T} \right]_{p, \mu} \quad V = + \left. \frac{\partial \Theta}{\partial p} \right]_{\mu, T} \quad N = - \left. \frac{\partial \Theta}{\partial \mu} \right]_{T, p} \quad E = \Theta + TS - pV + \mu N$$

NB : pour que ces équations soient utilisables, il est nécessaire que la fonction $I(\beta, \xi, \gamma)$ contienne un certain facteur d'échelle M (voir note de bas de page 7 relative aux équations 1.12 et 1.20).

1.4 Les ensembles statistiques

1.4.1 Introduction et conventions d'écriture

Le système étudié jusqu'à présent était caractérisé uniquement par les valeurs des variables intensives. Sa "taille" n'était donc pas fixée puisque toutes ses variables extensives étaient libres de fluctuer. On ne pourra utiliser ces résultats que si les données du problème imposent un certain facteur d'échelle M (voir équations 1.12 et 1.20).

C'est le cas en statistiques de réseaux (voir paragraphe 2.1) où le nombre M de sites disponibles fixe la taille du système et par conséquent les valeurs d'équilibre des variables extensives. Dans les autres cas pratiques, l'une au moins des variables extensives est fixée par les conditions expérimentales.

Dans un premier temps, nous envisageons les systèmes caractérisés par une variable extensive fixe, les deux autres étant libres de fluctuer. On imagine facilement un système constitué d'une partie de l'espace de volume V fixé où les particules sont libres d'entrer et sortir à leur guise (paragraphe 1.4.2). Ou un système constitué d'un nombre N fixé de particules pouvant occuper le volume qu'elles souhaitent (paragraphe 1.4.3).

Suivront ensuite les systèmes caractérisés par les deux variables N et V fixes (paragraphe 1.4.4), puis par les trois variables E , V , N fixes (paragraphe 1.4.5).

On peut aussi envisager des systèmes à énergie constante alors que N et/ou V sont libres de fluctuer (c'est possible avec des particules immatérielles comme les photons par exemple). Ces cas peu usités sont décrits au paragraphe 1.4.6.

Convention d'écriture : Les variables fluctuantes (à ensemble discret de valeurs) sont en indice, alors que les variables thermodynamiques sont mises entre parenthèses. On fera par exemple la distinction entre $Z_V(\beta, \gamma)$ où V est une variable discrète, et $Z(\beta, V, \gamma)$ où V est une variable thermodynamique, continue et macroscopique. On pourra dériver $Z(\beta, V, \gamma)$ par rapport à V , mais pas $Z_V(\beta, \gamma)$.

Une **mise en garde** au sujet de cette convention : à V constant, $Z_V(\beta, \gamma)$ devient $Z(\beta, V, \gamma)$; à N constant, $Z_N(\beta, \xi)$ devient $Z(\beta, \xi, N)$. Les deux fonctions obtenues portent le même symbole Z . Il est donc très important de faire figurer les variables dans les écritures : ce sont des fonctions différentes, qui portent d'ailleurs des noms différents.

Autre notation : soit une fonction $f(E, V, N)$ et la triple somme $\sum_{EVN} f(E, V, N)$.

Dans cette triple somme, on suppose implicitement que dans chaque terme de la somme, les trois valeurs E , V , N sont compatibles entre elles, autrement dit qu'elles peuvent être observées simultanément sur le système. Si on veut détailler dans quel ordre on effectue les trois sommations, on écrira par exemple :

$$\sum_{EVN} f(E, V, N) = \sum_N \sum_{V(N)} \sum_{E(V, N)} f(E(V, N), V(N), N)$$

La première sommation porte sur toutes les valeurs possibles de N , la deuxième sommation porte sur l'ensemble des volumes V possibles lorsque le nombre de particules vaut N , et la troisième sommation porte sur l'ensemble des énergies E possibles lorsque le nombre de particules vaut N et lorsque le volume vaut V .

1.4.2 L'ensemble grand canonique

C'est le cas où V est fixe : $V_j = V$ pour tout j . Le système est en contact avec un réservoir de chaleur et de particules fixant la température et le potentiel chimique, mais pas la pression.

On réécrit la fonction de partition totale 1.26 en posant $V_j = V$ et en supprimant la sommation sur V , sachant que les valeurs accessibles à E et N dépendent de la valeur de V choisie :

$$I(\beta, \xi, \gamma) = e^{-\xi V} \sum_{E(V), N(V)} \Omega_{EN}(V) e^{-\beta E(V)} e^{-\gamma N(V)} = e^{-\xi V} Z(\beta, V, \gamma)$$

On remarquera l'écriture $\Omega_{EN}(V)$ au lieu de $\Omega_{E(V), V, N(V)}$, signifiant que V est maintenant une variable thermodynamique.

$$\text{On en déduit : } \beta\Theta(\beta, \xi, \gamma) = -\ln I(\beta, \xi, \gamma) = \xi V - \ln Z(\beta, V, \gamma).$$

Cela signifie que $\beta\Theta(\beta, \xi, \gamma)$ est la transformée de Legendre de $\ln Z(\beta, V, \gamma)$ pour la variable V . En thermodynamique, $\beta\Theta(\beta, \xi, \gamma)$ est la transformée de Legendre de $-\beta\Psi(\beta, V, \gamma)$ pour la variable V .

$$\text{On a donc : } \beta\Psi(\beta, V, \gamma) = -\ln Z(\beta, V, \gamma).$$

La fonction $(\beta\Psi)$ est la fonction caractéristique de l'ensemble grand canonique. $Z(\beta, V, \gamma)$ est appelée grande fonction de partition :

$$Z(\beta, V, \gamma) = \sum_{E(V), N(V)} \Omega_{EN}(V) e^{-\beta E(V)} e^{-\gamma N(V)}$$

On pourra bien sûr utiliser le grand potentiel $\Psi(T, V, \mu)$:

$$\Psi(T, V, \mu) = -kT \ln Z(T, V, \mu)$$

$$\text{avec } Z(T, V, \mu) = \sum_{E(V), N(V)} \Omega_{EN}(V) e^{-E(V)/kT} e^{\mu N(V)/kT}$$

On peut maintenant écrire **les différentes probabilités** :

– d'observer le micro-état j :

$$P_j(\beta, V, \gamma) = \frac{e^{-\beta E_j(V)} e^{-\gamma N_j(V)}}{Z(\beta, V, \gamma)}$$

– de mesurer E et N simultanément :

$$P_{EN}(\beta, V, \gamma) = \frac{\Omega_{EN}(V) e^{-\beta E(V)} e^{-\gamma N(V)}}{Z(\beta, V, \gamma)}$$

– de mesurer E indépendamment de N :

$$P_E(\beta, V, \gamma) = \frac{Q_E(V, \gamma) e^{-\beta E(V)}}{Z(\beta, V, \gamma)} \quad \text{avec} \quad Q_E(V, \gamma) = \sum_{N(V)} \Omega_{EN}(V) e^{-\gamma N(V)}$$

- de mesurer N indépendamment de E :

$$P_N(\beta, V, \gamma) = \frac{Q_N(\beta, V) e^{-\gamma N(V)}}{Z(\beta, V, \gamma)} \quad \text{avec} \quad Q_N(\beta, V) = \sum_{E(V)} \Omega_{EN}(V) e^{-\beta E(V)}$$

Les principaux résultats :

- en utilisant $\beta\Psi(\beta, V, \gamma) = -\ln Z(\beta, V, \gamma)$:

$$E = - \left. \frac{\partial \ln Z}{\partial \beta} \right]_{V, \gamma} \quad \xi = + \left. \frac{\partial \ln Z}{\partial V} \right]_{\beta, \gamma} \quad N = - \left. \frac{\partial \ln Z}{\partial \gamma} \right]_{V, \beta}$$

$$\frac{S}{k} = \beta E + \gamma N + \ln Z = -\beta \left. \frac{\partial \ln Z}{\partial \beta} \right]_{V, \gamma} - \gamma \left. \frac{\partial \ln Z}{\partial \gamma} \right]_{V, \beta} + \ln Z$$

- en utilisant $\Psi(T, V, \mu) = -kT \ln Z(T, V, \mu)$:

$$S(T, V, \mu) = \left. \frac{\partial kT \ln Z}{\partial T} \right]_{V, \mu} = k \ln Z + kT \left. \frac{\partial \ln Z}{\partial T} \right]_{V, \mu}$$

$$p(T, V, \mu) = kT \left. \frac{\partial \ln Z}{\partial V} \right]_{T, \mu} \quad N(T, V, \mu) = kT \left. \frac{\partial \ln Z}{\partial \mu} \right]_{T, V}$$

$$E(T, V, \mu) = TS + \mu N - kT \ln Z = kT \left(\mu \left. \frac{\partial \ln Z}{\partial \mu} \right]_{T, V} + T \left. \frac{\partial \ln Z}{\partial T} \right]_{V, \mu} \right)$$

On se débrouille ensuite pour écrire $E(T, V, N)$, $V(T, p, N)$ et $H(T, p, N)$. On a ainsi accès aux capacités thermiques C_V et C_p , au coefficient de dilatation thermique isobare α et au module de compressibilité isotherme χ par les relations :

$$C_V = \left. \frac{\partial E}{\partial T} \right]_{V, N} \quad C_p = \left. \frac{\partial H}{\partial T} \right]_{p, N} \quad \alpha = \frac{1}{V} \left. \frac{\partial V}{\partial T} \right]_{p, N} \quad \chi = -\frac{1}{V} \left. \frac{\partial V}{\partial p} \right]_{T, N} \quad (1.27)$$

Une relation utile : $kT \frac{\partial}{\partial \mu} = -\frac{\partial}{\partial \gamma} = \lambda \frac{\partial}{\partial \lambda}$, avec $\lambda = e^{-\gamma} = e^{\mu/kT}$ (fugacité).

1.4.3 L'ensemble T-p

Le système est maintenant en contact avec un réservoir de chaleur muni d'un piston mobile fixant la température et la pression. C'est N qui est fixe : $N_j = N$ pour tout j . Le principe est rigoureusement le même que précédemment. Il y a juste permutation des rôles respectifs des variables V et N . On réécrit la fonction de partition totale 1.26 en remplaçant N_j par N (et en supprimant la sommation sur N), sachant que les valeurs accessibles à E et V dépendent de la valeur de N choisie :

$$I(\beta, \xi, \gamma) = e^{-\gamma N} \sum_{E(N), V(N)} \Omega_{EV}(N) e^{-\beta E(N)} e^{-\xi V(N)} = e^{-\gamma N} Z(\beta, \xi, N)$$

On en déduit : $\beta\Theta(\beta, \xi, \gamma) = -\ln I(\beta, \xi, \gamma) = \gamma N - \ln Z(\beta, \xi, N)$.

Cela signifie que $\beta\Theta(\beta, \xi, \gamma)$ est la transformée de Legendre de $\ln Z(\beta, \xi, N)$ pour la variable N . En thermodynamique, $\beta\Theta(\beta, \xi, \gamma)$ est la transformée de Legendre de $-\beta G(\beta, \xi, N)$ pour la variable N .

On a donc : $\beta G(\beta, \xi, N) = -\ln Z(\beta, \xi, N)$.

La fonction (βG) est la fonction caractéristique de l'ensemble T - p . $Z(\beta, \xi, N)$ est appelée fonction de partition T - p :

$$Z(\beta, \xi, N) = \sum_{E(N), V(N)} \Omega_{EV}(N) e^{-\beta E(N)} e^{-\xi V(N)}$$

Ce qui donne, en utilisant l'enthalpie libre $G(T, p, N)$:

$$G(T, p, N) = -kT \ln Z(T, p, N)$$

$$\text{avec } Z(T, p, N) = \sum_{E(N), V(N)} \Omega_{EV}(N) e^{-E(N)/kT} e^{-pV(N)/kT}$$

Les différentes probabilités :

– d'observer le micro-état j :

$$P_j(\beta, \xi, N) = \frac{e^{-\beta E_j(N)} e^{-\xi V_j(N)}}{Z(\beta, \xi, N)}$$

– de mesurer E et V simultanément :

$$P_{EV}(\beta, \xi, N) = \frac{\Omega_{EV}(N) e^{-\beta E(N)} e^{-\xi V(N)}}{Z(\beta, \xi, N)}$$

– de mesurer E indépendamment de V :

$$P_E(\beta, \xi, N) = \frac{Q_E(\xi, N) e^{-\beta E(N)}}{Z(\beta, \xi, N)} \quad \text{avec} \quad Q_E(\xi, N) = \sum_{V(N)} \Omega_{EV}(N) e^{-\xi V(N)}$$

– de mesurer V indépendamment de E :

$$P_V(\beta, \xi, N) = \frac{Q_V(\beta, N) e^{-\xi V(N)}}{Z(\beta, \xi, N)} \quad \text{avec} \quad Q_V(\beta, N) = \sum_{E(N)} \Omega_{EV}(N) e^{-\beta E(N)}$$

Les principaux résultats :

– en utilisant $\beta G(\beta, \xi, N) = -\ln Z(\beta, \xi, N)$:

$$E = - \left. \frac{\partial \ln Z}{\partial \beta} \right]_{\xi, N} \quad V = - \left. \frac{\partial \ln Z}{\partial \xi} \right]_{N, \beta} \quad \gamma = + \left. \frac{\partial \ln Z}{\partial N} \right]_{\beta, \xi}$$

$$\frac{S}{k} = \beta E + \xi V + \ln Z = -\beta \left. \frac{\partial \ln Z}{\partial \beta} \right]_{\xi, N} - \xi \left. \frac{\partial \ln Z}{\partial \xi} \right]_{N, \beta} + \ln Z$$

– en utilisant $G(T, p, N) = -kT \ln Z(T, p, N)$:

$$S(T, p, N) = \left. \frac{\partial kT \ln Z}{\partial T} \right]_{p, N} = k \ln Z + kT \left. \frac{\partial \ln Z}{\partial T} \right]_{p, N}$$

$$V(T, p, N) = -kT \left. \frac{\partial \ln Z}{\partial p} \right]_{T, N} \quad \mu(T, V, \mu) = -kT \left. \frac{\partial \ln Z}{\partial N} \right]_{T, p}$$

$$E(T, p, N) = TS - pV - kT \ln Z = kT \left(p \left. \frac{\partial \ln Z}{\partial p} \right]_{T, N} + T \left. \frac{\partial \ln Z}{\partial T} \right]_{p, N} \right)$$

Pour les utilisations de ces résultats, voir les équations 1.27.

Une relation utile : en posant $\eta = e^{-\xi V} = e^{-pV/kT}$ on a : $\frac{kT}{V} \frac{\partial}{\partial p} = \frac{1}{V} \frac{\partial}{\partial \xi} = -\eta \frac{\partial}{\partial \eta}$.

1.4.4 L'ensemble canonique

L'ensemble canonique est fabriqué à partir d'un système *isotherme*, *isochore*, *fermé* (T, V, N constants). On peut le fabriquer par exemple en partant de l'ensemble grand canonique. On réécrit $Z(\beta, V, \gamma)$ en ne considérant qu'une seule valeur de N . Les niveaux E accessibles deviennent fonction de V et N :

$$Z(\beta, V, \gamma) = e^{-\gamma N(V)} \sum_{E(V, N)} \Omega_E(V, N) e^{-\beta E(V, N)} = e^{-\gamma N(V)} Q(\beta, V, N)$$

On en déduit : $\beta \Psi(\beta, V, \gamma) = -\ln Z(\beta, V, \gamma) = \gamma N - \ln Q(\beta, V, N)$.

$\beta \Psi(\beta, V, \gamma)$ est donc la transformée de Legendre de $\ln Q(\beta, V, N)$ pour la variable N . En thermodynamique, $\beta \Psi(\beta, V, \gamma)$ est la transformée de Legendre de $-\beta F(\beta, V, N)$ pour la variable N . On a donc au choix :

$$\beta F(\beta, V, N) = -\ln Q(\beta, V, N) \quad \text{avec} \quad Q(\beta, V, N) = \sum_{E(V, N)} \Omega_E(V, N) e^{-\beta E(V, N)}$$

$$F(T, V, N) = -kT \ln Q(T, V, N) \quad \text{avec} \quad Q(T, V, N) = \sum_{E(V, N)} \Omega_E(V, N) e^{-E(V, N)/kT}$$

La fonction (βF) est la fonction caractéristique de l'ensemble canonique.

$Q(\beta, V, N)$ est appelée fonction de partition canonique.

Les différentes probabilités :

– d'observer le micro-état j :

$$P_j(\beta, V, N) = \frac{1}{Q(\beta, V, N)} e^{-\beta E_j(V, N)} \quad (1.28)$$

– de mesurer E :

$$P_E(\beta, V, N) = \frac{\Omega_E(V, N)}{Q(\beta, V, N)} e^{-\beta E(V, N)} \quad (1.29)$$

Les principaux résultats :

– en utilisant $\beta F(\beta, V, N) = -\ln Q(\beta, V, N)$:

$$E(\beta, V, N) = - \left. \frac{\partial \ln Q}{\partial \beta} \right]_{V, N} \quad \xi(\beta, V, N) = \left. \frac{\partial \ln Q}{\partial V} \right]_{\beta, N} \quad \gamma(\beta, V, N) = \left. \frac{\partial \ln Q}{\partial N} \right]_{\beta, V}$$

$$\frac{S(\beta, V, N)}{k} = \beta E + \ln Q = -\beta \left. \frac{\partial \ln Q}{\partial \beta} \right]_{V, N} + \ln Q$$

– en utilisant $F(T, V, N) = -kT \ln Q(T, V, N)$:

$$S(T, V, N) = \left. \frac{\partial kT \ln Q}{\partial T} \right]_{V, N} = k \ln Q + kT \left. \frac{\partial \ln Q}{\partial T} \right]_{V, N}$$

$$p(T, V, N) = kT \left. \frac{\partial \ln Q}{\partial V} \right]_{T, N} \quad \mu(T, V, N) = -kT \left. \frac{\partial \ln Q}{\partial N} \right]_{T, V}$$

$$E(T, V, N) = TS - kT \ln Q = kT^2 \left. \frac{\partial \ln Q}{\partial T} \right]_{V, N}$$

Pour les utilisations de ces résultats, voir les équations 1.27.

1.4.5 L'ensemble microcanonique

L'ensemble microcanonique est construit à partir d'un *système isolé* : aucune variable extensive n'est autorisée à fluctuer (E, V, N fixes). On peut par exemple le construire à partir de l'ensemble canonique. On réécrit la fonction $Q(\beta, V, N)$ en ne considérant qu'un seul niveau E . $\Omega_E(V, N)$ devient $\Omega(E, V, N)$, c'est-à-dire le nombre de façons d'observer E, V, N simultanément :

$$Q(\beta, V, N) = e^{-\beta E} \Omega(E, V, N)$$

$$\Rightarrow \beta F(\beta, V, N) = -\ln Q(\beta, V, N) = \beta E - \ln \Omega(E, V, N)$$

$\beta F(\beta, V, N)$ est donc la transformée de Legendre de $\ln \Omega(E, V, N)$ pour la variable E . Puisque en thermodynamique βF est la transformée de Legendre de S/k pour la variable E , on en déduit :

$$S(E, V, N) = k \ln \Omega(E, V, N) \quad \textbf{Équation de Boltzmann}$$

En partant de $S = -k \sum_j P_j \ln P_j$, on trouve la même chose.

Tous les états sont équiprobables, et il y en a Ω ; on a donc $P_j = \frac{1}{\Omega}$ pour tout j :

$$S = -k \sum_{j=1}^{\Omega} P_j \ln P_j = -k \sum_{j=1}^{\Omega} \frac{1}{\Omega} \ln \frac{1}{\Omega} = -k \Omega \times \frac{1}{\Omega} \ln \frac{1}{\Omega} = k \ln \Omega$$

On exprime maintenant les variables intensives en fonction des variables extensives :

$$\beta(E, V, N) = \left. \frac{\partial \ln \Omega}{\partial E} \right]_{V, N} \quad \xi(E, V, N) = \left. \frac{\partial \ln \Omega}{\partial V} \right]_{E, N} \quad \gamma(E, V, N) = \left. \frac{\partial \ln \Omega}{\partial N} \right]_{E, V}$$

Maintenant que nous connaissons l'expression de l'entropie en fonction des autres variables extensives, on peut reformuler de façon plus condensée la probabilité, dans chaque ensemble statistique, d'observer simultanément un jeu de valeurs données pour les variables fluctuantes.

Partons par exemple de la probabilité dans l'ensemble T - p - μ de mesurer E , V , N simultanément. On avait trouvé au paragraphe 1.3.4 :

$$P_{EVN}(\beta, \xi, \gamma) = \frac{\Omega_{EVN} e^{-\beta E} e^{-\xi V} e^{-\gamma N}}{I(\beta, \xi, \gamma)}$$

Soit $S_{EVN} = k \ln \Omega_{EVN}$. C'est l'entropie que posséderait le système si on limitait les microétats accessibles à ceux compatibles avec (E, V, N) . En remplaçant dans le résultat ci-dessus Ω_{EVN} par $e^{S_{EVN}/k}$ et $I(\beta, \xi, \gamma)$ par $e^{-\beta \Theta(\beta, \xi, \gamma)}$, on trouve :

$$P_{EVN}(\beta, \xi, \gamma) = \frac{e^{S_{EVN}/k - \beta E - \xi V - \gamma N}}{e^{-\beta \Theta(\beta, \xi, \gamma)}} = \frac{e^{-\beta \Theta_{EVN}(\beta, \xi, \gamma)}}{e^{-\beta \Theta(\beta, \xi, \gamma)}} = e^{-\beta \Delta \Theta_{EVN}}$$

$$\text{avec } \beta \Delta \Theta_{EVN} = \beta \Theta_{EVN}(\beta, \xi, \gamma) - \beta \Theta(\beta, \xi, \gamma)$$

$\Theta(\beta, \xi, \gamma)$ est le potentiel nul calculé à partir de la fonction de partition totale prenant en compte tous les états (E, V, N) accessibles au système. $\Theta_{EVN}(\beta, \xi, \gamma)$ est la valeur du potentiel nul lorsque seuls sont accessibles les microétats compatibles avec un jeu de valeurs (E, V, N) donné. Ce jeu de valeurs peut correspondre aux valeurs moyennes $(\bar{E}, \bar{V}, \bar{N})$, c'est-à-dire *l'état d'équilibre*. Si ce n'est pas le cas, on dira que le jeu (E, V, N) correspond à un *état excité* du système.

En opérant de même dans les autres ensembles, on trouve respectivement dans l'ensemble grand canonique, dans l'ensemble T - p et dans l'ensemble canonique :

$$P_{EN}(\beta, V, \gamma) = e^{-\beta \Delta \Psi_{EN}} \quad \text{avec} \quad \beta \Delta \Psi_{EN} = \beta \Psi_{EN}(\beta, V, \gamma) - \beta \Psi(\beta, V, \gamma)$$

$$P_{EV}(\beta, \xi, N) = e^{-\beta \Delta G_{EV}} \quad \text{avec} \quad \beta \Delta G_{EV} = \beta G_{EV}(\beta, \xi, N) - \beta G(\beta, \xi, N)$$

$$P_E(\beta, V, N) = e^{-\beta \Delta F_E} \quad \text{avec} \quad \beta \Delta F_E = \beta F_E(\beta, V, N) - \beta F(\beta, V, N)$$

Dans chaque cas, et on montre facilement que c'est valable quel que soit l'ensemble, la probabilité d'occupation d'un état excité est liée à la variation engendrée par l'excitation du potentiel thermodynamique caractéristique de l'ensemble statistique utilisé (celui prenant en compte toutes les excitations possibles du système).

Ce formalisme est à la base de nombreux modèles de la *physique statistique hors équilibre* dédiés à l'étude de l'évolution des systèmes vers un nouvel état d'équilibre lorsqu'on modifie les conditions expérimentales.

La théorie de l'information s'intéresse à l'information contenue dans un système probabiliste. Nous en donnons ici les grandes lignes dans le but de retrouver intuitivement l'équation de Boltzmann et de donner une définition plus générale de l'entropie.

Imaginons par exemple que j'ai une boule dans la main et que je la place dans l'une des N boîtes posées sur la table. Je ferme les boîtes et je pose la question : « Où est la boule ? ».

Hypothèses de départ :

- la boule est forcément dans une des N boîtes et dans une seule ;
- mon interlocuteur n'a pas vu dans quelle boîte j'ai mis la boule ;
- je ne suis ni menteur ni atteint de la maladie d'Alzheimer.

Étudions les différents cas possibles :

1. Toutes les boîtes sont identiques (rouges) sauf une (bleue), et je dis : « la boule est dans la boîte bleue ». La probabilité de trouver la boule vaut 1 ;
2. Parmi les N boîtes, p sont bleues et $N - p$ sont rouges, et je dis : « la boule est dans une boîte bleue ». La probabilité de trouver la boule vaut p/N .
3. Toutes les boîtes sont identiques : la probabilité de trouver la boule est $1/N$.

Il est clair qu'on a plus d'information dans le cas 1 que dans le cas 2, et plus d'information dans le cas 2 que dans le cas 3.

Dans cet exemple, on constate que la probabilité de trouver la boule augmente avec la quantité d'information que l'on possède. Pour quantifier l'information, on introduit une variable $\mathcal{I}$ représentant *l'information manquante*. On en donne ici les propriétés essentielles. Les N boîtes sont maintenant toutes identiques :

- L'information manquante augmente avec N : $N' > N \Rightarrow \mathcal{I}(N') > \mathcal{I}(N)$ (a)
- S'il n'y a qu'une boîte, l'information manquante est nulle : $\mathcal{I}(1) = 0$ (b)

Chacune des N boîtes est maintenant divisée en M compartiments. La question est : « dans quel compartiment est la boule ? ». Deux manières de la trouver :

- on a en tout $M \times N$ compartiments, d'où une information manquante $\mathcal{I}(M \times N)$;
- on cherche d'abord la boîte (information manquante $\mathcal{I}(N)$), puis le compartiment (information manquante $\mathcal{I}(M)$), soit au total $\mathcal{I}(N) + \mathcal{I}(M)$.

D'où la dernière propriété importante : $\mathcal{I}(M \times N) = \mathcal{I}(M) + \mathcal{I}(N)$ (c)

Les trois propriétés (a), (b), (c) ci-dessus sont caractéristiques des fonctions logarithmiques, ce qui permet d'écrire :

$$\mathcal{I}(N) = K \ln N$$

K est une constante arbitraire positive. Si on remplace la boule par un système thermodynamique isolé, et les N boîtes par les Ω micro-états accessibles à ce système, on obtient l'équation de Boltzmann $S(\Omega) = k \ln \Omega$. C'est pourquoi l'information manquante est appelée *entropie d'information*. Si on identifie S et $\mathcal{I}$, on dira que lorsque l'équilibre est atteint, l'entropie (l'information manquante) est maximale et que l'information (connue) est minimale. *L'entropie statistique* est une mesure de l'information manquante, autrement dit du *désordre* présent dans le système.

1.4.6 Les autres ensembles en variance 3

L'ensemble ξ - γ

Cet ensemble présente un intérêt pratique très limité (son nom n'a d'ailleurs rien d'officiel : les deux lettres grecques représentent les variables intensives entropiques fixes). Il permettra simplement d'introduire l'ensemble γ (un nom pas plus officiel que le premier) utilisable dans l'étude du corps noir.

Le système est en contact avec un réservoir de particules qui fixe ξ et γ . Les grandeurs V et N sont libres de fluctuer alors que E est fixe : $E_j = E$ pour tout j .

En partant de l'ensemble T - p - μ , on trouve :

$$Z(E, \xi, \gamma) = \sum_{V(E), N(E)} \Omega_{VN}(E) e^{-\xi V(E)} e^{-\gamma N(E)} \quad \beta E(E, \xi, \gamma) = \ln Z(E, \xi, \gamma)$$

$$\text{avec } \beta = \left. \frac{\partial \ln Z}{\partial E} \right]_{\xi, \gamma} \quad V = - \left. \frac{\partial \ln Z}{\partial \xi} \right]_{E, \gamma} \quad N = - \left. \frac{\partial \ln Z}{\partial \gamma} \right]_{E, \xi}$$

Attention : ξ et γ fixés ne signifie pas p et μ fixés. On ne peut donc pas travailler ici avec les variables p et μ .

L'ensemble γ

Le système étudié est maintenant tel que E et V sont constants, alors que le nombre de particules N est libre de fluctuer. C'est possible avec des particules immatérielles (comme les photons) qui ont la faculté d'apparaître ou disparaître à volonté. La seule variable fluctuante étant N , on notera $\Omega_N(E, V)$ le nombre de façons, pour E et V donnés, de placer N particules dans le système. Partant de l'ensemble ξ - γ , on trouve :

$$Q(E, V, \gamma) = \sum_{N(E, V)} \Omega_N(E, V) e^{-\gamma N(E, V)} \quad \beta H(E, V, \gamma) = \ln Q(E, V, \gamma)$$

$$\text{avec } \beta = \left. \frac{\partial \ln Q}{\partial E} \right]_{V, \gamma} \quad \xi = \left. \frac{\partial \ln Q}{\partial V} \right]_{E, \gamma} \quad N = - \left. \frac{\partial \ln Q}{\partial \gamma} \right]_{E, V}$$

Avec la même remarque que précédemment : en fixant γ , on ne fixe pas le potentiel chimique μ , mais uniquement μ/kT . On ne peut donc pas travailler avec la variable μ .

L'ensemble ξ

Pour que la liste soit complète, il faut ajouter un dernier ensemble (d'aucune utilité pratique) fabriqué à partir d'un système caractérisé par les variables E , ξ et N . La seule variable fluctuante étant V , on notera $\Omega_V(E, N)$ le nombre de façons, pour E et N donnés, d'observer le volume V . Partant de l'ensemble ξ - γ , on trouve :

$$Q(E, \xi, N) = \sum_{V(E, N)} \Omega_V(E, N) e^{-\xi V(E, N)} \quad \beta \Phi(E, \xi, N) = \ln Q(E, \xi, N)$$

$$\text{avec } \beta = \left. \frac{\partial \ln Q}{\partial E} \right]_{\xi, N} \quad V = - \left. \frac{\partial \ln Q}{\partial \xi} \right]_{E, N} \quad \gamma = \left. \frac{\partial \ln Q}{\partial N} \right]_{E, \xi}$$

1.4.7 Synthèse des résultats

On peut reprendre les résultats, en partant cette fois de l'ensemble microcanonique. Celui-ci est fabriqué à partir d'un système où toutes les variables extensives sont fixes.

À partir de l'ensemble microcanonique, on fabrique les ensembles à une variable extensive fluctuante (ensemble canonique, ensemble γ , ensemble ξ). La variable intensive entropique associée à la variable fluctuante est fixe. La *fonction caractéristique* (FC) de l'ensemble obtenu est une transformée de Legendre de S/k . La *fonction de partition* (FP) est une transformée de Laplace du nombre de complexions Ω . La variable transformée est dans les deux transformations la variable extensive fluctuante.

$$\text{À retenir : } (\text{FC} = \text{TdLegendre de } S/k) = -\ln (\text{FP} = \text{TdLaplace de } \Omega) \quad (1.30)$$

Les ensembles à deux variables extensives fluctuantes (ensemble grand-canonique, ensemble T - p , ensemble ξ - γ) se fabriquent par double ou simple transformée suivant l'ensemble de départ (attention aux changements de signe, voir relation 1.4).

Et enfin la dernière étape : la fabrication de l'ensemble T - p - μ , par triple, double ou simple transformée suivant l'ensemble de départ.

Dans les formules qui suivent, les transformées de Legendre (à gauche) et de Laplace (à droite) associées sont reliées par le symbole $\longleftrightarrow$.

Ensemble microcanonique

$$S(E, V, N)/k = \ln \Omega(E, V, N) \quad dS/k = d \ln \Omega = \beta dE + \xi dV + \gamma dN$$

Ensemble canonique

$$\begin{aligned} +\beta F(\beta, V, N) &= \beta E - S/k \longleftrightarrow Q(\beta, V, N) = \sum_E \Omega(E, V, N) e^{-\beta E} \\ +\beta F(\beta, V, N) &= -\ln Q(\beta, V, N) \quad d(+\beta F) = d(-\ln Q) = +E d\beta - \xi dV - \gamma dN \end{aligned}$$

Ensemble γ

$$\begin{aligned} -\beta H(E, V, \gamma) &= \gamma N - S/k \longleftrightarrow Q(E, V, \gamma) = \sum_N \Omega(E, V, N) e^{-\gamma N} \\ -\beta H(E, V, \gamma) &= -\ln Q(E, V, \gamma) \quad d(-\beta H) = d(-\ln Q) = -\beta dE - \xi dV + N d\gamma \end{aligned}$$

Ensemble ξ

$$\begin{aligned} -\beta \Phi(E, \xi, N) &= \xi V - S/k \longleftrightarrow Q(E, \xi, N) = \sum_V \Omega(E, V, N) e^{-\xi V} \\ -\beta \Phi(E, \xi, N) &= -\ln Q(E, \xi, N) \quad d(-\beta \Phi) = d(-\ln Q) = -\beta dE + V d\xi - \gamma dN \end{aligned}$$

Ensemble grand-canonique

$$\begin{aligned} +\beta \Psi(\beta, V, \gamma) &= \beta E + \gamma N - S/k \longleftrightarrow Z(\beta, V, \gamma) = \sum_{EN} \Omega(E, V, N) e^{-\beta E} e^{-\gamma N} \\ +\beta \Psi(\beta, V, \gamma) &= \gamma N - (-\beta F) \longleftrightarrow Z(\beta, V, \gamma) = \sum_N Q(\beta, V, N) e^{-\gamma N} \\ +\beta \Psi(\beta, V, \gamma) &= \beta E - (+\beta H) \longleftrightarrow Z(\beta, V, \gamma) = \sum_E Q(E, V, \gamma) e^{-\beta E} \\ +\beta \Psi(\beta, V, \gamma) &= -\ln Z(\beta, V, \gamma) \quad d(+\beta \Psi) = d(-\ln Z) = +E d\beta - \xi dV + N d\gamma \end{aligned}$$

Ensemble T-p

$$+\beta G(\beta, \xi, N) = \beta E + \xi V - S/k \longleftrightarrow Z(\beta, \xi, N) = \sum_{EV} \Omega(E, V, N) e^{-\beta E} e^{-\xi V}$$

$$+\beta G(\beta, \xi, N) = \xi V - (-\beta F) \longleftrightarrow Z(\beta, \xi, N) = \sum_V Q(\beta, V, N) e^{-\xi V}$$

$$+\beta G(\beta, \xi, N) = \beta E - (+\beta \Phi) \longleftrightarrow Z(\beta, \xi, N) = \sum_E Q(E, \xi, N) e^{-\beta E}$$

$$+\beta G(\beta, \xi, N) = -\ln Z(\beta, \xi, N) \quad d(+\beta G) = d(-\ln Z) = +E d\beta + V d\xi - \gamma dN$$

Ensemble ξ - γ

$$-\beta E(E, \xi, \gamma) = \xi V + \gamma N - S/k \longleftrightarrow Z(E, \xi, \gamma) = \sum_{VN} \Omega(E, V, N) e^{-\xi V} e^{-\gamma N}$$

$$-\beta E(E, \xi, \gamma) = \xi V - (+\beta H) \longleftrightarrow Z(E, \xi, \gamma) = \sum_V Q(E, V, \gamma) e^{-\xi V}$$

$$-\beta E(E, \xi, \gamma) = \gamma N - (+\beta \Phi) \longleftrightarrow Z(E, \xi, \gamma) = \sum_N Q(E, \xi, N) e^{-\gamma N}$$

$$-\beta E(E, \xi, \gamma) = -\ln Z(E, \xi, \gamma) \quad d(-\beta E) = d(-\ln Z) = -\beta dE + V d\xi + N d\gamma$$

Ensemble T-p- μ

$$+\beta \Theta(\beta, \xi, \gamma) = \beta E + \xi V + \gamma N - S/k \leftrightarrow I(\beta, \xi, \gamma) = \sum_{EVN} \Omega(E, V, N) e^{-\beta E} e^{-\xi V} e^{-\gamma N}$$

$$+\beta \Theta(\beta, \xi, \gamma) = \xi V + \gamma N - (-\beta F) \longleftrightarrow I(\beta, \xi, \gamma) = \sum_{VN} Q(\beta, V, N) e^{-\xi V} e^{-\gamma N}$$

$$+\beta \Theta(\beta, \xi, \gamma) = \beta E + \xi V - (+\beta H) \longleftrightarrow I(\beta, \xi, \gamma) = \sum_{EV} Q(E, V, \gamma) e^{-\beta E} e^{-\xi V}$$

$$+\beta \Theta(\beta, \xi, \gamma) = \beta E + \gamma N - (+\beta \Phi) \longleftrightarrow I(\beta, \xi, \gamma) = \sum_{EN} Q(E, \xi, N) e^{-\beta E} e^{-\gamma N}$$

$$+\beta \Theta(\beta, \xi, \gamma) = \xi V - (-\beta \Psi) \longleftrightarrow I(\beta, \xi, \gamma) = \sum_V Z(\beta, V, \gamma) e^{-\xi V}$$

$$+\beta \Theta(\beta, \xi, \gamma) = \gamma N - (-\beta G) \longleftrightarrow I(\beta, \xi, \gamma) = \sum_N Z(\beta, \xi, N) e^{-\gamma N}$$

$$+\beta \Theta(\beta, \xi, \gamma) = \beta E - (+\beta E) \longleftrightarrow I(\beta, \xi, \gamma) = \sum_E Z(E, \xi, \gamma) e^{-\beta E}$$

$$+\beta \Theta(\beta, \xi, \gamma) = -\ln Z(\beta, \xi, \gamma) \quad d(+\beta \Theta) = d(-\ln I) = +E d\beta + V d\xi + N d\gamma$$

1.4.8 Ensembles en variance supérieure à 3**Systèmes polyconstituants**

Dans un système polyconstituant, les particules sont réparties en p espèces. On introduit p nombres de particules $N_1 \dots N_p$, p potentiels chimiques $\mu_1 \dots \mu_p$ et p paramètres $\gamma_i = -\mu_i/kT$ ($i = 1$ à p).

Aux huit ensembles précédents (où on remplace les sommes sur N par p sommes sur $N_1 \dots N_p$), il faut ajouter les ensembles construits à partir de systèmes ouverts

pour certains constituants et fermés pour les autres. On donne ci-dessous les principaux ensembles relatifs à un système biconstituant.

L'ensemble canonique a pour fonction caractéristique $+\beta F(\beta, V, N_1, N_2)$ et pour fonction de partition $Q(\beta, V, N_1, N_2)$, avec les relations :

$$\begin{aligned} +\beta F(\beta, V, N_1, N_2) &= \beta E - S/k = -\ln Q(\beta, V, N_1, N_2) \\ d(+\beta F) &= d(-\ln Q) = +E d\beta - \xi dV - \gamma_1 dN_1 - \gamma_2 dN_2 \end{aligned}$$

$$Q(\beta, V, N_1, N_2) = \sum_E \Omega(E, V, N_1, N_2) e^{-\beta E}$$

L'ensemble grand-canonique a pour fonction caractéristique $+\beta \Psi(\beta, V, \gamma_1, \gamma_2)$ et pour fonction de partition $Z(\beta, V, \gamma_1, \gamma_2)$, avec les relations :

$$\begin{aligned} +\beta \Psi(\beta, V, \gamma_1, \gamma_2) &= \beta E + \gamma_1 N_1 + \gamma_2 N_2 - S/k = -\ln Z(\beta, V, \gamma_1, \gamma_2) \\ d(+\beta \Psi) &= d(-\ln Z) = +E d\beta - \xi dV + N_1 d\gamma_1 + N_2 d\gamma_2 \end{aligned}$$

$$Z(\beta, V, \gamma_1, \gamma_2) = \sum_{EN_1N_2} \Omega(E, V, N_1, N_2) e^{-\beta E} e^{-\gamma_1 N_1} e^{-\gamma_2 N_2}$$

L'ensemble T - μ_1 est construit à partir d'un système isotherme, ouvert pour le constituant 1, fermé pour le constituant 2. Soit $+\beta \Phi_1(\beta, V, \gamma_1, N_2)$ sa fonction caractéristique, et $\Delta_1(\beta, V, \gamma_1, N_2)$ sa fonction de partition :

$$\begin{aligned} +\beta \Phi_1(\beta, V, \gamma_1, N_2) &= \beta E + \gamma_1 N_1 - S/k = -\ln \Delta_1(\beta, V, \gamma_1, N_2) \\ d(+\beta \Phi_1) &= d(-\ln \Delta_1) = +E d\beta - \xi dV + N_1 d\gamma_1 - \gamma_2 dN_2 \end{aligned}$$

$$\Delta_1(\beta, V, \gamma_1, N_2) = \sum_{EN_1} \Omega(E, V, N_1, N_2) e^{-\beta E} e^{-\gamma_1 N_1}$$

L'ensemble T - μ_2 est construit à partir d'un système isotherme, fermé pour le constituant 1, ouvert pour le constituant 2. Soit $+\beta \Phi_2(\beta, V, N_1, \gamma_2)$ sa fonction caractéristique, et $\Delta_2(\beta, V, N_1, \gamma_2)$ sa fonction de partition :

$$\begin{aligned} +\beta \Phi_2(\beta, V, N_1, \gamma_2) &= \beta E + \gamma_2 N_2 - S/k = -\ln \Delta_2(\beta, V, N_1, \gamma_2) \\ d(+\beta \Phi_2) &= d(-\ln \Delta_2) = +E d\beta - \xi dV - \gamma_1 dN_1 + N_2 d\gamma_2 \end{aligned}$$

$$\Delta_2(\beta, V, N_1, \gamma_2) = \sum_{EN_2} \Omega(E, V, N_1, N_2) e^{-\beta E} e^{-\gamma_2 N_2}$$

Autre contribution au travail que pV

Même principe s'il y a une autre contribution au travail que pV . Soit (X, P) le couple de variables supplémentaires, avec $\phi = -P/kT$.

Aux huit ensembles précédents (où on remplace la somme sur V par une double somme sur V et X), il faut ajouter les ensembles construits à partir de systèmes à V et ϕ constants et de systèmes à X et ξ constants.

Ci-dessous les ensembles les plus utilisés.

L'ensemble canonique a pour fonction caractéristique $+\beta F(\beta, V, X, N)$ et pour fonction de partition $Q(\beta, V, X, N)$, avec les relations :

$$+\beta F(\beta, V, X, N) = \beta E - S/k = -\ln Q(\beta, V, X, N)$$

$$d(+\beta F) = d(-\ln Q) = +E d\beta - \xi dV - \phi dX - \gamma dN$$

$$Q(\beta, V, X, N) = \sum_E \Omega(E, V, X, N) e^{-\beta E}$$

L'ensemble T - p - P a pour fonction caractéristique $+\beta G(\beta, \xi, \phi, N)$ et pour fonction de partition $Z(\beta, \xi, \phi, N)$, avec les relations :

$$+\beta G(\beta, \xi, \phi, N) = \beta E + \xi V + \phi X - S/k = -\ln Z(\beta, \xi, \phi, N)$$

$$d(+\beta G) = d(-\ln Z) = +E d\beta + V d\xi + X d\phi - \gamma dN$$

$$Z(\beta, V, \xi, \phi, N) = \sum_{EVX} \Omega(E, V, X, N) e^{-\beta E} e^{-\xi V} e^{-\phi X}$$

L'ensemble T - p est construit à partir d'un système isotherme, isobare, avec X fixé. Soit $+\beta \Phi_V(\beta, \xi, X, N)$ sa fonction caractéristique, et $\Delta_V(\beta, \xi, X, N)$ sa fonction de partition :

$$+\beta \Phi_V(\beta, \xi, X, N) = \beta E + \xi V - S/k = -\ln \Delta_V(\beta, \xi, X, N)$$

$$d(+\beta \Phi_V) = d(-\ln \Delta_V) = +E d\beta + V d\xi - \phi dX - \gamma dN$$

$$\Delta_V(\beta, \xi, X, N) = \sum_{EV} \Omega(E, V, X, N) e^{-\beta E} e^{-\xi V}$$

L'ensemble T - P est construit à partir d'un système isotherme, isochore, en contact avec un réservoir qui fixe P . Soit $+\beta \Phi_X(\beta, V, \phi, N)$ sa fonction caractéristique, et $\Delta_X(\beta, V, \phi, N)$ sa fonction de partition :

$$+\beta \Phi_X(\beta, V, \phi, N) = \beta E + \phi X - S/k = -\ln \Delta_X(\beta, V, \phi, N)$$

$$d(+\beta \Phi_X) = d(-\ln \Delta_X) = +E d\beta - \xi dV + X d\phi - \gamma dN$$

$$\Delta_X(\beta, V, \phi, N) = \sum_{EX} \Omega(E, V, X, N) e^{-\beta E} e^{-\phi X}$$

1.5 Le deuxième principe et le troisième principe

1.5.1 Équivalence entre ensembles statistiques

Fluctuations de V autour de $\bar{V}$ dans l'ensemble T - p

Soit le volume moyen d'un système isotherme, isobare, fermé :

$$\bar{V}(\beta, \xi, N) = \sum_V V P_V(\beta, \xi, N) = \frac{1}{Z(\beta, \xi, N)} \sum_V V Q_V(\beta, N) e^{-\xi V}$$

Avec $Z(\beta, \xi, N) = \sum_V Q_V(\beta, N) e^{-\xi V}$, on obtient :

$$\bar{V}(\beta, \xi, N) \sum_V Q_V(\beta, N) e^{-\xi V} = \sum_V V Q_V(\beta, N) e^{-\xi V}$$

On dérive par rapport à ξ et on trouve (χ est la compressibilité) :

$$\Delta V^2 = \overline{V^2} - \bar{V}^2 = - \left. \frac{\partial \bar{V}}{\partial \xi} \right]_{\beta, N} = -kT \left. \frac{\partial \bar{V}}{\partial p} \right]_{T, N} = kT \chi \bar{V} \Rightarrow \frac{\Delta V}{\bar{V}} = \sqrt{\frac{kT \chi}{\bar{V}}}$$

Fluctuations de N autour de $\bar{N}$ dans l'ensemble grand-canonique

Soit le nombre moyen de particules dans un système isotherme, isochore, ouvert :

$$\bar{N}(\beta, V, \gamma) = \sum_N N P_N(\beta, V, \gamma) = \frac{1}{Z(\beta, V, \gamma)} \sum_N N Q_N(\beta, V) e^{-\gamma N}$$

Avec $Z(\beta, V, \gamma) = \sum_N Q_N(\beta, V) e^{-\gamma N}$, on obtient :

$$\bar{N}(\beta, V, \gamma) \sum_N Q_N(\beta, V) e^{-\gamma N} = \sum_N N Q_N(\beta, V) e^{-\gamma N}$$

On dérive par rapport à γ et on trouve :

$$\Delta N^2 = - \left. \frac{\partial \bar{N}}{\partial \gamma} \right]_{\beta, V} = kT \left. \frac{\partial \bar{N}}{\partial \mu} \right]_{T, V} \Rightarrow \frac{\Delta N}{\bar{N}} = \sqrt{\frac{kT}{\bar{N}^2} \left. \frac{\partial \bar{N}}{\partial \mu} \right]_{T, V}} = \sqrt{\frac{kT \chi}{\bar{V}}}$$

La dernière égalité a été obtenue en utilisant les déterminants de Jacobi et une relation de Maxwell.

Fluctuations de E autour de $\bar{E}$ dans l'ensemble canonique

Même principe que précédemment. On part de :

$$\bar{E}(\beta, V, N) = \sum_E E P_E(\beta, V, N) = \frac{1}{Q(\beta, V, N)} \sum_E E \Omega_E(V, N) e^{-\beta E}$$

Avec $Q(\beta, V, N) = \sum_E \Omega_E(V, N) e^{-\beta E}$, on obtient :

$$\bar{E}(\beta, V, N) \sum_E \Omega_E(V, N) e^{-\beta E} = \sum_E E \Omega_E(V, N) e^{-\beta E}$$

On dérive par rapport à β et on trouve :

$$\Delta E^2 = - \left. \frac{\partial \bar{E}}{\partial \beta} \right]_{V, N} = kT^2 \left. \frac{\partial \bar{E}}{\partial T} \right]_{V, N} = kT^2 C_V \Rightarrow \frac{\Delta E}{\bar{E}} = \frac{kT}{\bar{E}} \sqrt{\frac{C_V}{k}}$$

Équivalence entre ensembles

Le gaz parfait étant l'un des systèmes les plus compressibles qui soient, on peut l'utiliser pour évaluer une borne supérieure aux fluctuations. En utilisant les résultats du paragraphe 3.3.1, on trouve pour les trois exemples ci-dessus :

$$\frac{\Delta V}{\bar{V}} = \frac{1}{\sqrt{N}} \quad \frac{\Delta N}{\bar{N}} = \frac{1}{\sqrt{N}} \quad \frac{\Delta E}{\bar{E}} = \sqrt{\frac{2}{3N}}$$

Les fluctuations autour des valeurs moyennes sont extrêmement faibles, pour ainsi dire nulles

On peut montrer que, quel que soit l'ensemble statistique, et quelle que soit la variable X fluctuante, la quantité $\Delta X/\bar{X}$ varie toujours en $1/\sqrt{N}$. Les fluctuations sont si faibles qu'on peut les ignorer. C'est un résultat de la thermodynamique également : à l'équilibre tous les flux étant nuls, toutes les variables extensives sont constantes à l'équilibre, même si elles sont autorisées à varier. Un système possède à l'équilibre une énergie constante *même s'il n'est pas isolé*, une composition constante *même s'il est ouvert*, un volume constant *même s'il est expansible*, etc. Comme en thermodynamique où tous les potentiels sont équivalents :

Tous les ensembles statistiques sont équivalents thermodynamiquement

Le choix de l'ensemble statistique est arbitraire, non imposé par l'environnement du système. On peut indifféremment écrire : $S = k \ln \Omega$ ou $S = -k \sum_j P_j \ln P_j$.

Dans la dernière expression, quel que soit l'ensemble choisi pour dénombrer les états, on trouve le même résultat : les termes significatifs dans la somme sur j correspondent dans tous les cas à des états d'énergie, de volume et de nombre de particules très proches de $\bar{E}$, $\bar{V}$ et $\bar{N}$. C'est tout simplement une conséquence de l'approximation du terme maximum. Partant par exemple de l'ensemble T - p - μ :

$$\begin{aligned} S(\beta, \xi, \gamma) &= -k \sum_j P_j \ln P_j = -k \sum_{E\bar{V}\bar{N}} \sum_{j \in J_{E\bar{V}\bar{N}}} P_j \ln P_j \approx -k \sum_{j \in J_{\bar{E}\bar{V}\bar{N}}} P_j \ln P_j \\ &= -k \sum_{j \in J_{\bar{E}\bar{V}\bar{N}}} \frac{1}{\Omega_{\bar{E}\bar{V}\bar{N}}} \ln \frac{1}{\Omega_{\bar{E}\bar{V}\bar{N}}} = k \ln \Omega_{\bar{E}\bar{V}\bar{N}} = S(\bar{E}, \bar{V}, \bar{N}) \end{aligned}$$

Seule distinction suivant l'ensemble utilisé : l'entropie S sera fonction de E ou $\bar{E}$, V ou $\bar{V}$, N ou $\bar{N}$ suivant que telle ou telle variable est fixe ou fluctuante.

1.5.2 Le deuxième principe

Considérons une enceinte isolée. Une paroi la divise en deux compartiments, l'un vide, l'autre rempli d'un gaz sous une certaine pression. Le système ainsi constitué est supposé en équilibre. La pression n'étant pas homogène, ce n'est pas l'équilibre, mais un *équilibre sous contrainte*. Les états liés à la présence de gaz dans l'enceinte vide sont dits *inaccessibles*. On ne peut pas les qualifier d'impossibles puisque la suppression de la paroi les rendra *accessibles*.

Autre exemple : soit une enceinte isolée remplie d'un mélange de molécules d'hydrogène et d'azote. Si on suppose l'équilibre atteint, est-ce vraiment l'équilibre ? Non, puisque rigoureusement, il faut envisager par exemple la formation de molécules d'ammoniac NH_3 . Bien que la cinétique de réaction soit infiniment lente, les états liés à la présence d'ammoniac ne sont pas impossibles puisque l'ajout d'un catalyseur peut provoquer la réaction. Ce sont des états *inaccessibles*, que l'on peut

rendre *accessibles* par l'ajout du catalyseur. Dans cet exemple, la cinétique infiniment lente tient lieu de "paroi", et l'ajout du catalyseur de "suppression de la paroi".

La frontière entre états accessibles et états inaccessibles est assez floue. Si dans notre premier exemple la paroi laisse passer une molécule par jour, les états liés à la présence de gaz dans le compartiment vide peuvent être considérés comme accessibles ou inaccessibles suivant la durée de l'expérience.

En thermodynamique, le deuxième principe s'écrit : « *dans un système isolé et en équilibre sous contraintes, la suppression d'une contrainte entraîne un déplacement spontané de l'équilibre, accompagné obligatoirement d'un accroissement d'entropie* ».

Dans l'équation de Boltzmann $S = k \ln \Omega$, Ω est le nombre de micro-états accessibles au système. Lorsqu'on supprime une contrainte interne, Ω (donc S) ne peut qu'augmenter : les micro-états jusque là accessibles restent accessibles, et un certain nombre de micro-états jusque là inaccessibles deviennent accessibles. L'équation de Boltzmann est donc bien en accord avec le deuxième principe.

Autre manière de voir les choses : calculons la probabilité d'un processus spontané à $\Delta S < 0$. Un gaz parfait de N atomes occupe une enceinte de volume V . Quelle est la probabilité que, spontanément, tous les atomes se trouvent dans la même moitié de l'enceinte ? Réponse : $(1/2)^N$. Rien qu'avec $N = 100$, si vous voulez avoir une chance de l'observer pendant une picoseconde, il vous faudra attendre ... 40 milliards d'années. Avec $N = 10^{20}$, vous pouvez toujours attendre, ça n'arrivera JAMAIS.

1.5.3 Équivalence entre S maxi et E, F, G, H mini

Cette fois le système n'est plus isolé : les processus envisagés ne sont plus spontanés. On considère par exemple un système fermé, isotherme, isochore et soumis à une contrainte interne. Les grandeurs N et V sont fixées, alors que l'énergie interne fluctue autour d'une certaine valeur d'équilibre.

Le système possède un certain nombre de niveaux accessibles, et sur chaque niveau un certain nombre de micro-états. Lorsqu'on retire la contrainte, on peut avoir, comme pour le système isolé, augmentation niveau par niveau du nombre de micro-états accessibles, mais aussi apparition de nouveaux niveaux accessibles. Pour les niveaux initialement accessibles, on note $\Omega(E)$ et $\Omega'(E)$ les nombres de micro-états du niveau E avant et après suppression de la contrainte. Pour les niveaux initialement inaccessibles, même notation, avec $\Omega(E) = 0$.

Dans les deux cas, on peut écrire : $\Omega'(E) > \Omega(E)$ pour tout E . Ce qui donne pour les fonctions de partition :

$$Q'(T) = \sum_E \Omega'(E) e^{-E/kT} > Q(T) = \sum_E \Omega(E) e^{-E/kT}$$

Et pour l'énergie libre : $F'(T) = -kT \ln Q'(T) < F(T) = -kT \ln Q(T)$.

La suppression d'une contrainte dans un système isotherme ne peut que faire diminuer son énergie libre. L'équilibre est atteint lorsque toutes les contraintes sont

supprimées. C'est donc, pour les systèmes fermés isothermes isochores, l'état d'énergie libre minimal. Ceci est bien en accord avec la thermodynamique macroscopique.

Pour tout processus isotherme, on a donc : $\Delta F = \Delta E - T\Delta S < 0$. Deux cas sont à envisager :

- si $\Delta E > 0$, le processus n'est possible que si $\Delta S > 0$ (et même $\Delta S > \Delta E/T$) ;
- si $\Delta E < 0$, on peut avoir $\Delta S < 0$ (à condition qu'il soit supérieur à $\Delta E/T$).

Autrement dit, on ne peut avoir augmentation de l'énergie interne sans avoir simultanément augmentation de l'entropie. Inversement, on ne peut avoir diminution de l'entropie sans avoir en même temps diminution de l'énergie interne.

Aux très basses températures, le terme $T\Delta S$ a un poids très faible dans ΔF . C'est ΔE qui domine. Au zéro absolu, le système est dans son *état fondamental* qui par définition est l'état de plus basse énergie interne.

Aux très hautes températures, c'est $T\Delta S$ qui est prédominant. A température infinie, le système choisit son état d'entropie maximale, indépendamment de la valeur de l'énergie interne. Cet état d'entropie maximale, correspondant à une dissociation complète des constituants de la matière, est caractérisé par *l'équiprobabilité* d'observation de tous les états accessibles au système.

Aux températures intermédiaires, les deux termes sont à prendre en compte. L'état d'équilibre, fonction de T , résulte d'une compétition entre l'énergie interne qui tend à être la plus faible possible et l'entropie qui tend à être la plus forte possible.

Exemple. L'état d'énergie minimale pour l'hydrogène est l'état solide : tous les atomes étant liés entre eux, l'entropie est minimale, ce qui fait qu'on le trouve dans cet état aux très basses températures. Au fur et à mesure que la température augmente, il passe à l'état liquide, puis à l'état gazeux (molécules H_2). Si on augmente encore la température, on assiste à une dissociation des molécules ($H_2 \rightarrow H + H$), puis à une ionisation ($H \rightarrow \text{proton} + \text{électron}$). A chaque étape, il y a augmentation à la fois de l'énergie et de l'entropie. L'équilibre à chaque température est fixé par l'équation $F = E - TS$ minimale.

1.5.4 Notion de fluctuations aléatoires

A l'échelle macroscopique, les fluctuations autour des valeurs moyennes sont négligeables. Aucun appareil de mesures, aussi précis soit-il, n'est capable de les détecter. On en déduit, ce qui est conforme par ailleurs aux principes thermodynamiques, qu'à l'équilibre, toutes les grandeurs sont macroscopiquement constantes quel que soit l'environnement du système. Mais qu'en est-il à l'échelle microscopique ?

Prenons l'exemple d'une particule libre de masse m confinée dans un volume L^3 .

La théorie quantique donne l'écart entre deux niveaux d'énergie :

$$\delta\varepsilon = \frac{h^2}{2mL^2} \quad \text{avec} \quad h \approx 6,626 \cdot 10^{-34} \text{ J s}$$

C'est extrêmement faible, même à l'échelle microscopique : pour une mole de néon (masse atomique : $m = 3,32 \cdot 10^{-26} \text{ kg}$) dans les conditions habituelles ($L^3 = 22,4$ litres), on trouve : $\delta\varepsilon \approx 10^{-40} \text{ J}$. Si on tente de réaliser un système parfaitement isolé, on est amené à le protéger contre un certain nombre de perturbations dues à des fluctuations de paramètres extérieurs au système. Même en imaginant des conditions expérimentales très sophistiquées permettant de protéger le système contre toutes les agressions (parois parfaitement adiabatiques, parfaitement rigides, protections contre les ondes électromagnétiques, ...), on ne pourra jamais éliminer les forces gravitationnelles. Une masse M située à une distance D d'une molécule du gaz place celle-ci dans un champ de potentiel :

$$W = -\frac{G M m}{D} \quad \text{avec} \quad G = 6,673 \cdot 10^{-11} \text{ m}^2 \text{ s}^{-2} \text{ kg}^{-1}$$

Si M est immobile, cette énergie varie en fonction de la position instantanée de la molécule. D'une paroi à l'autre du récipient, la variation d'énergie vaut :

$$\Delta W = \frac{G M m}{D^2} L$$

Pour $M = 3 \text{ kg}$ et $D = 1 \text{ m}$, on trouve $\Delta W \approx 2 \cdot 10^{-36} \text{ J}$, c'est-à-dire environ 20 000 niveaux d'énergie!! Cela signifie qu'un chat passant à côté du récipient perturbe de façon considérable le système, à tel point que celui-ci n'est plus homogène : le champ de force créé par la présence du chat, même immobile, n'est pas uniforme, et deux particules situées chacune à une extrémité du récipient n'auront pas accès aux mêmes états quantiques, ou du moins pas avec les mêmes probabilités.

Un système n'est donc jamais parfaitement isolé à l'échelle microscopique. Il est soumis à de multiples perturbations incontrôlables, qui ont des effets à la fois parfaitement négligeables à l'échelle macroscopique et très importants à l'échelle microscopique. À cause de ces fluctuations aléatoires, un système macroscopique n'est jamais dans un état quantique pur (exception : condensation de Bose). Son état macroscopique, même sur un intervalle de temps très court, correspond à un état moyen, moyenne sur un grand nombre de micro-états. L'ensemble microcanonique dans lequel on n'envisage qu'un seul niveau d'énergie, n'est pas réaliste. Il n'a sa raison d'être que dans le cadre de l'approximation du terme maximum.

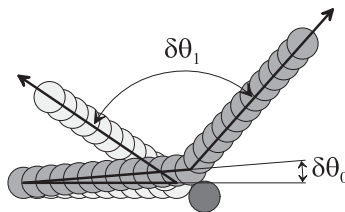


FIG. 1.3 – Amplification de déviation lors d'une collision

L'effet de la gravitation est plus spectaculaire encore en **mécanique classique**. La masse M est cette fois celle d'un neutron ($M \approx 1,66 \cdot 10^{-27} \text{ kg}$) situé aux confins de l'univers ($D \approx 2 \cdot 10^{26} \text{ m}$). La figure 1.3 montre qu'une déviation infime $\delta\theta_0$ de la trajectoire initiale d'une particule est très amplifiée après collision ($\delta\theta_1$). Si la

déviations initiales est causée par la présence du neutron, un calcul grossier donne, avec les valeurs courantes $d \sim 0,15 \mu\text{m}$ et $v \sim 500 \text{ m/s}$ (d est la distance parcourue par la molécule depuis sa dernière collision, et v sa vitesse) :

$$\delta\theta_0 \sim \frac{G M d^2}{D^3 v^2} \sim 10^{-135} \text{ radians}$$

C'est à dire trois fois rien. Après collision, la déviation est de l'ordre de :

$$\delta\theta_1 \sim \delta\theta_0 \frac{d}{r} \sim 10^3 \delta\theta_0 \quad (r = \text{rayon de la molécule} \sim 0,15 \text{ nm})$$

Au bout de n collisions, on peut estimer la déviation globale à $\delta\theta_n \sim 10^{3n} \delta\theta_0$. Elle devient de l'ordre de 1 radian au bout de 45 collisions, c'est-à-dire au bout de 13,5 nanosecondes!!!

Conclusion : une perturbation, aussi infime soit-elle, provoque un changement radical de la trajectoire initialement prévue de la molécule, et ce en un temps extrêmement court. Cela prouve la grande instabilité des équations dynamiques. Il est hors de question de prendre en compte tout ce qui est susceptible de perturber le système.

1.5.5 Le problème ergodique

Un système est dit ergodique si on peut lui appliquer le principe ergodique, c'est-à-dire si, pour toute variable mécanique, moyenne temporelle et moyenne sur l'ensemble sont identiques. Les systèmes **que l'on sait ergodiques** sont rares. Il est facile par exemple de montrer qu'un oscillateur harmonique à une dimension est un système ergodique. La démonstration est moins simple pour un "gaz de sphères dures" qui n'interagissent que lorsqu'elles sont en contact l'une avec l'autre. Il existe même des systèmes **que l'on sait non ergodiques**, mais ces systèmes sont peu réalistes physiquement : gaz rigoureusement parfait, système d'oscillateurs rigoureusement harmoniques et rigoureusement indépendants, etc. Pour un système quelconque, l'équivalence entre les deux moyennes n'est pas encore parfaitement démontrée à ce jour. Rien ne prouve que les microétats représentés dans l'ensemble statistique le soient avec des probabilités reflétant leurs fréquences moyennes d'apparition dans le système au cours du temps. Le principe ergodique reste toujours un postulat, encore jamais démenti par l'expérience.

À titre d'exemple, on peut traiter le cas relativement simple des **oscillateurs harmoniques** dans le cadre de la mécanique classique.

Soit un système constitué d'un seul oscillateur unidimensionnel classique, et supposons le parfaitement isolé du milieu extérieur. Son mouvement est décrit par l'équation $x(t) = a \cos(\omega t + \varphi)$. Les deux constantes (la phase φ et l'amplitude a) définissent un état possible de l'oscillateur. L'énergie totale de celui-ci ne dépend que de l'amplitude ($E = m\omega^2 a^2/2$). La valeur de a est donc fixée (le système est isolé). La valeur de φ quant à elle est fixée par les conditions initiales.

Soit f une fonction de x . La valeur moyenne de f peut être calculée, soit en intégrant sur t (moyenne temporelle $\bar{f}_{\text{tps}}$), soit par la méthode des ensembles (moyenne d'ensemble $\bar{f}_{\text{ens}}$). Pour $\bar{f}_{\text{ens}}$ on utilise le principe d'équiprobabilité : tous les états compatibles avec E sont équiprobables : $\bar{f}_{\text{ens}}$ se calcule en intégrant sur φ . La notation

$\theta = \omega t$ conduit à une expression de x où θ et φ jouent le même rôle : $x = a \cos(\theta + \varphi)$, avec θ et φ variant tous deux de 0 à 2π sur une période $T = 2\pi/\omega$. D'où :

$$\bar{f}_{\text{tps}} = \frac{1}{T} \int_0^T f(x(t)) dt = \frac{1}{2\pi} \int_0^{2\pi} f(x(a, \theta, \varphi)) d\theta = \frac{1}{2\pi} \int_0^{2\pi} f(x(a, \theta, \varphi)) d\varphi = \bar{f}_{\text{ens}}$$

Le système est ergodique puisque $\bar{f}_{\text{tps}} = \bar{f}_{\text{ens}}$. Cette valeur commune est fonction de a . On la notera $F(a)$.

Considérons maintenant un système constitué de deux oscillateurs harmoniques linéaires identiques et indépendants. Si on note a et b leurs amplitudes et φ et ψ leurs phases, la moyenne temporelle de f sur le système est simplement $\bar{f}_{\text{tps}} = F(a) + F(b)$.

C'est également ce qu'on obtient en intégrant sur φ et ψ . Mais attention, cette dernière intégration ne donne pas la moyenne d'ensemble. Tous les états compatibles avec E doivent être équiprobables. Comme $E = m\omega^2(a^2 + b^2)/2$, il faut envisager toutes les valeurs de a et b telles que $a^2 + b^2 = u^2$ avec $u^2 = 2E/(m\omega^2)$. Toutes les valeurs de a et b telles que $a = u \cos \alpha$ et $b = u \sin \alpha$ ($\alpha \in [0, \pi/2]$) conviennent. D'où :

$$\bar{f}_{\text{ens}} = \frac{2}{\pi} \int_0^{\pi/2} (F(u \cos \alpha) + F(u \sin \alpha)) d\alpha \neq \bar{f}_{\text{tps}}$$

Sauf cas particuliers, les deux moyennes sont différentes et le système n'est pas ergodique. Pour le rendre ergodique, il faut autoriser les échanges d'énergie entre les deux oscillateurs, c'est-à-dire les rendre anharmoniques et non indépendants.

Le deuxième principe de la thermodynamique et le principe d'équiprobabilité affirment chacun dans ses termes que, au bout d'un temps très court, tous les microétats compatibles avec l'énergie imposée au système sont équiprobables dans le temps. Dans un système rigoureusement isolé constitué d'un très grand nombre d'oscillateurs rigoureusement indépendants, rigoureusement harmoniques, si les oscillateurs sont lâchés au même instant en phase et avec la même amplitude, rien en théorie ne permet aux différents microétats d'apparaître de façon équiprobable. Les oscillateurs conservent éternellement leur phase et leur amplitude initiales. C'est en totale contradiction avec le deuxième principe. Ce n'est pas pour autant un échec du deuxième principe. Aucun système n'est à tel point parfait. Un tel système ne peut pas exister.

Les calculs montrent que la non ergodicité de ce système n'est pas due à l'harmonicité et l'indépendance des oscillateurs. En rendant les oscillateurs anharmoniques, les calculs montrent qu'il existe des conditions initiales telles que les phases et les amplitudes restent cantonnées au voisinage de leurs valeurs initiales. L'équiprobabilité n'est donc pas atteinte et le système n'est pas ergodique.

Les perturbations extérieures aléatoires sont donc nécessaires à l'établissement de l'équilibre. Ce sont elles qui en un temps extrêmement court rendent possible l'évolution vers l'état d'équilibre. Le système est dit **mélangeant**. C'est un terme plus fort qu'ergodique : tout système ergodique est mélangeant. Sans entrer dans les détails, un système est dit mélangeant s'il évolue spontanément vers son état d'équilibre quel que soit son état initial et quel que soit le traitement qu'on lui fait subir pour y parvenir. Le terme vient de l'exemple typique du mélange de deux liquides. On verse une goutte d'encre à la surface d'un verre d'eau. Au début, l'encre

est concentrée à l'endroit où on l'a déposée. Même en prenant toutes les précautions d'isolation possibles, le liquide devient au bout d'un certain temps uniformément coloré. Le processus est purement physique : la molécule d'encre ne montre aucune préférence à s'entourer de molécules d'eau. Autrement dit, le processus a lieu même si les énergies initiale et finale sont rigoureusement identiques.

L'étude du mouvement brownien (marche au hasard) montre que le chemin parcouru par une molécule subissant des chocs au cours desquels elle change de direction de façon aléatoire, forme une "pelote" dont la taille augmente indéfiniment au cours du temps (dans un volume infini, bien sûr). Le caractère mélangeant de ce système est naturel : il n'est pas nécessaire d'introduire d'éventuelles perturbations extérieures pour que le phénomène ait lieu. Pour le gaz de sphères dures ou pour le système d'oscillateurs harmoniques indépendants, le caractère mélangeant est dû à l'extrême sensibilité du système aux perturbations extérieures. Sans elles, les différents micro-états compatibles avec l'énergie totale n'apparaîtraient pas de façon équiprobable dans le temps et le principe ergodique (donc la méthode des ensembles) ne serait plus valable. En clair, ces deux systèmes ne sont pas ergodiques en théorie, mais le sont dans la pratique par le biais de leurs réactions "explosives" aux moindres perturbations extérieures. Le problème qui se pose est de savoir s'il n'existerait pas de systèmes non ergodiques dans la pratique. Pour l'instant on n'en connaît aucun, et on postule qu'il n'en existe pas.

Dernier point : il existe, aussi bien en mécanique classique qu'en mécanique quantique, des théorèmes en contradiction apparente avec le deuxième principe.

En mécanique classique, c'est le **théorème de Poincaré** : tout système finit toujours, si l'on attend suffisamment longtemps, par revenir à un état aussi voisin que l'on veut de son état initial. Si celui-ci est hors équilibre, ce théorème exclut toute évolution vers l'équilibre. C'est vrai en théorie, beaucoup moins dans la pratique : d'après une estimation de Boltzmann, un litre de gaz ordinaire revient à son état initial au bout de $10^{10\,000\,000\,000}$ années. De plus, dans un ensemble statistique, c'est un grand nombre de systèmes qui sont étudiés. Chacun peut revenir à son état initial, mais ils n'y reviendront pas tous en même temps. Si à un instant donné, l'un des systèmes est dans son état initial, cela n'aura aucune influence sur le calcul des probabilités. Cette éventualité est d'ailleurs permise par la théorie puisque l'état initial est lui aussi un micro-état accessible au système.

En mécanique quantique, l'évolution d'un système fini est **quasi périodique ou récurrente**. Si $\{\dots E_i \dots\}$ sont les niveaux d'énergie accessibles au système, la valeur moyenne de toute grandeur physique s'écrit comme une somme discrète de termes périodiques de fréquences $|E_i - E_j|/h$ (fréquences de Bohr). Le système finit toujours par revenir à son état initial. Mais pour un système macroscopique, les niveaux sont si serrés (d'autant plus que le système est toujours soumis à d'innombrables perturbations aléatoires) que la somme est quasi infinie. Le temps nécessaire au retour à l'état initial dépasse largement l'âge de l'univers. Notons qu'ici le problème est plus complexe car l'évolution récurrente concerne non seulement les états purs comme en mécanique classique, mais aussi les mélanges statistiques, autrement dit, non seulement un système particulier de l'ensemble, mais aussi l'ensemble statistique tout entier. Celui-ci reviendra tôt ou tard à son état initial. Plutôt tard que tôt ...

1.5.6 Le troisième principe

En thermodynamique, il s'énonce : *"Pour tout processus isotherme faisant intervenir des phases pures en équilibre interne, la variation d'entropie tend vers 0 quand T tend vers 0."* L'équilibre peut être sous contraintes internes, à condition que ces contraintes ne s'inhibent pas lorsque l'on passe de T à 0 K. On autorise par exemple la possibilité d'une réaction chimique infiniment lente, la présence d'isotopes instables mais à durée de vie très longue, etc. Lors d'un processus isotherme, où l'entropie passe de $S(T)$ à $S'(T)$, le troisième principe simplifie les écritures en posant :

$$\Delta S(T = 0) = S'(0) - S(0) = 0 \quad \text{et même} \quad S'(0) = S(0) = 0$$

C'est ce qu'on appelle une **entropie opérationnelle** : on la calcule en ne tenant compte que des degrés de liberté *excitables* dans le processus étudié. Si on cherchait à définir une **entropie absolue** $S_{abs}(T)$, il faudrait tenir compte de tous les degrés de liberté (liés aux noyaux, à la présence d'isotopes, etc) et on ne pourrait plus poser $S'(0) = S(0) = 0$, mais uniquement $S'(0) = S(0)$, et de toute façon $\Delta S(T = 0) = 0$.

La **thermodynamique statistique** donne une définition à l'**entropie absolue conventionnelle** :

$$S(T) = S_{abs}(T) - S_{abs}(0) = k \ln \Omega(T) - k \ln \Omega(0) = k \ln \frac{\Omega(T)}{\Omega(0)}$$

Pour l'entropie opérationnelle, on pose arbitrairement $\Omega(0) = 1$. On ignore simplement les degrés de liberté n'intervenant pas dans le processus étudié. Lors d'un processus isotherme où l'entropie passe de $S(T)$ à $S'(T)$, on écrira :

$$\Delta S(T) = S'(T) - S(T) = k \ln \frac{\Omega'(T)}{\Omega'(0)} - k \ln \frac{\Omega(T)}{\Omega(0)} = k \ln \frac{\Omega'(T)}{\Omega(T)}$$

(bien que non égaux à 1, $\Omega'(0)$ et $\Omega(0)$ sont néanmoins égaux)

On peut donc indifféremment travailler avec l'entropie absolue conventionnelle ou l'entropie opérationnelle. L'utilisation de cette dernière est plus simple mais il faut être sûr de prendre en compte tous les degrés de liberté nécessaires à l'étude.

Chapitre 2

Statistiques quantiques

2.1 Les statistiques de réseaux

Un réseau en physique statistique est un rassemblement de sites, chaque site étant une "boîte" pouvant accueillir un petit sous-système constitué de quelques particules.

On supposera que les sites sont **indépendants** : aucune interaction n'est envisagée entre particules occupant deux sites différents. Dans ce cas, la géométrie du réseau n'a aucune importance, si ce n'est dans l'expression du travail : on l'écrira $-pV$, $-\varphi A^{[1]}$ ou $-\tau L^{[2]}$ suivant que le réseau est tri-, bi- ou unidimensionnel.

2.1.1 Réseaux de sites tous identiques

Pour la mise en équations, on suppose que le système est entièrement décrit par les variables extensives E , V , N (on généralisera en fin de démonstration). Chaque état possible j d'un site sera caractérisé par une énergie ε_j , un volume v_j , un nombre n_j de particules. Si on note m_j le nombre de sites dans l'état j et M le nombre total de sites, on a :

$$\begin{aligned} (1) \quad \sum_j m_j &= M & \sum_j m_j \varepsilon_j &= E & (2) \\ (3) \quad \sum_j m_j v_j &= V & \sum_j m_j n_j &= N & (4) \end{aligned} \tag{2.1}$$

Pour la distribution $\{\dots m_j \dots\}$, il y a $\frac{M!}{\prod_j m_j!}$ configurations possibles.

Pour calculer le nombre de complexions $\Omega(M, E, V, N)$, on somme sur toutes les distributions vérifiant les conditions (1), (2), (3), (4) :

$$\Omega(M, E, V, N) = \sum_{\{\dots m_j \dots\}}^{(1,2,3,4)} \frac{M!}{\prod_j m_j!}$$

¹ A est l'aire de la surface et φ la tension de surface

² L est la longueur de la chaîne et τ la tension de ligne

Cette somme serait calculable s'il n'y avait que la condition (1). On se débarrasse des conditions gênantes en effectuant des sommations sur E , V , N . La sommation sur E revient à étendre la sommation sur les distributions (c'est l'approximation du terme maximum à l'envers). Elle a pour effet de supprimer la condition (2). En raisonnant de même pour N et V , toute sommation sur E , V , N s'écrit :

$$\sum_N \sum_V \sum_E \sum_{\{...m_j...\}}^{(1,2,3,4)} = \sum_N \sum_V \sum_{\{...m_j...\}}^{(1,3,4)} = \sum_N \sum_{\{...m_j...\}}^{(1,4)} = \sum_{\{...m_j...\}}^{(1)} \quad (2.2)$$

On est ainsi conduit à calculer la fonction de partition totale :

$$I(M, \beta, \xi, \gamma) = \sum_N \sum_V \sum_E \Omega(M, E, V, N) e^{-\beta E} e^{-\xi V} e^{-\gamma N}$$

On utilise 2.1 pour remplacer E par $\sum_j m_j \varepsilon_j$, V par $\sum_j m_j v_j$, N par $\sum_j m_j n_j$,

et donc $e^{-\beta E}$ par $\prod_j (e^{-\beta \varepsilon_j})^{m_j}$, $e^{-\xi V}$ par $\prod_j (e^{-\xi v_j})^{m_j}$, $e^{-\gamma N}$ par $\prod_j (e^{-\gamma n_j})^{m_j}$

Maintenant qu'il n'y a plus que la condition (1), la sommation sur les distributions est calculable :

$$I(M, \beta, \xi, \gamma) = \sum_{\{...m_j...\}}^{(1)} M! \prod_j \frac{(e^{-\beta \varepsilon_j} e^{-\xi v_j} e^{-\gamma n_j})^{m_j}}{m_j!} = (i(\beta, \xi, \gamma))^M$$

avec $i(\beta, \xi, \gamma) = \sum_j e^{-\beta \varepsilon_j} e^{-\xi v_j} e^{-\gamma n_j}$

La fonction $i(\beta, \xi, \gamma)$ est la fonction de partition totale relative à un site. On est ainsi ramené à l'étude d'un seul site :

$$\beta \Theta(M, \beta, \xi, \gamma) = -\ln I(M, \beta, \xi, \gamma) = -M \ln i(\beta, \xi, \gamma)$$

$$E = M \bar{\varepsilon} = -M \left[\frac{\partial \ln i}{\partial \beta} \right]_{\xi, \gamma} \quad V = M \bar{v} = -M \left[\frac{\partial \ln i}{\partial \xi} \right]_{\beta, \gamma} \quad N = M \bar{n} = -M \left[\frac{\partial \ln i}{\partial \gamma} \right]_{\beta, \xi}$$

$$\frac{S}{k} = M(\beta \bar{\varepsilon} + \xi \bar{v} + \gamma \bar{n} + \ln i)$$

Le nombre M de sites fait ici office de facteur d'échelle (voir équations 1.12 et 1.20).

Il est en général plus aisé de calculer $i(\beta, \xi, \gamma)$, non pas en sommant sur les états mais en sommant sur les valeurs possibles de ε , v , n ^[3]. En notant $\omega_{\varepsilon v n}$ le nombre d'états d'énergie $\varepsilon(v, n)$, on peut écrire $i(\beta, \xi, \gamma)$ de façon similaire à 1.26 :

$$i(\beta, \xi, \gamma) = \sum_{\varepsilon v n} \omega_{\varepsilon v n} e^{-\beta \varepsilon} e^{-\xi v} e^{-\gamma n}$$

³ Par convention, les lettres majuscules sont réservées aux caractéristiques du réseau, et les lettres minuscules aux caractéristiques d'un site.

Bien entendu, s'il y a plusieurs espèces en présence, la somme sur n est à remplacer par une somme multiple sur $n_1, n_2, \dots$. De même, s'il est nécessaire d'introduire un couple (X, P) pour décrire l'état thermodynamique du système, on a, avec $\phi = -P/kT$:

$$i(\beta, \xi, \phi, \gamma_1, \gamma_2, \dots) = \sum_{\varepsilon v x n_1 n_2, \dots} \omega_{\varepsilon v x n_1 n_2 \dots} e^{-\beta \varepsilon} e^{-\xi v} e^{-\phi x} e^{-\gamma_1 n_1} e^{-\gamma_2 n_2} \dots$$

2.1.2 Réseaux à différents types de sites

Soit un réseau $\mathcal{R}$ constitué de deux sous-réseaux $\mathcal{R}'$ et $\mathcal{R}''$, le premier contenant M' sites identiques, et le deuxième contenant M'' sites identiques. La généralisation à un nombre quelconque de sous-réseaux s'obtient facilement par récurrence.

La méthode la plus simple consiste à étudier les deux sous-systèmes séparément. L'équilibre mutuel se traduit simplement par le fait qu'ils possèdent les mêmes grandeurs intensives β, ξ, γ . Cette méthode permet de calculer séparément (E', V', N') et (E'', V'', N'') .

Chaque site de $\mathcal{R}'$ est caractérisé par des valeurs possibles (ε', v', n') pour l'énergie, le volume et le nombre de particules, chaque triplet apparaissant $\omega_{\varepsilon' v' n'}$ fois dans la liste. La fonction de partition de ce sous-réseau s'écrit donc :

$$I'(M', \beta, \xi, \gamma) = (i'(\beta, \xi, \gamma))^{M'}$$

$$\text{avec } i'(\beta, \xi, \gamma) = \sum_{\varepsilon' v' n'} \omega_{\varepsilon' v' n'} e^{-\beta \varepsilon'} e^{-\xi v'} e^{-\gamma n'}$$

$$E' = -M' \left. \frac{\partial \ln i'}{\partial \beta} \right]_{\xi, \gamma} \quad V' = -M' \left. \frac{\partial \ln i'}{\partial \xi} \right]_{\beta, \gamma} \quad N' = -M' \left. \frac{\partial \ln i'}{\partial \gamma} \right]_{\beta, \xi}$$

$$\frac{S'}{k} = \beta E' + \xi V' + \gamma N' + M' \ln i'$$

De même pour le deuxième sous-réseau $\mathcal{R}''$:

$$I''(M'', \beta, \xi, \gamma) = (i''(\beta, \xi, \gamma))^{M''}$$

$$\text{avec } i''(\beta, \xi, \gamma) = \sum_{\varepsilon'' v'' n''} \omega_{\varepsilon'' v'' n''} e^{-\beta \varepsilon''} e^{-\xi v''} e^{-\gamma n''}$$

$$E'' = -M'' \left. \frac{\partial \ln i''}{\partial \beta} \right]_{\xi, \gamma} \quad V'' = -M'' \left. \frac{\partial \ln i''}{\partial \xi} \right]_{\beta, \gamma} \quad N'' = -M'' \left. \frac{\partial \ln i''}{\partial \gamma} \right]_{\beta, \xi}$$

$$\frac{S''}{k} = \beta E'' + \xi V'' + \gamma N'' + M'' \ln i''$$

On utilise ensuite la propriété d'additivité des grandeurs extensives :

$$E = E' + E'' \quad V = V' + V'' \quad N = N' + N'' \quad S = S' + S''$$

$$\beta \Theta = \beta \Theta' + \beta \Theta'' \quad \Rightarrow \quad I(M', M'', \beta, \xi, \gamma) = I'(M', \beta, \xi, \gamma) \times I''(M'', \beta, \xi, \gamma)$$

2.1.3 Recherche de la fonction adaptée au problème

On revient au cas où tous les sites sont identiques. Il n'est pas toujours nécessaire de calculer la fonction de partition totale pour résoudre un problème de statistiques de réseaux : souvent, le nombre de variables extensives intervenant dans l'étude est restreint.

Voici les cas les plus courants. On part de la fonction de partition totale relative à un site d'un réseau monoconstituant :

$$i(\beta, \xi, \gamma) = \sum_{\varepsilon v n} \omega_{\varepsilon v n} e^{-\beta \varepsilon} e^{-\xi v} e^{-\gamma n}$$

On peut réécrire cette fonction de partition de différentes manières :

$$i(\beta, \xi, \gamma) = \sum_v z(\beta, v, \gamma) e^{-\xi v} \quad \text{avec} \quad z(\beta, v, \gamma) = \sum_{\varepsilon n} \omega_{\varepsilon n}(v) e^{-\beta \varepsilon} e^{-\gamma n} \quad (2.3)$$

$$i(\beta, \xi, \gamma) = \sum_n z(\beta, \xi, n) e^{-\gamma n} \quad \text{avec} \quad z(\beta, \xi, n) = \sum_{\varepsilon v} \omega_{\varepsilon v}(n) e^{-\beta \varepsilon} e^{-\xi v} \quad (2.4)$$

$$i(\beta, \xi, \gamma) = \sum_{v n} q(\beta, v, n) e^{-\xi v} e^{-\gamma n} \quad \text{avec} \quad q(\beta, v, n) = \sum_{\varepsilon} \omega_{\varepsilon}(v, n) e^{-\beta \varepsilon} \quad (2.5)$$

1. Partant de 2.3, on fixe v .

Il n'y a donc plus de sommation sur v : $i(\beta, \xi, \gamma) = z(\beta, v, \gamma) e^{-\xi v}$. On peut alors utiliser l'ensemble grand canonique ($V = M v$ est fixe) :

$$Z(M, \beta, V, \gamma) = (z(\beta, v, \gamma))^M \quad \beta \Psi(M, \beta, V, \gamma) = -M \ln z(\beta, v, \gamma)$$

$$E = -M \left[\frac{\partial \ln z}{\partial \beta} \right]_{v, \gamma} \quad \xi = \left[\frac{\partial \ln z}{\partial v} \right]_{\beta, \gamma} \quad N = -M \left[\frac{\partial \ln z}{\partial \gamma} \right]_{\beta, v}$$

2. Partant de 2.4, on fixe n .

Il n'y a donc plus de sommation sur n : $i(\beta, \xi, \gamma) = z(\beta, \xi, n) e^{-\gamma n}$. On peut alors utiliser l'ensemble T - p ($N = M n$ est fixe) :

$$Z(M, \beta, \xi, N) = (z(\beta, \xi, n))^M \quad \beta G(M, \beta, \xi, N) = -M \ln z(\beta, \xi, n)$$

$$E = -M \left[\frac{\partial \ln z}{\partial \beta} \right]_{\xi, n} \quad V = -M \left[\frac{\partial \ln z}{\partial \xi} \right]_{\beta, n} \quad \gamma = \left[\frac{\partial \ln z}{\partial n} \right]_{\beta, \xi}$$

3. Partant de 2.5, on fixe v et n .

Il n'y a plus de sommation ni sur v ni sur n : $i(\beta, \xi, \gamma) = q(\beta, v, n) e^{-\xi v} e^{-\gamma n}$. On peut alors utiliser l'ensemble canonique ($V = M v$ et $N = M n$ sont fixes) :

$$Q(M, \beta, V, N) = (q(\beta, v, n))^M \quad \beta F(M, \beta, V, N) = -M \ln q(\beta, v, n)$$

$$E = -M \left[\frac{\partial \ln q}{\partial \beta} \right]_{\xi, n} \quad \xi = \left[\frac{\partial \ln q}{\partial v} \right]_{\beta, n} \quad \gamma = \left[\frac{\partial \ln q}{\partial n} \right]_{\beta, v}$$

On généralise : pour savoir quelle fonction de partition utiliser, on répertorie pour un site donné, toutes les variables fixes et toutes les variables fluctuantes.

La fonction de partition d'un site sera, par transformée de Laplace de ω pour les variables fluctuantes, fonction des variables fixes et des variables intensives entropiques associées aux variables fluctuantes.

La fonction de partition du réseau (composé de M sites) est la puissance $M^{\text{ième}}$ de la fonction de partition d'un site ; on lui associe la transformée de Legendre de S/k pour les mêmes variables. Si on note $\beta\Phi$ cette transformée de Legendre et Δ la transformée de Laplace, on a $\beta\Phi = -\ln \Delta$.

Si le réseau est constitué d'un nombre quelconque de sous-réseaux différents et indépendants, on retrouve, par des arguments analogues à ceux du paragraphe 2.1.2, le même résultat : la fonction de partition du système est égale au produit des fonctions de partition relatives aux sous-réseaux.

2.1.4 Emploi d'autres fonctions de partition

Exemple 1 : l'ensemble canonique au lieu du grand canonique

Considérons un réseau de M sites, chacun pouvant soit être vide soit être occupé par un atome. On suppose qu'un site vide est à l'énergie $\varepsilon_0 = 0$ (choix d'origine des énergies), et on suppose également $\omega_0 = 1$ (choix d'origine des entropies). La fonction de partition d'un site vide s'écrit donc : $q_0 = q(\beta, v, 0) = \omega_0 e^{-\beta\varepsilon_0} = 1$. La fonction de partition $q(\beta, v) = q(\beta, v, 1)$ d'un site occupé sera supposée connue. Le volume de chaque site est indépendant de son occupation ($V = Mv$ avec v constant). On peut par conséquent travailler avec la grande fonction de partition (ensemble grand canonique) :

$$Z(M, \beta, V, \gamma) = (z(\beta, v, \gamma))^M \text{ avec } z(\beta, v, \gamma) = \sum_{n=0}^1 q(\beta, v, n) e^{-\gamma n} = 1 + q(\beta, v) e^{-\gamma}$$

La fonction $z(\beta, v, \gamma)$ ne contient que deux termes ; on peut réexprimer $Z(\beta, V, \gamma)$ en utilisant la formule du développement du binôme à la puissance M :

$$Z(M, \beta, V, \gamma) = (1 + q(\beta, v) e^{-\gamma})^M = \sum_{N=0}^M \frac{M!}{N! (M-N)!} q(\beta, v)^N e^{-\gamma N}$$

Et comme $Z(M, \beta, V, \gamma)$ s'écrit $\sum_{N=0}^M Q(M, \beta, V, N) e^{-\gamma N}$, on en déduit :

$$Q(M, \beta, V, N) = \frac{M!}{N! (M-N)!} q(\beta, v)^N$$

On peut dans cet exemple utiliser au choix la grande fonction de partition ou la fonction de partition canonique, pour obtenir, bien sûr les mêmes résultats.

Exemple 2 : l'ensemble microcanonique au lieu du grand canonique

Considérons un réseau de M sites, chacun pouvant accueillir 0, 1 ou 2 particules. On suppose que les particules ou groupes de particules n'ont accès qu'à leurs niveaux d'énergie fondamentaux, non dégénérés. L'énergie d'un site vide étant nulle, on notera ε_1 l'énergie d'un site occupé par une particule, et $\varepsilon_2 \neq 2\varepsilon_1$ l'énergie d'un site occupé par deux particules. Le volume n'intervient pas dans l'étude. On note m_0 , m_1 et m_2 les nombres de sites contenant 0, 1 et 2 particules. Pour M , E et N donnés, ces trois nombres doivent vérifier :

$$\begin{aligned} (1) \quad & M = m_0 + m_1 + m_2 \\ (2) \quad & E = \varepsilon_1 m_1 + \varepsilon_2 m_2 \\ (3) \quad & N = m_1 + 2m_2 \end{aligned}$$

Il y a autant de m_i que de variables macroscopiques. Le système ci-dessus (de déterminant non nul) est inversible. On trouve :

$$m_0 = M - \frac{N(\varepsilon_2 - \varepsilon_1) - E}{\varepsilon_2 - 2\varepsilon_1} \quad m_1 = \frac{N\varepsilon_2 - 2E}{\varepsilon_2 - 2\varepsilon_1} \quad m_2 = \frac{E - N\varepsilon_1}{\varepsilon_2 - 2\varepsilon_1}$$

Il est donc possible d'exprimer $\Omega(m_0, m_1, m_2)$ en fonction des grandeurs macroscopiques (M, E, N) . Ce qui permet d'utiliser l'ensemble microcanonique :

$$\Omega(M, E, N) = \frac{(m_0 + m_1 + m_2)!}{m_0! m_1! m_2!} = \frac{M!}{\left(M - \frac{N(\varepsilon_2 - \varepsilon_1) - E}{\varepsilon_2 - 2\varepsilon_1}\right)! \left(\frac{N\varepsilon_2 - 2E}{\varepsilon_2 - 2\varepsilon_1}\right)! \left(\frac{E - N\varepsilon_1}{\varepsilon_2 - 2\varepsilon_1}\right)!}$$

Un peu lourd, mais utilisable. On peut calculer :

$$\beta = \left. \frac{\partial \ln \Omega}{\partial E} \right]_{M, V} \quad \text{et} \quad \gamma = \left. \frac{\partial \ln \Omega}{\partial N} \right]_{M, E}$$

$$\text{On trouve : } \beta(\varepsilon_2 - 2\varepsilon_1) = 2 \ln \frac{m_1}{m_0} - \ln \frac{m_2}{m_0} \quad \text{et} \quad \gamma(\varepsilon_2 - 2\varepsilon_1) = -\varepsilon_2 \ln \frac{m_1}{m_0} + \varepsilon_1 \ln \frac{m_2}{m_0}$$

On vérifiera à l'aide de l'équation (1) que ces résultats sont équivalents à ceux trouvés par la méthode ci-dessous.

Une méthode moins lourde consiste à utiliser les équations (1), (2), (3) pour effectuer un changement de variables dans l'équation de Gibbs. Dans cette optique, il est nécessaire d'associer à M une grandeur intensive entropique ϕ :

$$\frac{dS}{k} = \phi dM + \beta dE + \gamma dN = \phi dm_0 + (\phi + \beta\varepsilon_1 + \gamma) dm_1 + (\phi + \beta\varepsilon_2 + 2\gamma) dm_2$$

$$\text{avec : } \frac{S}{k}(m_0, m_1, m_2) = \ln \Omega(m_0, m_1, m_2) = \ln \frac{(m_0 + m_1 + m_2)!}{m_0! m_1! m_2!}$$

$$\text{On en déduit : } \phi = \left. \frac{\partial \ln \Omega}{\partial m_0} \right]_{m_1, m_2} = \ln \frac{M}{m_0} \quad \phi + \beta\varepsilon_1 + \gamma = \left. \frac{\partial \ln \Omega}{\partial m_1} \right]_{m_0, m_2} = \ln \frac{M}{m_1}$$

$$\text{et} \quad \phi + \beta\varepsilon_2 + 2\gamma = \left. \frac{\partial \ln \Omega}{\partial m_2} \right]_{m_0, m_1} = \ln \frac{M}{m_2}$$

De ces trois équations, on tire ϕ , β , γ en fonction de m_0 , m_1 , m_2 . On se débarrasse de ϕ en utilisant l'équation (1), ce qui permet d'exprimer m_0 , m_1 , m_2 puis E et N (à l'aide des équations 2 et 3) en fonction de M , β et γ . On trouve :

$$m_0 = M \frac{1}{z(\beta, v, \gamma)} \quad m_1 = M \frac{e^{-\beta\varepsilon_1 - \gamma}}{z(\beta, v, \gamma)} \quad m_2 = M \frac{e^{-\beta\varepsilon_2 - 2\gamma}}{z(\beta, v, \gamma)}$$

$$E = M \frac{\varepsilon_1 e^{-\beta\varepsilon_1 - \gamma} + \varepsilon_2 e^{-\beta\varepsilon_2 - 2\gamma}}{z(\beta, v, \gamma)} \quad N = M \frac{e^{-\beta\varepsilon_1 - \gamma} + 2e^{-\beta\varepsilon_2 - 2\gamma}}{z(\beta, v, \gamma)}$$

avec $z(\beta, v, \gamma) = 1 + e^{-\beta\varepsilon_1 - \gamma} + e^{-\beta\varepsilon_2 - 2\gamma}$

Le volume intervient dans ε_1 et ε_2 . C'est ce qu'on aurait trouvé en utilisant l'ensemble grand canonique. On remarquera d'ailleurs que le dénominateur commun aux résultats est la grande fonction de partition relative à un site.

2.2 Particules identiques et indépendantes

Il n'est plus question ici de réseau. On considère un système constitué de N particules identiques et indépendantes partageant la même partie de l'espace, la même "boîte". Comme elles sont indépendantes, chacune d'elles se comporte comme si elle était seule dans la boîte. On est alors tenté de dire que tout se passe comme si chacune occupait une boîte qui lui est propre, autrement dit un site de réseau.

C'est ce que nous ferons dans un premier temps (paragraphe 2.2.1), **sachant qu'il ne faut pas**. Ce raisonnement ne tient pas compte de l'**indiscernabilité** des particules, propriété fondamentale (décrite au paragraphe 2.2.2) des particules quantiques.

2.2.1 Si les particules quantiques étaient discernables ...

On suppose dans ce paragraphe que les particules, toutes identiques, se comportent comme si chacune d'elle occupait un site de réseau.

À température et volume constants, on peut utiliser la fonction de partition canonique. On a vu au paragraphe 2.1.3 que, pour un réseau de M sites de même volume v fixe et tous occupés par n particules, on pouvait écrire $Q(M, \beta, V, N) = (q(\beta, v, n))^M$. Ici, il faut remplacer M par N (il y a autant de boîtes que de particules), v par V et n par 1. On écrit donc, avec la notation $q(\beta, V) = q(\beta, V, 1)$:

$$Q(\beta, V, N) = (q(\beta, V))^N \quad \Rightarrow \quad \beta F(\beta, V, N) = N \ln q(\beta, V) = N \beta f(\beta, V)$$

Soient $\{\dots \varepsilon_i \dots\}$ les niveaux d'énergie accessibles à une particule et $\{\dots \omega_i \dots\}$ les dégénérescences correspondantes. La fonction de partition individuelle s'écrit :

$$q(\beta, V) = \sum_i \omega_i e^{-\beta\varepsilon_i}$$

Soit i_j le niveau d'énergie occupé par la particule j ($j = 1$ à N) et calculons la probabilité d'observer le système à l'énergie :

$$E_{i_1 \dots i_N} = \varepsilon_{i_1} + \dots + \varepsilon_{i_N} = \sum_{j=1}^N \varepsilon_{i_j}$$

Ce niveau est $\Omega_{i_1 \dots i_N} = \omega_{i_1} \times \dots \times \omega_{i_N} = \prod_{j=1}^N \omega_{i_j}$ fois dégénéré.

Avec $e^{-\beta E_{i_1 \dots i_N}} = \prod_{j=1}^N e^{-\beta \varepsilon_{i_j}}$, la probabilité recherchée vaut (voir éq. 1.29) :

$$P_{i_1 \dots i_N} = \frac{\Omega_{i_1 \dots i_N} e^{-\beta E_{i_1 \dots i_N}}}{Q(\beta, V, N)} = \frac{1}{(q(\beta, V))^N} \prod_{j=1}^N \omega_{i_j} e^{-\beta \varepsilon_{i_j}} = \prod_{j=1}^N \frac{\omega_{i_j} e^{-\beta \varepsilon_{i_j}}}{q(\beta, V)}$$

On en déduit intuitivement (et c'est confirmé par le calcul) la probabilité qu'une particule donnée soit à l'énergie ε_i et le nombre moyen de particules à l'énergie ε_i :

$$\eta(\varepsilon_i) = \frac{\omega_i e^{-\beta \varepsilon_i}}{q(\beta, V)} \quad \text{et} \quad \bar{n}(\varepsilon_i) = N \eta(\varepsilon_i) = N \frac{\omega_i e^{-\beta \varepsilon_i}}{q(\beta, V)}$$

Pour calculer les probabilités par état, on divise simplement le résultat trouvé par ω_i . La probabilité qu'une particule donnée soit dans l'un des ω_i états d'énergie ε_i , et le nombre moyen correspondant s'écrivent :

$$\eta_i = \frac{e^{-\beta \varepsilon_i}}{q(\beta, V)} \quad \text{et} \quad \bar{n}_i = N \eta_i = N \frac{e^{-\beta \varepsilon_i}}{q(\beta, V)} \quad (2.6)$$

Ces probabilités permettent de calculer l'énergie moyenne d'une particule et l'énergie moyenne du système :

$$\bar{\varepsilon} = \sum_{\text{niveaux } i} \varepsilon_i \eta(\varepsilon_i) = \sum_{\text{états } j} \varepsilon_j \eta_j \quad \text{et} \quad \bar{E} = \sum_{\text{niveaux } i} \varepsilon_i \bar{n}(\varepsilon_i) = \sum_{\text{états } j} \varepsilon_j \bar{n}_j$$

À un facteur d'échelle N près, une particule et le système entier semblent avoir le même comportement. C'est au niveau des fluctuations autour des valeurs moyennes que se trouve la différence. On a calculé au paragraphe 1.5.1 la variance de E pour N particules :

$$\Delta E^2 = kT^2 C_V = N kT^2 c_V \quad (c_V : \text{capacité thermique d'une particule})$$

$$\text{Avec } N = 1, \text{ on obtient : } \Delta \varepsilon^2 = kT^2 c_V = \frac{kT^2 C_V}{N}$$

$$\text{Ce qui donne, avec } E \sim NkT, C_V \sim Nk, \bar{\varepsilon} \sim kT : \frac{\Delta E}{E} \sim \frac{1}{\sqrt{N}} \quad \text{et} \quad \frac{\Delta \varepsilon}{\bar{\varepsilon}} \sim 1.$$

Conclusion :

Des fluctuations négligeables autour de $\bar{E}$, et très étendues autour de $\bar{\varepsilon}$.

Quelques résultats intéressants tirés de l'expression de $\bar{n}_j$:

1. *Lors d'une élévation de température, les niveaux d'énergie supérieure à $\bar{\varepsilon}$ se peuplent au détriment de ceux d'énergie inférieure.*

Ce résultat s'obtient en dérivant $\bar{n}_j$ (equation 2.6) par rapport à T .

En remarquant que $\frac{dq}{d\beta} = -q\bar{\varepsilon}$, on trouve :

$$\left. \frac{\partial \bar{n}_j}{\partial T} \right]_{V,N} = \frac{\bar{n}_j(\varepsilon_j - \bar{\varepsilon})}{kT^2} = \frac{N}{kT^2} \eta_j(\varepsilon_j - \bar{\varepsilon}) \quad \left\{ \begin{array}{l} < 0 \text{ si } \varepsilon_j < \bar{\varepsilon} \\ > 0 \text{ si } \varepsilon_j > \bar{\varepsilon} \end{array} \right. \quad \text{CQFD} \quad (2.7)$$

2. *Un apport de chaleur modifie les populations des niveaux d'énergie, et non les niveaux eux-mêmes : les valeurs des $\bar{n}_j$ sont modifiées, mais pas celles des ε_j .*

L'apport de chaleur est obtenu par une élévation de température dT à V , N constants. C'est $C_V dT$. Partant de l'expression de $\Delta\varepsilon^2$, on trouve :

$$C_V = \frac{N\Delta\varepsilon^2}{kT^2} = \frac{N}{kT^2} (\bar{\varepsilon}^2 - \varepsilon^2) = \frac{N}{kT^2} \left(\sum_j \varepsilon_j^2 \eta_j - \bar{\varepsilon} \sum_j \varepsilon_j \eta_j \right) = \sum_j \frac{\bar{n}_j(\varepsilon_j - \bar{\varepsilon})\varepsilon_j}{kT^2}$$

Ce qui donne, en utilisant le résultat 2.7 : $C_V = \sum_j \varepsilon_j \left. \frac{\partial \bar{n}_j}{\partial T} \right]_{V,N}$.

Ce sont bien les populations qui sont modifiées.

Attention : lorsqu'on fournit un travail, on modifie les valeurs ε_j des niveaux d'énergie, et donc en même temps les populations. Il y a donc en même temps apport de chaleur : la proposition tentante « un apport de travail modifie les niveaux mais pas les populations » est fausse.

3. *La capacité thermique tend vers 0 quand T tend vers 0.*

Au zéro absolu, le système est dans son état fondamental : toutes les particules sont à l'énergie la plus basse ε_1 . À température légèrement supérieure, on peut considérer que le plupart sont à l'énergie ε_1 , et que quelques unes sont à l'énergie ε_2 (premier *niveau excité*). Les autres niveaux seront considérés inaccessibles :

$$C_V = \sum_i \frac{\bar{n}_i(\varepsilon_i - \bar{\varepsilon})\varepsilon_i}{kT^2} = \frac{N}{kT^2} (\eta_1(\varepsilon_1 - \bar{\varepsilon})\varepsilon_1 + \eta_2(\varepsilon_2 - \bar{\varepsilon})\varepsilon_2)$$

avec $\eta_1 = \frac{e^{-\beta\varepsilon_1}}{e^{-\beta\varepsilon_1} + e^{-\beta\varepsilon_2}} \quad \eta_2 = \frac{e^{-\beta\varepsilon_2}}{e^{-\beta\varepsilon_1} + e^{-\beta\varepsilon_2}} \quad \bar{\varepsilon} = \eta_1\varepsilon_1 + \eta_2\varepsilon_2$

On trouve, avec la notation $Y = \frac{\varepsilon_2 - \varepsilon_1}{2kT}$:

$$C_V = Nk \left(\frac{Y}{\cosh Y} \right)^2 \longrightarrow 0 \text{ quand } T \longrightarrow 0 \quad \text{CQFD}$$

2.2.2 ... mais les particules quantiques sont indiscernables

Dans ce paragraphe, la condition "particules indépendantes" n'est pas requise. Les propriétés qui y sont décrites sont des propriétés générales des particules quantiques.

L'indiscernabilité

Pour expliquer cette propriété, considérons un récipient constitué de deux compartiments séparés par une cloison. On place une particule dans chaque compartiment. Les deux particules sont numérotées (1) et (2) en fonction du compartiment qu'elles occupent. Elles sont alors **discernables**, c'est-à-dire qu'on sait les reconnaître à chaque instant. Lorsqu'on supprime la cloison qui sépare les deux compartiments, les deux particules mettent en commun l'ensemble du récipient. Si on replace la cloison en faisant en sorte de les séparer à nouveau, il est impossible de savoir laquelle avait reçu le numéro 1, laquelle le numéro 2. Sans la cloison, les particules sont dites **indiscernables**.

La situation est différente en **mécanique classique** : si à l'instant initial les particules sont numérotées, on peut suivre chacune d'elles le long de sa **trajectoire** qui lui est propre. On sait à chaque instant reconnaître une particule en fonction de sa position et de sa vitesse : que les particules soient identiques ou non, elles sont, en mécanique classique, discernables.

L'indiscernabilité est une propriété fondamentale des particules quantiques. Le fait qu'on les considère dans certains cas comme discernables vient du fait que chacune évolue dans une région de l'espace qui lui est propre. Ce n'est pas par nature mais uniquement pour des raisons géométriques (particules en réseaux par exemple^[4]). Autrement dit :

Deux particules sont indiscernables si elles ont en commun un certain nombre d'états possibles, autrement dit si elles sont identiques et si leurs probabilités de présence sont simultanément non nulles dans un même domaine de l'espace.

Postulat de symétrisation

Pour les particules identiques, il existe un postulat spécifique appelé *postulat de symétrisation*. Considérons un système constitué de N particules indiscernables. Après permutation de deux particules, l'état du système reste inchangé : les deux vecteurs d'état $|\varphi_0\rangle$ et $|\varphi_1\rangle$ représentant l'état du système avant et après permutation sont donc identiques à un *facteur multiplicatif* près. Les fonctions d'ondes étant normalisées, ce facteur est un complexe de norme 1 noté $e^{i\theta}$: $|\varphi_1\rangle = e^{i\theta}|\varphi_0\rangle$. Si on permute à nouveau les deux particules, on multiplie $|\varphi_1\rangle$ par $e^{i\theta}$: $|\varphi_2\rangle = e^{i\theta}|\varphi_1\rangle = e^{2i\theta}|\varphi_0\rangle$. Puisqu'on a retrouvé la situation de départ, $|\varphi_2\rangle = |\varphi_0\rangle$ et donc :

$$e^{2i\theta} = 1 \quad \Rightarrow \quad e^{i\theta} = \pm 1$$

⁴ si chaque particule est confinée sur un site, avec interdiction de le quitter ; dès qu'on introduit un site vide permettant aux particules de migrer, celles-ci deviennent alors indiscernables : chacune a une probabilité non nulle d'occuper chacun des sites du réseau et il devient alors impossible d'identifier telle ou telle particule.

On distingue ainsi deux grandes classes de particules quantiques :

- les **bosons** dont le vecteur d'état est **symétrique** vis-à-vis de la permutation de deux particules ($e^{i\theta} = +1$) ;
- Les **fermions** dont le vecteur d'état est **antisymétrique** vis-à-vis de la permutation de deux particules ($e^{i\theta} = -1$).

Mais ce raisonnement n'est pas rigoureux : rien n'oblige le deuxième $e^{i\theta}$ à être identique au premier. C'est pourquoi on l'énonce en temps que postulat, dit **postulat de symétrisation** :

Lorsqu'un système comprend plusieurs particules identiques, son état ne peut être décrit que par certains vecteurs d'état : suivant la nature de ces particules (bosons ou fermions), les vecteurs d'états sont, soit symétriques, soit antisymétriques par rapport à la permutation de deux particules identiques.

Conséquence immédiate pour les fermions : si deux fermions sont dans le même état, leur permutation laisse inchangée l'état du système (ils sont indiscernables) et par conséquent son vecteur d'état. Celui-ci ne peut donc pas changer de signe.

C'est le **principe d'exclusion de Pauli** :

Deux fermions ne peuvent être dans le même état.

Le spin

Le spin est une propriété purement quantique des constituants de la matière. C'est un moment cinétique intrinsèque de la particule qui, suivant la particule, peut être *entier* ou *demi entier* (en unité $\hbar$). Le microétat d'une particule est caractérisé par un *état orbital* (spatial) et un *état de spin* : ces deux sous-états seront supposés indépendants (pas d'interaction spin-orbite). Une particule de spin s (en unité $\hbar$) possède $2s + 1$ **états de spin possibles** ; il faudra en tenir compte dans le calcul des probabilités.

Deux cas sont à envisager :

- en l'absence de champ magnétique, tous les états de spin sont équiprobables : dans les calculs statistiques, il faudra simplement multiplier les fonctions de partition par $2s + 1$;
- sous champ magnétique, l'énergie du système dépendra de l'état de spin des particules qui le constituent (qui dit moment cinétique dit forcément moment magnétique associé).

L'expérience montre que les particules de *spin entier* sont des *bosons*, et que les particules de *spin demi-entier* sont des *fermions*. Électrons, protons et neutrons sont des fermions ($s = 1/2$) ; les photons sont des bosons ($s = 1$). Un corps formé de plusieurs particules est un boson ou un fermion suivant le nombre de fermions qui le constituent. L'hélium 3 par exemple est un fermion, alors que l'hélium 4 est un boson. Ces deux isotopes ont des comportements fondamentalement différents à basse température.

Conséquences sur le nombre de complexion

Pour bien montrer la différence entre bosons, fermions et particules discernables, considérons un système constitué de N particules identiques et indépendantes, chacune ayant accès à ω microétats. Répartir N particules sur ω états disponibles revient à distribuer N objets dans ω boîtes. Pour les fermions, on a nécessairement $\omega \geq N$.

Étudions les différents cas possibles.

- Si les particules étaient **discernables**, chacune choisirait l'un des ω états indépendamment du choix des autres particules.

Il y a donc $\Omega_D = \omega^N$ possibilités.

- Pour les **fermions**, les seules distributions possibles sont celles où il n'y a qu'un seul objet par boîte : parmi les ω boîtes, N sont occupées et $(\omega - N)$ sont vides.

Ce qui donne $\Omega_F = \frac{\omega!}{N! (\omega - N)!}$ possibilités.

- Pour les **bosons**, c'est plus compliqué. Si on veut employer la même méthode que pour les fermions, il faut définir le nombre ω_0 de boîtes vides, le nombre ω_1 de boîtes occupées par une particule, le nombre ω_2 de boîtes occupées par 2 particules, etc. Les distributions $[\omega_n] = \{\omega_0 \dots \omega_n \dots \omega_N\}$ doivent vérifier :

$$(1) \quad \sum_{n=0}^N \omega_n = \omega \qquad (2) \quad \sum_{n=0}^N n \omega_n = N$$

$\Omega_B = \sum_{[\omega_n]}^{(1,2)} \frac{\omega!}{\omega_0! \dots \omega_n! \dots \omega_N!}$ est incalculable à cause de la condition (2).

Mais il existe une autre méthode qui consiste à remplacer les ω boîtes par une seule découpée en ω compartiments séparés par $(\omega - 1)$ cloisons. On passe d'une configuration à une autre en permutant un des N bosons et une des $(\omega - 1)$ cloisons.

On a donc $\Omega_B = \frac{(N + \omega - 1)!}{N! (\omega - 1)!}$ possibilités.

Résultat important au sujet des nombres d'états possibles :

$$\Omega_F = \frac{1}{N!} \frac{\omega!}{(\omega - N)!} = \frac{\omega^N}{N!} \left(1 - \frac{1}{\omega}\right) \left(1 - \frac{2}{\omega}\right) \dots \left(1 - \frac{N-1}{\omega}\right)$$

$$\Omega_B = \frac{1}{N!} \frac{(\omega + N - 1)!}{(\omega - 1)!} = \frac{\omega^N}{N!} \left(1 + \frac{1}{\omega}\right) \left(1 + \frac{2}{\omega}\right) \dots \left(1 + \frac{N-1}{\omega}\right)$$

Avec $\Omega_D = \omega^N$, on constate qu'on a toujours : $\Omega_F \leq \frac{\Omega_D}{N!} \leq \Omega_B$.

À très haute température, le nombre ω d'états individuels possibles devient extrêmement grand : $\omega \longrightarrow \infty$ quand $T \longrightarrow \infty$. On en déduit :

$$\Omega_F \longrightarrow \frac{\Omega_D}{N!} \longleftarrow \Omega_B \quad \text{quand} \quad T \longrightarrow \infty$$

Bosons et fermions ont donc le même comportement à haute température, celui des **particules classiques** (*i.e.* décrites par la mécanique classique). Un calcul simple montre que la fonction de partition canonique du système suit un comportement analogue aux nombres de complexions. Quand $T \longrightarrow \infty$, on a :

$$Q_F(\beta, V, N) \longrightarrow Q(\beta, V, N) = \frac{Q_D(\beta, V, N)}{N!} = \frac{q(\beta, V, N)^N}{N!} \longleftarrow Q_B(\beta, V, N)$$

Remarque : si le système est multiconstituant, on dit simplement que les différents sous-systèmes sont discernables. Ce qui donne à haute température :

$$Q(\beta, V, N_1 \dots N_p) = \frac{q_1(\beta, V, N_1 \dots N_p)^{N_1}}{N_1!} \times \dots \times \frac{q_p(\beta, V, N_1 \dots N_p)^{N_p}}{N_p!}$$

De même, si il existe un autre couple de variables (X, P) , on a avec les notations du paragraphe 1.4.8 :

$$\begin{aligned} - \text{ à } X \text{ constant : } Q(\beta, V, X, N) &= \frac{q(\beta, V, X, N)^N}{N!} \\ - \text{ à } P \text{ constant : } \Delta_X(\beta, V, \phi, N) &= \frac{\delta_X(\beta, V, \phi, N)^N}{N!} \end{aligned}$$

Si les particules sont des **bosons**, bien qu'il soit possible d'en trouver plusieurs simultanément dans le même état, ω est tellement grand que cette éventualité est hautement improbable. Il est donc normal que Ω_B tende vers $\Omega_D/N!$: toutes les situations (il y en a $N!$) se déduisant l'une l'autre par permutation des N particules représentent le même état.

On peut se demander pourquoi les **fermions** n'ont pas ce comportement classique à toute température, puisque deux fermions ne peuvent se trouver dans le même état. Réponse : même si les fermions sont "classiquement" indépendants, il existe entre eux des *interactions quantiques* : le fait qu'un fermion soit dans un état donné affecte le comportement des autres dans le sens que l'état qu'il occupe n'est plus disponible aux autres. Cet effet devient négligeable à haute température, à cause du nombre extrêmement grand d'états disponibles.

À très basse température, bosons et fermions ont un comportement radicalement différent. Si on range les états par ordre d'énergie croissante, les probabilités d'occupation convergent vers zéro d'autant plus rapidement que la température est basse. Dans la limite $T = 0\text{K}$, si les particules sont des **bosons**, seul le niveau fondamental est occupé par tous les bosons ($\omega = 1$) : c'est ce qu'on appelle la *condensation de Bose*. Les **fermions** par contre occupent les N états de plus basse énergie, à raison de un par état ($\omega = N$). Le niveau occupé de plus haute énergie est appelé *niveau de Fermi*. Dans les deux cas, Ω_F et Ω_B tendent vers 1 quand T tend vers 0.

2.3 Systèmes de particules indiscernables

2.3.1 Limite des hautes températures

On se place en situation où on peut écrire : $Q(\beta, V, N) = \frac{q(\beta, V, N)^N}{N!}$

C'est le passage à la limite infinie pour la température.

Comme au paragraphe 2.2.1, calculons la probabilité d'observer le système à l'énergie $E_{i_1 \dots i_N}$. Le calcul est ici plus délicat : un microétat du système est caractérisé par un multiplet $\{i_1 \dots i_N\}$ ou par toute permutation de celui-ci. Il y a donc $N!$ fois moins de microétats (la température est suffisamment élevée pour qu'il soit improbable que deux particules se trouvent dans le même état). Il est donc normal que l'on trouve des probabilités $N!$ fois plus importantes :

$$P_{i_1 \dots i_N} = \frac{\Omega_{i_1 \dots i_N} e^{-\beta E_{i_1 \dots i_N}}}{Q(\beta, V, N)} = \frac{N!}{(q(\beta, V, N))^N} \prod_{j=1}^N \omega_{i_j} e^{-\beta \varepsilon_{i_j}} = N! \prod_{j=1}^N \frac{\omega_{i_j} e^{-\beta \varepsilon_{i_j}}}{q(\beta, V, N)}$$

Pour calculer η_{i_1} par exemple, on somme $P_{i_1 \dots i_N}$ sur *tous* les $\{i_2 \dots i_N\}$ et on divise par $(N-1)!$. On divise encore par N , car le résultat obtenu est la probabilité que *n'importe quelle* particule soit dans l'état i_1 . Moyennant quoi, on trouve :

$$\begin{aligned} \eta_{i_1} &= \frac{1}{N} \frac{1}{(N-1)!} \sum_{i_2} \dots \sum_{i_N} N! \frac{\omega_{i_1} e^{-\beta \varepsilon_{i_1}}}{q(\beta, V, N)} \frac{\omega_{i_2} e^{-\beta \varepsilon_{i_2}}}{q(\beta, V, N)} \dots \frac{\omega_{i_N} e^{-\beta \varepsilon_{i_N}}}{q(\beta, V, N)} \\ &= \frac{\omega_{i_1} e^{-\beta \varepsilon_{i_1}}}{q(\beta, V, N)^N} \left(\sum_{i_2} \omega_{i_2} e^{-\beta \varepsilon_{i_2}} \right) \times \dots \times \left(\sum_{i_N} \omega_{i_N} e^{-\beta \varepsilon_{i_N}} \right) = \frac{\omega_{i_1} e^{-\beta \varepsilon_{i_1}}}{q(\beta, V, N)} \end{aligned}$$

(chaque somme entre parenthèses vaut $q(\beta, V, N)$ et il y en a $N-1$)

Même résultat que pour les particules discernables. Et donc les deux premiers résultats tirés de l'expression de $\bar{n}_j$ (voir paragraphe 2.2.1) sont encore valables ici (pas le troisième puisqu'il a été obtenu à basse température).

Un autre résultat intéressant est obtenu lorsque les ε_j (et donc la fonction q) ne dépendent pas de N .

Calculons à partir de $Q(\beta, V, N) = \frac{(q(\beta, V))^N}{N!}$ la grande fonction de partition :

$$Z(\beta, V, \gamma) = \sum_{N=0}^{\infty} \frac{(q(\beta, V))^N}{N!} e^{-\gamma N} = \sum_{N=0}^{\infty} \frac{(\lambda q(\beta, V))^N}{N!} \quad \text{avec } \lambda = e^{-\gamma}$$

Comme $q(\beta, V)$ est indépendante de N , $Z(\beta, V, \gamma)$ est la série entière d'une fonction exponentielle : $Z(\beta, V, \gamma) = e^{\lambda q(\beta, V)} \Rightarrow -\beta \Psi(\beta, V, \gamma) = \ln Z(\beta, V, \gamma) = \lambda q(\beta, V)$.

On en déduit à l'aide de la relation $\Psi = E - TS - \mu N = -pV$:

$$N = - \left. \frac{\partial \ln Z}{\partial \gamma} \right]_{\beta, V} = \lambda \left. \frac{\partial \ln Z}{\partial \lambda} \right]_{\beta, V} = \lambda q = \ln Z = -\beta \Psi = -\frac{\Psi}{kT} = \frac{pV}{kT} \Rightarrow pV = NkT$$

Le simple fait que les niveaux d'énergie ne dépendent pas de N (autrement dit les particules se comportent comme si elles étaient seules dans le récipient) suffit à établir l'équation d'état des gaz parfaits. À retenir donc :

$$\varepsilon_j = \varepsilon_j(V) \quad \forall j \quad (\text{indépendants de } N) \quad \Longleftrightarrow \quad pV = NkT \quad (\text{gaz parfait}) \quad (2.8)$$

2.3.2 Fonction de distribution à température quelconque

À température quelconque (non infinie), on doit pour calculer $Q(\beta, V, N)$ sommer sur les états symétriques (bosons), ou sur les états antisymétriques (fermions).

De ce fait le résultat $Q(\beta, V, N) = \frac{q(\beta, V, N)^N}{N!}$, n'est plus valable.

Soit $[n] = \{n_1 \dots n_j \dots\}$ une distribution des particules sur les états j accessibles.^[5] Pour calculer $Q(\beta, V, N)$, on fixe la valeur de N . On doit donc avoir :

$$(1) \quad \sum_j n_j = N$$

Les n_j particules dans l'état j constituent un sous système à un seul niveau (non dégénéré) d'énergie $E_j = n_j \varepsilon_j$. Sa fonction de partition canonique s'écrit donc :

$$Q_j = e^{-\beta E_j} = e^{-\beta n_j \varepsilon_j} = q_j^{n_j} \quad \text{avec} \quad q_j = e^{-\beta \varepsilon_j}$$

Le système à N particules est la réunion de tels sous-systèmes. Pour une distribution $[n]$ donnée, la fonction de partition $Q_{[n]}$ est le produit de tous les Q_j . La fonction de partition canonique du système s'obtient en sommant sur toutes les distributions vérifiant la condition (1) :

$$Q_{[n]} = \prod_j Q_j = \prod_j q_j^{n_j} \quad \text{et} \quad Q(\beta, V, N) = \sum_{[n]}^{(1)} Q_{[n]} = \sum_{[n]}^{(1)} \prod_j q_j^{n_j}$$

La condition (1) est gênante. Sans elle, la somme serait factorisable (voir calcul de Z ci-dessous). On la fait disparaître en sommant sur N , c'est-à-dire en calculant la grande fonction de partition :

$$\begin{aligned} Z(\beta, V, \gamma) &= \sum_{N=0}^{\infty} Q(\beta, V, N) e^{-\gamma N} = \sum_{N=0}^{\infty} \sum_{[n]}^{(1)} \prod_j q_j^{n_j} e^{-\gamma N} \\ &= \sum_{[n]} \prod_j (\lambda q_j)^{n_j} = \sum_{n_1} \dots \sum_{n_j} \dots ((\lambda q_1)^{n_1} \dots (\lambda q_j)^{n_j} \dots) \end{aligned}$$

$$\text{On a utilisé } \sum_{N=0}^{\infty} \sum_{[n]}^{(1)} = \sum_{[n]}^{\text{sans condition}} \quad \text{et} \quad e^{-\gamma N} = \exp\left(-\gamma \sum_j n_j\right) = \prod_j (e^{-\gamma})^{n_j} = \prod_j \lambda^{n_j}.$$

⁵ On raisonne ici sur les états : il n'y a pas de ω_j dans les expressions. Si un niveau i est ω_i fois dégénéré, il y aura dans la liste ω_i états de même énergie ε_i . Autre détail important : on ne restreint pas l'étude aux particules indépendantes (les ε_i peuvent éventuellement dépendre de N). Pour éviter les lourdeurs, toutes les variables ont été retirées des expressions.

Maintenant qu'il n'y a plus la condition (1), cette multiple somme est factorisable (dernière écriture : tous les indices muets n_j ont été remplacés par l'indice n) :

$$Z(\beta, V, \gamma) = \sum_{n_1} (\lambda q_1)^{n_1} \times \cdots \times \sum_{n_j} (\lambda q_j)^{n_j} \times \cdots = \prod_j \sum_n (\lambda q_j)^n$$

$Z(\beta, V, \gamma)$ est un produit de grandes fonctions de partition, chacune représentant toutes les populations possibles d'un état j :

$$Z(\beta, V, \gamma) = \prod_j \sum_n q_j^n e^{-\gamma n} = \prod_j \sum_n Q_j e^{-\gamma n} = \prod_j Z_j(\beta, V, \gamma)$$

Les valeurs possibles de n sont 0 et 1 pour les fermions (jamais deux fermions dans le même état), et 0 à l'infini pour les bosons :

$$\begin{aligned} \text{fermions : } Z(\beta, V, \gamma) &= \prod_j \sum_{n=0}^1 (\lambda q_j)^n = \prod_j (1 + \lambda q_j) \\ \text{bosons : } Z(\beta, V, \gamma) &= \prod_j \sum_{n=0}^{\infty} (\lambda q_j)^n = \prod_j \frac{1}{1 - \lambda q_j} \quad \text{si } \lambda q_j < 1 \quad \forall j \end{aligned}$$

La condition $\lambda q_j < 1 \quad \forall j$ pour les bosons s'écrit aussi $\mu < \varepsilon_j \quad \forall j$. En rangeant les états par ordre d'énergies croissantes, elle s'écrit : $\mu < \varepsilon_1$.

On peut condenser les deux résultats en une seule écriture :

$$Z(\beta, V, \gamma) = \prod_j (1 + \sigma \lambda q_j)^\sigma \quad \text{avec} \quad \begin{cases} q_j = e^{-\beta \varepsilon_j} \\ \lambda = e^{-\gamma} \end{cases} \quad \text{et} \quad \sigma = \begin{cases} +1 & (\text{fermions}) \\ -1 & (\text{bosons}) \end{cases}$$

Pour calculer le nombre moyen $\bar{n}_j$ de particules dans l'état j , on applique l'un des résultats du paragraphe 1.4.2 à la fonction $Z_j(\beta, V, \gamma) = (1 + \sigma \lambda q_j)^\sigma$:

$$\bar{n}_j = - \left. \frac{\partial \ln Z_j}{\partial \gamma} \right]_{\beta, V} = \lambda \left. \frac{\partial \ln Z_j}{\partial \lambda} \right]_{\beta, V} = \frac{\lambda q_j}{1 + \sigma \lambda q_j} = \frac{1}{\frac{1}{\lambda q_j} + \sigma} = \frac{1}{e^{(\varepsilon_j - \mu)/kT} + \sigma}$$

Ces **fonctions de distribution** sont à la base des **statistiques quantiques** :

- celle de **Fermi-Dirac** pour les fermions ($\sigma = +1$),
- celle de **Bose-Einstein** pour les bosons ($\sigma = -1$).

Elles permettent de calculer toutes les **propriétés thermodynamiques** du système :

$$\begin{aligned} Z(\beta, V, \gamma) &= \prod_j (1 + \sigma \lambda q_j)^\sigma \quad \Rightarrow \quad \ln Z(\beta, V, \gamma) = \sigma \sum_j \ln(1 + \sigma \lambda q_j) \\ N &= - \left. \frac{\partial \ln Z}{\partial \gamma} \right]_{V, \beta} = \sum_j \bar{n}_j = \sum_j \frac{1}{e^{(\varepsilon_j - \mu)/kT} + \sigma} \\ E &= - \left. \frac{\partial \ln Z}{\partial \beta} \right]_{V, \gamma} = \sum_j \bar{n}_j \varepsilon_j = \sum_j \frac{\varepsilon_j}{e^{(\varepsilon_j - \mu)/kT} + \sigma} \\ pV &= -\Psi = kT \ln Z = \sigma kT \sum_j \ln(1 + \sigma e^{(\mu - \varepsilon_j)/kT}) \end{aligned} \tag{2.9}$$

Dans la pratique, le nombre N de particules est fixé. L'expression de N sert alors à calculer μ . Celui-ci est ensuite remplacé dans les expressions de E et de pV .

2.3.3 Distributions de Fermi-Dirac et de Bose-Einstein

Pour les systèmes macroscopiques, les niveaux d'énergie sont si serrés qu'on peut considérer le spectre d'énergies comme continu (**approximation du continuum**). Nous raisonnerons donc dans ce contexte en étudiant les fonctions continues :

$$f_{FD}(\varepsilon, T, \mu) = \frac{1}{e^{(\varepsilon-\mu)/kT} + 1} \quad f_{BE}(\varepsilon, T, \mu) = \frac{1}{e^{(\varepsilon-\mu)/kT} - 1} \quad (2.10)$$

distribution de **Fermi-Dirac**

distribution de **Bose-Einstein**

La fonction de distribution $f_{FD}(\varepsilon, T, \mu)$ de Fermi-Dirac

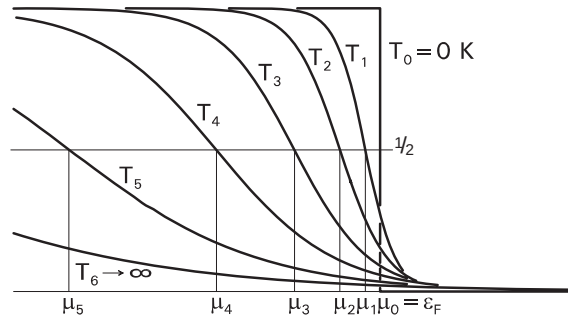


FIG. 2.1 – Distribution de Fermi-Dirac

Au zéro absolu, c'est une "marche" (voir figure 2.1) : elle vaut 1 si ε est inférieure à μ , et 0 pour les valeurs supérieures (le système choisit son état d'énergie minimale où les N particules occupent les N états de plus basse énergie, à raison de une par état). Le potentiel chimique est donc le dernier niveau occupé au zéro absolu. On le note ε_F et on l'appelle **niveau de Fermi** :

$$\mu(T) \longrightarrow \varepsilon_F \quad \text{et même} \quad \frac{\mu(T) - \varepsilon_F}{kT} \longrightarrow 0 \quad \text{quand} \quad T \rightarrow 0$$

(μ tend vers ε_F "plus vite" que T vers 0)

Quelle que soit T , la courbe est symétrique par rapport au point $(\mu, 1/2)$:

$$f_{FD}(\mu + \varepsilon, T, \mu) + f_{FD}(\mu - \varepsilon, T, \mu) = 1 \quad \forall \varepsilon \quad \text{et donc} \quad f_{FD}(\mu, T, \mu) = 1/2$$

La fonction de distribution $f_{BE}(\varepsilon, T, \mu)$ de Bose-Einstein

Elle n'est définie que pour les énergies supérieures au potentiel chimique μ . Celui-ci est donc toujours inférieur à tout ε_j , donc à ε_1 .

Au zéro absolu, c'est un pic de Dirac (voir figure 2.2) : elle vaut N pour $\varepsilon = \varepsilon_1$ et 0 ailleurs (le système choisit son état d'énergie minimale où toutes les particules sont sur le niveau fondamental).

Au voisinage du zéro absolu, on peut écrire :

$$f_{BE}(\varepsilon_1, T, \mu) \longrightarrow N \quad \text{quand} \quad T \rightarrow 0 \quad \Rightarrow \quad \mu(T) \approx \varepsilon_1 - kT \ln(1 + 1/N) \approx \varepsilon_1 - \frac{kT}{N}$$

On en déduit :

$$\mu(T) \longrightarrow \varepsilon_1^- \quad \text{et même} \quad \frac{\mu(T) - \varepsilon_1}{kT} \longrightarrow 0 \quad \text{quand} \quad T \rightarrow 0$$

(μ tend vers ε_1 "plus vite" que T vers 0)

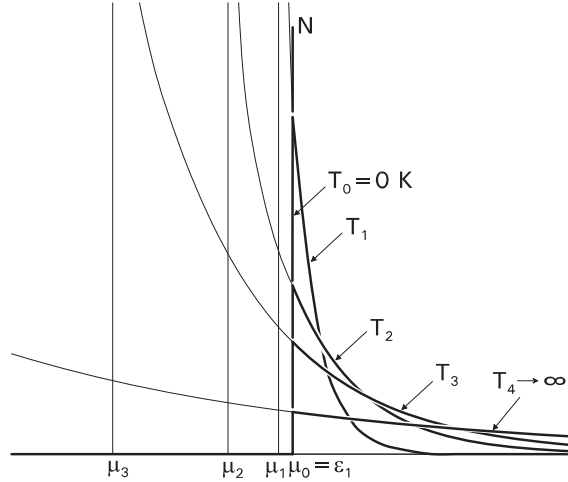


FIG. 2.2 – Distribution de Bose-Einstein

À très haute température

Que ce soient des **bosons** ou des **fermions**, le nombre d'états accessibles aux particules est extrêmement grand, beaucoup plus grand que N , de sorte que le taux moyen d'occupation par état devient extrêmement faible. Le système choisit son état d'entropie maximale, indépendamment de la valeur de l'énergie. Il y a équiprobabilité entre les différents états individuels, et le taux d'occupation de tous les états devient quasi nul. On en déduit :

$$\mu(T) \longrightarrow -\infty \quad \text{et même} \quad \frac{\mu(T)}{kT} \longrightarrow -\infty \quad \text{quand} \quad T \rightarrow +\infty$$

(μ tend vers $-\infty$ "plus vite" que T vers $+\infty$)

2.4 Le corps noir et le gaz de photons

2.4.1 Théorie électromagnétique

Lorsqu'on maintient une enceinte à une température T , il s'établit spontanément à l'intérieur un système d'ondes électromagnétiques. Cette cavité fermée en équilibre thermique avec un thermostat est appelée **corps noir** ou **radiateur intégral**. On peut observer le rayonnement émis par un trou de taille négligeable percé dans la paroi. Historiquement, c'est ce qui amena Planck à formuler l'hypothèse des quanta d'énergie et les principaux postulats de la mécanique quantique.

Point de départ : les équations de Maxwell dans le vide :

$$\vec{\nabla} \times \vec{E} = -\frac{\partial \vec{B}}{\partial t} \quad \vec{\nabla} \cdot \vec{B} = 0 \quad \vec{\nabla} \times \vec{B} = \mu_0 \varepsilon_0 \frac{\partial \vec{E}}{\partial t} \quad \vec{\nabla} \cdot \vec{E} = 0$$

Elles conduisent aux équations de propagation :

$$\Delta \vec{E} = \frac{1}{c^2} \frac{\partial^2 \vec{E}}{\partial t^2} \quad \text{et} \quad \Delta \vec{B} = \frac{1}{c^2} \frac{\partial^2 \vec{B}}{\partial t^2} \quad \text{avec} \quad \frac{1}{c^2} = \mu_0 \varepsilon_0 \quad (2.11)$$

Les ondes planes $\vec{E} = \vec{E}_0 e^{i(\vec{k} \cdot \vec{r} - \omega t)}$ et $\vec{B} = \vec{B}_0 e^{i(\vec{k} \cdot \vec{r} - \omega t)}$ sont solutions de ces équations. En remplaçant dans les diverses équations, on trouve :

– que la norme k du **vecteur d'onde**^[6] est liée à la pulsation ω par la relation :

$$k = \frac{\omega}{c}$$

– que les vecteurs $(\vec{k}, \vec{E}, \vec{B})$ sont perpendiculaires entre eux en tout point et pour tout t et qu'ils forment un trièdre direct.

Les équations 2.11 étant linéaires, la solution générale, qui dépend des conditions aux limites, est une superposition linéaire de ces solutions élémentaires monochromatiques. Si on considère une enceinte vide, supposée cubique, on trouve, par une démarche similaire à l'établissement de 2.24 plus loin, que les composantes du vecteur d'onde doivent être des nombres entiers de $2\pi/L$, L étant l'arête du cube :

$$\vec{k} = \frac{2\pi}{L} (n_x, n_y, n_z) \quad \text{avec} \quad n_x, n_y, n_z \text{ entiers positifs ou négatifs}$$

Les solutions monochromatiques possédant de tels vecteurs d'onde sont appelés **modes propres** de la cavité. Pour caractériser entièrement un mode propre, il faut également définir la direction de $\vec{E}$, autrement dit un **état de polarisation**. On peut définir deux états de polarisation indépendants en décomposant $\vec{E}$ sur deux axes orthogonaux contenus dans le plan perpendiculaire à $\vec{k}$.

Un champ électromagnétique quelconque sera toujours une combinaison linéaire des divers modes propres ainsi définis. Les dimensions de l'enceinte étant macroscopiques, l'ensemble des modes propres forme un ensemble quasi continu. On peut donc définir une **densité de modes propres** $g_\omega(\omega)$: $g_\omega(\omega) d\omega$ est le nombre de modes dont la pulsation est comprise entre ω et $\omega + d\omega$. Cette densité s'obtient en calculant le nombre de modes dont le module du vecteur d'onde est compris entre k et $k + dk$, puis en remplaçant k par ω/c et dk par $d\omega/c$. En tenant compte des deux états de polarisation possibles, on trouve :

$$g_\omega(\omega) d\omega = 2 \frac{V}{8\pi^3} 4\pi k^2 dk = \frac{V \omega^2}{\pi^2 c^3} d\omega \quad \Rightarrow \quad g_\omega(\omega) = \frac{V \omega^2}{\pi^2 c^3}$$

⁶ Attention : ici, k n'est pas la constante de Boltzmann.

2.4.2 Les photons

Si on s'en tient au résultat donné par la théorie classique de l'électromagnétisme, l'amplitude de chacun des modes propres (E_0 ou B_0) peut prendre une valeur quelconque. On trouverait par conséquent un spectre d'énergie continu de 0 à $+\infty$.

En réalité, l'énergie d'un champ électromagnétique, comme celle d'un système matériel, est quantifiée. À une onde électromagnétique de vecteur d'onde $\vec{k}$ et de pulsation ω , on associe des particules appelées **photons**. Chaque photon possède une impulsion et une énergie données par les *équations de Planck-Einstein* :

$$\vec{p} = \hbar \vec{k} \quad \text{et} \quad \varepsilon = \hbar \omega \quad \text{avec} \quad \hbar = \frac{h}{2\pi} \approx 1,055 \cdot 10^{-34} \text{ J s}$$

On relie l'énergie d'un photon à celle de l'onde qui lui est associée en disant que cette énergie vaut $N_\omega \hbar \omega$, où N_ω est le nombre de photons de pulsation ω . L'énergie de l'onde est ainsi quantifiée puisque N_ω est un entier positif ou nul.

En utilisant $k = \omega/c$ et les équations de Planck–Einstein, on en déduit $\varepsilon = pc$. Si on compare ce résultat à celui, tout à fait général, donné par la mécanique relativiste ($\varepsilon^2 = p^2 c^2 + m^2 c^4$), on en déduit que **le photon est une particule de masse nulle**.

L'onde possède deux états de polarisation possibles. Les photons correspondants auront donc **deux états de spin possibles**. Mais attention : il ne faut pas en conclure pour autant que leur spin vaut $1/2$. Une étude plus approfondie montre qu'en fait le spin des photons vaut $s = 1$, mais que la valeur $m_s = 0$ est interdite, de la même façon que l'état de polarisation longitudinale de l'onde qui lui est associée est interdit par les équations de Maxwell. Conclusion : **les photons sont des bosons** de spin 1, mais il faudra remplacer $(2s + 1)$ non pas par 3, mais par 2.

Pour terminer le passage onde particules, on associe un état individuel possible pour les photons (une pulsation ω et un état de spin) à chaque mode propre classique (un vecteur d'onde et un état de polarisation). C'est ce qui nous permet de calculer la densité d'états individuels pour les photons. Il suffit de remplacer dans $g_\omega(\omega) d\omega$ la pulsation ω par $\varepsilon/\hbar$ et $d\omega$ par $d\varepsilon/\hbar$:

$$g(\varepsilon) d\varepsilon = g_\omega(\omega) d\omega = \frac{V \omega^2 d\omega}{\pi^2 c^3} = \frac{V \varepsilon^2}{\pi^2 c^3 \hbar^3} d\varepsilon \quad \Rightarrow \quad g(\varepsilon) = \frac{V \varepsilon^2}{\pi^2 c^3 \hbar^3}$$

2.4.3 Distribution de Planck, loi de Planck

Contrairement aux particules matérielles pour lesquelles il faut négliger les interactions entre particules, les interactions entre photons sont rigoureusement nulles. On a donc affaire à un gaz de bosons rigoureusement parfait. La fonction de distribution de Bose-Einstein (avec $\varepsilon = \hbar \omega$) permet de calculer le nombre moyen de photons de pulsation ω :^[7]

$$N_\omega = \frac{1}{e^{(\hbar \omega - \mu)/kT} - 1}$$

⁷ À partir de maintenant, k redevient la constante de Boltzmann.

Les photons interagissent avec les parois. C'est ce qui rend possible l'établissement de l'équilibre thermique avec le thermostat. Pour un système isolé à l'équilibre thermique, les nombres de photons émis ou absorbés peuvent être différents, à partir du moment où l'énergie est conservée. C'est un système ouvert, bien que l'enceinte soit fermée ... Le fait que le nombre de photons ne soit pas conservé permet de dire que **le potentiel chimique des photons est nul**. D'où :

$$N_\omega = \frac{1}{e^{\hbar\omega/kT} - 1} \quad \text{Distribution de Planck.}$$

Attention : la situation n'a rien à voir avec celle d'un système ouvert échangeant des particules avec un réservoir. Lorsqu'on étudie un système ouvert, le nombre de particules du système fluctue mais le nombre total de particules (système + réservoir) reste constant. Dans le cas d'un système de photons, le *nombre total* de photons n'est pas conservé. Les parois émettent ou absorbent des photons, elles les créent ou les annihilent.

En conséquence de ce potentiel chimique nul, toutes les grandeurs thermodynamiques pourront être exprimées uniquement en fonction de T et V . Thermodynamiquement, il n'y a plus que deux termes dans l'équation fondamentale, ce qui restreint le nombre de transformées de Legendre possibles. À partir de S/k :

$$S/k(E, V, N) = \beta E + \xi V \quad \beta = \left. \frac{\partial S/k}{\partial E} \right]_{V,N} \quad \xi = \left. \frac{\partial S/k}{\partial V} \right]_{E,N}$$

$(+\beta F)$ est la Transformée de Legendre de S/k pour la variable E :

$$\beta F(\beta, V, N) = \beta E - S/k \quad E = \left. \frac{\partial \beta F}{\partial \beta} \right]_{V,N} \quad \xi = - \left. \frac{\partial \beta F}{\partial V} \right]_{\beta,N}$$

$(-\beta \Phi)$ est la Transformée de Legendre de S/k pour la variable V :

$$\beta \Phi(E, \xi, N) = S/k - \xi V \quad \beta = \left. \frac{\partial \beta \Phi}{\partial E} \right]_{\xi,N} \quad V = - \left. \frac{\partial \beta \Phi}{\partial \xi} \right]_{E,N}$$

$(+\beta G)$ est la Transformée de Legendre de S/k pour les variables E et V :

$$\beta G(\beta, \xi, N) = \beta E + \xi V - S/k \quad E = \left. \frac{\partial \beta G}{\partial \beta} \right]_{\xi,N} \quad V = \left. \frac{\partial \beta G}{\partial \xi} \right]_{\beta,N}$$

βH , $\beta \Psi$, βE , et $\beta \Theta$ sont respectivement égaux à S/k , βF , $\beta \Phi$ et βG .

La loi de Planck exprime la **densité spectrale d'énergie électromagnétique** $u(\omega, T)$. Par définition, à T donnée, $u(\omega, T) d\omega$ est l'énergie par unité de volume des ondes de pulsation comprise entre ω et $\omega + d\omega$. C'est, à un facteur multiplicatif près, l'intensité mesurée en fonction de ω du rayonnement émis par un trou dans la paroi. C'est la loi la plus importante. Historiquement mais aussi parce qu'elle permet de déduire toutes les autres.

Pour calculer la densité spectrale $u(\omega, T)$, on écrit de deux manières différentes l'énergie dE que possèdent les photons de pulsation comprise entre ω et $\omega + d\omega$:

- d’une part en utilisant la définition de $u(\omega, T)$: $dE_\omega = V u(\omega, T) d\omega$
- d’autre part, c’est le nombre de modes propres de pulsation entre ω et $\omega + d\omega$ multiplié par l’énergie emmagasinée par les N_ω phonons d’énergie $\hbar\omega$:

$$dE_\omega = g_\omega(\omega) d\omega \times N_\omega \hbar\omega = \frac{V \omega^2}{\pi^2 c^3} d\omega \frac{1}{e^{\hbar\omega/kT} - 1} \hbar\omega$$

Ce qui donne **la loi de Planck** (figure 2.3) : $u(\omega, T) = \frac{\hbar\omega^3}{\pi^2 c^3 (e^{\hbar\omega/kT} - 1)}$

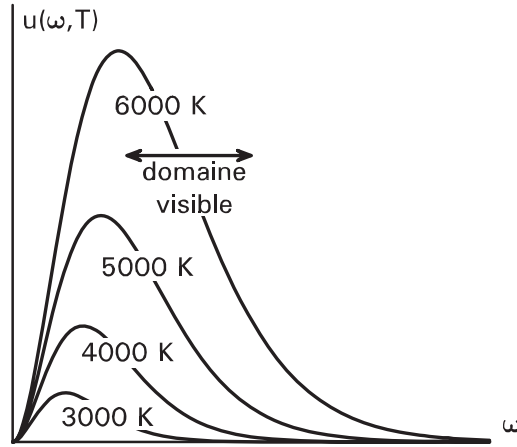


FIG. 2.3 – Loi de Planck

Historiquement, cette loi a une importance considérable. Elle fut obtenue par Planck à partir de mesures expérimentales. A noter que théorie et expérience sont en excellent accord dans tout le domaine de température accessible, et quelle que soit la pulsation.

À basse fréquence, on obtient :

$$e^{\hbar\omega/kT} - 1 \approx \frac{\hbar\omega}{kT} \quad \Rightarrow \quad u(\omega, T) \approx \frac{kT \omega^2}{\pi^2 c^3}$$

C’est la **formule de Rayleigh-Jeans** aussi appelée **formule classique**, car c’est le résultat que donne la théorie électromagnétique classique (on peut constater d’ailleurs que la constante de Planck a disparu). Selon la théorie classique, la densité spectrale est une fonction constamment croissante de la pulsation. Ce n’est physiquement pas possible, puisque l’énergie totale (l’aire sous la courbe) contenue dans l’enceinte serait infinie. Expérimentalement, ce résultat est correct à basse fréquence, mais on constate que, à fréquence plus élevée, la densité spectrale passe par un maximum pour décroître rapidement vers 0 à haute fréquence : ce phénomène logique physiquement (connu sous le nom de **catastrophe ultraviolette**) n’avait aucune justification théorique à l’époque.

La position ω_m du maximum varie proportionnellement à T :^[8]

$$\omega_m = 2,8214 \frac{kT}{\hbar}$$

⁸ 2,8214 est solution de l’équation $x = 3(1 - e^{-x})$.

Ce résultat, déduit initialement des courbes expérimentales, est connu sous le nom de **loi du déplacement de Wien**. La mesure de ω_m conduit à la température de la source. Cette technique est utilisée pour mesurer la température des étoiles.

La valeur u_m que prend ce maximum varie en T^3 :

$$u_m = u(\omega_m, T) = 1,4212 \frac{k^3 T^3}{\pi^2 \hbar^2 c^3}$$

Aux températures usuelles (lampe à incandescence : 3000 K, arc électrique : 4000 K, surface du soleil : 5500 K, ...), ω_m est dans l'infrarouge. Il atteint le domaine visible à une température d'environ 6400 K. A noter que le passage de $T = 2800$ K (lampe ordinaire) à $T = 3200$ K (lampe halogène) multiplie l'intensité par un facteur variant de 2 à 5 dans le domaine visible.

2.4.4 Thermodynamique du rayonnement thermique

La loi de Planck permet d'accéder à toutes les propriétés importantes du rayonnement. Les résultats seront exprimés en fonction de la **constante de Stephan** :

$$\sigma = \frac{\pi^2 k^4}{60 \hbar^3 c^2} = 5,6705054 \cdot 10^{-8} \text{ kg s}^{-3} \text{ K}^{-4}$$

Deux intégrales numériques sont également utiles :

$$\int_0^\infty \frac{x^2 dx}{e^x - 1} \approx 2,404 \quad \text{et} \quad \int_0^\infty \frac{x^3 dx}{e^x - 1} = \frac{\pi^4}{15}$$

Énergie totale et capacité thermique : il suffit d'intégrer $u(\omega, T)$.

$$E(T, V) = \int_0^\infty V u(\omega, T) d\omega = \frac{V \hbar}{\pi^2 c^3} \int_0^\infty \frac{\omega^3 d\omega}{e^{\hbar\omega/kT} - 1} = \frac{V k^4 T^4}{\pi^2 \hbar^3 c^3} \int_0^\infty \frac{x^3 dx}{e^x - 1}$$

$$E(T, V) = \frac{4\sigma}{c} V T^4 \quad (\text{Loi de Stephan}) \quad C_V = \left. \frac{\partial E}{\partial T} \right|_V = \frac{16\sigma}{c} V T^3$$

Énergie libre et grand potentiel :

$$F = \Psi = kT \int_0^\infty g(\varepsilon) \ln(1 - e^{-\varepsilon/kT}) d\varepsilon = \frac{V kT}{\pi^2 \hbar^3 c^3} \int_0^\infty \varepsilon^2 \ln(1 - e^{-\varepsilon/kT}) d\varepsilon$$

En intégrant par parties on trouve $F = \Psi = -\frac{E}{3}$, et on en déduit :

$$p = -\frac{\Psi}{V} = \frac{E}{3V} = \frac{4\sigma}{3c} T^4 \quad \text{et} \quad S = \frac{E - F}{T} = \frac{4E}{3T} = \frac{C_V}{3} = \frac{16\sigma}{3c} V T^3$$

La relation $pV = E/3$ (au lieu de $pV = 2E/3$, voir équation 3.1) est due au comportement relativiste des photons.

La pression de radiation est en général négligeable, mais pas toujours :

- À 300 K, la pression vaut $2 \cdot 10^{-6} \text{ Pa} = 2 \cdot 10^{-11} \text{ bars}$. C'est en général parfaitement négligeable, sauf dans les expériences en ultravide, où la pression de radiation peut perturber les appareils de mesures (les pressions recherchées peuvent parfois descendre jusqu'à 10^{-17} bars).
- Lors d'une explosion nucléaire, la température peut localement frôler les 10^6 K . La pression de radiation peut alors atteindre plusieurs milliers de bars.

Remarque : dans une naine blanche, la température est de l'ordre de 10^7 K . On trouve alors une pression de radiation de $2,5 \cdot 10^7 \text{ bars}$. Cela peut paraître énorme, mais pas du tout. L'une des relations 5.9 du paragraphe 5.3.3 permet d'estimer la contribution des électrons à la pression : on trouve environ 10^{17} bars . Et il faut encore ajouter la contribution des noyaux présents dans la naine blanche. Compte tenu de la densité en noyaux ($\sim 10^{35} \text{ par m}^3$) et de leur masse moyenne, on peut à 10^7 K assimiler le système de noyaux à un gaz parfait. L'équation d'état des gaz parfaits donne pour les noyaux une pression de l'ordre de 10^{14} bars .

Conclusion : dans une naine blanche, la pression de radiation, pourtant énorme, est parfaitement négligeable.

La distribution de Planck permet également de calculer la densité de photons :

$$N = \int_0^\infty \frac{g(\varepsilon) d\varepsilon}{e^{\hbar\omega/kT} - 1} = \frac{V}{\pi^2 c^3} \int_0^\infty \frac{\omega^2 d\omega}{e^{\hbar\omega/kT} - 1} = 2,404 \frac{V k^3 T^3}{\pi^2 \hbar^3 c^3}$$

Ce qui donne :

$$E = 2,7012 NkT \quad pV = 0,9004 NkT \quad C_V = 10,8048 Nk \quad S = 3,6016 Nk$$

Une mise en garde au sujet de ces dernières relations. Elles sont données dans le seul but de comparer le gaz de photons à un gaz de particules non relativistes. Le nombre N de photons étant fonction de la température (loi en T^3), il ne faut surtout pas déduire de $E = 2,7012 NkT$ que E suit une loi linéaire en T . D'ailleurs, C_V n'est pas égal à $2,7012 Nk$.

Petite anecdote historique :

En 1964, deux radio-astronomes (Penzias et Wilson) découvrent tout à fait par hasard le **rayonnement cosmique**, quelques mois avant que les astrophysiciens en prédisent l'existence à partir de la fameuse théorie du big bang. Les mesures hors atmosphère qui suivirent permirent de conclure que l'Univers est rempli à l'heure actuelle d'un rayonnement qu'émettrait un corps noir porté à une température voisine de 3 K.

En voici l'explication. Selon la **théorie du big bang**, juste après la grande explosion, la température était extrêmement élevée, de sorte que l'Univers contenait essentiellement des photons, des électrons libres, et des noyaux atomiques nus. L'interaction entre les photons et les particules chargées étant très forte, les photons étaient piégés par la matière, sans arrêt diffusés, émis et absorbés. C'est ainsi que, malgré son expansion et la baisse continue de sa température, l'Univers était

à chaque instant en équilibre et que l'énergie des photons était distribuée suivant la loi de Planck du corps noir. L'abaissement de température favorise la formation d'atomes, d'où une disparition progressive des particules chargées libres et une interaction de plus en plus faible entre la matière et les photons. Lorsque la température atteint, selon les calculs, une température d'environ 3000 K, l'Univers devint transparent aux photons (ce qu'on appelle le découplage du rayonnement et de la matière). Le rayonnement fossile que nous recevons actuellement provient de cet instant.

Pourquoi correspond-il à un corps noir porté à 3 K ? C'est la **théorie de l'expansion de l'Univers** qui apporte la réponse. Pendant un intervalle de temps donné, toutes les distances intergalactiques s'accroissent dans les mêmes proportions. Ce qu'il est important de noter, c'est que les longueurs d'onde s'accroissent également dans les mêmes proportions. On estime que les distances se sont accrues d'un facteur 1000 depuis l'instant où la température était de 3000 K. Les longueurs d'onde que nous observons actuellement étaient 1000 fois plus faibles à l'origine, et les pulsations 1000 fois plus élevées. Le spectre que nous enregistrons actuellement correspond donc à une température 1000 fois plus faible que la température initiale, c'est-à-dire environ 3 K.

2.5 L'oscillateur harmonique et le gaz de phonons

2.5.1 L'oscillateur harmonique linéaire quantique

En mécanique quantique, un oscillateur linéaire est typiquement une particule de masse m dans un champ de potentiel $U(x)$ présentant un minimum, par exemple en $x = 0$: $U'(0) = 0$. La valeur de $U(0)$ n'a aucune importance : on la choisira nulle (choix d'origine des énergies). En développant $U(x)$ au voisinage de 0, on trouve :

$$U(x) = U(0) + U'(0)x + \frac{U''(0)}{2}x^2 + \frac{U'''(0)}{6}x^3 + \dots = \frac{U''(0)}{2}x^2 + \frac{U'''(0)}{6}x^3 + \dots$$

Dans l'approximation de l'oscillateur harmonique, on ne conserve que le premier terme non nul :

$$U(x) \simeq \frac{m\omega^2}{2}x^2 \quad \text{avec } \omega = \sqrt{\frac{U''(0)}{m}}$$

En mécanique quantique, l'énergie d'un oscillateur harmonique linéaire est quantifiée. On montre que les niveaux d'énergie, non dégénérés, sont donnés par :

$$\varepsilon_n = \hbar\omega\left(n + \frac{1}{2}\right) \quad \text{avec } n \text{ entier positif ou nul}$$

Calculons la fonction de partition : $q(T) = \sum_{n=0}^{\infty} e^{-\varepsilon_n/kT} = e^{-\hbar\omega/2kT} \sum_{n=0}^{\infty} \left(e^{-\hbar\omega/kT}\right)^n$

La somme sur n est une série géométrique : $q(T) = \frac{e^{-\hbar\omega/2kT}}{1 - e^{-\hbar\omega/kT}} = \frac{1}{2 \sinh \frac{\hbar\omega}{2kT}}$

Cette fonction de partition permet de calculer les propriétés ci-dessous.

Probabilité d'observer un état n :

$$P_n = \frac{e^{-\varepsilon_n/kT}}{q(T)} = (1 - e^{-\hbar\omega/kT}) e^{-n\hbar\omega/kT}$$

Nombre quantique moyen :

$$\bar{n} = \sum_{n=0}^{\infty} n P_n = (1 - e^{-\hbar\omega/kT}) \sum_{n=0}^{\infty} n (e^{-\hbar\omega/kT})^n = \frac{1}{e^{\hbar\omega/kT} - 1} \quad (2.12)$$

On a utilisé $\sum_{n=0}^{\infty} n x^n = \frac{x}{(1-x)^2}$.

Énergie moyenne, capacité thermique :

$$\bar{\varepsilon} = \hbar\omega(\bar{n} + 1/2) = \frac{\hbar\omega}{2} \frac{e^{\hbar\omega/kT} + 1}{e^{\hbar\omega/kT} - 1} = \frac{\hbar\omega}{2} \coth \frac{\hbar\omega}{2kT} \quad c = \frac{\partial \bar{\varepsilon}}{\partial T} = k \left(\frac{\frac{\hbar\omega}{2kT}}{\sinh \frac{\hbar\omega}{2kT}} \right)^2$$

Énergie libre, entropie :

$$f = -kT \ln q(T) = kT \ln \left(2 \sinh \frac{\hbar\omega}{2kT} \right) \quad s = \frac{\bar{\varepsilon} - f}{T}$$

On peut constater que la distribution $\bar{n}$ (équation 2.12) suit une loi de Bose-Einstein, avec $\mu = 0$ et $\varepsilon = \hbar\omega$. L'oscillateur dans l'état n peut donc être assimilé à une population de n bosons de potentiel chimique nul et d'énergie $\hbar\omega$. Ces pseudo-particules sont appelés **phonons**. Ce sont des particules de même nature que les photons, à la différence près que les photons existent aussi bien dans la matière que dans le vide (ils ont une identité propre), alors que les phonons ne peuvent exister sans la matière. C'est pourquoi ils sont classés parmi les pseudo-particules.

La théorie quantique permet également de calculer l'extension spatiale Δx_n de l'oscillateur dans l'état n (à un facteur $\sqrt{2}$ près, c'est l'équivalent de l'amplitude d'oscillation de la mécanique classique) :

$$\Delta x_n = \sqrt{(n + 1/2) \frac{\hbar}{m\omega}}$$

En remplaçant n par $\bar{n}$, on obtient l'extension spatiale moyenne :

$$\Delta x(T) = \sqrt{(\bar{n}(T) + 1/2) \frac{\hbar}{m\omega}} = \sqrt{\frac{\bar{\varepsilon}(T)}{m\omega^2}} = \sqrt{\frac{\hbar}{2m\omega} \coth \frac{\hbar\omega}{2kT}}$$

La longueur moyenne $\ell(T)$ occupée par l'oscillateur valant $2\Delta x(T)$, il est possible de calculer un coefficient de dilatation thermique. On trouve :

$$\alpha(T) = \frac{1}{\ell} \frac{d\ell}{dT} = \frac{1}{T} \frac{\frac{\hbar\omega}{kT}}{\sinh \frac{\hbar\omega}{kT}}$$

Considérons maintenant un oscillateur confiné sur un segment de longueur a . Autrement dit une particule de masse m dans un champ d'énergie potentielle présentant deux branches infinies en $-a/2$ et $+a/2$ lui interdisant toute possibilité de

quitter le voisinage de $x = 0$. Lorsque a varie, on suppose que le graphe de $U(x)$ varie de façon homothétique (autrement dit, U est fonction de x/a uniquement). On en déduit que la dérivée seconde $U''(0)$ varie en $1/a^2$, et donc que la pulsation varie en $1/a$ (on suppose toujours que le profil est parabolique au voisinage de 0). Autrement dit :

$$\frac{d\omega}{\omega} = -\frac{da}{a}$$

On peut maintenant calculer la tension de ligne en dérivant l'énergie libre f par rapport à a :

$$\tau = -\left[\frac{\partial f}{\partial a}\right]_T = -\left[\frac{d\omega}{da} \frac{\partial f}{\partial \omega}\right]_T = \frac{\hbar\omega}{2a} \coth \frac{\hbar\omega}{2kT}$$

Pour calculer le coefficient de dilatation thermique α (à τ constant) et le module de compressibilité χ (à T constant), il faut connaître l'expression de a en fonction de T et de τ . Il est malheureusement impossible d'extraire $a(T, \tau)$ de l'expression de τ ci-dessus (ω dépend de a). On est alors conduit à écrire :

$$\begin{aligned} \alpha = \frac{1}{a} \left[\frac{\partial a}{\partial T}\right]_{\tau} \quad \text{et} \quad \chi = -\frac{1}{a} \left[\frac{\partial a}{\partial \tau}\right]_T &\Rightarrow da = a(\alpha dT - \chi d\tau) \\ \Rightarrow d\tau = \frac{\alpha}{\chi} dT - \frac{1}{a\chi} da &\Rightarrow \frac{\alpha}{\chi} = \left[\frac{\partial \tau}{\partial T}\right]_a \quad \text{et} \quad \frac{1}{\chi} = -a \left[\frac{\partial \tau}{\partial a}\right]_T \end{aligned}$$

Les calculs donnent :

$$\alpha = \frac{1}{T} \frac{\frac{\hbar\omega}{2kT}}{\sinh \frac{\hbar\omega}{kT} - \frac{\hbar\omega}{2kT}} \quad \text{et} \quad \chi = \frac{2a}{\hbar\omega} \frac{\sinh^2 \frac{\hbar\omega}{2kT}}{\sinh \frac{\hbar\omega}{kT} - \frac{\hbar\omega}{2kT}}$$

Ci-dessous les principaux résultats (on a posé $u = \hbar\omega/2kT$) :

$$\begin{aligned} \bar{\varepsilon} &= \frac{\hbar\omega}{2} \coth u & c_a &= k \frac{u^2}{\sinh^2 u} & f &= kT \ln(2 \sinh u) & s &= \frac{\bar{\varepsilon} - f}{T} \\ \tau &= \frac{1}{a} \frac{\hbar\omega}{2} \coth u & \alpha &= \frac{1}{T} \frac{u}{\sinh(2u) - u} & \chi &= \frac{2a}{\hbar\omega} \frac{\sinh^2 u}{\sinh(2u) - u} \end{aligned}$$

A haute température ($kT \gg \hbar\omega$), on obtient au premier ordre des résultats tout-à-fait corrects :

$$\begin{aligned} \bar{\varepsilon} &\simeq kT \left(1 + \frac{1}{3}u^2\right) & c &\simeq k \left(1 - \frac{1}{3}u^2\right) & \tau &\simeq \frac{kT}{a} \left(1 + \frac{1}{3}u^2\right) \\ \alpha &\simeq \frac{1}{T} \left(1 - \frac{4}{3}u^2\right) & \chi &\simeq \frac{a}{kT} (1 - u^2) \simeq \frac{1}{\tau} \left(1 - \frac{2}{3}u^2\right) \end{aligned}$$

2.5.2 Vibrations de réseau : modèle d'Einstein

Dans un cristal parfait, les atomes sont rangés de façon tri-périodique sur un réseau de sites, avec impossibilité de les quitter. N'ayant aucun degré de liberté en translation, chacun possède trois degrés de liberté de vibration.

Physiquement, le système peut être décrit comme un ensemble de masses ponctuelles (les noyaux) baignant dans un nuage d'électrons. Ce nuage d'électrons fait office de champ d'énergie potentielle constant (dans le temps) et non homogène. Les positions d'équilibre des noyaux rendent cette énergie potentielle minimale : ses dérivées premières par rapport aux coordonnées de position sont nulles, et ce sont les dérivées secondes qui, pour des vibrations de faible amplitude, vont déterminer les fréquences de vibration.

Mécaniquement, il existe plusieurs modèles permettant de décrire le système. Nous en étudierons deux : le modèle d'Einstein et le modèle de Debye. Le modèle d'Einstein, peu performant à basse température, a l'avantage d'être simple. Son intérêt pratique est limité, mais c'est historiquement le premier. Il a permis à Debye de construire le sien, beaucoup plus performant, mais requérant des connaissances avancées en cristallographie.

Dans le modèle d'Einstein, les atomes, bien que liés entre eux par des forces interatomiques importantes, sont supposés indépendants. On supposera pour simplifier que les N atomes sont tous identiques et qu'ils sont arrangés sur un réseau cubique simple. Chacun occupe une "boîte" de volume $v = a^3$, identique pour tous les atomes. En conséquence, ils vibrent *tous à la même fréquence* indépendamment les uns des autres. Cette fréquence sera notée ω_E . La loi de variation de ω_E en fonction de v sera supposée telle que :

$$\frac{d\omega_E}{\omega_E} = -\frac{da}{a} = -\frac{1}{3} \frac{dv}{v}$$

Chaque atome dans sa boîte constitue un oscillateur harmonique tridimensionnel que l'on peut décomposer en trois oscillateurs harmoniques linéaires et indépendants. Le système en contient donc $3N$. Puisqu'on a affaire à un réseau, on peut écrire :

$$Q(T, V, N) = q(T, v)^{3N} \quad \text{avec} \quad V = N v \quad \text{et} \quad q(T, v) = \frac{1}{2 \sinh \frac{\hbar \omega_E(v)}{2kT}}$$

Les résultats qui suivent ont été trouvés selon la démarche du paragraphe précédent. A quelques nuances près (facteur d'échelle $3N$, changement de variable $a \rightarrow v$), ils sont tout à fait similaires. On a posé $u = \hbar \omega_E / 2kT$:

$$E = 3N \frac{\hbar \omega_E}{2} \coth u \quad C_V = 3Nk \frac{u^2}{\sinh^2 u} \quad F = 3NkT \ln(2 \sinh u) \quad S = \frac{E - F}{T}$$

$$p = \frac{1}{v} \frac{\hbar \omega_E}{2} \coth u \quad \alpha = \frac{1}{T} \frac{3u}{2 \sinh(2u) - u} \quad \chi = \frac{6v}{\hbar \omega_E} \frac{\sinh^2 u}{2 \sinh(2u) - u}$$

A haute température, ces résultats deviennent :

$$E \simeq 3NkT \left(1 + \frac{1}{3}u^2\right) \quad C_V \simeq 3Nk \left(1 - \frac{1}{3}u^2\right) \quad p \simeq \frac{kT}{v} \left(1 + \frac{1}{3}u^2\right)$$

$$\alpha \simeq \frac{1}{T} \left(1 - \frac{8}{9}u^2\right) \quad \chi \simeq \frac{v}{kT} \left(1 - \frac{5}{9}u^2\right) \simeq \frac{1}{p} \left(1 - \frac{20}{9}u^2\right)$$

A basse température, la capacité thermique s'écrit :

$$C_V \simeq 3Nk \left(\frac{\hbar\omega_E}{kT} \right)^2 e^{-\hbar\omega_E/kT}$$

Ce résultat n'est pas satisfaisant : expérimentalement, les courbes $C_V^{(\text{vib})}(T)$ suivent à basse température des lois en T^3 .

2.5.3 Vibrations de réseau : modèle de Debye

Modélisation, équations du mouvement

On supposera, pour simplifier les écritures, que les N atomes ($\alpha = 1$ à N) sont identiques (de masse M) et qu'ils occupent à l'équilibre les nœuds d'un réseau cubique simple : dans un repère $(O, \vec{e}_1, \vec{e}_2, \vec{e}_3)$ orthonormé, les positions d'équilibre s'écrivent :

$$\vec{R}_\alpha = a(n_{\alpha 1}\vec{e}_1 + n_{\alpha 2}\vec{e}_2 + n_{\alpha 3}\vec{e}_3) \quad (2.13)$$

avec $n_{\alpha 1}, n_{\alpha 2}, n_{\alpha 3}$ entiers relatifs (a est appelé paramètre de maille).

Point de départ du modèle mécanique : l'énergie potentielle est fonction de $3N$ coordonnées de position. Son développement de Taylor au voisinage de la configuration d'équilibre doit donc faire apparaître toutes les dérivées secondes.

En plaçant l'origine des énergies au minimum de l'énergie potentielle, et sachant que les dérivées premières sont nulles à l'équilibre, on doit donc écrire :

$$U = -\frac{1}{2} \sum_{\alpha=1}^N \sum_{\beta=1}^N \sum_{i=1}^3 \sum_{j=1}^3 K_{\alpha i}^{\beta j} x_{\alpha i} x_{\beta j} \quad \text{avec} \quad K_{\alpha i}^{\beta j} = - \left. \frac{\partial^2 U}{\partial x_{\alpha i} \partial x_{\beta j}} \right|_{\text{équil}} \quad (2.14)$$

Le développement ci-dessus est écrit en termes de déplacements : $(x_{\alpha i}, i = 1 \text{ à } 3)$ et $(x_{\beta j}, j = 1 \text{ à } 3)$ sont les composantes des déplacements de l'atome α et de l'atome β par rapport à leurs positions d'équilibre.

Les équations du mouvement s'écrivent :

$$M\ddot{x}_{\alpha i} = F_{\alpha i} = - \frac{\partial U}{\partial x_{\alpha i}} = \sum_{\beta=1}^N \sum_{j=1}^3 K_{\alpha i}^{\beta j} x_{\beta j} \quad (\alpha = 1 \text{ à } N \text{ et } i = 1 \text{ à } 3) \quad (2.15)$$

Avant d'exploiter ces équations, montrons que le système est mécaniquement identique à une population de N masses ponctuelles reliées entre elles par des ressorts.

On utilise dans un premier temps l'invariance par translation. Si on déplace tous les atomes d'une distance unité dans la même direction, aucune distance relative entre atomes n'est modifiée. Toutes les forces sont nulles, et les équations 2.15 s'écrivent :

$$0 = \sum_{\beta=1}^N \sum_{j=1}^3 K_{\alpha i}^{\beta j} \quad \Rightarrow \quad K_{\alpha i}^{\alpha j} = - \sum_{\beta \neq \alpha}^N \sum_{j=1}^3 K_{\alpha i}^{\beta j} \quad \forall \alpha, i, j \quad (2.16)$$

On remplace dans 2.15 :

$$M\ddot{x}_{\alpha i} = \sum_{\beta \neq \alpha} \sum_{j=1}^3 K_{\alpha i}^{\beta j} (x_{\beta j} - x_{\alpha j}) \quad (\alpha = 1 \text{ à } N \text{ et } i = 1 \text{ à } 3) \quad (2.17)$$

Dans un deuxième temps, on étudie un système constitué de N masses ponctuelles reliées entre elles par des ressorts de constantes de raideur C_{α}^{β} . Lorsqu'on déplace l'atome α d'un vecteur $\vec{r}_{\alpha}$ et l'atome β d'un vecteur $\vec{r}_{\beta}$, le ressort " $\alpha\beta$ " s'allonge au premier ordre de $\delta\ell_{\alpha\beta} = \vec{e}_{\alpha}^{\beta} \cdot (\vec{r}_{\beta} - \vec{r}_{\alpha})$, $\vec{e}_{\alpha}^{\beta}$ étant le vecteur unitaire porté par la ligne du ressort au repos. L'énergie potentielle de ce système s'écrit donc :

$$U = \frac{1}{4} \sum_{\alpha=1}^N \sum_{\beta \neq \alpha} C_{\alpha}^{\beta} \delta\ell_{\alpha\beta}^2 = \frac{1}{4} \sum_{\alpha=1}^N \sum_{\beta \neq \alpha} C_{\alpha}^{\beta} \left(\vec{e}_{\alpha}^{\beta} \cdot (\vec{r}_{\beta} - \vec{r}_{\alpha}) \right)^2 \quad (2.18)$$

Le facteur $1/4$ au lieu de $1/2$ vient du fait que dans la somme chaque couple (α, β) a été compté deux fois. Les équations du mouvement de ce système s'écrivent :

$$M\ddot{x}_{\alpha i} = \sum_{\beta \neq \alpha} \sum_{j=1}^3 C_{\alpha}^{\beta} e_{\alpha i}^{\beta} e_{\alpha j}^{\beta} (x_{\beta j} - x_{\alpha j}) \quad (\alpha = 1 \text{ à } N \text{ et } i = 1 \text{ à } 3) \quad (2.19)$$

Comparant 2.19 à 2.17, on en déduit que les deux systèmes sont équivalents si on a $K_{\alpha i}^{\beta j} = C_{\alpha}^{\beta} e_{\alpha i}^{\beta} e_{\alpha j}^{\beta}$, les $e_{\alpha j}^{\beta}$ étant les cosinus directeurs de la ligne du ressort " $\alpha\beta$ ".

Solutions particulières : ondes progressives

Revenons au système 2.15. Les $3N$ équations sont linéaires, mais chacune d'elles met en jeu toutes les variables.

C'est un système d'équations gigantesque que l'on ne peut pas résoudre sans utiliser la périodicité du réseau. Celle-ci permet de ramener l'étude à celle du contenu d'une maille (ici un seul atome).

On recherche des solutions particulières de la forme $x_{\alpha i} = u_i e^{i(\vec{k} \cdot \vec{R}_{\alpha} - \omega t)}$ (ondes progressives). En remplaçant dans 2.15, on trouve en utilisant 2.16 :

$$\omega^2 u_i - \sum_{j=1}^3 \Delta_{ij}(\vec{k}) u_j = 0 \quad (2.20)$$

$$\text{avec } \Delta_{ij}(\vec{k}) = -\frac{1}{M} \sum_{\beta=1}^N K_{\alpha i}^{\beta j} e^{i\vec{k} \cdot \vec{R}_{\alpha\beta}} = \frac{1}{M} \sum_{\beta \neq \alpha} K_{\alpha i}^{\beta j} (1 - e^{i\vec{k} \cdot \vec{R}_{\alpha\beta}}) \quad (2.21)$$

Quelques remarques au sujet des $\Delta_{ij}(\vec{k})$:

- Les $\Delta_{ij}(\vec{k})$ sont indépendants de α : $K_{\alpha i}^{\beta j}$ ne dépend que des $\vec{R}_{\alpha\beta} = \vec{R}_{\beta} - \vec{R}_{\alpha}$, c'est-à-dire des positions relatives des atomes β par rapport à l'atome α . Le choix de l'atome α utilisé pour le calcul est tout à fait arbitraire (c'est ici qu'on utilise la périodicité).

- Les $\Delta_{ij}(\vec{k})$ sont tripériodiques. En utilisant 2.13, on peut écrire $\vec{R}_{\alpha\beta}$ sous la forme $a(n_1\vec{e}_1 + n_2\vec{e}_2 + n_3\vec{e}_3)$. On en déduit ($n_1, n_2, n_3, m_1, m_2, m_3$, sont des entiers) :

$$\Delta_{ij}(\vec{k} + \vec{K}_m) = \Delta_{ij}(\vec{k}) \quad \text{pour tout } \vec{K}_m = \frac{2\pi}{a}(m_1\vec{e}_1 + m_2\vec{e}_2 + m_3\vec{e}_3)$$

Conclusion : les vecteurs $\vec{k}$ significatifs pointent dans un cube d'arête $2\pi/a$ centré sur l'origine (c'est ce qu'on appelle la **zone de Brillouin**). Chaque composante de chaque vecteur $\vec{k}$ est comprise entre $-\pi/a$ et $+\pi/a$.

Pour calculer les $\Delta_{ij}(\vec{k})$, on choisit arbitrairement un atome α , et on effectue les sommations sur ses premiers voisins, deuxièmes voisins, etc. Le voisinage de cet atome est choisi en fonction de la précision voulue en décidant que les interactions sont négligeables au delà d'une certaine distance entre atomes.

Quoiqu'il en soit, le système à résoudre est maintenant constitué de trois équations :

$$\begin{aligned} (\omega^2 - \Delta_{11})u_1 - \Delta_{12}u_2 - \Delta_{13}u_3 &= 0 \\ \Delta_{21}u_1 - (\omega^2 - \Delta_{22})u_2 - \Delta_{23}u_3 &= 0 \\ \Delta_{31}u_1 - \Delta_{32}u_2 - (\omega^2 - \Delta_{33})u_3 &= 0 \end{aligned} \quad (2.22)$$

Les déplacements nuls étant sans intérêt, il est nécessaire que le déterminant de ce système soit nul. C'est ce qu'on appelle l'équation séculaire qui permet de trouver les trois pulsations propres ω_ℓ du système ($\ell = 1$ à 3). Chacune sera fonction du vecteur d'onde $\vec{k}$ choisi. Les fonctions $\omega_\ell(\vec{k})$ constituent ce qu'on appelle la loi de dispersion. Cette loi de dispersion est ici constituée de trois *branches* $\ell = 1$ à 3 . On peut remarquer que tous les Δ_{ij} sont nuls pour $\vec{k} = \vec{0}$. On en déduit qu'on a pour les trois branches $\omega_\ell(\vec{0}) = 0$ (les trois branches sont dites *acoustiques*).

Une fois les trois pulsations propres trouvées, on les remplace une à une dans le système 2.22. On trouve pour chaque $\omega_\ell(\vec{k})$ un vecteur $\vec{u}_\ell(\vec{k})$ (que l'on choisit normé) de composantes $(u_{i\ell}(\vec{k}), i = 1 \text{ à } 3)$. Un théorème mathématique bien connu nous assure l'orthogonalité de ces trois vecteurs. C'est pourquoi on les appelle *modes propres* ou *modes orthogonaux* ou encore *modes normaux* de vibration :

$$\vec{u}_\ell(\vec{k}) \cdot \vec{u}_{\ell'}(\vec{k}) = \delta_{\ell\ell'} \quad \text{pour tout } \vec{k} \quad (2.23)$$

Pour trouver les vecteurs $\vec{k}$ possibles, on utilise les conditions aux limites. Il nous faut N vecteurs $\vec{k}$, de façon à ce que le nombre de pulsations propres soit égal au nombre de degrés de liberté du système, c'est-à-dire $3N$.

Les conditions aux limites dépendent de la forme de l'échantillon. Cette forme est arbitraire : on étudie le système pour ses propriétés macroscopiques et celles-ci sont indépendantes de la forme de l'échantillon. Usuellement, on utilise les conditions aux limites périodiques : on choisit un échantillon cubique de côté $L = na$ (avec $n^3 = N$), et on exige que L soit un nombre (positif ou négatif) de longueurs d'onde $\lambda = 2\pi/k$ de façon à ce que l'onde traverse l'échantillon sans variation d'énergie (le système est supposé isolé). Cette condition s'écrit :

$$\vec{k} = \frac{2\pi}{na}(m_1, m_2, m_3) \quad \text{avec } m_1, m_2, m_3 \text{ entiers relatifs} \quad (2.24)$$

Pour que $\vec{k}$ pointe dans la zone de Brillouin, on a, en supposant n impair :

$$-\frac{n-1}{2} \leq m_1, m_2, m_3 \leq +\frac{n-1}{2} \quad (2.25)$$

On a bien n valeurs possibles pour chacune des composantes, et donc au total $N = n^3$ vecteurs $\vec{k}$ possibles. On dispose donc maintenant, non pas de $3N$, mais de $6N$ solutions particulières de l'équation du mouvement 2.15. En effet, remplacer $e^{-i\omega_\ell(\vec{k})t}$ par $e^{+i\omega_\ell(\vec{k})t}$ ne change rien dans tous les calculs qui précèdent :

$$\begin{aligned} \vec{r}_{\alpha\ell}(\vec{k}) &= \vec{u}_\ell(\vec{k}) e^{i\vec{k}\cdot\vec{R}_\alpha} e^{-i\omega_\ell(\vec{k})t} \quad \text{et} \quad \vec{r}'_{\alpha\ell}(\vec{k}) = \vec{u}_\ell(\vec{k}) e^{i\vec{k}\cdot\vec{R}_\alpha} e^{+i\omega_\ell(\vec{k})t} \\ (\ell = 1 \text{ à } 3 \text{ et } \vec{k} \text{ définis par 2.24 et 2.25}) \end{aligned}$$

Solutions générales

On fabrique la solution générale par combinaisons linéaires **réelles** des solutions particulières. Soient $\Lambda_\ell(\vec{k})$ et $\Lambda'_\ell(\vec{k})$ les coefficients de ces combinaisons linéaires :

$$\vec{r}_\alpha = \sum_\ell \sum_{\vec{k} \in \text{ZdB}}^{(2.24)} \Lambda_\ell(\vec{k}) \vec{r}_{\alpha\ell}(\vec{k}) + \sum_\ell \sum_{\vec{k} \in \text{ZdB}}^{(2.24)} \Lambda'_\ell(\vec{k}) \vec{r}'_{\alpha\ell}(\vec{k})$$

Les sommes sur $\vec{k}$ portent sur les vecteurs vérifiant 2.24 et pointant dans la zone de Brillouin (équation 2.25). Parmi ces $6N$ termes, $3N$ sont redondants. Pour trouver $3N$ termes indépendants donnant une somme **réelle**, on doit découper la zone de Brillouin en deux demi-zones Z_1 et Z_2 , avec : $\forall \vec{k} \in Z_1, -\vec{k} \in Z_2$.

La première sommation sur $\vec{k}$ s'effectue sur Z_1 , et la deuxième sur Z_2 :

$$\vec{r}_\alpha = \sum_\ell \sum_{\vec{k} \in Z_1}^{(2.24)} \Lambda_\ell(\vec{k}) e^{-i\omega_\ell(\vec{k})t} \vec{u}_\ell(\vec{k}) e^{i\vec{k}\cdot\vec{R}_\alpha} + \sum_\ell \sum_{\vec{k} \in Z_2}^{(2.24)} \Lambda'_\ell(\vec{k}) e^{+i\omega_\ell(\vec{k})t} \vec{u}_\ell(\vec{k}) e^{i\vec{k}\cdot\vec{R}_\alpha}$$

$$\text{Et on pose : } \begin{cases} q_\ell(\vec{k}, t) = \sqrt{NM} \Lambda_\ell(\vec{k}) e^{-i\omega_\ell(\vec{k})t} & \forall \vec{k} \in Z_1 \\ q_\ell(\vec{k}, t) = \sqrt{NM} \Lambda'_\ell(\vec{k}) e^{+i\omega_\ell(\vec{k})t} & \forall \vec{k} \in Z_2 \end{cases}$$

avec $\Lambda'_\ell(-\vec{k}) = \Lambda_\ell^*(\vec{k})$ pour que $\vec{r}_\alpha(t)$ soit réel.

En résumé, on pose :

$$\vec{r}_\alpha(t) = \frac{1}{\sqrt{NM}} \sum_\ell \sum_{\vec{k} \in \text{ZdB}} q_\ell(\vec{k}, t) \vec{u}_\ell(\vec{k}) e^{i\vec{k}\cdot\vec{R}_\alpha} \quad \text{avec} \quad q_\ell(-\vec{k}, t) = q_\ell^*(\vec{k}, t)$$

$$\text{et} \quad \vec{u}_\ell(-\vec{k}) = \vec{u}_\ell(\vec{k}) \quad \omega_\ell(-\vec{k}) = \omega_\ell(\vec{k}) \quad \vec{u}_\ell(\vec{k}) \cdot \vec{u}_{\ell'}(\vec{k}) = \delta_{\ell\ell'}$$

A partir de maintenant, dans le souci de condenser les écritures, la variable t sera omise. De plus, le vecteur $\vec{k}$ passe en indice (sans flèche), et le vecteur $-\vec{k}$ sera noté $\bar{k}$. Et enfin, dans les sommations sur ℓ et $\vec{k}$, les bornes sont supprimées.

Énergie cinétique

$$T = \frac{M}{2} \sum_{\alpha=1}^N \sum_{i=1}^3 \dot{x}_{\alpha i}^2 = \frac{M}{2} \sum_{\alpha=1}^N \sum_{i=1}^3 \left[\frac{1}{\sqrt{NM}} \sum_{\ell k} q_{\ell k} u_{i\ell k} e^{i\vec{k} \cdot \vec{R}_{\alpha}} \right] \left[\frac{1}{\sqrt{NM}} \sum_{\ell' k'} q_{\ell' k'} u_{i\ell' k'} e^{i\vec{k}' \cdot \vec{R}_{\alpha}} \right]$$

On permute judicieusement les signes sommes :

$$T = \frac{1}{2N} \sum_{\ell k} \sum_{\ell' k'} \dot{q}_{\ell k} \dot{q}_{\ell' k'} \sum_{i=1}^3 u_{i\ell k} u_{i\ell' k'} \sum_{\alpha=1}^N e^{i(\vec{k} + \vec{k}') \cdot \vec{R}_{\alpha}}$$

Compte tenu de 2.24 et 2.25, la somme sur α vaut $N\delta_{\vec{k}\vec{k}'}$ (N si $\vec{k}' = -\vec{k}$ et 0 sinon). Une fois $\vec{k}'$ remplacé par $-\vec{k}$, la somme sur i vaut $\vec{u}_{\ell k} \cdot \vec{u}_{\ell' \bar{k}} = \vec{u}_{\ell k} \cdot \vec{u}_{\ell' k} = \delta_{\ell\ell'}$:

$$T = \frac{1}{2} \sum_{\ell k} \sum_{\ell'} \dot{q}_{\ell k} \dot{q}_{\ell' \bar{k}} \sum_{i=1}^3 u_{i\ell k} u_{i\ell' \bar{k}} = \frac{1}{2} \sum_{\ell k} \dot{q}_{\ell k} \dot{q}_{\ell \bar{k}} = \frac{1}{2} \sum_{\ell k} \dot{q}_{\ell k} \dot{q}_{\ell k}^*$$

Énergie potentielle

On repart de l'équation initiale 2.14 :

$$U = -\frac{1}{2} \sum_{\alpha=1}^N \sum_{i=1}^3 \sum_{\beta=1}^N \sum_{j=1}^3 K_{\alpha i}^{\beta j} x_{\alpha i} x_{\beta j} \quad \text{avec} \quad \begin{cases} \dot{x}_{\alpha i} = \frac{1}{\sqrt{NM}} \sum_{\ell k} q_{\ell k} u_{i\ell k} e^{i\vec{k} \cdot \vec{R}_{\alpha}} \\ \dot{x}_{\beta j} = \frac{1}{\sqrt{NM}} \sum_{\ell' k'} q_{\ell' k'} u_{j\ell' k'} e^{i\vec{k}' \cdot \vec{R}_{\beta}} \end{cases}$$

En remarquant que $\vec{k} \cdot \vec{R}_{\alpha} + \vec{k}' \cdot \vec{R}_{\beta} = (\vec{k} + \vec{k}') \cdot \vec{R}_{\alpha} + \vec{k}' \cdot (\vec{R}_{\beta} - \vec{R}_{\alpha})$:

$$U = -\frac{1}{2NM} \sum_{\ell k} \sum_{\ell' k'} q_{\ell k} q_{\ell' k'} \sum_{i=1}^3 u_{i\ell k} \sum_{\alpha=1}^N e^{i(\vec{k} + \vec{k}') \cdot \vec{R}_{\alpha}} \sum_{j=1}^3 u_{j\ell' k'} \sum_{\beta=1}^N K_{\alpha i}^{\beta j} e^{i\vec{k}' \cdot \vec{R}_{\alpha\beta}}$$

La somme sur β vaut $-M\Delta_{ij}(\vec{k}') = -M\Delta_{ijk'}$ (équation 2.21) :

$$U = \frac{1}{2N} \sum_{\ell k} \sum_{\ell' k'} q_{\ell k} q_{\ell' k'} \sum_{i=1}^3 u_{i\ell k} \sum_{\alpha=1}^N e^{i(\vec{k} + \vec{k}') \cdot \vec{R}_{\alpha}} \sum_{j=1}^3 \Delta_{ijk'} u_{j\ell' k'}$$

La somme sur j vaut $\omega_{\ell' k'}^2 u_{i\ell' k'}$ (équation 2.20) :

$$U = \frac{1}{2N} \sum_{\ell k} \sum_{\ell' k'} \omega_{\ell' k'}^2 q_{\ell k} q_{\ell' k'} \sum_{i=1}^3 u_{i\ell k} u_{i\ell' k'} \sum_{\alpha=1}^N e^{i(\vec{k} + \vec{k}') \cdot \vec{R}_{\alpha}}$$

Compte tenu de 2.24 et 2.25, la somme sur α vaut $N\delta_{\vec{k}\vec{k}'}$ (N si $\vec{k}' = -\vec{k}$ et 0 sinon). Une fois $\vec{k}'$ remplacé par $-\vec{k}$, la somme sur i vaut $\vec{u}_{\ell k} \cdot \vec{u}_{\ell' \bar{k}} = \vec{u}_{\ell k} \cdot \vec{u}_{\ell' k} = \delta_{\ell\ell'}$:

$$U = \frac{1}{2} \sum_{\ell k} \sum_{\ell'} \omega_{\ell' \bar{k}}^2 q_{\ell k} q_{\ell' \bar{k}} \sum_{i=1}^3 u_{i\ell k} u_{i\ell' \bar{k}} = \frac{1}{2} \sum_{\ell k} \omega_{\ell k}^2 q_{\ell k} q_{\ell \bar{k}} = \frac{1}{2} \sum_{\ell k} \omega_{\ell k}^2 q_{\ell k} q_{\ell k}^*$$

Lagrangien, hamiltonien, nouvelles équations du mouvement

Pour le calcul des momentaides $p_{\ell k}$ conjugués aux $q_{\ell k}$, ne pas oublier que chaque vecteur $\vec{k}$ apparaît deux fois dans la somme :

$$L = T - U = \frac{1}{2} \sum_{\ell k} (\dot{q}_{\ell k} \dot{q}_{\ell k}^* - \omega_{\ell k}^2 q_{\ell k} q_{\ell k}^*) = \frac{1}{2} \sum_{\ell k} (\dot{q}_{\ell k} \dot{q}_{\ell \bar{k}} - \omega_{\ell k}^2 q_{\ell k} q_{\ell \bar{k}})$$

$$p_{\ell k} = \frac{\partial L}{\partial \dot{q}_{\ell k}} = \dot{q}_{\ell \bar{k}} = \dot{q}_{\ell k}^* \quad \text{et} \quad p_{\ell \bar{k}} = \frac{\partial L}{\partial \dot{q}_{\ell \bar{k}}} = \dot{q}_{\ell k} = p_{\ell k}^*$$

$$H = T + U = \frac{1}{2} \sum_{\ell k} (p_{\ell k} p_{\ell k}^* + \omega_{\ell k}^2 q_{\ell k} q_{\ell k}^*) = \frac{1}{2} \sum_{\ell k} (p_{\ell k} p_{\ell \bar{k}} + \omega_{\ell k}^2 q_{\ell k} q_{\ell \bar{k}})$$

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{q}_{\ell \bar{k}}} - \frac{\partial L}{\partial q_{\ell \bar{k}}} = 0 \quad \Rightarrow \quad \ddot{q}_{\ell k} + \omega_{\ell k}^2 q_{\ell k} = 0$$

Étude statistique, propriétés thermodynamiques

Grâce au changement de variables $x_{\alpha i} \rightarrow q_{\ell k}$, le système est maintenant décrit comme une population de $3N$ oscillateurs harmoniques unidimensionnels indépendants. On peut donc écrire comme au paragraphe 2.5.1 (la constante de Boltzmann sera ici notée k_B) :

$$E = \sum_{\ell k} \hbar \omega_{\ell k} (\bar{n}_{\ell k} + 1/2) \quad \text{avec} \quad \bar{n}_{\ell k} = \frac{1}{e^{\hbar \omega_{\ell k}/k_B T} - 1} \quad (2.26)$$

Il est clair que chacun des $3N$ oscillateurs n'est plus *localisé* sur un atome particulier du cristal. Ce sont des *oscillations collectives, délocalisées*, qui concernent le cristal entier. Le système est maintenant équivalent à un **gaz parfait de phonons**.

Avant d'étudier le système en fonction de la température, répertorions les propriétés utiles de la loi de dispersion $\omega_{\ell k} = \omega_{\ell}(\vec{k})$:

- Elle est constituée de trois branches dites *acoustiques* ($\omega_{\ell}(\vec{0}) = 0$). Pour un système plus quelconque que celui que nous avons étudié (maille non cubique, n_a atomes par maille), la loi de dispersion comporte $3n_a$ branches, dont 3 branches acoustiques et $3n_a - 3$ branches dites *optiques* ($\omega_{\ell}(\vec{0}) \neq 0$).
- Au voisinage de l'origine (grandes longueurs d'onde), les trois branches acoustiques sont linéaires ; la pente $v_{\ell} = d\omega_{\ell}/dk$ dans une direction $\vec{e}_k = \vec{k}/k$ donnée est par définition la vitesse de groupe de l'onde dans cette direction (c'est la vitesse du son). De plus, si pour chaque $\omega_{\ell}(\vec{k})$ on calcule le jeu de déplacements instantanés des atomes, on constate qu'à l'une des trois branches acoustiques sont associés des déplacements longitudinaux (parallèles à $\vec{k}$) et qu'aux deux autres sont associés des déplacements transverses (perpendiculaires à $\vec{k}$).

Dans un réseau cubique (optiquement isotrope), les $\omega_{\ell}(\vec{k})$ ne dépendent que du module de $\vec{k}$: chaque vitesse v_{ℓ} ($\ell = 1$ à 3) est indépendante de la direction $\vec{e}_k$. Toujours pour des raisons d'isotropie, les deux branches transverses sont identiques.

- Les vecteurs $\vec{k}$ significatifs pointent dans la zone de Brillouin ; c'est une portion de l'espace réciproque (l'espace où l'on représente les vecteurs $\vec{k}$) qui, pour un réseau cubique simple, a la forme d'un cube d'arête $2\pi/a$ centré sur l'origine. Chaque extrémité de vecteurs $\vec{k}$ y occupe un élément de volume $8\pi^3/V$, N fois plus petit que celui de la zone de Brillouin. Ainsi, l'ensemble des vecteurs $\vec{k}$ possibles pourra être considéré comme continu, autrement dit on pourra le cas échéant remplacer les signes sommes par des intégrales.

A **haute température** ($k_B T \gg$ tous les $\omega_{\ell k}$), on peut dans 2.26 développer l'exponentielle au deuxième ordre. On trouve :

$$\frac{1}{e^x - 1} + \frac{1}{2} = \frac{1}{2} \coth \frac{x}{2} \approx \frac{1}{x} \Rightarrow E \approx \sum_{\ell k} \hbar \omega_{\ell k} \frac{k_B T}{\hbar \omega_{\ell k}} = 3Nk_B T \quad \text{et} \quad C_V \approx 3Nk_B$$

Si le cristal est plus quelconque (n_a atomes par maille), ces résultats sont à multiplier par n_a . Dans tous les cas, la capacité thermique est bien trouvée constante à haute température : c'est k_B multiplié par le nombre de degrés de liberté du système.

A **basse température**, les hypothèses simplificatrices sont les suivantes :

- Linéarité et isotropie des branches acoustiques ($\omega_{\ell}(\vec{k}) \approx v_{\ell} k$) :

$$E - E_0 \approx \sum_{\ell=1}^3 \sum_{\vec{k} \in \text{ZdB}} \frac{\hbar v_{\ell} k}{e^{\hbar v_{\ell} k / k_B T} - 1} \quad \text{avec} \quad E_0 = \sum_{\ell=1}^3 \sum_{\vec{k} \in \text{ZdB}} \frac{\hbar \omega_{\ell}(\vec{k})}{2}$$

Ceci n'est pas valable sur toute la zone de Brillouin, mais on considère qu'à température suffisamment basse la fonction sommée prend des valeurs négligeables au delà du domaine où les trois branches acoustiques sont linéaires. Mieux encore : on peut se permettre de sommer sur *tous* les vecteurs $\vec{k}$ de l'espace réciproque, et ainsi d'intégrer de 0 à $+\infty$ (voir ci-dessous).

- Approximation du continuum : on remplace $\sum_{\vec{k} \in \text{ZdB}}$ par $\frac{V}{8\pi^3} \int_0^{\infty} 4\pi k^2 dk$.

Ce qui donne finalement, en posant $u = \hbar v_{\ell} k / k_B T$

$$E - E_0 \approx \frac{V}{8\pi^3} \sum_{\ell=1}^3 \int_0^{\infty} \frac{4\pi \hbar v_{\ell} k^3 dk}{e^{\hbar v_{\ell} k / k_B T} - 1} = \frac{V}{2\pi^2} \frac{(k_B T)^4}{\hbar^3} \sum_{\ell=1}^3 \frac{1}{v_{\ell}^3} \int_0^{\infty} \frac{u^3 du}{e^u - 1}$$

L'intégrale vaut $\pi^4/15$:

$$E - E_0 \approx \frac{\pi^2}{10} V \frac{(k_B T)^4}{\hbar^3 \tilde{v}^3} \quad \text{avec} \quad \frac{1}{\tilde{v}^3} = \frac{1}{3} \sum_{\ell=1}^3 \frac{1}{v_{\ell}^3} = \frac{1}{3} \left(\frac{1}{v_L^3} + \frac{2}{v_T^3} \right)$$

La vitesse "moyenne" $\tilde{v}$ est l'inverse de la racine cubique de la valeur moyenne des $1/v_{\ell}^3$ (v_L est la vitesse de groupe de l'onde longitudinale et v_T la vitesse de groupe des deux ondes transverses).

Et enfin, la fameuse loi en T^3 de Debye :

$$C_V(T) = \left. \frac{\partial E}{\partial T} \right]_{V,N} \approx \frac{2\pi^2}{5} V k_B \left(\frac{k_B T}{\hbar \tilde{v}} \right)^3$$

Aux **températures intermédiaires**, on affine simplement le modèle précédent en limitant l'intégrale à la zone de Brillouin. Celle-ci est assimilée à une sphère. En partant d'un cube, c'est une approximation assez grossière. Pour que la sphère soit de même volume que la zone de Brillouin, on définit le vecteur d'onde de Debye k_D :

$$\frac{4}{3}\pi k_D^3 = \frac{8\pi^3}{a^3} \Rightarrow k_D = \frac{(6\pi^2)^{1/3}}{a}$$

Autre approximation grossière : on suppose toujours la linéarité et l'isotropie des trois branches, bien que ce ne soit pas justifié pour les grandes valeurs de k .

Moyennant quoi, on écrit maintenant : $\sum_{\vec{k} \in \text{ZdB}} \approx \frac{V}{8\pi^3} \int_0^{k_D} 4\pi k^2 dk$

Et $E - E_0$ devient (θ_ℓ^D est la *température de Debye* pour la branche ℓ) :

$$E - E_0 \approx \frac{V}{2\pi^2} \frac{(k_B T)^4}{\hbar^3} \sum_{\ell=1}^3 \frac{1}{v_\ell^3} \int_0^{\theta_\ell^D/T} \frac{u^3 du}{e^u - 1} \quad \text{avec} \quad k_B \theta_\ell^D = \hbar \omega_\ell^D = \hbar v_\ell k_D$$

On peut réécrire $E - E_0$ en utilisant la fonction $f(y) = \frac{15}{\pi^4} \frac{1}{y^3} \int_0^y \frac{u^3 du}{e^u - 1}$:

$$E - E_0 \approx \frac{\pi^2}{30} V k_D^3 k_B T \sum_{\ell=1}^3 f(\theta_\ell^D/T) = \frac{3\pi^4}{5} N k_B T \times \frac{1}{3} \sum_{\ell=1}^3 f(\theta_\ell^D/T)$$

On simplifie encore en ne considérant qu'une seule température de Debye. On définit *la* température de Debye θ_D telle que :

$$f(\theta_D/T) = \frac{1}{3} \sum_{\ell=1}^3 f(\theta_\ell^D/T) \Rightarrow E - E_0 \approx \frac{3\pi^4}{5} N k_B T f(\theta_D/T)$$

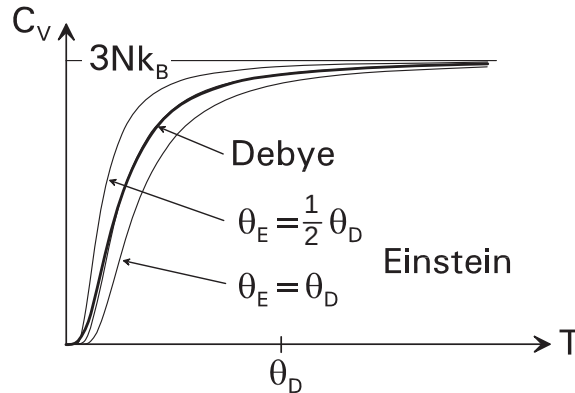


FIG. 2.4 – Vibrations de réseau : capacité thermique

En dérivant par rapport à T , on obtient (f_D est la fonction de Debye) :

$$C_V(T) \approx 3Nk_B f_D(\theta_D/T) \quad \text{avec} \quad f_D(y) = \frac{3}{y^3} \int_0^y \frac{u^4 e^u du}{(e^u - 1)^2}$$

Les calculs numériques conduisent à la courbe représentée figure 2.4, comparée à celles obtenues par Einstein pour différentes valeurs de $\theta_E = \hbar\omega_E/k_B$. Celle de Debye est en bon accord avec les résultats expérimentaux. La loi en T^3 des basses températures est valable jusqu'à environ $\frac{1}{10}\theta_D$. On peut remarquer (au voisinage de l'origine) la présence d'une courbe qui "colle" parfaitement à celle de Debye pour $T > \frac{1}{5}\theta_D$. C'est la courbe obtenue avec le modèle d'Einstein en prenant $\theta_E = \frac{3}{4}\theta_D$. On peut vérifier que cette valeur correspond à une fréquence moyenne du spectre de Debye pondérée par une densité de modes variant en ω^2 . Le modèle d'Einstein, malgré sa simplicité, est donc tout à fait correct au delà de $\frac{1}{5}\theta_D$.

Valeurs numériques : θ_D vaut en général quelques centaines de kelvins (Fe : 455 K, Al : 385 K, Cu : 320 K, Ag : 215 K, Au : 185 K, etc.). Le diamant bat tous les records avec $\theta_D = 2230$ K.

Chapitre 3

Gaz parfait de masses ponctuelles

3.1 Quelques ordres de grandeur

D'abord, un petit tableau rassemblant les préfixes multiplicatifs (l'absence d'accent sur "peta" n'est pas une faute de frappe) :

préfixe	kilo	méga	giga	téra	peta	exa	zetta	yotta
symbole	k	M	G	T	P	E	Z	Y
valeur	10^{+3}	10^{+6}	10^{+9}	10^{+12}	10^{+15}	10^{+18}	10^{+21}	10^{+24}
préfixe	milli	micro	nano	pico	femto	atto	zepto	yocto
symbole	m	μ	n	p	f	a	z	y
valeur	10^{-3}	10^{-6}	10^{-9}	10^{-12}	10^{-15}	10^{-18}	10^{-21}	10^{-24}

3.1.1 Les grandeurs énergétiques

Dans un gaz parfait, l'énergie cinétique moyenne de translation d'une molécule est, aux températures habituelles, proportionnelle à kT . Il en est de même de son énergie cinétique de rotation. Il existe ainsi de nombreux exemples où la quantité kT constitue une énergie de référence à laquelle on compare d'autres grandeurs énergétiques.

La quantité kT est une grandeur microscopique : on l'exprime en atto-, zepto- ou yocto- Joule. Ces préfixes ne suffisent même plus aux cryogénistes qui parfois parlent en nano-Kelvins. C'est pourquoi l'électronvolt ($1 \text{ eV} \approx 0,16 \text{ aJ}$), bien qu'il soit aboli par les instances officielles, est encore largement utilisé : à 1 nK correspond environ $0,1$ picoélectronvolt ; jusqu'au yoctoélectronvolt, il y a encore de la marge ...

T (K)		1		10		100		1 000		10 000		100 000	
kT (zJ)	U	0,014	T	0,14	B	1,4		14	H	140	T	1 400	U
v (m/s)	B	28	B	87	T	280		870	T	2 800	H	8 700	H
ω (THz)	T	0,063	T	0,2		0,63		2		6,3	T	20	T
E _{vib} (zJ)		23		23		23		25		140		1 400	

Le tableau ci-dessus indique la correspondance entre énergie et température. En troisième et quatrième lignes sont indiquées les valeurs moyennes de la vitesse v

d'une molécule d'azote et de sa vitesse angulaire ω de rotation ($1 \text{ Hz} = 1 \text{ tr/s}$). En cinquième ligne, on trouve l'énergie de vibration de la molécule (fréquence : 70 THz).

Dans la gamme des ultra basses températures (UBT), il est possible d'atteindre le milli-Kelvin en réfrigération continue et le micro-kelvin et même le nano-kelvin avec des techniques telles que la désaimantation adiabatique et le refroidissement optique. Le domaine TBT est celui de la liquéfaction de l'hélium (4,2 K). Dans le domaine BT se trouvent les températures de solidification et de liquéfaction de tous les autres gaz.

Dans le domaine HT, on trouve la température de fusion et de vaporisation de la plupart des métaux, ainsi que la température superficielle du soleil (5500 K). Pour les ultra hautes températures (UHT), on peut citer la température au centre d'une explosion thermonucléaire (10^5 à 10^6 K), celle régnant au cœur d'une étoile (10^6 à 10^7 K), et la température exigée pour l'amorçage d'une réaction de fusion thermonucléaire (10^8 K).

3.1.2 Densités et pressions

NB : toutes les densités sont exprimées en molécules par mètre cube ; elles ont été arrondies à la puissance de 10 la plus proche.

Il faut ici distinguer deux domaines bien distincts : l'état gazeux d'une part, les milieux denses (liquides ou solides) d'autre part. Dans ce document, seuls les gaz sont abordés. Au sujet des milieux denses on dira simplement que les densités varient entre 10^{27} et 10^{30} (distance entre deux particules : $d \approx \rho^{-1/3} \sim 0,1$ à 1 nm).

La densité ρ de l'air habituel ($p = 0,1 \text{ MPa}$, $T = 293 \text{ K}$) est d'environ 10^{25} , ce qui donne $d \sim 5 \text{ nm}$. En laboratoire ou dans l'industrie, il existe pour les basses pressions une classification liée à la méthode d'obtention.

Les ***vides moyens ou primaires*** ($p \sim 0,1$ à 100 Pa , $\rho \sim 10^{19}$ à 10^{22} , $d \sim 50$ à 500 nm) sont obtenus au moyen de pompes volumétriques (pompes à palettes mobiles, pompes à bec, ...) qui évacuent le gaz dans l'atmosphère.

Les ***vides poussés ou secondaires*** ($p \sim 10 \mu\text{Pa}$ à $0,1 \text{ Pa}$, $\rho \sim 10^{16}$ à 10^{19} , $d \sim 0,5$ à $5 \mu\text{m}$) sont obtenus au moyen de pompes cinétiques (pompes à diffusion d'huile, pompes turbomoléculaires, ...) qui évacuent le gaz dans une enceinte elle-même vidée par une pompe primaire. Pour atteindre les $10 \mu\text{Pa}$, on équipe une pompe turbomoléculaire d'une pompe dite fermée : le gaz résiduel est fixé par physisorption (adsorption réversible) sur du charbon actif ou une zéolite. Pour que l'adsorption soit suffisante, on doit travailler en dessous de 20 K.

Pour atteindre ***l'ultravide*** ($p \sim 1 \text{ nPa}$ à $10 \mu\text{Pa}$, $\rho \sim 10^{12}$ à 10^{16} , $d \sim 5$ à $100 \mu\text{m}$), la chimisorption (adsorption irréversible) sur du titane est nécessaire.

La plus performante de ces pompes (la pompe ionique de Penning) permet d'atteindre le ***vide extrême*** ($p \sim 1 \text{ pPa}$ à 10 nPa , $\rho \sim 10^9$ à 10^{12} , $d \sim 0,1$ à 1 mm).

Remarque. Le milieu intersidéral n'est pas vide. Il est principalement constitué d'hydrogène et d'hélium. Il n'est pas homogène non plus. Suivant les régions, la température varie de 10 à 100 K. La densité des régions les plus froides est d'environ 10^{15} ($d \sim 10 \mu\text{m}$, $p \sim 0,1 \mu\text{Pa}$). Celle des régions les plus chaudes bat tous les records : $\rho \sim 10^8$ ($d \sim 2 \text{ mm}$, $p \sim 0,1 \text{ pPa}$).

3.2 Particule libre dans une boîte cubique

3.2.1 Résultats de la mécanique quantique

Dans un gaz parfait, les N particules qui le constituent sont indépendantes (sans interactions). On peut considérer que chacune est seule dans le récipient. On est ainsi conduit à étudier un système constitué d'une seule particule. Les propriétés macroscopiques du gaz étant indépendantes de la forme du récipient, on choisit la forme la plus simple qui soit : un cube de volume $V = L^3$, un carré d'aire $A = L^2$ ou un segment de longueur L , suivant la dimension du système étudié. On montre en mécanique quantique qu'une particule de masse m confinée sur un segment de longueur L a accès à des niveaux d'énergie non dégénérés définis par :^[1]

$$\varepsilon_{n_1} = \frac{h^2}{2mL^2} n_1^2 \quad \text{avec } n_1 \text{ entier positif ou négatif}$$

De même, si elle est confinée sur un carré d'aire $A = L^2$, les niveaux d'énergie sont définis par deux nombres entiers :

$$\varepsilon_{n_1 n_2} = \frac{h^2}{2mL^2} (n_1^2 + n_2^2) \quad \text{avec } n_1 \text{ et } n_2 \text{ entiers positifs ou négatifs}$$

Et enfin, si le récipient est cubique ($V = L^3$), trois nombres entiers sont nécessaires :

$$\varepsilon_{n_1 n_2 n_3} = \frac{h^2}{2mL^2} (n_1^2 + n_2^2 + n_3^2) \quad \text{avec } n_1, n_2 \text{ et } n_3 \text{ entiers positifs ou négatifs}$$

La fonction de partition canonique est, si on ne tient pas compte du spin, celle d'une particule à 1, 2 ou 3 degrés de liberté indépendants :

$$\begin{aligned} q(\beta, L) &= \sum_{n=-\infty}^{+\infty} e^{-\beta h^2 n^2 / (2mL^2)} && \text{sur un segment} \\ q(\beta, A) &= \left[\sum_{n=-\infty}^{+\infty} e^{-\beta h^2 n^2 / (2mA)} \right]^2 && \text{sur un carré} \\ q(\beta, V) &= \left[\sum_{n=-\infty}^{+\infty} e^{-\beta h^2 n^2 / (2mV^{2/3})} \right]^3 && \text{dans un cube} \end{aligned}$$

Ces fonctions de partition sont impossibles à calculer rigoureusement, mais l'approximation suivante peut être considérée comme quasi-rigoureuse.

3.2.2 L'approximation du continuum et la densité d'états

L'approximation du continuum consiste à remplacer dans les équations statistiques les sommes discrètes par des intégrales. Elle est justifiée lorsque l'écart entre deux niveaux d'énergie est extrêmement faible par rapport à l'énergie moyenne d'une particule. C'est le cas ici.

¹ $h \approx 6,626 \cdot 10^{-34}$ Js est la constante de Planck.

Soit un système quelconque dont les niveaux d'énergie $\{\varepsilon_1 \dots \varepsilon_j \dots \varepsilon_n\}$ sont rangés par ordre croissant. On notera $\{\omega_1 \dots \omega_j \dots \omega_n\}$ les dégénérescences correspondantes. *L'approximation du continuum* consiste à rechercher une fonction g telle que pour toute fonction f continue on ait :

$$\sum_{j=1}^n \omega_j f(\varepsilon_j) \approx \int_{\varepsilon_1}^{\varepsilon_n} g(\varepsilon) f(\varepsilon) d\varepsilon$$

La fonction g est par définition *la densité d'états* : $g(\varepsilon) d\varepsilon$ est le nombre d'états d'énergie comprise entre ε et $\varepsilon + d\varepsilon$. Pour calculer $g(\varepsilon)$, on calcule d'abord $G(\varepsilon)$, nombre d'états d'énergie *inférieure* à ε . Ensuite, on écrit :

$$g(\varepsilon) d\varepsilon = G(\varepsilon + d\varepsilon) - G(\varepsilon) = \frac{dG}{d\varepsilon} d\varepsilon \quad \Rightarrow \quad g(\varepsilon) = \frac{dG}{d\varepsilon}$$

Pour les **gaz parfaits tridimensionnels**, $G(\varepsilon)$ est le nombre d'états dont le vecteur $\vec{n}$ de composantes (n_1, n_2, n_3) pointe à l'intérieur de la sphère de rayon et de volume :

$$n = \sqrt{\frac{2m V^{2/3} \varepsilon}{h^2}} \quad \text{et} \quad V^* = \frac{4}{3} \pi \left(\frac{2m V^{2/3} \varepsilon}{h^2} \right)^{3/2}$$

Chaque état y occupe un volume $\delta V^* = 1$. En tenant compte du spin, on trouve :

$$G(\varepsilon) = (2s+1) \frac{V^*}{\delta V^*} = (2s+1) \frac{4\pi V}{3} \left(\frac{2m \varepsilon}{h^2} \right)^{3/2} \Rightarrow g(\varepsilon) = (2s+1) 2\pi V \left(\frac{2m}{h^2} \right)^{3/2} \sqrt{\varepsilon}$$

Pour les **gaz parfaits bidimensionnels**, $G(\varepsilon)$ est le nombre d'états dont le vecteur $\vec{n}$ de composantes (n_1, n_2) pointe à l'intérieur du cercle de rayon et d'aire :

$$n = \sqrt{\frac{2m A \varepsilon}{h^2}} \quad \text{et} \quad A^* = \pi \frac{2m A}{h^2} \varepsilon$$

Chaque état y occupe une aire $\delta A^* = 1$. On en déduit :

$$G(\varepsilon) = (2s+1) \frac{A^*}{\delta A^*} = (2s+1) \pi A \frac{2m \varepsilon}{h^2} \Rightarrow g(\varepsilon) = (2s+1) \pi A \frac{2m}{h^2}$$

Pour les **gaz parfaits unidimensionnels**, $G(\varepsilon)$ est le nombre d'états dont le vecteur $\vec{n}$ de composante (n_1) pointe à l'intérieur du segment de longueur :

$$L^* = 2 \sqrt{\frac{2m L^2 \varepsilon}{h^2}} \quad \text{puisque} \quad - \sqrt{\frac{2m L^2 \varepsilon}{h^2}} \leq n_1 \leq + \sqrt{\frac{2m L^2 \varepsilon}{h^2}}$$

Chaque état y occupe une longueur $\delta L^* = 1$. On en déduit :

$$G(\varepsilon) = (2s+1) \frac{L^*}{\delta L^*} = (2s+1) 2L \sqrt{\frac{2m \varepsilon}{h^2}} \Rightarrow g(\varepsilon) = (2s+1) L \sqrt{\frac{2m}{h^2}} \frac{1}{\sqrt{\varepsilon}}$$

3.2.3 Fonctions thermodynamiques

On réécrit d'abord les résultats 2.9 obtenus au paragraphe 2.3.2 dans l'approximation du continuum :

$$N = \int_0^\infty \frac{g(\varepsilon) d\varepsilon}{e^{(\varepsilon-\mu)/kT} + \sigma} \quad E = \int_0^\infty \frac{\varepsilon g(\varepsilon) d\varepsilon}{e^{(\varepsilon-\mu)/kT} + \sigma}$$

$$pV = \sigma kT \int_0^\infty \ln(1 + \sigma e^{(\mu-\varepsilon)/kT}) g(\varepsilon) d\varepsilon$$

On peut remarquer que $\sigma kT \frac{d}{d\varepsilon} \ln(1 + \sigma e^{(\mu-\varepsilon)/kT}) = -\frac{1}{e^{(\varepsilon-\mu)/kT} + \sigma}$.

Ceci permet d'intégrer pV par parties. Avec $\int g(\varepsilon) d\varepsilon = G(\varepsilon)$ et $G(0) = 0$, on trouve :

$$pV = kT \lim_{\varepsilon \rightarrow +\infty} G(\varepsilon) e^{(\mu-\varepsilon)/kT} + \int_0^\infty \frac{G(\varepsilon) d\varepsilon}{e^{(\varepsilon-\mu)/kT} + \sigma}$$

Lorsque la densité d'états est de la forme $g(\varepsilon) = a \varepsilon^\alpha$, on a :

$$G(\varepsilon) = \frac{a \varepsilon^{\alpha+1}}{\alpha+1} = \frac{\varepsilon g(\varepsilon)}{\alpha+1} \quad \text{et} \quad \lim_{\varepsilon \rightarrow +\infty} G(\varepsilon) e^{(\mu-\varepsilon)/kT} = 0$$

Et donc $pV = \frac{E}{\alpha+1}$ *quelle que soit la statistique*. Pour un gaz parfait :

$$\begin{array}{c} \text{tridimensionnel} \\ \alpha = 1/2 : pV = \frac{2}{3} E \end{array} \left| \begin{array}{c} \text{bidimensionnel} \\ \alpha = 0 : \varphi A = E \end{array} \right| \begin{array}{c} \text{unidimensionnel} \\ \alpha = -1/2 : \tau L = 2 E \end{array} \quad (3.1)$$

Pour les systèmes bi- et uni-dimensionnels, pV doit être remplacé respectivement par φA et τL , où φ est la pression ou tension de surface et τ la tension de ligne.

3.3 Gaz parfait classique de masses ponctuelles

3.3.1 Limite hautes températures de la théorie quantique

Le gaz parfait sera **tridimensionnel** et décrit dans **l'approximation du continuum**. Les variables choisies comme fixées sont T , V , N , en fonction desquelles on exprimera les autres grandeurs (E , S , p , μ , C_V). On rappelle les équations de base :

$$N = \int_0^\infty \frac{g(\varepsilon) d\varepsilon}{e^{(\varepsilon-\mu)/kT} + \sigma} \quad E = \frac{3}{2} pV = \int_0^\infty \frac{\varepsilon g(\varepsilon) d\varepsilon}{e^{(\varepsilon-\mu)/kT} + \sigma}$$

On sait que $\frac{\mu}{kT}$ tend vers $-\infty$ quand T tend vers $+\infty$. Dans les fonctions de distribution 2.10, il est donc possible à haute température de négliger $\sigma = \pm 1$:

$$f_{FD}(\varepsilon, T, \mu) \approx f_{BE}(\varepsilon, T, \mu) \approx e^{(\mu-\varepsilon)/kT} = \lambda e^{-\varepsilon/kT} = f_{MB}(\varepsilon, T, \mu)$$

Cette fonction de distribution, dite de **Maxwell-Boltzmann**, décrit le comportement des particules à haute température. On appelle la statistique correspondante **statistique classique** car plus aucun phénomène quantique n'intervient dans l'étude.

Les équations de base deviennent, avec $g(\varepsilon) = (2s + 1) 2\pi V \left(\frac{2m}{h^2}\right)^{3/2} \sqrt{\varepsilon}$:

$$N = (2s + 1) 2\pi V \left(\frac{2m}{h^2}\right)^{3/2} \lambda \int_0^\infty \sqrt{\varepsilon} e^{-\varepsilon/kT} d\varepsilon$$

$$E = \frac{3}{2} pV = (2s + 1) 2\pi V \left(\frac{2m}{h^2}\right)^{3/2} \lambda \int_0^\infty \varepsilon^{3/2} e^{-\varepsilon/kT} d\varepsilon$$

On calcule le **potentiel chimique** μ à partir de l'expression de N . Sachant que $\int_0^\infty \sqrt{u} e^{-u} du = \frac{\sqrt{\pi}}{2}$, on trouve, en posant $\rho = \frac{N}{V}$:

$$\lambda = e^{-\gamma} = e^{\mu/kT} = \frac{\rho \Lambda^3}{2s + 1} \quad \text{avec} \quad \Lambda(T) = \frac{h}{\sqrt{2\pi m kT}}$$

$\Lambda(T)$ est la **longueur d'onde thermique de de Broglie**. C'est une longueur caractéristique du gaz, qui dépend de T et de la masse des particules. La **fugacité** λ permet de déterminer les conditions de validité du modèle. C'est une grandeur sans dimension qui définit entièrement le gaz (la masse et le spin d'une particule et la densité et la température du gaz).

On remplace maintenant λ dans **l'énergie interne**.

Sachant $\int_0^\infty u^{3/2} e^{-u} du = \frac{3\sqrt{\pi}}{4}$, on trouve :

$$E = \frac{3}{2} NkT \qquad pV = NkT \qquad C_V = \frac{3}{2} Nk$$

Loi de **Boyle-Mariotte** Loi de **Dulong et Petit**

Ce sont les **résultats classiques**. On les appelle ainsi car on peut les obtenir à partir de la mécanique classique, sans utilisation de la théorie quantique. On remarquera d'ailleurs l'absence de la constante de Planck et du spin dans les équations. Par contre, les résultats qui suivent ne peuvent être obtenus sans un minimum de concepts quantiques. On y trouve d'ailleurs la constante de Planck (dans $\Lambda^3(T)$) et le spin :

$$\mu = kT \ln \frac{\rho \Lambda^3(T)}{2s + 1} \qquad S = Nk \left(\frac{5}{2} - \ln \frac{\rho \Lambda^3(T)}{2s + 1} \right)$$

Formule de **Sackur-Tetrode**

Outre $pV = 2E/3$ et $\mu = kT \ln \lambda$, on a utilisé les relations générales :

$$C_V = \left[\frac{\partial E}{\partial T} \right]_{V,N} \qquad F = E - TS = -pV + N\mu \qquad S = - \left[\frac{\partial F}{\partial T} \right]_{V,N}$$

L'équation $pV = NkT$ permet d'exprimer le potentiel chimique uniquement en fonction de T et p :

$$\mu(T, p) = kT \ln \frac{p \Lambda^3(T)}{(2s + 1)kT} = \mu_T^0 + kT \ln \frac{p}{p_0} \quad \text{avec} \quad \mu_T^0 = kT \ln \frac{p_0 \Lambda^3(T)}{(2s + 1)kT}$$

Un résultat bien connu en thermodynamique : μ_T^0 et le potentiel chimique standard. En thermodynamique, on le mesure en fonction de la température à la pression

atmosphérique p_0 . La physique statistique permet de l'exprimer en fonction de T et des caractéristiques du gaz (la masse et le spin des atomes).

Condition de validité du modèle haute température.

C'est la condition ayant permis d'utiliser la distribution de Maxwell-Boltzmann :

$$\lim_{T \rightarrow +\infty} \frac{\mu}{kT} = -\infty \quad \Rightarrow \quad \lambda \ll 1 \quad \text{ou} \quad \rho \Lambda^3(T) \ll 1 \quad \text{ou} \quad \ell \gg \Lambda(T)$$

($\ell \sim \rho^{-1/3}$ est la distance moyenne entre deux particules)

Dans la pratique, trois paramètres sont à prendre en compte pour savoir si le modèle décrit correctement le comportement du gaz : la **température** et la **densité** du gaz, et la **masse** de la particule. Si ces particules sont des atomes, on peut exprimer λ en fonction de ρ , T et du nombre A de nucléons de l'atome. On trouve :

$$\lambda \ll 1 \quad \Rightarrow \quad \rho \ll 10^{27} A^{3/2} T^{3/2} \quad \text{ou} \quad T \gg 3 \times 10^{-18} \frac{\rho^{2/3}}{A} \quad (3.2)$$

Dans les conditions habituelles de densité ($\rho \sim 10^{21}$ à 10^{27} atomes/m³), le modèle est parfaitement correct jusqu'à des températures de l'ordre du Kelvin. Pour les atomes et molécules, on peut donc utiliser le modèle classique quelles que soient les conditions.^[2]

3.3.2 Fonction de partition canonique, paradoxe de Gibbs

les grandeurs sont exprimées de façon simple en fonction de T , V , N , variables naturelles de l'énergie libre F . On peut calculer celle-ci à partir de pV et de μ , et ainsi remonter à la fonction de partition canonique. Étant donné que $N \ln N - N = \ln(N!)$, on trouve :

$$\begin{aligned} F(T, V, N) &= E - TS = -pV + \mu N = -NkT + NkT \ln \frac{N \Lambda^3(T)}{(2s+1)V} \\ &= kT (N \ln N - N) - kT \ln \left(\left(\frac{(2s+1)V}{\Lambda^3(T)} \right)^N \right) = -kT \ln \left(\frac{1}{N!} \left(\frac{(2s+1)V}{\Lambda^3(T)} \right)^N \right) \end{aligned}$$

On en déduit :

$$Q(T, V, N) = \frac{1}{N!} (q(T, V))^N \quad \text{avec} \quad q(T, V) = \frac{(2s+1)V}{\Lambda^3(T)}$$

On vérifie qu'on a bien $q(T, V) = \sum_{\varepsilon} \omega_{\varepsilon} e^{-\varepsilon/kT} \approx \int_0^{\infty} g(\varepsilon) e^{-\varepsilon/kT} d\varepsilon$:

$$\int_0^{\infty} g(\varepsilon) e^{-\varepsilon/kT} d\varepsilon = (2s+1) 2\pi V \left(\frac{2m}{h^2} \right)^{3/2} \int_0^{\infty} \sqrt{\varepsilon} e^{-\varepsilon/kT} d\varepsilon = \frac{(2s+1)V}{\Lambda^3(T)} = q(T, V)$$

² L'hélium à très basse température est une exception (voir chapitre 5).

Vérifions maintenant les hypothèses.

- L'approximation du continuum est justifiée si l'écart entre deux niveaux d'énergie est très inférieur à l'énergie moyenne :

$$\frac{h^2}{2mL^2} \ll \frac{3}{2}kT \quad \Rightarrow \quad \Lambda(T) \ll L\sqrt{\frac{3}{2\pi}} \quad \Rightarrow \quad \rho\Lambda^3(T) \ll N$$

Cette condition est largement vérifiée puisqu'elle est beaucoup moins restrictive que $\rho\Lambda^3(T) \ll 1$.

- Pour l'utilisation de l'équation $Q = q^N/N!$, il faut que N soit beaucoup plus petit que le nombre ω d'états accessibles à une particule. Rigoureusement ce nombre est infini, en pratique il correspond aux états d'énergie inférieure à environ $5kT$ (voir figure 3.1 paragraphe 3.4.1) :

$$\omega(T) \approx \int_0^{5kT} g(\varepsilon) d\varepsilon = \frac{20}{3} \sqrt{\frac{5}{\pi}} (2s+1) \frac{V}{\Lambda^3(T)} \approx 8,41 (2s+1) \frac{V}{\Lambda^3(T)}$$

$$\omega(T) \gg N \quad \Rightarrow \quad \rho\Lambda^3(T) \ll 8,41 (2s+1)$$

C'est bien une condition équivalente à $\rho\Lambda^3(T) \ll 1$.

On peut étendre ces résultats aux gaz parfaits multiconstituants. Par exemple pour deux constituants, les deux sous-systèmes étant indépendants et discernables :

$$Q(T, V, N_1, N_2) = \frac{q_1^{N_1}}{N_1!} \times \frac{q_2^{N_2}}{N_2!} \quad \text{avec} \quad \begin{cases} q_1 = (2s_1+1) \frac{\Lambda_1^3}{V} \text{ et } \Lambda_1 = \frac{h}{\sqrt{2\pi m_1 kT}} \\ q_2 = (2s_2+1) \frac{\Lambda_2^3}{V} \text{ et } \Lambda_2 = \frac{h}{\sqrt{2\pi m_2 kT}} \end{cases}$$

On en déduit facilement :

$$E = \frac{3}{2} (N_1 + N_2) kT \quad pV = (N_1 + N_2) kT \quad C_V = \frac{3}{2} (N_1 + N_2) k$$

$$S = \frac{5}{2} (N_1 + N_2) k - N_1 k \ln \frac{(2s_1+1) V}{N_1 \Lambda_1^3} - N_2 k \ln \frac{(2s_2+1) V}{N_2 \Lambda_2^3}$$

$$\mu_1 = -kT \ln \frac{(2s_1+1) V}{N_1 \Lambda_1^3} \quad \mu_2 = -kT \ln \frac{(2s_2+1) V}{N_2 \Lambda_2^3}$$

Le paradoxe de Gibbs est à l'origine de la notion d'indiscernabilité dont découle toute la théorie quantique. Les physiciens du XIX^e siècle ont développé la mécanique statistique sur des concepts classiques, c'est-à-dire en considérant les particules comme discernables. Dans ce contexte, ils arrivèrent au résultat (à cette époque, le spin était inconnu) :

$$Q(T, V, N) = q(T, V)^N \quad \text{avec} \quad q(T, V) = \frac{V}{\Lambda^3(T)}$$

Ils en déduisirent :

$$S(T, V, N) = Nk \left(\frac{3}{2} - \ln \frac{V}{\Lambda^3} \right)$$

Cette entropie a le gros défaut de ne pas être additive. Un système de $2N$ particules dans un volume $2V$ n'a pas d'après cette formule une entropie $2S(T, V, N)$:

$$S(T, 2V, 2N) = 2Nk \left(\frac{3}{2} - \ln \frac{2V}{\Lambda^3} \right) = 2S(T, V, N) - 2Nk \ln 2 \neq 2S(T, V, N)$$

La difficulté fut contournée en divisant Q par $N!$, c'est-à-dire en décrétant, sans justification théorique à l'époque, de ne pas compter comme distincts deux états du système obtenus l'un l'autre par permutation des N particules. On obtient alors la formule de Sackur-Tetrode (sans spin) et :

$$S(T, 2V, 2N) = 2Nk \left(\frac{5}{2} - \ln \frac{2V}{2N\Lambda^3} \right) = 2Nk \left(\frac{5}{2} - \ln \frac{V}{N\Lambda^3} \right) = 2S(T, V, N)$$

Et l'entropie devient miraculeusement additive. C'est ce paradoxe et cet artifice de calcul qui mirent pour la première fois en évidence la notion d'indiscernabilité.

3.4 Éléments de théorie cinétique des gaz parfaits

La théorie cinétique des gaz est une théorie **classique** : elle n'est valable qu'à haute température et/ou à basse densité. On suppose dans ce paragraphe que le gaz est parfait et que les particules qui le constituent sont ponctuelles. La généralisation aux gaz quelconques fera l'objet du paragraphe 3.6.3.

3.4.1 Distribution de Maxwell-Boltzmann (des énergies)

Elle donne la fraction de particules dont l'énergie est comprise entre ε et $\varepsilon + d\varepsilon$. C'est aussi la probabilité qu'une particule prise au hasard ait une énergie comprise entre ε et $\varepsilon + d\varepsilon$. Soit $W(\varepsilon)d\varepsilon$ cette probabilité ($W(\varepsilon)$ est une densité de probabilité exprimée en J^{-1}). Avec nos arguments quantiques, on peut écrire :

$$W(\varepsilon) d\varepsilon = \frac{\omega(\varepsilon) e^{-\varepsilon/kT}}{\sum_{\varepsilon} \omega(\varepsilon) e^{-\varepsilon/kT}} = \frac{g(\varepsilon) e^{-\varepsilon/kT} d\varepsilon}{\int_0^{\infty} g(\varepsilon) e^{-\varepsilon/kT} d\varepsilon} \quad (3.3)$$

Ces arguments quantiques n'étaient pas connus à l'époque, mais on savait de la mécanique classique que $g(\varepsilon)$ variait en $\sqrt{\varepsilon}$, ce qui est suffisant pour écrire :

$$W(\varepsilon) d\varepsilon = \frac{\sqrt{\varepsilon} e^{-\varepsilon/kT} d\varepsilon}{\int_0^{\infty} \sqrt{\varepsilon} e^{-\varepsilon/kT} d\varepsilon} \quad \text{et donc} \quad W(\varepsilon) = \frac{2}{\sqrt{\pi}} \frac{1}{kT} \sqrt{\frac{\varepsilon}{kT}} e^{-\varepsilon/kT}$$

C'est la **distribution de Maxwell-Boltzmann** représentée figure 3.1.

On rappelle les valeurs des intégrales $J_n = \int_0^{\infty} x^n e^{-x^2/2\sigma^2} dx$:

$$J_0 = \sigma \sqrt{\frac{\pi}{2}} \quad J_1 = \sigma^2 \quad J_2 = \sigma^3 \sqrt{\frac{\pi}{2}} \quad J_3 = 2\sigma^4 \quad J_4 = 3\sigma^5 \sqrt{\frac{\pi}{2}} \quad J_5 = 8\sigma^6$$

À l'aide desquelles on peut vérifier qu'on a bien :

$$\int_0^{\infty} W(\varepsilon) d\varepsilon = 1 \quad \text{et} \quad \bar{\varepsilon} = \int_0^{\infty} \varepsilon W(\varepsilon) d\varepsilon = \frac{3}{2} kT$$

Voici les principales caractéristiques de cette distribution :

- elle est indépendante de la nature du gaz (la masse et le spin) ;
- elle est indépendante des grandeurs purement quantiques h et s ; des arguments classiques suffisent pour l'obtenir ;
- elle prend des valeurs négligeables au delà de $5 kT$;
- elle atteint son maximum pour $\varepsilon = \frac{kT}{2}$; cette valeur la plus probable ne correspond pas à la valeur moyenne ;
- la dispersion autour de la valeur moyenne est très grande ; on peut d'ailleurs calculer l'écart type :

$$\overline{\varepsilon^2} = \int_0^\infty \varepsilon^2 W(\varepsilon) d\varepsilon = \frac{15}{4} (kT)^2 \quad \Rightarrow \quad \Delta\varepsilon = \sqrt{\overline{\varepsilon^2} - \bar{\varepsilon}^2} = kT \sqrt{\frac{3}{2}} = \bar{\varepsilon} \sqrt{\frac{2}{3}}$$

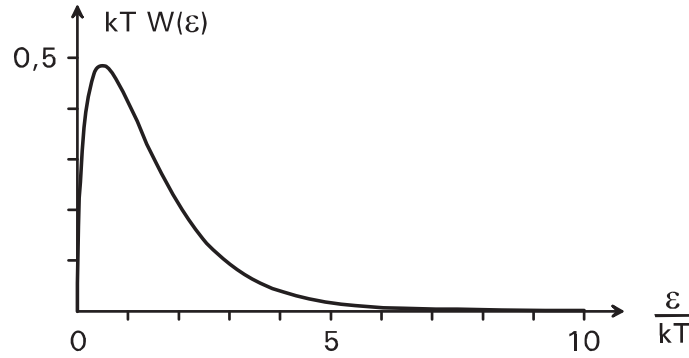


FIG. 3.1 – Distribution des énergies de Maxwell-Boltzmann

3.4.2 Distribution de Maxwell (des vitesses)

Calculons la fraction de particules dont le module du vecteur vitesse est compris entre v et $v + dv$. Il suffit de remplacer ε par $\frac{1}{2} m v^2$ et $d\varepsilon$ par $m v dv$ dans $W(\varepsilon) d\varepsilon$. On trouve :

$$W(v) dv = \left(\frac{m}{2\pi kT} \right)^{3/2} e^{-mv^2/2kT} 4\pi v^2 dv \quad (3.4)$$

C'est la **distribution de Maxwell** représentée figure 3.2.

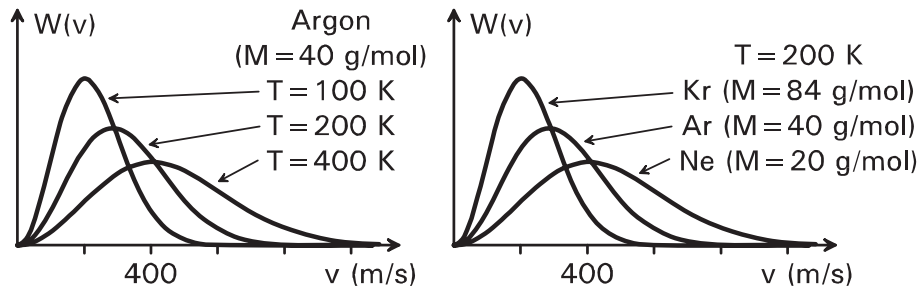


FIG. 3.2 – Distribution des vitesses de Maxwell

Maxwell découvrit cette loi (1860) à partir d'observations expérimentales et de notions de symétrie. Boltzmann la démontra ensuite (1872) à partir d'une étude statistique des collisions élastiques entre molécules.

À noter que cette distribution est valable pour tout fluide, que ce soit un gaz parfait, un gaz réel ou un liquide (seules contraintes : homogénéité et isotropie du milieu).

Pour les calculs, on écrira :

$$W(v) = \frac{4\pi v^2 e^{-v^2/2\sigma^2}}{(2\pi\sigma^2)^{3/2}} \quad \text{avec} \quad \sigma = \sqrt{\frac{kT}{m}}$$

Le graphe de $W(v)$ présente un maximum en $v^* = \sigma\sqrt{2} = \sqrt{\frac{2kT}{m}}$.

Cette valeur la plus probable diffère de la valeur moyenne :

$$\bar{v} = \int_0^\infty v W(v) dv = \frac{J_3(\sigma)}{J_2(\sigma)} = \sigma \sqrt{\frac{8}{\pi}} = \sqrt{\frac{8kT}{\pi m}} \quad (3.5)$$

Et enfin, on peut calculer la vitesse quadratique moyenne et l'écart type :

$$\overline{v^2} = \int_0^\infty v^2 W(v) dv = 3\sigma^2 = \frac{3kT}{m} \quad (3.6)$$

$$\Delta v = \sqrt{\overline{v^2} - \bar{v}^2} = \sigma \sqrt{3 - 8/\pi} \approx 0,4220 \bar{v}$$

On remarquera qu'on a, non pas $\bar{\varepsilon} = \frac{1}{2} m \bar{v}^2$, mais $\bar{\varepsilon} = \frac{1}{2} m \overline{v^2} = \frac{3}{2} kT$.

Voici quelques ordres de grandeur des vitesses moyennes en m/s :

masse molaire	10 K	30 K	100 K	300 K	1000 K	3000 K
4 g (He)	230	398	728	1260	2301	3985
20 g (Ne)	103	178	325	564	1029	1782
40 g (Ar)	73	126	230	398	728	1260
84 g (Kr)	50	87	159	275	502	870
131 g (Xe)	40	70	127	220	402	696

$4\pi v^2 dv$ est l'élément de volume contenant les extrémités de tous les vecteurs vitesse de module compris entre v et $v + dv$. Le gaz étant homogène et isotrope, on déduit de l'expression de $W(v) dv$ la densité de probabilité (les indices 1, 2, 3 sont relatifs à trois axes orthogonaux) :

$$W(\vec{v}) = W(v_1, v_2, v_3) = \frac{e^{-v^2/2\sigma^2}}{(2\pi\sigma^2)^{3/2}}$$

$W(v_1, v_2, v_3) dv_1 dv_2 dv_3$ est la fraction de particules dont les trois composantes cartésiennes de vitesse sont respectivement comprises entre v_1 et $v_1 + dv_1$, v_2 et $v_2 + dv_2$, v_3 et $v_3 + dv_3$. Avec $v^2 = v_1^2 + v_2^2 + v_3^2$, on obtient :

$$W(v_1, v_2, v_3) dv_1 dv_2 dv_3 = \frac{e^{-v_1^2/2\sigma^2} dv_1}{\sigma\sqrt{2\pi}} \times \frac{e^{-v_2^2/2\sigma^2} dv_2}{\sigma\sqrt{2\pi}} \times \frac{e^{-v_3^2/2\sigma^2} dv_3}{\sigma\sqrt{2\pi}}$$

Toujours pour des raisons d'isotropie, on en déduit :

$$W(v_i) = \frac{e^{-v_i^2/2\sigma^2}}{\sigma\sqrt{2\pi}} = \sqrt{\frac{m}{2\pi kT}} e^{-mv_i^2/2kT} \quad (i = 1, 2, 3)$$

$W(v_i) dv_i$ est la probabilité que la $i^{\text{ème}}$ composante de vitesse ait une valeur comprise entre v_i et $v_i + dv_i$ indépendamment des valeurs des deux autres composantes.

$$\text{On en déduit facilement : } \bar{v}_i = 0 \quad \overline{|v_i|} = \sqrt{\frac{2kT}{\pi m}} \quad \overline{v_i^2} = \frac{kT}{m}$$

Le gaz parfait bidimensionnel se rencontre très souvent dans les modèles. On donne ci-dessous les différents résultats relatifs à de tels systèmes.

$$W(\varepsilon) = \frac{1}{kT} e^{-\varepsilon/kT} \quad \bar{\varepsilon} = kT \quad \overline{\varepsilon^2} = 2(kT)^2 \quad \Delta\varepsilon = kT \quad \sigma = \sqrt{\frac{kT}{m}} \quad (3.7)$$

$$W(v) = \frac{e^{-v^2/2\sigma^2}}{2\pi\sigma^2} 2\pi v \quad v^* = \sigma \quad \bar{v} = \sigma\sqrt{\pi/2} \quad \overline{v^2} = 2\sigma^2 \quad \Delta v = \sigma\sqrt{2 - \pi/2}$$

$$W(v_i) = \frac{e^{-v_i^2/2\sigma^2}}{\sigma\sqrt{2\pi}} \quad \bar{v}_i = 0 \quad \overline{|v_i|} = \frac{\sigma}{\sqrt{2\pi}} \quad \overline{v_i^2} = \sigma^2 \quad \Delta v_i = \sigma \quad (i = 1 \text{ et } 2)$$

Il existe différentes méthodes expérimentales de mesure de la distribution de vitesses des molécules d'un gaz. On peut étudier l'effet Doppler sur les raies spectrales émises par les molécules (la fréquence du rayonnement émis par une molécule est fonction de sa vitesse), ou étudier le phénomène d'effusion (on perce un trou dans la paroi d'un récipient, et on compte en fonction du temps les molécules qui s'échappent en fonction de leur vitesse).

Ce qui a définitivement validé la théorie cinétique des gaz, c'est le calcul qui suit de la pression cinétique. Pour les gaz parfaits, cette pression cinétique est égale à la pression thermodynamique du système.^[3]

3.4.3 Calcul de la pression cinétique

On place une paroi fictive au milieu du récipient (extrêmement grand par rapport au libre parcours moyen des molécules).^[4] Soit un élément ΔS de la paroi fictive.

On choisit l'origine des coordonnées au milieu de ΔS et les axes tels que la paroi soit dans le plan xOy ($z = 0$). Le vecteur unitaire porté par Oz (perpendiculaire à la paroi) sera noté $\vec{n}$.

On étudie la force qu'exercent sur ΔS les molécules situées dans le demi-espace inférieur ($z < 0$). Soit une molécule de vitesse $\vec{v}(v_x, v_y, v_z)$ venant percuter ΔS . Le choc est supposé élastique : la molécule repart avec une vitesse $\vec{v}'(v_x, v_y, -v_z)$.

³ Pour un gaz réel ou un liquide, il faut, pour obtenir la pression thermodynamique, ajouter à la pression cinétique une pression dynamique, contribution due aux forces intermoléculaires.

⁴ Le libre parcours moyen est la distance que parcourt en moyenne une molécule sans subir de collisions. Pour le calcul de ce libre parcours moyen, voir paragraphe 3.5.4.

Sa quantité de mouvement au moment du choc a donc varié de $m(\vec{v}' - \vec{v})$, c'est-à-dire qu'elle a transféré à ΔS une impulsion :

$$\delta \vec{P} = -m(\vec{v}' - \vec{v}) = 2m v_z \vec{n}$$

Toute molécule ayant la vitesse $\vec{v}$ transfère l'impulsion $\delta \vec{P}$ lorsqu'elle percute ΔS . Pour connaître la force exercée par ces molécules sur ΔS , il suffit de compter le nombre de chocs que subit ΔS dans un intervalle de temps Δt donné. Ce nombre multiplié par $\delta \vec{P}$ donnera l'impulsion totale transférée à ΔS pendant Δt par ces molécules. On intègre ensuite sur tous les vecteurs vitesse possibles. Le transfert d'impulsion $\Delta \vec{P}$ obtenu permet de calculer la force exercée sur ΔS par l'ensemble des molécules, et enfin la pression :

$$\vec{F} = \frac{\Delta \vec{P}}{\Delta t} \quad \text{et} \quad p = \frac{F_z}{\Delta S} = \frac{\vec{F} \cdot \vec{n}}{\Delta S}$$

On raisonne donc d'abord sur toutes les molécules possédant la vitesse $\vec{v}$. A un instant t donné, celles qui vont frapper ΔS entre t et $t + \Delta t$ ont une distance inférieure à $v \Delta t$ à parcourir avant le choc. Elles sont donc contenues dans le cylindre oblique de base ΔS , de longueur $v \Delta t$, et de génératrice parallèle à $\vec{v}$. Si θ est l'angle entre $\vec{v}$ et $\vec{n}$, le volume de ce cylindre vaut $\Delta V = v \Delta t \Delta S \cos \theta$. Le nombre total de particules contenues dans ΔV est $\Delta N = N \Delta V / V$. Parmi elles, $\Delta N \times W(\vec{v}) d^3 v$ possèdent un vecteur vitesse dont l'extrémité est dans l'élément $d^3 v = v^2 dv \sin \theta d\theta d\varphi$ autour de celle de $\vec{v}$. Elles transfèrent à ΔS pendant Δt une impulsion $N \Delta V W(\vec{v}) d^3 v \delta \vec{P} / V$, c'est-à-dire :

$$\frac{N}{V} (v \Delta t \Delta S \cos \theta) \times \left(\frac{m}{2\pi kT} \right)^{3/2} e^{-m v^2 / 2kT} \times v^2 dv \sin \theta d\theta d\varphi \times 2m v \cos \theta \vec{n}$$

On intègre sur θ de 0 à $\pi/2$, sur φ de 0 à 2π , et sur v de 0 à $+\infty$:

$$\Delta \vec{P} = \vec{n} \frac{2m N \Delta t \Delta S}{V} \left(\frac{m}{2\pi kT} \right)^{3/2} \int_0^{2\pi} d\varphi \int_0^{\pi/2} \cos^2 \theta \sin \theta d\theta \int_0^{+\infty} e^{-m v^2 / 2kT} v^4 dv$$

$$\text{On trouve } \Delta \vec{P} = \frac{NkT}{V} \Delta t \Delta S \vec{n} \quad \text{et donc} \quad \vec{F} = \frac{\Delta \vec{P}}{\Delta t} = \frac{NkT}{V} \Delta S \vec{n}$$

$$\text{Et le résultat tant attendu : } p = \frac{\vec{F} \cdot \vec{n}}{\Delta S} = \frac{NkT}{V} \quad \text{et donc} \quad pV = NkT$$

3.5 Le mouvement brownien

3.5.1 Historique et hypothèses

En 1827, le botaniste Robert Brown (1773-1858), observant sous un microscope du pollen en suspension dans de l'eau, remarqua que les grains microscopiques étaient soumis à un mouvement continu et irrégulier. Il crut découvrir "la molécule primitive" responsable de la vie, mais s'aperçut plus tard que l'on pouvait observer ce phénomène avec toutes sortes de particules. Un grand nombre de physiciens ont

cherché à expliquer ce phénomène en mettant en cause tantôt la lumière incidente, tantôt des forces électriques ...

C'est Delsaux (1877) qui suggéra pour la première fois que ce mouvement résultait de l'ensemble des chocs exercés par les molécules de solvant sur la particule observée, propulsant celle-ci de façon aléatoire tantôt dans une direction, tantôt dans une autre. Des développements théoriques furent alors entrepris, entre autres par F. Exner (1900). Albert Einstein (1905) et indépendamment M. von Smoluchowski (1906) proposèrent une théorie complète du mouvement brownien. Les résultats sont en parfait accord avec l'expérience. Ce fut l'argument décisif présenté aux derniers opposants à la théorie atomique.

Le mouvement brownien existe dans tous les fluides, mais il est particulièrement observable sur des particules en suspension dans les liquides (eau, solvant organique, ...). Les conditions idéales d'observation sont caractéristiques des suspensions colloïdales, où la taille des particules (de quelques dixièmes de micromètre à quelques micromètres) est suffisamment grande pour être observée sous un microscope, et suffisamment petite pour que les chocs avec les molécules du solvant soient efficaces. Partant d'une situation où toutes les particules sont concentrées au milieu du récipient, on constate qu'en moyenne la "pelote" formée par les déplacements aléatoires reste centrée sur son point de départ mais qu'elle grandit au cours du temps proportionnellement à $\sqrt{t}$. Cela signifie que pour chaque particule, la valeur moyenne du déplacement est nulle et que la moyenne du carré du déplacement augmente linéairement avec le temps.

Le mouvement brownien n'est pas propre aux particules en suspension. Dans un gaz par exemple, les molécules sont sans cesse en mouvement, s'entrechoquent, et décrivent une marche aléatoire dans l'espace. L'étude statistique de ces mouvements est indispensable à la description de nombreuses propriétés de la matière. Citons à titre d'exemple le cas de la conductivité électrique d'un gaz parfait de particules chargées électriquement. Placée dans un champ électrique constant et uniforme, chaque particule, si elle était seule dans le récipient, acquerrait un mouvement uniformément accéléré. Non seulement la résistance du milieu serait nulle (la notion de résistance électrique serait alors inexplicable), mais de plus la vitesse de chaque particule tendrait vers l'infini (une aberration). Grâce au mouvement brownien, on peut introduire la notion de "libre parcours moyen", distance moyenne (entre deux chocs) le long de laquelle la particule est effectivement libre.

Les hypothèses sont celles de la théorie cinétique des gaz. Les molécules sont des "*sphères dures*" qui se déplacent librement sauf lors de *chocs élastiques* où elles changent aléatoirement de direction. La trajectoire d'une molécule est une succession de "pas" linéaires de longueur variable, chaque pas étant indépendant du précédent, tant en longueur qu'en direction (le processus est dit *markovien*). Question : quelle est la position de l'objet au bout de N sauts ? Si le fluide est dans un champ de forces, la probabilité peut être fonction de la direction choisie par la molécule après un choc.

On procédera par étapes :

- les vecteurs déplacement possibles formeront d'abord un ensemble fini. C'est typiquement la diffusion de particules sur un réseau cristallin. Un objet se déplace sur un réseau de sites en effectuant, de façon aléatoire, des sauts de

longueur et de direction bien déterminées. Le réseau sera dans un premier temps unidimensionnel afin d'aborder plus facilement le problème tridimensionnel.

- les vecteurs déplacement formeront ensuite un ensemble continu. Ils décrivent les mouvements de molécules au sein d'un fluide. En première approche, on étudiera le cas où la longueur du pas est fixe et où la direction est quelconque. L'extension au pas de longueur variable constitue le mouvement brownien proprement dit.

3.5.2 Diffusion sur un réseau unidimensionnel

Étude statistique

Un objet se déplace aléatoirement le long d'une droite par sauts de longueur ℓ . Chaque saut est indépendant du précédent (processus markovien), et est dirigé soit dans le sens positif, soit dans le sens négatif, avec des probabilités P_+ et P_- . Pour l'étude statistique, on étudie dans un premier temps une situation donnée : après un nombre N de sauts, l'objet placé initialement à l'origine se trouve à l'abscisse $x = n\ell$. Pour arriver à cette position, l'objet a effectué N_+ sauts positifs et N_- sauts négatifs. Les relations suivantes sont évidentes :

$$(N_+ + N_- = N \quad \text{et} \quad N_+ - N_- = n) \quad \Rightarrow \quad \left(N_+ = \frac{N+n}{2} \quad \text{et} \quad N_- = \frac{N-n}{2} \right)$$

avec N et n de même parité $-N \leq n \leq +N$ et $0 \leq N_+$ et $N_- \leq N$

Seuls N et n sont connus. On ne connaît pas le trajet qu'a effectué l'objet pour y parvenir : tous les trajets comportant les mêmes nombres N_+ et N_- de sauts positifs et négatifs ont été possibles. Les distributions de probabilités suivent une loi binomiale :

$$P(N_+, N_-) = \frac{(N_+ + N_-)!}{N_+! N_-!} P_+^{N_+} P_-^{N_-} \quad \Rightarrow \quad P(N, n) = \frac{N!}{\frac{N+n}{2}! \frac{N-n}{2}!} P_+^{\frac{N+n}{2}} P_-^{\frac{N-n}{2}}$$

Un calcul simple utilisant 1.21 donne la valeur moyenne de n et son écart type :

$$\bar{n} = N \delta P \quad \text{et} \quad \Delta n = \sqrt{N(1 - \delta P)} \quad \text{avec} \quad \delta P = P_+ - P_-$$

Considérons maintenant un très grand nombre M d'objets partant tous en même temps de l'origine.^[5] Après un grand nombre N de sauts, on obtient une distribution gaussienne des objets de part et d'autre de la valeur moyenne de n . En notant $m(M, n)$ le nombre d'objets à l'abscisse $x = n\ell$, on a dans la limite des grands nombres :

$$m(N, n) = \frac{M}{\sqrt{2\pi \Delta n^2}} \exp \left(-\frac{1}{2} \frac{(n - \bar{n})^2}{\Delta n^2} \right) \quad (3.8)$$

⁵ C'est concrètement réalisable ici (on construit un faisceau de M droites parallèles) mais ce ne sera plus le cas en trois dimensions. On dira alors qu'on construit un ensemble statistique constitué d'un très grand nombre M de systèmes (d'objets).

Profil de concentration

L'objectif maintenant est de relier ces résultats statistiques aux grandeurs macroscopiques mesurables (concentration, flux de matière, coefficient de diffusion).

Les objets sont des particules indépendantes, c'est-à-dire sans interactions mutuelles, mais aussi libres de tout mouvement, sans gêne mutuelle. On peut imaginer par exemple un faisceau de M droites toutes parallèles à Ox , les M origines étant placées sur le même front (plan perpendiculaire à Ox). À l'instant $t = 0$, on place les M objets sur les M origines. Chaque particule est autorisée à se déplacer uniquement sur sa propre droite.

On introduit ensuite x et t dans les équations en remplaçant n par x/ℓ et N par t/τ ($1/\tau$ est pour une particule donnée la fréquence moyenne de sauts).

Chaque front occupe un volume $\ell \times \Delta S$, où ΔS est la section du cylindre formé par le faisceau de droites. Dans ce volume, on observe $m(N, n)$ objets. La concentration moyenne d'un front (en particules par unité de volume) s'écrit donc :

$$c = \frac{m}{\ell \Delta S} = c_0 \frac{m}{M} \quad \text{avec} \quad c_0 = \frac{M}{\ell \Delta S} \quad (3.9)$$

(c_0 est la concentration du front origine à $t = 0$)

En pratique, ℓ se mesure en nanomètres et t en nanosecondes. M étant en outre très grand, on peut à l'échelle macroscopique considérer les grandeurs x , t , c comme continues. En remplaçant dans l'expression 3.8 de $m(n, N)$, on obtient :

$$c(x, t) = \frac{c_0 \ell}{\sqrt{2\pi} \Delta x^2} \exp\left(-\frac{1}{2} \frac{(x - \bar{x})^2}{\Delta x^2}\right) \Rightarrow \int_{-\infty}^{+\infty} c(x, t) dx = c_0 \ell \quad (3.10)$$

avec $\bar{x} = \ell \frac{t}{\tau} \delta P$ et $\Delta x = \ell \sqrt{\frac{t}{\tau} (1 - \delta P)}$

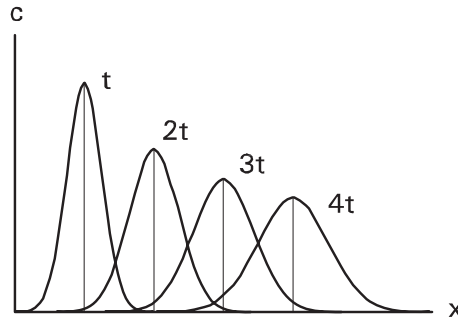


FIG. 3.3 – Profil de concentration en fonction de x et de t

Deux résultats importants illustrés sur la figure 3.3 :

- $\bar{x}$ est proportionnel à t : la vitesse de déplacement du pic de concentration est constante ;
- la dispersion autour de $\bar{x}$ varie en $\sqrt{t}$: le pic de concentration s'étale de plus en plus autour de $\bar{x}$ au fur et à mesure de sa progression (Δx est une mesure de la taille de la "pelote" évoquée au paragraphe 3.5.1).

On remarquera également que plus δP est important, plus la vitesse moyenne est élevée, et plus la dispersion autour de la position moyenne est faible.

Coefficient de diffusion, loi de Fick

Dorénavant, on s'occupe du cas particulier important où $P_+ = P_- = 1/2$. On a dans ce cas $\bar{x} = 0$ et $\Delta x^2 = \ell^2 t / \tau$, et donc :

$$c(x, t) = c_0 \sqrt{\frac{\tau}{2\pi t}} \exp\left(-\frac{1}{2} \frac{x^2}{\ell^2} \frac{\tau}{t}\right) \quad (3.11)$$

Aucune force extérieure n'agit sur le système. La force motrice est ici "interne" : *le gradient de concentration*, ou plus généralement *le gradient de potentiel chimique*, est la cause de l'évolution du système. D'un état initial hors équilibre (particules concentrées en $x = 0$), le système évolue spontanément vers un état d'équilibre où les particules sont réparties de façon homogène dans le cylindre. Celui-ci étant supposé de longueur infinie, on ne s'étonnera pas de trouver une concentration d'équilibre nulle lorsque t tend vers l'infini.

Le gradient de concentration engendre un flux de particules des régions les plus denses vers celles de moindre densité. Le coefficient de diffusion D (unité : m²/s) relie au signe près le flux J de particules (par m² et par seconde) au gradient de concentration. Dans notre problème, nous n'envisageons que des déplacements suivant Ox . On a donc par définition de D :

$$J = -D \frac{\partial c}{\partial x} \quad (3.12)$$

Dans un problème de diffusion réel, D dépend de c et éventuellement de x :

- de c parce que les particules partagent les mêmes sites de transit : plus la concentration est élevée (à gradient identique), plus la gêne mutuelle entre particules est élevée, et plus le flux est faible. Ce n'est pas le cas dans notre problème puisque les particules se déplacent chacune sur sa propre droite.
- de x si le milieu n'est pas homogène. On suppose ici que tous les sites sont identiques (c'est un cas de diffusion en milieu homogène).

Dans notre problème donc, D est indépendant de x , aussi bien de façon explicite que par l'intermédiaire de la concentration.

Le bilan de matière qui suit a pour but de transformer l'équation 3.12 en une équation aux dérivées partielles ne faisant apparaître que la concentration.

Considérons à l'abscisse x une tranche de cylindre de section ΔS et d'épaisseur dx . Pendant l'intervalle de temps dt , le nombre de particules entrant à gauche vaut $J(x, t) \Delta S dt$, et le nombre de particules sortant à droite vaut $J(x + dx, t) \Delta S dt$. Le nombre de particules dans la tranche dx a donc varié de :

$$(J(x, t) - J(x + dx, t)) \Delta S dt = -\frac{\partial J}{\partial x} dx \Delta S dt \quad (3.13)$$

En raisonnant sur la concentration, on peut également dire que le nombre de particules de cette même tranche au cours du même intervalle de temps a varié de :

$$(c(x, t + dt) - c(x, t)) \Delta S dx = \frac{\partial c}{\partial t} dt \Delta S dx \quad (3.14)$$

Comparant 3.14 à 3.13 on déduit : $\frac{\partial c}{\partial t} = -\frac{\partial J}{\partial x}$, puis en utilisant 3.12 :

$$\frac{\partial c}{\partial t} = \frac{\partial}{\partial x} \left(D \frac{\partial c}{\partial x} \right) = D \frac{\partial^2 c}{\partial x^2} \quad \textbf{Loi de Fick} \quad (3.15)$$

La deuxième expression n'est valable que si D est indépendant de c (déplacements indépendants) et de x (milieux homogènes).

On vérifiera que 3.11 est solution de 3.15 si $\tau = \frac{\ell^2}{2D}$.

On relie ainsi le paramètre statistique τ (temps moyen entre deux sauts) à la grandeur macroscopique et mesurable D . On retiendra donc :

$$c(x, t) = \frac{c_0 \ell}{\sqrt{4\pi D t}} \exp\left(-\frac{x^2}{4 D t}\right) \quad \Delta x = \sqrt{2 D t} \quad |\overline{x}| = 2\sqrt{\frac{D t}{\pi}} \quad (3.16)$$

La distance moyenne à l'origine des particules varie en $\sqrt{t}$. Ce résultat est très bien vérifié expérimentalement.

3.5.3 Diffusion sur un réseau tridimensionnel

Étude statistique

L'objet se déplace cette fois d'un site à l'autre sur un réseau tridimensionnel. Chaque site est entouré d'un certain nombre v , appelé *coordinnence*, de sites premiers voisins. On se restreint aux réseaux où tous les sites sont cristallographiquement identiques : les déplacements possibles, et les probabilités correspondantes, sont indépendants de la position de l'objet.

Soient $\Delta = \{\vec{\rho}_i, i = 1 \text{ à } v\}$ l'ensemble de ces déplacements, et $\Pi = \{P_i, i = 1 \text{ à } v\}$ les probabilités correspondantes. Le fait que tous les sites soient identiques sous-entend qu'il y a sur chaque site un centre de symétrie (si $\vec{\rho}_i \in \Delta$ alors $-\vec{\rho}_i \in \Delta$). La coordinnence v est donc paire et on pose $v = 2q$. Les écritures de Δ et Π deviennent :

$$\Delta = \{\vec{\rho}_i, -\vec{\rho}_i, i = 1 \text{ à } q\} \quad \text{et} \quad \Pi = \{P_{+i}, P_{-i}, i = 1 \text{ à } q\}$$

On utilisera les notations : $P_i = P_{+i} + P_{-i}$ et $\delta P_i = P_{+i} - P_{-i}$. Soit N le nombre total de sauts effectués par l'objet. On notera N_{+i} et N_{-i} les nombres de sauts dans les différentes directions, et $[N]$ la distribution $\{N_{+i}, N_{-i}, i = 1 \text{ à } q\}$. La probabilité d'atteindre la position $\vec{r}(x, y, z)$ en N sauts vaut :

$$P(N, \vec{r}) = \sum_{[N]}^{(1,2,3,4)} P_{[N]} \quad \text{avec} \quad P_{[N]} = N! \prod_{i=1}^q \frac{P_{+i}^{N_{+i}} P_{-i}^{N_{-i}}}{N_{+i}! N_{-i}!}$$

Soit $\vec{\rho}_i(u_i, v_i, w_i)$. La somme porte sur les distributions vérifiant :

$$(1) \sum_{i=1}^q N_i = N \quad (2) \sum_{i=1}^q u_i \delta N_i = x \quad (3) \sum_{i=1}^q v_i \delta N_i = y \quad (4) \sum_{i=1}^q w_i \delta N_i = z$$

On a posé $N_i = N_{+i} + N_{-i}$ et $\delta N_i = N_{+i} - N_{-i}$. Cette somme n'est pas calculable rigoureusement. La méthode qui suit (quasi rigoureuse) consiste à ne retenir dans la somme que les termes rendant $\ln P_{[N]}$ maximum sous les contraintes (1, 2, 3, 4). On introduit donc quatre multiplicateurs de Lagrange $\alpha, \lambda_x, \lambda_y$ et λ_z et on cherche $[N]^*$ telle que :

$$\frac{\partial}{\partial N_{+j}} \left(\ln P_{[N]} - \alpha \sum_{i=1}^q N_i - \lambda_x \sum_{i=1}^q u_i \delta N_i - \lambda_y \sum_{i=1}^q v_i \delta N_i - \lambda_z \sum_{i=1}^q w_i \delta N_i \right) = 0$$

$$\frac{\partial}{\partial N_{-j}} \left(\ln P_{[N]} - \alpha \sum_{i=1}^q N_i - \lambda_x \sum_{i=1}^q u_i \delta N_i - \lambda_y \sum_{i=1}^q v_i \delta N_i - \lambda_z \sum_{i=1}^q w_i \delta N_i \right) = 0$$

On trouve, avec la formule de Stirling et en définissant le vecteur $\vec{\lambda}(\lambda_x, \lambda_y, \lambda_z)$:

$$\begin{aligned} N_{+j}^* &= N P_{+j} e^{-\alpha} e^{-\lambda_x u_j - \lambda_y v_j - \lambda_z w_j} = N P_{+j} e^{-\alpha} e^{-\vec{\lambda} \cdot \vec{\rho}_j} \\ N_{-j}^* &= N P_{-j} e^{-\alpha} e^{+\lambda_x u_j + \lambda_y v_j + \lambda_z w_j} = N P_{-j} e^{-\alpha} e^{+\vec{\lambda} \cdot \vec{\rho}_j} \end{aligned}$$

On remplace ensuite les solutions dans les quatre contraintes.

L'équation (1) donne $e^\alpha = \sum_{i=1}^q (P_{+i} e^{-\vec{\lambda} \cdot \vec{\rho}_i} + P_{-i} e^{+\vec{\lambda} \cdot \vec{\rho}_i})$, et donc :

$$\begin{aligned} N_{+j}^* &= N \frac{P_{+j} e^{-\vec{\lambda} \cdot \vec{\rho}_j}}{\sum_{i=1}^q (P_{+i} e^{-\vec{\lambda} \cdot \vec{\rho}_i} + P_{-i} e^{+\vec{\lambda} \cdot \vec{\rho}_i})} \quad \text{et} \quad N_{-j}^* = N \frac{P_{-j} e^{+\vec{\lambda} \cdot \vec{\rho}_j}}{\sum_{i=1}^q (P_{+i} e^{-\vec{\lambda} \cdot \vec{\rho}_i} + P_{-i} e^{+\vec{\lambda} \cdot \vec{\rho}_i})} \\ N_j^* &= N \frac{P_{+j} e^{-\vec{\lambda} \cdot \vec{\rho}_j} + P_{-j} e^{+\vec{\lambda} \cdot \vec{\rho}_j}}{\sum_{i=1}^q (P_{+i} e^{-\vec{\lambda} \cdot \vec{\rho}_i} + P_{-i} e^{+\vec{\lambda} \cdot \vec{\rho}_i})} \quad \text{et} \quad \delta N_j^* = N \frac{P_{+j} e^{-\vec{\lambda} \cdot \vec{\rho}_j} - P_{-j} e^{+\vec{\lambda} \cdot \vec{\rho}_j}}{\sum_{i=1}^q (P_{+i} e^{-\vec{\lambda} \cdot \vec{\rho}_i} + P_{-i} e^{+\vec{\lambda} \cdot \vec{\rho}_i})} \end{aligned}$$

En remplaçant dans les équations (2, 3, 4), on trouve :

$$\begin{aligned} \sum_{j=1}^q \left((N u_j - x) P_{+j} e^{-\vec{\lambda} \cdot \vec{\rho}_j} - (N u_j + x) P_{-j} e^{+\vec{\lambda} \cdot \vec{\rho}_j} \right) &= 0 \\ \sum_{j=1}^q \left((N v_j - y) P_{+j} e^{-\vec{\lambda} \cdot \vec{\rho}_j} - (N v_j + y) P_{-j} e^{+\vec{\lambda} \cdot \vec{\rho}_j} \right) &= 0 \\ \sum_{j=1}^q \left((N w_j - z) P_{+j} e^{-\vec{\lambda} \cdot \vec{\rho}_j} - (N w_j + z) P_{-j} e^{+\vec{\lambda} \cdot \vec{\rho}_j} \right) &= 0 \end{aligned}$$

Un système de 3 équations à 3 inconnues $(\lambda_x, \lambda_y, \lambda_z)$ dont la résolution est simplifiée grâce à la constatation suivante : le vecteur $\vec{\lambda} = \vec{0}$ est solution pour $(x, y, z) = (N\bar{u}, N\bar{v}, N\bar{w})$ avec :

$$\bar{u} = \sum_{j=1}^q u_j \delta P_j \quad \bar{v} = \sum_{j=1}^q v_j \delta P_j \quad \bar{w} = \sum_{j=1}^q w_j \delta P_j$$

Autrement dit, si on n'impose pas les contraintes (2, 3, 4) ($\lambda_x = \lambda_y = \lambda_z = 0$), l'objet se trouve statistiquement au bout de N sauts à la position moyenne attendue.

Or, pour les grandes valeurs de N , les probabilités sont appréciables uniquement pour les positions très proches de la position moyenne. Pour ces positions les produits scalaires $\vec{\lambda} \cdot \vec{\rho}_j$ sont vraisemblablement très petits. On vérifiera cette hypothèse *a posteriori*. On pose donc $|\vec{\lambda} \cdot \vec{\rho}_j| \ll 1$ et :

$$e^{-\vec{\lambda} \cdot \vec{\rho}_j} \approx 1 - \lambda_x u_j - \lambda_y v_j - \lambda_z w_j \quad \text{et} \quad e^{+\vec{\lambda} \cdot \vec{\rho}_j} \approx 1 + \lambda_x u_j + \lambda_y v_j + \lambda_z w_j \quad (3.17)$$

Et on remplace dans les trois équations. On trouve :

$$\begin{aligned}(N\bar{u}^2 - \bar{u}x)\lambda_x + (N\bar{v}\bar{u} - \bar{v}x)\lambda_y + (N\bar{w}\bar{u} - \bar{w}x)\lambda_z &= N\bar{u} - x \\ (N\bar{u}\bar{v} - \bar{v}y)\lambda_x + (N\bar{v}^2 - \bar{v}y)\lambda_y + (N\bar{w}\bar{v} - \bar{w}y)\lambda_z &= N\bar{v} - y \\ (N\bar{u}\bar{w} - \bar{w}z)\lambda_x + (N\bar{v}\bar{w} - \bar{v}z)\lambda_y + (N\bar{w}^2 - \bar{w}z)\lambda_z &= N\bar{w} - z\end{aligned}$$

$$\begin{aligned}\text{avec } \bar{u}^2 &= \sum_{j=1}^q P_j u_j^2 & \bar{v}^2 &= \sum_{j=1}^q P_j v_j^2 & \bar{w}^2 &= \sum_{j=1}^q P_j w_j^2 \\ \text{et } \bar{u}\bar{v} &= \bar{v}\bar{u} = \sum_{j=1}^q P_j u_j v_j & \bar{v}\bar{w} &= \bar{w}\bar{v} = \sum_{j=1}^q P_j v_j w_j & \bar{w}\bar{u} &= \bar{u}\bar{w} = \sum_{j=1}^q P_j w_j u_j\end{aligned}$$

Autres résultats :

$$e^\alpha = 1 \quad N_{+j}^* = N P_{+j}(1 - \vec{\lambda} \cdot \vec{\rho}_j) \quad N_{-j}^* = N P_{-j}(1 + \vec{\lambda} \cdot \vec{\rho}_j)$$

Quoiqu'un peu lourde, la résolution de ce système de trois équations linéaires ne pose aucun problème.

Sans forces extérieures ($P_{+j} = P_{-j} = P_j/2$ pour tout j), c'est un peu moins lourd car on a $\bar{u} = \bar{v} = \bar{w} = 0 \Rightarrow \bar{u}^2 = \Delta u^2, \bar{v}^2 = \Delta v^2, \bar{w}^2 = \Delta w^2$, et donc :

$$\begin{aligned}\Delta u^2 \lambda_x + \bar{v}\bar{u} \lambda_y + \bar{w}\bar{u} \lambda_z &= -x/N \\ \bar{u}\bar{v} \lambda_x + \Delta v^2 \lambda_y + \bar{w}\bar{v} \lambda_z &= -y/N \\ \bar{u}\bar{w} \lambda_x + \bar{v}\bar{w} \lambda_y + \Delta w^2 \lambda_z &= -z/N\end{aligned}$$

Dans la plupart des cas pratiques, on a $\bar{u}\bar{v} = \bar{v}\bar{w} = \bar{w}\bar{u} = 0$. Sinon, on choisit les axes pour qu'il en soit ainsi (axes principaux). D'où :

$$\lambda_x = -\frac{x}{N \Delta u^2} \quad \lambda_y = -\frac{y}{N \Delta v^2} \quad \lambda_z = -\frac{z}{N \Delta w^2} \quad (3.18)$$

$$N_{+j}^* = N \frac{P_j}{2} \left(1 + \frac{u_j x}{N \Delta u^2} + \frac{v_j y}{N \Delta v^2} + \frac{w_j z}{N \Delta w^2} \right)$$

$$N_{-j}^* = N \frac{P_j}{2} \left(1 - \frac{u_j x}{N \Delta u^2} - \frac{v_j y}{N \Delta v^2} - \frac{w_j z}{N \Delta w^2} \right)$$

On écrit ensuite l'approximation du terme maximum :

$$P(N, \vec{r}) = \sum_{[N]}^{(1,2,3,4)} P_{[N]} \approx A P_{[N^*]} \quad \text{avec} \quad P_{[N^*]} = N! \prod_{i=1}^q \frac{P_{+i}^{N_{+i}^*} P_{-i}^{N_{-i}^*}}{N_{+i}^{*!} N_{-i}^{*!}} = N! \prod_{i=1}^q \frac{\left(\frac{P_i}{2}\right)^{N P_i}}{N_{+i}^{*!} N_{-i}^{*!}}$$

où A est une constante de normalisation indépendante au premier ordre de x, y, z .

Avec la formule de Stirling, on trouve une densité de probabilité :

$$W(N, \vec{r}) = A \exp \left(- \sum_{j=1}^q N \frac{P_j}{2} \left(\frac{u_j x}{\Delta x^2} + \frac{v_j y}{\Delta y^2} + \frac{w_j z}{\Delta z^2} \right)^2 \right)$$

$$\text{avec } \Delta x^2 = N \Delta u^2 \quad \Delta y^2 = N \Delta v^2 \quad \Delta z^2 = N \Delta w^2$$

Et enfin, on développe la parenthèse au carré en utilisant $\overline{uv} = \overline{vw} = \overline{wu} = 0$. Après normalisation, on trouve :

$$W(N, \vec{r}) = \frac{\exp\left(-\frac{1}{2N}\left(\frac{x^2}{\Delta u^2} + \frac{y^2}{\Delta v^2} + \frac{z^2}{\Delta w^2}\right)\right)}{\sqrt{8\pi^3 N^3 \Delta u^2 \Delta v^2 \Delta w^2}} \quad (3.19)$$

On vérifie maintenant la condition $|\vec{\lambda} \cdot \vec{\rho}_j| \ll 1$ ayant permis d'écrire 3.17. Soit ℓ la longueur moyenne d'un saut (la moyenne des $||\vec{\rho}_i||$) : $\Delta u^2, \Delta v^2$ et Δw^2 sont de l'ordre de ℓ^2 . De 3.18 on tire :

$$\vec{\lambda} \sim -\frac{\vec{r}}{N\ell^2} \quad \Rightarrow \quad |\vec{\lambda} \cdot \vec{\rho}_j| \sim \frac{|\vec{r} \cdot \vec{\rho}_j|}{N\ell^2} < \frac{||\vec{r}|| \ ||\vec{\rho}_j||}{N\ell^2} \sim \frac{r}{N\ell}$$

On constate que l'hypothèse $|\vec{\lambda} \cdot \vec{\rho}_i| \ll 1$ est valide si $r = ||\vec{r}|| \ll N\ell$. Aucune difficulté pour vérifier cette condition. Les valeurs significatives de x sont dans un intervalle de quelques $\Delta x = \Delta u\sqrt{N} \sim \ell\sqrt{N}$ autour de la valeur moyenne nulle. Même chose pour y et z . On en déduit $r \sim \ell\sqrt{N}$, ce qui est bien très inférieur à $N\ell$.

Profil de concentration

Pour passer à l'échelle macroscopique, même démarche que dans le cas unidimensionnel. On considère un grand nombre M de particules partant toutes à l'instant $t = 0$, du même site origine, et se déplaçant sans gêne mutuelle. Inutile de dire que c'est plutôt difficile à réaliser expérimentalement.

On calcule $c(\vec{r}, t)$ en multipliant $W(N, \vec{r})$ par M et remplaçant N par t/τ . De plus, on relie M à la concentration initiale à l'origine en écrivant $c_0 = M/\Omega$, où Ω est le volume d'un site :

$$c(\vec{r}, t) = c_0 \beta \left(\frac{\tau}{2\pi t}\right)^{3/2} \exp\left(-\frac{\tau}{2t}\left(\frac{x^2}{\Delta u^2} + \frac{y^2}{\Delta v^2} + \frac{z^2}{\Delta w^2}\right)\right) \quad (3.20)$$

où $\beta = \frac{\Omega}{\Delta u \Delta v \Delta w}$ est un facteur géométrique dépendant du réseau.

À chaque instant, les surfaces de concentration constante sont des ellipsoïdes. C'est ce que prédit également le modèle macroscopique suivant.

Coefficient de diffusion, loi de Fick

Par définition, le coefficient de diffusion D_{xy} relie au signe près le flux de particules suivant Ox lorsqu'il y a variation de concentration suivant Oy . Avec des définitions similaires pour les autres coefficients, on a :

$$\begin{aligned} J_x &= -D_{xx} \frac{\partial c}{\partial x} - D_{xy} \frac{\partial c}{\partial y} - D_{xz} \frac{\partial c}{\partial z} \\ J_y &= -D_{yx} \frac{\partial c}{\partial x} - D_{yy} \frac{\partial c}{\partial y} - D_{yz} \frac{\partial c}{\partial z} \\ J_z &= -D_{zx} \frac{\partial c}{\partial x} - D_{zy} \frac{\partial c}{\partial y} - D_{zz} \frac{\partial c}{\partial z} \end{aligned} \quad \text{en écriture tensorielle : } \vec{J} = -[D] \vec{\nabla} c \quad (3.21)$$

Comme dans le modèle unidimensionnel, on supposera $[D]$ indépendant de x, y, z aussi bien de façon explicite que par l'intermédiaire de la concentration. Le bilan de matière se fait ici sur un petit élément $dV = dx dy dz$. On calcule la variation du nombre de particules présentes dans dV en étudiant le nombre de particules entrant et sortant par chacune des six faces de dV . On trouve de façon similaire à 3.13 :

$$-\left(\frac{\partial J_x}{\partial x} + \frac{\partial J_y}{\partial y} + \frac{\partial J_z}{\partial z}\right) dx dy dz = -\nabla \vec{J} dx dy dz dt \quad (3.22)$$

En raisonnant sur la concentration, on peut également dire que le nombre de particules de cette même tranche au cours du même intervalle de temps a varié de :

$$(c(\vec{r}, t + dt) - c(\vec{r}, t)) dx dy dz = \frac{\partial c}{\partial t} dt dx dy dz \quad (3.23)$$

Comparant 3.23 à 3.22 on déduit : $\frac{\partial c}{\partial t} = -\nabla \vec{J}$, puis en utilisant 3.21 :

$$\frac{\partial c}{\partial t} = \nabla([D] \vec{\nabla} c) = [D] \Delta c \quad \textbf{Loi de Fick} \quad (3.24)$$

La deuxième expression n'est valable que si $[D]$ est indépendant de c (déplacements indépendants) et de $\vec{r}$ (milieux homogènes).

On vérifiera que 3.20 est solution si :

$$\tau = \frac{\Delta u^2}{2 D_{xx}} = \frac{\Delta v^2}{2 D_{yy}} = \frac{\Delta w^2}{2 D_{zz}} \quad \text{et} \quad D_{xy} = D_{yx} = D_{yz} = D_{zy} = D_{zx} = D_{xz} = 0$$

$$\Rightarrow c(\vec{r}, t) = \frac{c_0 \Omega}{\sqrt{4\pi D_{xx} t} \sqrt{4\pi D_{yy} t} \sqrt{4\pi D_{zz} t}} \exp\left(-\frac{1}{4t} \left(\frac{x^2}{D_{xx}} + \frac{y^2}{D_{yy}} + \frac{z^2}{D_{zz}}\right)\right)$$

Dans les **cristaux à symétrie cubique**, les polyèdres de coordination les plus courants sont représentés figure 3.4.

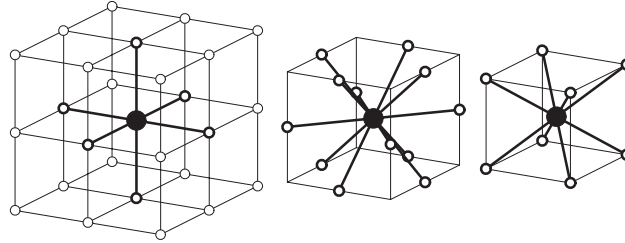


FIG. 3.4 – Polyèdres de coordination dans les cristaux cubiques

Dans les trois cas, les pas ont même longueur. Si on note ℓ cette longueur, on trouve :

$$\Delta u^2 = \Delta v^2 = \Delta w^2 = \frac{\ell^2}{3} \quad \Rightarrow \quad D_{xx} = D_{yy} = D_{zz} = \frac{\ell^2}{6\tau} = D$$

La concentration ne dépend plus que de t et de $r^2 = x^2 + y^2 + z^2$:

$$c(r, t) = c_0 \beta \left(\frac{\tau}{2\pi t}\right)^{3/2} \exp\left(-\frac{3 r^2 \tau}{2 \ell^2 t}\right) = \frac{c_0 \Omega}{(4\pi D t)^{3/2}} \exp\left(-\frac{r^2}{4 D t}\right)$$

La distance moyenne à l'origine varie toujours en $\sqrt{t}$:

$$\Delta r = \ell \sqrt{\frac{t}{3\tau}} = \sqrt{2Dt} \quad \text{et} \quad \bar{r}(t) = \frac{\int_0^\infty r c(r, t) 4\pi r^2 dr}{\int_0^\infty c(r, t) 4\pi r^2 dr} = \ell \sqrt{\frac{8t}{3\pi\tau}} = 4\sqrt{\frac{Dt}{\pi}}$$

3.5.4 Mouvement de particules dans un fluide

L'objet (la particule) se déplace toujours par pas successifs de longueur fixe, mais cette fois, à partir d'une position donnée, toutes les directions sont possibles. Si toutes les directions sont équiprobables (pas de champ de forces extérieur, pas de gradient de concentration), le résultat trouvé au paragraphe précédent reste valable. En effet, aucune hypothèse n'a été faite sur le nombre de directions possibles à partir d'un site donné. Rien ne nous interdit de faire tendre celui-ci vers l'infini. Il faut simplement vérifier que les valeurs moyennes et variances sont identiques aux précédentes.

On place l'objet à l'origine, et on cherche à savoir à quelle distance il se trouve de celle-ci au bout de N pas (ou au bout d'un temps $t = N\tau$). À partir d'un point quelconque de la trajectoire, la probabilité pour que l'objet choisisse une direction située dans l'angle solide $\sin\theta d\theta d\varphi$ autour de (θ, φ) vaut :

$$dP(\theta, \varphi) = \frac{\sin\theta d\theta d\varphi}{4\pi}$$

Il nous faut les valeurs moyennes des coordonnées cartésiennes et de leurs carrés. On trouve en intégrant $x dP$, $y dP$, $z dP$, $x^2 dP$, $y^2 dP$, $z^2 dP$ (avec $x = \ell \sin\theta \cos\varphi$, $y = \ell \sin\theta \sin\varphi$, $z = \ell \cos\theta$, $\varphi = 0$ à 2π , $\theta = 0$ à π) :

$$\bar{x} = \bar{y} = \bar{z} = 0 \quad \Delta x^2 = \overline{x^2} = \Delta y^2 = \overline{y^2} = \Delta z^2 = \overline{z^2} = \frac{\ell^2}{3} \quad (3.25)$$

On en déduit la densité de probabilité de trouver l'objet à l'instant t à une distance r de l'origine (mêmes résultats) :

$$W(r, t) = \frac{1}{(4\pi Dt)^{3/2}} \exp\left(-\frac{r^2}{4Dt}\right) \quad \text{avec} \quad D = \frac{\ell^2}{6\tau} \quad (3.26)$$

puis les différentes valeurs moyennes :

$$\bar{r}(t) = 4\sqrt{\frac{Dt}{\pi}} \quad \overline{r^2} = 6Dt \quad \frac{\Delta r}{\bar{r}} = \sqrt{\frac{3\pi}{8}} - 1 \approx 0,422 \quad (3.27)$$

Conclusion : une distance moyenne à l'origine variant en $\sqrt{t}$, et les fluctuations très étendues autour de $\bar{r}$.

Si maintenant on prend en compte le fait que les pas sont de longueur λ variable, on introduit la probabilité $dP(\lambda, \theta, \varphi)$ pour que l'objet choisisse un pas de longueur comprise entre λ et $\lambda + d\lambda$ dans une direction située dans l'angle solide $\sin\theta d\theta d\varphi$ autour de (θ, φ) :

$$dP(\lambda, \theta, \varphi) = w(\lambda) d\lambda \frac{\sin\theta d\theta d\varphi}{4\pi}$$

Les relations 3.25, 3.26 et 3.27 restent identiques, mis à part que maintenant ℓ est le **libre parcours moyen** $\ell = \int_0^\infty \lambda w(\lambda) d\lambda$, et ℓ^2 le libre parcours quadratique moyen $\ell^2 = \int_0^\infty \lambda^2 w(\lambda) d\lambda$.

Calcul du libre parcours moyen.

On considère ici un **gaz parfait de sphères dures**. Les molécules qui le constituent se déplacent en ligne droite, ou plus exactement sur des lignes brisées constituées de segments rectilignes : chaque fois qu'une molécule en rencontre une autre, il se produit une collision provoquant un changement brusque de direction. On n'envisage ici que les collisions les plus courantes aux températures habituelles : chaque choc est supposé binaire (ne faisant intervenir que deux molécules) et élastique (pas de modification des structures internes des molécules, et donc conservation de la quantité de mouvement et de l'énergie).

Le libre parcours moyen est la distance moyenne que parcourt une molécule entre deux collisions. Soient deux molécules A et B de masses m_A et m_B convergeant toutes deux vers un point O à des vitesses $\vec{v}_A$ et $\vec{v}_B$ telles qu'elles vont y arriver en même temps. Leur centre de masse G converge également vers O. Il est facile de voir sur la figure 3.5 qu'on peut ramener le problème à deux particules convergeant vers G, la molécule A à la vitesse $\vec{v}_A - \vec{v}_G$, la molécule B à la vitesse $\vec{v}_B - \vec{v}_G$.

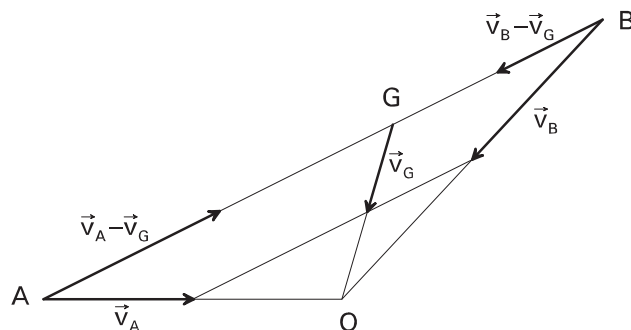


FIG. 3.5 – Collision : vitesses par rapport au centre de masse

Le mouvement du centre de masse n'a aucune influence sur la collision. On est ainsi conduit à n'étudier que les chocs frontaux : en mouvement relatif, les deux molécules partagent maintenant la même trajectoire. En termes d'énergie cinétique, on a tout simplement séparé le mouvement de translation du centre de masse du mouvement des molécules par rapport au centre de masse :

$$T = \frac{1}{2}m_A\vec{v}_A^2 + \frac{1}{2}m_B\vec{v}_B^2 = \frac{1}{2}M\vec{v}_G^2 + \frac{1}{2}m_A(\vec{v}_A - \vec{v}_G)^2 + \frac{1}{2}m_B(\vec{v}_B - \vec{v}_G)^2$$

$$\text{avec } M = m_A + m_B \quad \text{et} \quad M\vec{v}_G = m_A\vec{v}_A + m_B\vec{v}_B$$

Il faut maintenant prendre en compte le fait que les molécules ne sont pas ponctuelles. On peut avoir collision frontale si les deux trajectoires, toujours parallèles, sont séparées par une distance p (appelée paramètre d'impact). Si r_A et r_B sont les rayons des deux molécules, il y aura collision si $p < r_A + r_B$ (figure 3.6).

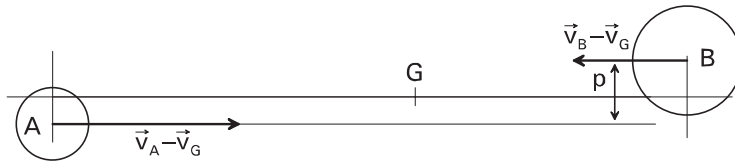


FIG. 3.6 – Collision frontale : paramètre d'impact

En remplaçant $\vec{v}_G$ par $(m_A \vec{v}_A + m_B \vec{v}_B)/M$ dans les deux derniers termes de l'énergie cinétique, on trouve :

$$T = \frac{1}{2} M \vec{v}_G^2 + \frac{1}{2} m_{AB} (\vec{v}_A - \vec{v}_B)^2 \quad \text{avec} \quad m_{AB} = \frac{m_A m_B}{m_A + m_B} \quad (\text{masse réduite})$$

C'est la réduction du problème à deux corps qui permet ici de remplacer les deux molécules mobiles A et B par un objet ponctuel et sans masse et un objet mobile de masse m_{AB} , de rayon $r_{AB} = r_A + r_B$ et de vitesse $\vec{v}_{AB} = \vec{v}_A - \vec{v}_B$.

Pour calculer le nombre moyen de collisions subies par une molécule A de la part des molécules B au cours de l'intervalle de temps δt , il suffit de compter le nombre de molécules B présentes dans un cylindre de section πr_{AB}^2 et de longueur $\bar{v}_{AB} \delta t$. En divisant ensuite par δt , on obtient la fréquence de collisions A-B par atome A. On trouve, en utilisant 3.5 :

$$Z_A^B = \rho_B \pi r_{AB}^2 \bar{v}_{AB} = \rho_B \pi (r_A + r_B)^2 \sqrt{\frac{8kT}{\pi m_{AB}}}$$

Si le gaz est multiconstituant (molécules A, B, C, etc), on obtient la **fréquence de collisions** pour une molécule A en ajoutant les différentes contributions :

$$Z_A = Z_A^A + Z_A^B + Z_A^C \dots$$

Et pour son **libre parcours moyen** il suffit d'écrire :

$$\ell_A = \frac{\bar{v}_A}{Z_A} = \frac{1}{Z_A} \sqrt{\frac{8kT}{\pi m_A}} = \frac{1}{4\pi r_A^2 \rho_A \sqrt{2} + \pi (r_A + r_B)^2 \rho_B \sqrt{1 + \frac{m_A}{m_B}} + \dots}$$

Le libre parcours moyen est indépendant de la température. Seules les densités, les rayons moléculaires et les rapports de masse interviennent dans l'expression.

Pour un gaz monoconstituant, on a (molécules de masse m et de rayon r) :

$$Z = 4\pi r^2 \rho \sqrt{\frac{16kT}{\pi m}} \quad \text{et} \quad \ell = \frac{1}{4\pi r^2 \rho \sqrt{2}}$$

Voici quelques valeurs numériques du libre parcours moyen, comparé à la distance moyenne d entre molécules, pour l'hélium ($r = 128$ pm), le néon ($r = 154$ pm), l'argon ($r = 188$ pm), le krypton ($r = 202$ pm) et le xénon ($r = 216$ pm).^[6]

⁶ VE : vide extrême, UV : ultravide, VS : vide secondaire, VP : vide primaire, GO : gaz ordinaire

On constate que le libre parcours moyen est dans un gaz ordinaire 1,2 à 7500 fois plus grand (15 à 40 dans les conditions standard) que la distance moyenne entre molécules, et qu'il peut atteindre des valeurs considérables en vide extrême (oui, ce sont bien des giga-mètres) : un atome d'hélium du milieu intersidéral peut parcourir plus de 30 millions de kilomètres (un an et demi à 100 K) sans subir aucune collision. Ce même atome dans les conditions standard subirait environ 300 millions de collisions par seconde.

ρ	10^8	V	10^{12}	U	10^{16}	V	10^{19}	V	10^{22}	G	10^{27}
d	2 mm	E	100 μm	V	5 μm	S	500 nm	P	50 nm	O	1 nm
He	34,3 Gm		3434 km		343 m		343 mm		343 μm		3,43 nm
Ne	23,7 Gm		2373 km		237 m		237 mm		237 μm		2,37 nm
Ar	15,9 Gm		1592 km		159 m		159 mm		159 μm		1,59 nm
Kr	13,8 Gm		1379 km		138 m		138 mm		138 μm		1,38 nm
Xe	12,1 Gm		1206 km		121 m		121 mm		121 μm		1,21 nm

Lorsque le libre parcours moyen devient supérieur aux dimensions du récipient, ce ne sont plus les collisions entre molécules qui définissent les propriétés du gaz, mais les chocs molécule – paroi. On parle alors de "régime moléculaire". *A contrario*, lorsque le libre parcours moyen est très inférieur aux dimensions de l'enceinte, on parle de "régime fluide" ou de "régime de fluide quasi continu".

3.6 Éléments de mécanique statistique classique

Dans ce paragraphe, il n'est plus question de gaz parfait. Les particules envisagées sont ponctuelles ou non, indépendantes ou non ... Seule hypothèse : elles sont indiscernables. Ce n'est pas un traitement rigoureux de la mécanique statistique classique : les différents éléments seront introduits sur des exemples simples, puis généralisés, le but étant de retrouver la distribution de Maxwell-Boltzmann 3.3 et la distribution de Maxwell 3.4 (paragraphe 3.4).

3.6.1 Systèmes à une particule

L'oscillateur harmonique linéaire

Ce premier exemple est le plus simple qui soit, puisqu'il n'y a qu'un seul degré de liberté, et que la théorie quantique permet de résoudre exactement le problème à toute température. Considérons une particule de masse m oscillant à la pulsation ω le long d'un axe Ox autour de sa position d'équilibre $x = 0$. On connaît le résultat quantique : les niveaux d'énergie, non dégénérés, accessibles au système sont donnés par $\varepsilon_n = \hbar \omega (n + 1/2)$ avec $n \geq 0$, et la fonction de partition canonique est calculable sans approximations :

$$q(T) = \sum_{n=0}^{\infty} e^{-\hbar \omega (n+1/2)/kT} = \frac{1}{2 \sinh \frac{\hbar \omega}{2kT}} \sim \frac{kT}{\hbar \omega} \quad \text{quand } T \rightarrow \infty$$

En mécanique classique, l'équation du mouvement conduit à :

$$x(t) = a \cos(\omega t + \phi) \quad \text{et} \quad \dot{x}(t) = -a\omega \sin(\omega t + \phi)$$

et à une énergie fonction de la pulsation et de l'amplitude :

$$\varepsilon = \frac{1}{2} m \dot{x}^2 + \frac{1}{2} m \omega^2 x^2 = \frac{1}{2} m \omega^2 a^2$$

En posant $q = x \sqrt{m\omega}$, le lagrangien devient : $L = \frac{1}{2} m \dot{x}^2 - \frac{m\omega^2}{2} x^2 = \frac{\dot{q}^2 - \omega^2 q^2}{2\omega}$

Ce qui conduit au hamiltonien (c'est aussi l'énergie totale du système) :

$$H(p, q) = \frac{\omega}{2} (q^2 + p^2) \quad \text{avec} \quad q = x \sqrt{m\omega} \quad \text{et} \quad p = \frac{\partial L}{\partial \dot{q}} = \frac{\dot{q}}{\omega} = \dot{x} \sqrt{\frac{m}{\omega}}$$

Un couple de variables (q, p) représente un état possible du système. Dans le raisonnement qui suit, ces deux variables sont considérées comme indépendantes. Elles peuvent toutes deux varier continûment de $-\infty$ à $+\infty$. On peut représenter les états possibles dans un espace à deux dimensions (q, p) appelé espace des phases du système. Dans cet espace, l'ensemble des états d'énergie E (ou d'amplitude a) est un cercle d'équation :

$$q^2 + p^2 = \frac{2E}{\omega} \quad \text{avec} \quad E = \frac{m\omega^2 a^2}{2}$$

Au cours du temps, le point représentatif de l'état du système se déplace le long du cercle. Chaque point traversé représente un état possible : *l'ensemble des états possibles est continu*, alors qu'en mécanique quantique c'est un ensemble discret.

Pour calculer la fonction de partition classique, il faut donc remplacer la somme discrète par une intégrale en introduisant une *densité d'états* $g(q, p)$:

$$q(T) = \sum_{n=0}^{\infty} \omega_n e^{-\varepsilon_n/kT} \quad \longrightarrow \quad q_{\text{class}} = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} g(q, p) e^{-H(q, p)/kT} dq dp$$

La densité d'états $g(q, p)$ est nécessaire car en mécanique quantique, seuls quelques cercles sont permis (ceux correspondant aux énergies $\varepsilon_n = \hbar \omega (n + 1/2)$). De plus, on sait que les niveaux sont non dégénérés : en mécanique classique, chaque point d'un cercle représente un état possible, alors qu'en mécanique quantique on ne compte *qu'un seul état par cercle*. Celui-ci est *délocalisé* sur le cercle. Autrement dit, à l'état quantique n , il faut faire correspondre les états classiques (q, p) situés entre les deux cercles correspondant aux énergies $\hbar \omega n$ et $\hbar \omega (n + 1)$. L'aire de cette surface vaut $2\pi \hbar = h$.

Conclusion : à une surface d'aire h correspond *un* état quantique. D'où :

$$g(q, p) = \frac{1}{h} \quad \text{et donc} \quad q_{\text{class}} = \frac{1}{h} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} e^{-H(q, p)/kT} dq dp \quad (3.28)$$

Vérification :

$$q_{\text{class}} = \frac{1}{h} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} e^{-\omega (q^2 + p^2)/2kT} dq dp = \frac{1}{h} \int_{-\infty}^{+\infty} e^{-\omega q^2/2kT} dq \int_{-\infty}^{+\infty} e^{-\omega p^2/2kT} dp$$

À l'aide de J_0 donnée au paragraphe 3.4.1, on trouve :

$$q_{\text{class}} = \frac{1}{h} \left(2\sqrt{\frac{kT}{\omega}} \sqrt{\frac{\pi}{2}} \right)^2 = \frac{2\pi kT}{h\omega} = \frac{kT}{\hbar\omega}$$

C'est bien la limite haute température de la fonction de partition quantique.

Masse ponctuelle dans une boîte cubique

Encore un problème où la limite haute température de la fonction de partition est bien connue. Le raisonnement classique doit en principe conduire à :

$$q_{\text{class}} = \frac{V}{\Lambda^3} \quad \text{avec} \quad \Lambda = \frac{h}{\sqrt{2\pi m kT}}$$

La particule a ici trois degrés de liberté. L'espace des phases $\mathcal{E}_6$ est à six dimensions : $\mathcal{E}_6 = \mathcal{E}_3 \times \mathcal{E}_3^*$, où $\mathcal{E}_3$ est l'espace affine des positions $\vec{r}$ (espace direct) et $\mathcal{E}_3^*$ est l'espace affine des impulsions $\vec{p}$ (espace réciproque). Pour une particule confinée dans une boîte de volume V , on a :

$$q_{\text{class}} = \iiint d^3r \iiint d^3p \, g(\vec{r}, \vec{p}) e^{-H(\vec{r}, \vec{p})/kT}$$

ou encore en explicitant les variables (pour une boîte cubique de volume $V = L^3$) :

$$q_{\text{class}} = \int_0^L dx \int_0^L dy \int_0^L dz \int_{-\infty}^{+\infty} dp_x \int_{-\infty}^{+\infty} dp_y \int_{-\infty}^{+\infty} dp_z \, g(x, y, z, p_x, p_y, p_z) e^{-H(x, y, z, p_x, p_y, p_z)/kT}$$

Le hamiltonien H est indépendant de x, y, z . Il est de plus à variables séparables :

$$H = \frac{p_x^2}{2m} + \frac{p_y^2}{2m} + \frac{p_z^2}{2m} \Rightarrow e^{-H/kT} = e^{-p_x^2/(2m kT)} \times e^{-p_y^2/(2m kT)} \times e^{-p_z^2/(2m kT)}$$

Quant à la densité d'états, il faut revenir au modèle quantique de la particule libre dans une boîte cubique. Les composantes possibles de l'impulsion sont données par :

$$p_x = \frac{h}{L} n_x \quad p_y = \frac{h}{L} n_y \quad p_z = \frac{h}{L} n_z \quad \text{avec } n_x, n_y, n_z \text{ entiers}$$

La densité d'états quantiques $g(\vec{r}, \vec{p})$ est le nombre d'états par unité de volume direct et par unité de volume réciproque. Dans l'espace réciproque, on compte un état dans chaque élément de volume réciproque $\delta V^* = h^3/L^3$. On en déduit :

$$g(\vec{r}, \vec{p}) = \frac{1}{V} \times \frac{1}{\delta V^*} = \frac{1}{h^3}$$

La fonction de partition q_{class} devient :

$$q_{\text{class}} = \frac{1}{h^3} \iiint d^3r \iiint d^3p \, e^{-H(\vec{r}, \vec{p})/kT} = \frac{V}{h^3} \iiint d^3p \, e^{-p^2/(2m kT)} \quad (3.29)$$

On obtient bien la limite haute température du modèle quantique :

$$q_{\text{class}} = \frac{V}{h^3} \int_{-\infty}^{+\infty} dp_x \, e^{-p_x^2/(2m kT)} \int_{-\infty}^{+\infty} dp_y \, e^{-p_y^2/(2m kT)} \int_{-\infty}^{+\infty} dp_z \, e^{-p_z^2/(2m kT)} = \frac{V}{\Lambda^3} \quad (3.30)$$

Particule unique quelconque

Une particule matérielle (non ponctuelle) possède des degrés de liberté qui ne peuvent pas être traités par la mécanique classique. Ce sont les degrés de liberté liés au nuage électronique et à la structure interne du noyau et, si c'est une molécule, les degrés de liberté de vibration. On sait que seuls les degrés de liberté de translation et de rotation peuvent être traités classiquement (et encore, pas toujours pour la rotation).

On supposera donc le hamiltonien séparable en deux termes, l'un regroupant tous les degrés de liberté à traiter par la mécanique quantique, l'autre regroupant ceux pour lesquels le traitement classique est suffisant. On notera n le nombre de degrés de liberté à traiter classiquement :

$$H = H_{\text{quant}} + H_{\text{class}}((q_i, p_i)_{(i=1 \text{ à } n)})$$

La fonction de partition s'écrit comme le produit de deux fonctions de partition (le terme $1/h^n$ provient d'une généralisation des résultats 3.28 et 3.29) :

$$q = q_{\text{quant}} \times q_{\text{class}} \text{ avec } q_{\text{class}} = \frac{1}{h^n} \int \cdots \int dq_1 \dots dq_n dp_1 \dots dp_n e^{-H_{\text{class}}/kT} \quad (3.31)$$

On définit **l'intégrale de phase** comme étant :

$$I(T, V) = \int \cdots \int dq_1 \dots dq_n dp_1 \dots dp_n e^{-H_{\text{class}}/kT}$$

et on énonce :

La fonction de partition "classique" d'une particule à n degrés de liberté "classiques" est égale à l'intégrale de phase divisée par h^n .

A noter que l'intégrale de phase peut prendre en compte la présence éventuelle d'un champ de potentiel extérieur.

3.6.2 Systèmes à N particules

Masses ponctuelles identiques et indépendantes dans une boîte cubique

Le hamiltonien du système est la somme de hamiltoniens individuels :

$$H = \sum_{i=1}^N \frac{p_i^2}{2m} = \sum_{i=1}^N \frac{p_{ix}^2 + p_{iy}^2 + p_{iz}^2}{2m}$$

Il suffit donc ici d'élever le résultat 3.30 à la puissance N , **et diviser par $N!$** :

$$Q = \frac{1}{N!} \left[\frac{V}{\Lambda^3} \right]^N$$

Le facteur $N!$ vient du fait que les particules sont indiscernables.

Pour comprendre sa présence, il faut noter que $(V/\Lambda^3)^N$ est l'intégrale de phase $I(T, V)$ divisée par h^{3N} . En effet, le calcul 3.29 devient ici :

$$\frac{I(T, V)}{h^{3N}} = \frac{1}{h^{3N}} \iiint d^3r_1 d^3p_1 \dots \iiint d^3r_N d^3p_N e^{-H/kT} = \left[\frac{V}{\Lambda^3} \right]^N$$

Lorsqu'on calcule I , on effectue N intégrales sextuples identiques et indépendantes, chacune sur l'élément $d\tau = d^3r d^3p$.

Prenons par exemple les particules 1 et 2, et deux éléments de volume $d\tau$ et $d\tau'$. Le calcul prend en compte, d'une part la présence de la particule 1 dans $d\tau$ et de la particule 2 dans $d\tau'$, d'autre part la présence de la particule 1 dans $d\tau'$ et de la particule 2 dans $d\tau$. On sait que ces deux configurations en mécanique quantique représentent le même état : on doit diviser le produit $I_1 \times I_2$ par 2. Si on généralise aux N particules, on doit diviser I par $N!$, c'est-à-dire par le nombre de permutations sur les N particules réparties dans N éléments $d\tau_1 \dots d\tau_N$.

Particules interactives ou dans un potentiel non uniforme

On généralise maintenant aux systèmes quelconques :

- Les particules ne sont plus ponctuelles : elles possèdent des degrés de liberté internes à traiter quantiquement (noyaux, électrons, vibrations) et des degrés de liberté pouvant être traités classiquement (translation, rotation). De plus on étend le raisonnement aux systèmes multiconstituants : on notera n_i ($i = 1$ à p) le nombre de degrés de liberté classiques de l'espèce i , et N_i le nombre de molécules i . La somme de tous les $n_i N_i$ sera notée n : c'est le nombre total de degrés de liberté classiques.
- Les particules ne sont plus indépendantes. Il existe entre elles des forces attractives ou répulsives qui en général ne dépendent que des positions relatives des particules entre elles. Mais le résultat 3.32 ci-dessous autorise également la présence de forces d'interaction faisant intervenir les vitesses.
- On autorise la présence d'un potentiel d'origine extérieur. Comme les forces d'interaction, ce potentiel extérieur peut éventuellement dépendre des vitesses.

La fonction de partition canonique s'écrit maintenant :

$$Q = \frac{Q_{\text{quant}}}{\prod_{i=1}^p (N_i! h^{n_i N_i})} \iint dq_1 dp_1 \dots \iint dq_n dp_n e^{-H_{\text{class}}/kT} \quad (3.32)$$

Un cas particulier fréquent est celui des systèmes constitués d'atomes tous identiques (sans structure interne à traiter classiquement) plongés dans un potentiel (extérieur et/ou d'interaction) indépendant des vitesses. On trouve alors :

$$H = H_{\text{quant}} + \sum_{i=1}^{3N} \frac{p_i^2}{2m} + U(q_1, \dots, q_{3N}) \quad \Rightarrow \quad Q = \frac{Q_{\text{quant}}}{N! \Lambda^{3N}} Z_N \quad (3.33)$$

avec $Z_N = \int dq_1 \dots \int dq_{3N} e^{-U/kT}$ **intégrale de configuration classique**

On considère un système macroscopique isolé, de volume V , et contenant des molécules appartenant à une ou plusieurs espèces. Soient (q_i, p_i) ($i = 1$ à n) l'ensemble des coordonnées et moments généralisés nécessaires à la description d'un microétat du système : l'espace des phases est à $2n$ dimensions. L'état du système à l'instant t est entièrement défini par la position de son point phase dans l'espace des phases. Au cours du temps, le système suit une certaine "trajectoire" dans l'espace des phases. Bien entendu cette trajectoire est impossible à calculer. On fabrique donc un ensemble statistique de $\mathcal{N}$ systèmes isolés, tous macroscopiquement identiques au système de départ. Le microétat de chacun des systèmes est représenté par un point phase, et l'ensemble des $\mathcal{N}$ points phase forme un nuage de points dans l'espace des phases. Dans la limite où $\mathcal{N}$ est infini, ce nuage est continu et possède une certaine densité.

Le théorème de Liouville dit que, pour un système isolé (E constant), cette densité reste constante le long de la trajectoire suivie par le système dans l'espace des phases. C'est tout simplement l'écriture classique du deuxième postulat (équiprobabilité des états de même énergie). Cela signifie concrètement que dans l'espace des phases, la densité est constante dans la région de l'espace des phases comprise entre la surface définie par E et celle définie par $E + dE$. D'après l'écriture de Q_{class} , on peut dire que cette densité est proportionnelle à $e^{-H_{\text{class}}/kT}$.

On termine par l'écriture de l'hypothèse ergodique classique : le point représentatif d'un système isolé passe autant de temps (sur une longue période) dans les différents petits volumes de l'espace des phases contenus dans la région comprise entre les surfaces d'énergies E et $E + dE$.

3.6.3 Distributions de Maxwell et de Maxwell-Boltzmann

Ces distributions ont été établies dans le cas du gaz parfait. Le raisonnement qui suit généralise ce résultat aux fluides quelconques. La seule supposition qui est faite ici est que les degrés de liberté classiques et quantiques sont séparables. On raisonnera pour simplifier sur un système monoconstituant.

On sait que les degrés de liberté de translation sont séparables de tous les autres. Il y en a $3N$ (les coordonnées de N centres de masse). On notera n le nombre de degrés de liberté "internes" pouvant être traités classiquement. Les coordonnées cartésiennes des centres de masse seront notées $(q_1 \dots q_{3N})$, et leurs moments conjugués $(p_1 \dots p_{3N})$. Les degrés de liberté internes seront représentés par les variables $(\xi_1 \dots \xi_n)$ et les momentôides conjugués $(\eta_1 \dots \eta_n)$:

$$Q = \frac{Q_{\text{quant}}}{N! h^{3N} h^n} \iiint \dots \iiint dq_1 \dots dq_{3N} dp_1 \dots dp_{3N} d\xi_1 \dots d\xi_n d\eta_1 \dots d\eta_n e^{-H_{\text{class}}/kT}$$

$$\text{avec } H_{\text{class}} = \sum_{i=1}^{3N} \frac{p_i^2}{2m} + T_{\text{int}}(\xi_1 \dots \xi_n, \eta_1 \dots \eta_n) + U(q_1 \dots q_{3N}, \xi_1 \dots \xi_n)$$

Dans H_{class} , le premier terme est l'énergie cinétique de translation, le deuxième l'énergie cinétique liée aux degrés de liberté internes classiques, et le troisième une énergie potentielle d'interaction supposée indépendante des momentôides.

On cherche à calculer la fraction de molécules dont le centre de masse possède une impulsion comprise entre p et $p + dp$ (en module).

On procède par étapes.

- Probabilité que le point phase soit, indépendamment de l'état quantique des molécules, dans l'élément $(dq_1 \dots dq_{3N} dp_1 \dots dp_{3N} d\xi_1 \dots d\xi_n d\eta_1 \dots d\eta_n)$:

$$\frac{dq_1 \dots dq_{3N} dp_1 \dots dp_{3N} d\xi_1 \dots d\xi_n d\eta_1 \dots d\eta_n e^{-H_{\text{class}}/kT}}{\iiint \dots \iiint dq_1 \dots dq_{3N} dp_1 \dots dp_{3N} d\xi_1 \dots d\xi_n d\eta_1 \dots d\eta_n e^{-H_{\text{class}}/kT}}$$

- Probabilité que le système possède, indépendamment des autres contributions à l'énergie, une certaine énergie cinétique de translation T_t .

On intègre le numérateur sur les $(dq_1 \dots dq_{3N} d\xi_1 \dots d\xi_n d\eta_1 \dots d\eta_n)$. Cette intégrale se simplifie avec le dénominateur :

$$\frac{dp_1 \dots dp_{3N} e^{-T_t/kT}}{\iiint \dots \iiint dp_1 \dots dp_{3N} e^{-T_t/kT}} = \frac{dp_1 \dots dp_{3N} e^{-T_t/kT}}{\prod_{i=1}^{3N} \int_{-\infty}^{+\infty} e^{-p_i^2/(2m kT)} dp_i} = \frac{dp_1 \dots dp_{3N} e^{-T_t/kT}}{(2\pi m kT)^{3N/2}}$$

- Probabilité de trouver, indiscernablement, n'importe quelle molécule dans un certain état (dp_x, dp_y, dp_z) autour de (p_x, p_y, p_z) .

Il suffit d'intégrer l'expression ci-dessus $3N - 3$ fois. Les variables d'intégration choisies sont sans importance : les trois variables restantes sont simplement renommées (p_x, p_y, p_z) . Soit $W(p_x, p_y, p_z) dp_x dp_y dp_z$ cette probabilité :

$$W(p_x, p_y, p_z) dp_x dp_y dp_z = \frac{dp_x dp_y dp_z e^{-(p_x^2 + p_y^2 + p_z^2)/(2m kT)}}{(2\pi m kT)^{3/2}} = \frac{d^3p e^{-p^2/(2m kT)}}{(2\pi m kT)^{3/2}}$$

On passe maintenant en coordonnées sphériques ($d^3p = p^2 dp \sin \theta d\theta d\varphi$) et on intègre sur θ et φ . On obtient la fraction de molécules dont le centre de masse possède une impulsion comprise entre p et $p + dp$:

$$W(p) dp = \frac{e^{-p^2/(2m kT)} p^2 dp}{(2\pi m kT)^{3/2}} \int_0^{2\pi} d\varphi \int_0^\pi \sin \theta d\theta = \frac{e^{-p^2/(2m kT)} 4\pi p^2 dp}{(2\pi m kT)^{3/2}}$$

D'où les différents résultats :

- **densité d'impulsion** : $\frac{e^{-p^2/(2m kT)}}{(2\pi m kT)^{3/2}} 4\pi p^2$
- **densité de vitesse** ($p = m v$) : $\left(\frac{m}{2\pi kT}\right)^{3/2} 4\pi v^2 e^{-m v^2/(2 kT)}$
- **densité d'énergie cinétique** ($\varepsilon = \frac{p^2}{2m}$) : $\frac{2}{\sqrt{\pi}} \frac{1}{kT} \sqrt{\frac{\varepsilon}{kT}} e^{-\varepsilon/kT}$

Ce sont bien les distributions trouvées aux paragraphes 3.3 et 3.4.

Chapitre 4

Gaz parfait d'atomes ou molécules

4.1 Les gaz parfaits classiques d'atomes

4.1.1 Degrés de liberté électroniques et nucléaires

Dans l'étude des gaz, le noyau est considéré comme une seule particule dotée d'une structure interne complexe. Dans les conditions habituelles, les états nucléaires et les états électroniques peuvent être considérés comme indépendants. On pourra donc écrire la fonction de partition "interne" $q^{(\text{int})}$ comme un produit d'une fonction de partition nucléaire $q^{(\text{nucl})}$ par une fonction de partition électronique $q^{(\text{elec})}$:

$$q^{(\text{int})}(T) = q^{(\text{nucl})}(T) \times q^{(\text{elec})}(T)$$

Pour les ordres de grandeur (on ignore les dégénérescences), on adoptera le raisonnement suivant. Soient ε_0 et ε_1 les énergies du niveau fondamental et du premier niveau excité, et $\Delta\varepsilon = \varepsilon_1 - \varepsilon_0$. Soient η_0 et η_1 les fractions d'atomes respectivement dans l'état fondamental et dans le premier état excité. Sachant que dans tous les cas pratiques les autres états excités sont quasiment inaccessibles, on a $\eta_1 = \eta_0 e^{-\Delta\varepsilon/kT}$. Lorsque T vaut $10^4 \Delta\varepsilon$, avec $\Delta\varepsilon$ exprimé en attojoules, on trouve $\eta_1 \approx 0,007$. On dira que η_1 est non négligeable devant η_0 si T est supérieure à $10^4 \Delta\varepsilon$.

Pour les noyaux, $\Delta\varepsilon$ est de l'ordre de 10^5 aJ. Pour peupler de façon non négligeable le premier niveau excité nucléaire, il faut atteindre des températures supérieures à 10^9 K. Dans tous les cas pratiques, seul le niveau fondamental nucléaire est accessible. Il est en général dégénéré. Si on note ω_n sa dégénérescence, et si on place l'origine des énergies sur le niveau fondamental nucléaire, on a :

$$q^{(\text{nucl})} = \omega_n \quad \text{à toute température}$$

Pour les électrons, même chose mais il y a quelques exceptions :

- Les gaz rares (He, Ne, Ar, Kr, Xe) : l'écart $\Delta\varepsilon$ est de l'ordre de 1 à 3 aJ. On devra prendre en compte le premier niveau excité lorsque $T > 10\,000$ K.
- Les alcalins (Li, Na, K, Rb) : $\Delta\varepsilon$ est de l'ordre de 0,1 à 0,3 aJ. Il faudra prendre en compte le premier niveau excité lorsque $T > 1\,000$ K.
- Les halogènes (F, Cl, Br, I) : $\Delta\varepsilon$ vaut respectivement 0,008, 0,018, 0,074 et 0,150 aJ. Pour le fluor et le chlore, le premier niveau excité est à prendre en compte même à l'ambiante. Pour le brome, on peut l'ignorer jusqu'à environ 500 K, et pour l'iode jusqu'à 1 500 K.

Si ω_0 et ω_1 sont les dégénérescences du niveau fondamental et du premier niveau excité, on a suivant que l'on doit prendre en compte ou non le premier niveau excité :

$$q^{(\text{élec})} = \omega_0 e^{-\varepsilon_0/kT} \quad \text{ou} \quad q^{(\text{élec})} = \omega_0 e^{-\varepsilon_0/kT} + \omega_1 e^{-\varepsilon_1/kT}$$

4.1.2 Conséquences sur les propriétés thermodynamiques

Jusqu'à présent, parmi tous les degrés de liberté internes, seul le spin était pris en compte : la fonction $q^{(\text{int})}(T)$ qui valait $(2s+1)$ doit maintenant être remplacée par $q^{(\text{nucl})} \times q^{(\text{élec})}(T)$.

Si seul le niveau fondamental électronique est à prendre en compte :

La fonction de partition canonique $Q(T, V, N)$ n'est que légèrement modifiée :

$$Q(T, V, N) = \frac{1}{N!} \left(\frac{(2s+1)V}{\Lambda^3(T)} \right)^N \rightarrow Q(T, V, N) = \frac{1}{N!} \left(\frac{\omega_n \omega_0 e^{-\varepsilon_0/kT} V}{\Lambda^3(T)} \right)^N$$

Cette fonction de partition canonique permet de calculer les principales propriétés du gaz :

- Les résultats classiques (pV , E et C_V) ne sont pas affectés par la prise en compte de la structure interne de l'atome, mise à part l'ajout d'un terme constant à E (qui correspond à un simple changement d'origine des énergies, celle-ci se trouvant maintenant sur le niveau fondamental nucléaire) :

$$pV = NkT \quad E = \frac{3}{2} NkT + N\varepsilon_0 \quad C_V = \frac{3}{2} Nk$$

- Dans S et μ , il faut remplacer $(2s+1)$ par $\omega_n \omega_0$ et ajouter ε_0 à μ (changement de l'origine des énergies) :

$$S = \frac{5}{2} NkT - Nk \ln \frac{\rho \Lambda^3(T)}{\omega_n \omega_0} \quad \mu = \varepsilon_0 + kT \ln \frac{\rho \Lambda^3(T)}{\omega_n \omega_0} \quad \text{avec} \quad \frac{\rho \Lambda^3(T)}{\omega_n \omega_0} \ll 1$$

Si on doit prendre en compte le premier niveau excité :

On a, en posant $\Delta\varepsilon = \varepsilon_1 - \varepsilon_0$ et $\omega = \omega_1/\omega_0$:

$$Q(T, V, N) = \frac{1}{N!} \left(\frac{\omega_n \omega_0 e^{-\varepsilon_0/kT} V}{\Lambda^3(T)} (1 + \omega e^{-\Delta\varepsilon/kT}) \right)^N$$

On trouve que, par rapport aux résultats précédents, il faut ajouter :

$$\text{à } E : N\Delta\varepsilon \frac{\omega e^{-\Delta\varepsilon/kT}}{1 + \omega e^{-\Delta\varepsilon/kT}} \quad \text{à } S : Nk \ln(1 + \omega e^{-\Delta\varepsilon/kT}) + Nk \frac{\Delta\varepsilon}{kT} \frac{\omega e^{-\Delta\varepsilon/kT}}{1 + \omega e^{-\Delta\varepsilon/kT}}$$

$$\text{à } C_V : Nk \left(\frac{\Delta\varepsilon}{kT} \right)^2 \frac{\omega e^{-\Delta\varepsilon/kT}}{(1 + \omega e^{-\Delta\varepsilon/kT})^2} \quad \text{à } \mu : -kT \ln(1 + \omega e^{-\Delta\varepsilon/kT})$$

4.1.3 Gaz adsorbé sur une surface

On considère la surface plane d'un solide constitué d'atomes Y en contact avec un gaz contenant des atomes X. À l'échelle microscopique, la surface est un réseau plan ordonné d'atomes Y. Lorsque des atomes X d'un gaz se déposent sur cette surface, on parle d'**adsorption**. On dit que les atomes X sont **adsorbés** sur la surface.

Le système thermodynamique choisi est constitué des atomes X adsorbés. On donne ci-dessous les caractéristiques de ce système et les hypothèses de travail.

- **c'est un système bidimensionnel.** Si on note A l'aire de la surface et φ la tension de surface, l'équation de Gibbs s'écrit : $dE = T dS - \varphi dA + \mu dN$.
- **c'est un système ouvert.** Il est en contact avec le gaz qui fait office de thermostat et de réservoir de particules : il fixe la température et le potentiel chimique. Ceci sous-entend que l'enceinte a des dimensions telles que les propriétés du gaz sont indépendantes du nombre d'atomes adsorbés. On supposera que le gaz est parfait, classique quelle que soit la température, et uniquement composé d'atomes X. La condition d'équilibre entre le système et le gaz s'écrit $\mu_{\text{sys}} = \mu_{\text{gaz}} (= \mu)$:

$$\mu = \varepsilon_0 + kT \ln \frac{\rho \Lambda^3(T)}{\omega_n \omega_0} \quad \text{avec} \quad \rho = \frac{N_{\text{gaz}}}{V_{\text{gaz}}} \quad (4.1)$$

Étant donné $p = \rho kT$, p étant la pression du gaz, on constate que " T et μ fixés" signifie également " p et ρ fixés".

- **c'est un système isochore.** Le volume du solide (et donc l'aire de sa surface) réagit à T et p par l'intermédiaire de son coefficient de dilatation thermique isobare et de son module de compressibilité isotherme. En travaillant à température et pression constantes, et compte tenu des deux hypothèses qui suivent, le volume du solide et l'aire A de la surface sont fixés.
- **Les liaisons X–Y sont faibles.** Ce ne sont pas des liaisons chimiques : le phénomène d'adsorption est d'origine purement électrostatique. On parle alors de *physisorption* et non de *chimisorption*. Deux conséquences importantes : les propriétés du solide sont indépendantes du nombre d'atomes adsorbés et les degrés de liberté électroniques et nucléaires des atomes X ne sont pas modifiés après adsorption. On peut donc réécrire la relation 4.1 en plaçant l'origine des énergies sur le niveau d'énergie électronique fondamental ($\varepsilon_0 = 0$) et en travaillant avec une entropie conventionnelle telle que $\omega_n \omega_0 = 1$. Les atomes X deviennent tout simplement des masses ponctuelles sans spin :

$$p = \rho kT \quad \mu = kT \ln (\rho \Lambda^3(T)) = kT \ln \frac{p \Lambda^3(T)}{kT} \quad (4.2)$$

- **Les liaisons X–X sont encore plus faibles.** Aucune interaction entre atomes X adsorbés n'est envisagée. La géométrie de la surface n'a aucune importance.

Le solide joue ainsi le rôle de champ de potentiel $U(x, y, z)$ pour les atomes X adsorbés (on choisira l'axe z normal à la surface). Dans le plan xy , pour une valeur de z donnée, le diagramme d'équipotentiels a l'allure d'une "boîte à œufs" : les sites d'adsorption sont des bassins reliés entre eux par des cols.

Les hautes températures

A haute température, les variations d'énergie potentielle sur la surface sont extrêmement faibles face à l'énergie cinétique des atomes.

C'est le modèle de **l'adsorption mobile**. On considère alors la surface comme parfaitement lisse, et l'énergie potentielle comme indépendante de x et y : chaque atome adsorbé possède deux degrés de liberté de translation. Le degré de liberté suivant l'axe z est vibrationnel : la variation d'énergie potentielle par rapport à z est supposée localement d'allure parabolique. On notera U_0 la valeur de l'énergie potentielle en $z = 0$, et ω_z la pulsation propre de l'oscillateur.

La fonction de partition d'un atome de spin nul s'écrit (A est l'aire de la surface) :

$$Q(T, A, N) = \frac{q(T, A)^N}{N!} \quad \text{avec} \quad q(T, A) = \frac{A}{\Lambda^2(T)} \frac{e^{-U_0/kT}}{2 \sinh u_z} \quad \text{et} \quad u_z = \frac{\hbar\omega_z}{2kT}$$

Par dérivation de $\ln Q$, on obtient le potentiel chimique :

$$\mu = -kT \left. \frac{\partial \ln Q}{\partial N} \right]_{T,A} = U_0 + kT \ln (2\sigma \Lambda^2(T) \sinh u_z) \quad \text{avec} \quad \sigma = \frac{N}{A}$$

que l'on identifie à celui du réservoir (équation d'équilibre 4.2). On en déduit la densité superficielle σ en fonction de T et de p :

$$\mu = kT \ln \frac{p \Lambda^3(T)}{kT} \quad \Rightarrow \quad \sigma(T, p) = \frac{p \Lambda(T) e^{-U_0/kT}}{2kT \sinh u_z}$$

On peut également calculer l'énergie interne et la capacité thermique C_A à aire constante :

$$E = kT^2 \left. \frac{\partial \ln Q}{\partial T} \right]_{A,N} = NkT + NU_0 + N \frac{\hbar\omega_z}{2} \coth u_z$$

$$C_A = \left. \frac{\partial E}{\partial T} \right]_{A,N} = Nk + Nk \left(\frac{u_z}{\sinh u_z} \right)^2$$

Le premier terme est celui d'un gaz parfait bidimensionnel, et le deuxième la contribution des vibrations perpendiculaires à la surface.

Les basses températures

A basse température, les barrières de potentiel sont supposées infranchissables. C'est un modèle **d'adsorption localisée** : chaque atome X est confiné sur un site et ne peut le quitter que par vaporisation. Nous traitons ici le modèle de Langmuir : on suppose que tous les sites sont identiques et que chaque site ne peut accueillir qu'un seul atome (adsorption monocouche). Dans ce modèle, les trois degrés de liberté sont vibrationnels. L'un des trois modes normaux est tout simplement z , et les deux autres correspondent à deux directions particulières de la surface qui dépendent de la géométrie de celle-ci. On y place les deux axes x et y . Chaque atome adsorbé est ainsi réduit à un jeu de trois oscillateurs harmoniques indépendants de pulsations ω_x , ω_y et ω_z . On notera U_0 la valeur de l'énergie potentielle au fond du puits. Les

liaisons X–Y étant supposées très faibles, on fait l'hypothèse que l'aire d'un site est indépendante de son occupation. On posera donc $A = Ma$, M étant le nombre de sites.

La fonction de partition d'un système de N atomes répartis sur M sites s'écrit (voir paragraphe 2.1.4) :

$$Q(M, T, A, N) = \frac{M!}{N!(M-N)!} q(T, a)^N \quad \text{avec}$$

$$q(T, a) = \frac{e^{-U_0/kT}}{8 \sinh u_x \sinh u_y \sinh u_z} \quad \text{et} \quad u_x = \frac{\hbar\omega_x}{2kT} \quad u_y = \frac{\hbar\omega_y}{2kT} \quad u_z = \frac{\hbar\omega_z}{2kT}$$

Le potentiel chimique est trouvé fonction de T et $\theta = \frac{N}{M}$ (*taux de recouvrement*) :

$$\mu = U_0 + kT \ln \frac{\theta}{1-\theta} + kT \ln (8 \sinh u_x \sinh u_y \sinh u_z)$$

La condition d'équilibre 4.2 permet de calculer $\theta(T, p)$:

$$\theta(T, p) = \frac{p \chi(T)}{1 + p \chi(T)} \quad \text{avec} \quad \chi(T) = \frac{\Lambda^3(T) e^{-U_0/kT}}{8 kT \sinh u_x \sinh u_y \sinh u_z}$$

L'énergie interne et la capacité thermique C_A ont pour expressions :

$$E = kT^2 \left. \frac{\partial \ln Q}{\partial T} \right]_{A, N} = NU_0 + N \frac{\hbar\omega_x}{2} \coth u_x + N \frac{\hbar\omega_y}{2} \coth u_y + N \frac{\hbar\omega_z}{2} \coth u_z$$

$$C_A = \left. \frac{\partial E}{\partial T} \right]_{A, N} = Nk \left(\frac{u_x}{\sinh u_x} \right)^2 + Nk \left(\frac{u_y}{\sinh u_y} \right)^2 + Nk \left(\frac{u_z}{\sinh u_z} \right)^2$$

Le calcul de la tension de surface nécessite la loi de variation des fréquences de vibration en fonction de l'aire de la surface. Dans de nombreux modèles, on considère que la fréquence d'un oscillateur linéaire confiné sur un segment de longueur ℓ varie proportionnellement à $1/\ell$ (voir paragraphe 2.5.1). On supposera donc ici que la pulsation ω_z est indépendante de a et que les pulsations ω_x et ω_y varient en $1/\sqrt{a}$:

$$\frac{d\omega_x}{da} = -\frac{\omega_x}{2a} \quad \frac{d\omega_y}{da} = -\frac{\omega_y}{2a} \quad \frac{d\omega_z}{da} = 0$$

$$\text{Ce qui donne avec } \sigma = \frac{\theta}{a} : \varphi = kT \left. \frac{\partial \ln Q}{\partial A} \right]_{T, N} = \sigma \frac{\hbar\omega_x}{4} \coth u_x + \sigma \frac{\hbar\omega_y}{4} \coth u_y$$

Les températures intermédiaires

Les atomes adsorbés ont maintenant la possibilité de changer de site moyennant le passage d'une barrière de potentiel. Pour simplifier, on choisira un réseau carré, avec des axes x et y portés par les directions de diffusion, et pour fixer les idées, on choisit pour l'énergie potentielle un profil sinusoïdal :

$$U(\vec{r}) = U(x, y, z) = U_0 + \frac{m\omega_z^2}{2} z^2 + \frac{U^*}{2} \left(1 - \cos \frac{2\pi x}{\ell} \right) + \frac{U^*}{2} \left(1 - \cos \frac{2\pi y}{\ell} \right) \quad (4.3)$$

Les sites stables (ou puits) sont à l'énergie U_0 et séparés suivant x et y d'une distance ℓ . Les positions de col sont à mi-distance entre deux sites stables (suivant x et y), et les barrières de potentiel à franchir valent toutes $\Delta U^* = U^* - U_0$.

Les pulsations ω_x et ω_y au fond d'un puits sont identiques. On peut remarquer qu'elles varient en $1/\ell$ comme dans le modèle précédent :

$$\omega_x = \sqrt{\frac{1}{m} \frac{\partial^2 U}{\partial x^2}} \Big|_{\vec{r}=\vec{0}} \quad \text{et} \quad \omega_y = \sqrt{\frac{1}{m} \frac{\partial^2 U}{\partial y^2}} \Big|_{\vec{r}=\vec{0}} \quad \Rightarrow \quad \omega_x = \omega_y = \frac{2\pi}{\ell} \sqrt{\frac{U^*}{m}}$$

L'énergie potentielle est à variables séparables. La fonction de partition d'un atome est donc un produit de fonctions de partition :

$$q = e^{-U_0/kT} q_z q_x q_y = \frac{e^{-U_0/kT}}{2 \sinh u_z} q_x q_y \quad \text{avec} \quad u_z = \frac{\hbar \omega_z}{2kT} \quad (4.4)$$

Vue la symétrie du problème, il est clair que les deux fonctions de partition q_x et q_y sont identiques.

Calculons dans un premier temps q_x comme si elle était classique. On utilise pour cela l'équation 3.31 avec ici $n = 1$. En posant $A = L^2$ et sachant $L = \ell\sqrt{M}$:

$$q_x^{\text{class}} = \frac{1}{\Lambda} \int_0^L e^{-U_x/kT} dx = \frac{\sqrt{M}}{\Lambda} \int_0^\ell e^{-U_x/kT} dx \quad \text{avec} \quad U_x = \frac{U^*}{2} \left(1 - \cos \frac{2\pi x}{\ell}\right)$$

Cette fonction de partition tend bien vers L/Λ quand T tend vers l'infini. A basse température, on doit trouver, mais ce n'est malheureusement pas le cas :

$$q_x \longrightarrow q_{\text{oh}} = \frac{1}{2 \sinh u_x} \quad \text{avec} \quad u_x = \frac{\hbar \omega_x}{2kT} = \frac{\pi \hbar}{\ell kT} \sqrt{\frac{U^*}{m}}$$

L'astuce consiste à multiplier q_x^{class} par q_{oh} et, pour que la fonction de partition soit toujours satisfaisante à haute température, à diviser le résultat par la limite haute température de q_{oh} , c'est-à-dire $kT/\hbar \omega_x$. On trouve, en posant $u^* = U^*/2kT$ et $\phi = 2\pi x/\ell$ (I_0 est la fonction de Bessel modifiée de première espèce et d'indice 0) :

$$q_x = \frac{L}{\Lambda} \frac{u_x}{\sinh u_x} e^{-u^*} I_0(u^*) \quad \text{avec} \quad I_0(u^*) = \frac{1}{2\pi} \int_0^{2\pi} e^{u^* \cos \phi} d\phi$$

Ce qui donne finalement en remplaçant $q_x = q_y$ dans 4.4 :

$$q(T, A) = 4\pi \frac{A}{\ell^2} q_{\text{quant}} u^* e^{-2u^*} I_0^2(u^*) \quad \text{avec} \quad q_{\text{quant}} = \frac{e^{-u_0}}{8 \sinh u_x \sinh u_y \sinh u_z}$$

$$u_0 = \frac{U_0}{kT} \quad u_x = u_y = \frac{\pi \hbar}{\ell kT} \sqrt{\frac{U^*}{m}} \quad u_z = \frac{\hbar \omega_z}{2kT} \quad u^* = \frac{U^*}{2kT}$$

Pour calculer la fonction de partition de N atomes sur la surface, il faut tenir compte du fait que les atomes ont accès à tous les sites présents sur la surface. Ils sont donc indiscernables. Si la température est suffisamment élevée, on peut écrire :

$$Q(T, A, N) = \frac{q(T, A)^N}{N!} \quad \text{puis} \quad E = kT^2 \left. \frac{\partial \ln Q}{\partial T} \right]_{A, N} \quad \text{et} \quad C_A = \left. \frac{\partial E}{\partial T} \right]_{A, N}$$

On donne figure 4.1 l'allure des courbes $C_A(T)$ obtenues pour différentes valeurs de la barrière ΔU^* . Les courbes ont été numérotées dans le sens " ΔU^* croissant".

Pourquoi à température suffisamment élevée ? Pour deux raisons qui finalement n'en font qu'une :

- Le fait d'avoir considéré dès le départ un seul atome sous-entend que la densité de la phase adsorbée est suffisamment faible pour négliger la gêne mutuelle entre atomes au cours de leurs déplacements. En dessous d'une certaine température, la phase adsorbée devient trop dense et le modèle n'est plus valable.
- Pour pouvoir écrire $Q = q^N/N!$, il faut que la température soit suffisamment élevée et la densité suffisamment faible.

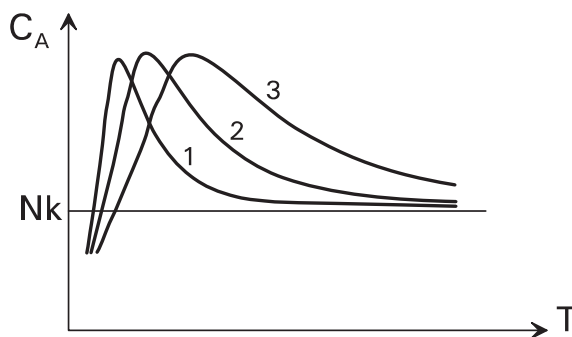


FIG. 4.1 – Translation entravée sur une surface : capacité thermique

4.2 Les degrés de liberté d'une molécule

4.2.1 Indépendance des degrés de liberté de translation

Dans un gaz parfait, chaque molécule se comporte comme si elle était seule dans le récipient. La fonction de partition $q(T, V)$ relative à une molécule est donc toujours indépendante de N . D'après 2.8 au paragraphe 2.3.1, l'équation d'état des gaz parfait reste inchangée. On a toujours à haute température et pour un gaz parfait supposé mécaniquement isolé du monde extérieur :

$$Q(T, V, N) = \frac{q(T, V)^N}{N!} \Rightarrow pV = NkT$$

En mécanique, le *théorème de Kœnig* s'énonce : "*L'énergie cinétique d'un système est égale à l'énergie cinétique qu'aurait la masse totale M du système rapportée en G additionnée de l'énergie cinétique du système dans son mouvement autour de G* ".

Pour une molécule, système constitué de quelques noyaux et d'un grand nombre d'électrons, cela signifie que les trois degrés de liberté de translation de son centre de masse affecté de la masse totale M sont indépendants des degrés de liberté internes de la molécule. La fonction de partition $q(T, V)$ d'une molécule s'écrit donc comme le produit d'une fonction de partition de translation $q^{(\text{tr})}(T, V)$ et d'une fonction de partition interne $q^{(\text{int})}(T)$. Le facteur $(2s + 1)$ fait maintenant partie de cette dernière :

$$q(T, V) = q^{(\text{int})}(T)q^{(\text{tr})}(T, V) \quad \text{avec} \quad q^{(\text{tr})}(T, V) = \frac{V}{\Lambda^3(T)} \quad \text{et} \quad \Lambda(T) = \frac{h}{\sqrt{2\pi M kT}}$$

4.2.2 Séparation des variables par approximations

Soit n est le nombre de noyaux de la molécule et ν le nombre d'électrons. Dans l'expression du hamiltonien H qui suit, les lettres majuscules et les indices i et i' se rapportent aux noyaux, les lettres minuscules et les indices k et k' aux électrons.

L'énergie potentielle, provenant uniquement des interactions entre les particules, n'est fonction que des positions relatives des particules entre elles. C'est la somme de trois termes, chacun regroupant respectivement les interactions noyau-noyau U_{nn} , les interactions électron-électron U_{ee} et les interactions noyau-électron U_{ne} :

$$H = \sum_{i=1}^n \frac{P_i^2}{2M_i} + \sum_{k=1}^{\nu} \frac{p_k^2}{2m} + U_{nn}((\vec{R}_{i'} - \vec{R}_i)_{i \text{ et } i'=1 \text{ à } n}) + U_{ee}((\vec{r}_{k'} - \vec{r}_k)_{k \text{ et } k'=1 \text{ à } \nu}) \\ + U_{ne}((\vec{R}_i - \vec{r}_k)_{i=1 \text{ à } n, k=1 \text{ à } \nu})$$

U_{nn} et U_{ee} sont des interactions répulsives. U_{ne} est la seule interaction de nature attractive : c'est elle qui assure la cohésion de la molécule. On remarquera que l'énergie potentielle est indépendante de l'orientation de la molécule. On dira que, tout comme la translation, la rotation est libre. Translation et rotation sont les deux mouvements laissant toutes les positions relatives des particules constantes.

Mais attention : les degrés de liberté de rotation ne sont pas rigoureusement séparables des autres degrés de liberté internes : l'énergie cinétique de rotation dépend de la géométrie de la molécule, donc des positions instantanées des particules.

Les deux approximations suivantes sont justifiées par le fait que les électrons sont beaucoup plus légers et beaucoup plus rapides que les noyaux. Elles vont permettre d'étudier séparément le mouvement du système de noyaux et le mouvement du nuage électronique.

Pour l'étude des électrons, l'approximation utilisée consiste à considérer les noyaux fixes. La translation et la rotation de la molécule n'ont quasiment aucune influence sur eux (dans *l'approximation adiabatique*, les électrons réagissent instantanément au mouvement des noyaux). Par contre, l'influence des vibrations est plus importante, mais on la néglige dans les calculs.

Dans l'expression de H retenue pour l'étude des électrons, les P_i deviennent nuls, les positions $\vec{R}_i$ deviennent des paramètres : les seules variables restantes sont les positions des électrons et leurs moments associés. On résout l'équation aux valeurs propres de Schrödinger et on trouve, dans les cas les plus simples, un certain nombre de niveaux d'énergie dépendant des différents $R_{ii'} = ||\vec{R}_{i'} - \vec{R}_i||$. Dans les cas plus complexes, on fabrique des potentiels empiriques à partir de données expérimentales (géométrie d'équilibre, énergie de dissociation, fréquences propres de vibration, etc).

Dans une molécule diatomique, les niveaux d'énergie sont fonctions d'une seule distance internucléaire ($R = R_{12}$). Si on trace les niveaux d'énergie $U_0(R)$, $U_1(R)$, $U_2(R)$, etc, on trouve des courbes présentant un minimum en (R_0, ε_0) , (R_1, ε_1) , (R_2, ε_2) , etc (figure 4.2 à gauche). En général, seul le niveau fondamental est accessible aux températures habituelles. Mais il arrive parfois que l'on ait à prendre en compte le premier niveau excité (jamais le deuxième dans la pratique).

Dans une molécule polyatomique, les niveaux d'énergie $U_0(R_{j=1 \text{ à } p})$, $U_1(R_{j=1 \text{ à } p})$, $U_2(R_{j=1 \text{ à } p})$, etc sont fonctions de p distances internucléaires. Pour un niveau donné, on représente les surfaces équipotentielles dans un espace à p dimensions dont les axes

portent les R_j . Dans tous les cas pratiques, seul le niveau fondamental $U_0(R_{j=1 \text{ à } p})$ est à prendre en compte. Celui-ci présente un minimum ε_0 correspondant à la configuration d'équilibre. Le cas le plus simple est celui de la molécule triatomique linéaire : deux distances $R_1 = R_{12}$ et $R_2 = R_{23}$ suffisent à la décrire et un graphe à deux dimensions est possible (figure 4.2 à droite).

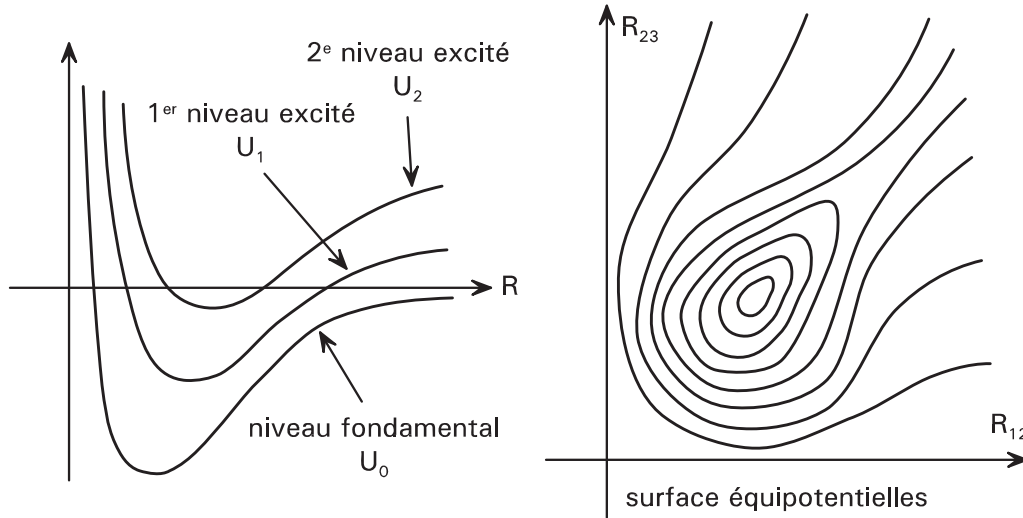


FIG. 4.2 – Niveaux d'énergie, surfaces équipotentielles

Pour l'étude du mouvement des noyaux, l'approximation consiste à remplacer dans H tous les termes faisant intervenir les positions et impulsions des électrons par $U_0(R_{j=1 \text{ à } p})$ ou éventuellement par $U_1(R_{j=1 \text{ à } p})$: les noyaux ne sont pas sensibles aux mouvements instantanés des électrons mais seulement à leur répartition moyenne dans l'espace. Le nuage d'électrons joue ainsi le rôle de champ de potentiel pour les noyaux. Dans H ne subsistent alors que les positions et impulsions des noyaux.

Une fois le problème résolu, on trouve en principe des niveaux d'énergie correspondant aux mouvements des noyaux, c'est-à-dire aux mouvements de rotation de la molécule et de vibration des noyaux au voisinage de leurs positions d'équilibre.

Moyennant ces approximations, et dans les cas où seul le niveau fondamental électronique est accessible, les degrés de liberté électroniques sont séparables des autres degrés de liberté internes (rotation et vibration du système de noyaux) :

$$q^{(\text{int})}(T) = \omega_n q_0^{(\text{élec})} \times q^{(\text{rot,vib})} \quad \text{avec} \quad q_0^{(\text{élec})} = \omega_0 e^{-\varepsilon_0/kT}$$

Ce n'est plus le cas lorsqu'il faut prendre en compte le premier niveau excité électronique. Les niveaux d'énergie correspondant aux mouvements des noyaux vont en principe dépendre du niveau électronique utilisé dans le calcul :

$$q^{(\text{int})}(T) = \omega_n (q_0^{(\text{élec})} \times q_0^{(\text{rot,vib})} + q_1^{(\text{élec})} \times q_1^{(\text{rot,vib})}) \quad \text{avec} \quad q_1^{(\text{élec})} = \omega_1 e^{-\varepsilon_1/kT}$$

Pour rendre les degrés de liberté électroniques séparables, on doit supposer que les états rotovibrationnels sont les mêmes quel que soit l'état quantique du nuage

électronique (autrement dit que les puits de potentiel des courbes de la figure 4.2 à gauche ont localement tous le même profil) :

$$q_1^{(\text{rot,vib})} = q_0^{(\text{rot,vib})} \Rightarrow q^{(\text{int})}(T) = \omega_n (q_0^{(\text{élec})} + q_1^{(\text{élec})}) q_0^{(\text{rot,vib})}$$

Remarque : les trois expressions de $q^{(\text{int})}(T)$ ci-dessus supposent que les degrés de libertés rotationnels sont séparables des degrés de liberté nucléaires. On verra au paragraphe 4.3.5 que ce n'est pas toujours le cas.

4.3 Les gaz parfaits de molécules diatomiques

4.3.1 Réduction du problème à deux corps

La molécule possède dans ce cas deux noyaux : sa géométrie est décrite par une seule distance internucléaire, que l'on notera r . On exprimera donc les coordonnées (x_1, y_1, z_1) et (x_2, y_2, z_2) des deux noyaux en fonction de r , et également en fonction des coordonnées (x, y, z) du centre de masse puisqu'on sait que ces trois variables sont indépendantes des autres. Pour les deux dernières variables qu'il nous manque, il est tout naturel de prendre les deux angles θ et φ caractérisant l'orientation de l'axe de la molécule. Une fois le changement de variables effectué, on trouve :

$$L = \left\{ \frac{1}{2} M (\dot{x}^2 + \dot{y}^2 + \dot{z}^2) \right\} + \left\{ \frac{1}{2} m \dot{r}^2 + \frac{1}{2} m r^2 (\dot{\theta}^2 + \dot{\varphi}^2 \sin^2 \theta) - U_0(r) \right\}$$

$$\text{avec } M = m_1 + m_2 \text{ (masse totale)} \quad \text{et} \quad m = \frac{m_1 m_2}{m_1 + m_2} \text{ (masse réduite)}$$

Le premier terme entre accolades dans L est relatif au mouvement translatatoire du centre de masse affecté de la masse totale. Le deuxième est relatif aux mouvements internes laissant le centre de masse fixe. Ces deux mouvements sont bien entendu indépendants. Le changement de variables effectué est connu sous le nom de "réduction du problème à deux corps" : le problème initial (deux particules en interaction) est ramené à l'étude d'une particule de masse m dans un potentiel central autour du centre de masse.

Le mouvement translatatoire du centre de masse ayant déjà été étudié, on peut maintenant travailler à centre de masse fixe, et écrire :

$$L(r, \theta, \dot{r}, \dot{\theta}, \dot{\varphi}) = \left\{ \frac{1}{2} m \dot{r}^2 - U_0(r) \right\} + \left\{ \frac{1}{2} m r^2 (\dot{\theta}^2 + \dot{\varphi}^2 \sin^2 \theta) \right\}$$

La première accolade est le lagrangien d'un oscillateur dans un champ central $U_0(r)$. La deuxième n'est pas tout-à-fait le lagrangien d'un rotateur libre : il prend bien en compte les variations de θ et φ , mais il dépend aussi de la variable r . Les mouvements de rotation et de vibration ne sont donc pas rigoureusement indépendants. L'approximation du rotateur rigide consiste à considérer r constant lorsqu'on étudie le mouvement de rotation. En posant $r = r_0$ dans le terme correspondant à la rotation, on obtient deux termes indépendants :

$$L(r, \theta, \varphi, \dot{r}, \dot{\theta}, \dot{\varphi}) = L^{(\text{vib})}(r, \dot{r}) + L^{(\text{rot})}(\theta, \dot{\theta}, \dot{\varphi}) \Rightarrow q^{(\text{rot,vib})} = q^{(\text{vib})} \times q^{(\text{rot})}$$

4.3.2 Mouvements de vibration

Développons $U_0(r)$ au voisinage de la distance d'équilibre r_0 :

$$U_0(r) = U_0(r_0) + U'_0(r_0)(r - r_0) + \frac{1}{2} U''_0(r_0)(r - r_0)^2 + \frac{1}{6} U'''_0(r_0)(r - r_0)^3 + \dots$$

Le premier terme $U_0(r_0)$ est nul : le fond du puits de potentiel vaut ε_0 et on l'a pris en compte dans la fonction de partition électronique $q^{(\text{elec})}$. On posera donc ici $U_0(r_0) = 0$. Le deuxième terme $U'_0(r_0)(r - r_0)$ est nul également par définition de r_0 (position d'équilibre $\Rightarrow U'_0(r_0) = 0$). Dans l'approximation de l'oscillateur harmonique, on ne conserve que le troisième terme :

$$L^{(\text{vib})}(r, \dot{r}) = \frac{1}{2} m \dot{r}^2 - \frac{m \omega^2}{2} (r - r_0)^2 \quad \text{avec} \quad \omega = \sqrt{\frac{U''_0(r_0)}{m}}$$

En mécanique quantique, les niveaux d'énergie (non dégénérés) d'un oscillateur harmonique sont donnés par $\varepsilon_n = \hbar \omega (n + 1/2)$, avec n entier positif ou nul. Sa fonction de partition a été calculée au paragraphe 2.5.1 (θ_V est la température caractéristique de vibration de la molécule) :

$$q^{(\text{vib})}(T) = \frac{e^{-\theta_V/2T}}{1 - e^{-\theta_V/T}} = \frac{1}{2 \sinh \frac{\theta_V}{2T}} \quad \text{avec} \quad \theta_V = \frac{\hbar \omega}{k}$$

On en déduit, pour un gaz parfait constitué de N molécules diatomiques, les contributions des vibrations aux différentes grandeurs thermodynamiques.

- Énergie libre : $F^{(\text{vib})}(T, N) = -NkT \ln q^{(\text{vib})}(T) = NkT \ln \left(2 \sinh \frac{\theta_V}{2T} \right)$
- Entropie : $S^{(\text{vib})}(T, N) = - \left. \frac{\partial F^{(\text{vib})}}{\partial T} \right]_N = Nk \frac{\theta_V}{2T} \coth \frac{\theta_V}{2T} - Nk \ln \left(2 \sinh \frac{\theta_V}{2T} \right)$
- Potentiel chimique : $\mu^{(\text{vib})}(T) = \left. \frac{\partial F^{(\text{vib})}}{\partial N} \right]_T = kT \ln \left(2 \sinh \frac{\theta_V}{2T} \right)$
- Énergie interne : $E^{(\text{vib})}(T, N) = F^{(\text{vib})} + T S^{(\text{vib})} = Nk \frac{\theta_V}{2} \coth \frac{\theta_V}{2T}$
- Capacité thermique : $C_V^{(\text{vib})}(T, N) = \left. \frac{\partial E^{(\text{vib})}}{\partial T} \right]_N = Nk \left(\frac{\frac{\theta_V}{2T}}{\sinh \frac{\theta_V}{2T}} \right)^2$

La courbe $C_V^{(\text{vib})}(T)$ est représentée figure 4.3. On constate :

- qu'au dessus d'environ $2\theta_V$, la capacité thermique est quasi constante. C'est le domaine dit "classique" où l'approximation $\sinh x \approx x$ est possible. Les principales grandeurs thermodynamiques deviennent :

$$E^{(\text{vib})}(T, N) \approx NkT \quad C_V^{(\text{vib})}(T, N) \approx Nk \quad S^{(\text{vib})}(T, N) \approx Nk + Nk \ln \frac{T}{\theta_V}$$

- qu'en dessous de $0,1\theta_V$, la capacité thermique est très faible. C'est le domaine où il est possible de négliger les populations des états excités. Pour les calculs, il faut tout de même prendre en compte le premier niveau excité. On dira qu'en dessous de $0,1\theta_V$, tous les oscillateurs sont dans leur état fondamental

sauf quelques uns se trouvant dans leur premier état excité. Par une série d'approximations du premier ordre, on trouve :

$$E^{(\text{vib})}(T, N) \approx Nk \frac{\theta_V}{2} + Nk\theta_V e^{-\theta_V/T}$$

$$C_V^{(\text{vib})}(T, N) \approx Nk \frac{\theta_V^2}{T^2} e^{-\theta_V/T} \quad S^{(\text{vib})}(T, N) \approx Nk \left(1 + \frac{\theta_V}{T}\right) e^{-\theta_V/T}$$

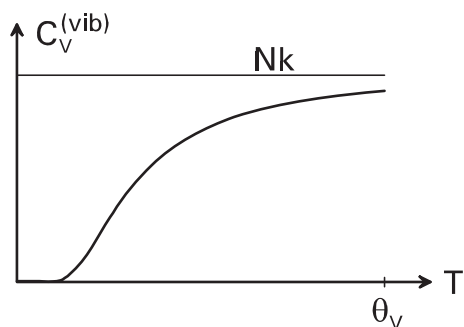


FIG. 4.3 – Capacité thermique : contribution des vibrations

Dans le tableau ci-dessous, on donne quelques valeurs de θ_V et les fractions α de molécules dans un état vibrationnel excité à 300 K et à 1 000 K^[1].

Pour le calcul de $\alpha(T)$, on dit tout simplement que les atomes dans les états excités ne sont pas dans l'état fondamental d'énergie $\hbar \omega/2 = k \theta_V/2$:

$$\alpha(T) = 1 - \frac{e^{-\theta_V/2T}}{q^{(\text{vib})}(T)} = 1 - \frac{1 - e^{-\theta_V/T}}{e^{-\theta_V/2T}} e^{-\theta_V/2T} = e^{-\theta_V/T}$$

Cette fraction $\alpha(T)$ est tout à fait négligeable à 300 K (sauf pour les trois plus lourds), mais l'est beaucoup moins, et même souvent pas du tout, à 1 000 K.

Vues les valeurs élevées de θ_V , le modèle simplifié des hautes températures (traitement classique) est dans tous les cas insuffisant (sauf à la rigueur pour les plus lourds Cl_2 , Br_2 , I_2 à haute température).

Conclusion : **Le traitement quantique est obligatoire.** Dans ce traitement, il arrivera souvent de supposer (soit parce que c'est justifié, soit parce qu'on cherche uniquement un ordre de grandeur) que seul le niveau fondamental est peuplé.

moléc	θ_V	$\alpha_{300 \text{ K}}$	$\alpha_{1\,000 \text{ K}}$	moléc	θ_V	$\alpha_{300 \text{ K}}$	$\alpha_{1\,000 \text{ K}}$
H_2	6210 K	10^{-9}	0,002	N_2	3340 K	10^{-5}	0,035
H D	5380 K	10^{-8}	0,005	C O	3070 K	10^{-5}	0,046
D_2	4390 K	10^{-7}	0,012	N O	2690 K	10^{-4}	0,068
				O_2	2230 K	10^{-4}	0,108
H Cl	4140 K	10^{-6}	0,016	Cl_2	810 K	0,067	0,445
H Br	3700 K	10^{-6}	0,025	Br_2	470 K	0,209	0,625
H I	3200 K	10^{-5}	0,040	I_2	310 K	0,356	0,733

¹ D est le deutérium ^2H

4.3.3 Mouvements de rotation

Le traitement quantique du rotateur rigide libre conduit à des niveaux d'énergie ε_j et à des dégénérescences ω_j donnés par :

$$\varepsilon_j = j(j+1) k \theta_R \quad \text{et} \quad \omega_j = 2j+1 \quad \text{avec} \quad \theta_R = \frac{h^2}{8\pi^2 m r_0^2 k}$$

(θ_R est la température caractéristique de rotation)

Pour calculer la fonction de partition, on somme sur toutes les valeurs de j de 0 à l'infini. Si la molécule est symétrique, le résultat doit être divisé par deux car les deux noyaux constitutifs de la molécule sont alors indiscernables : deux configurations se déduisant l'une de l'autre par permutation des deux noyaux (i.e. par une rotation de π) représentent le même état. On introduit un *facteur de symétrie* σ et on écrit :

$$q^{(\text{rot})}(T) = \frac{1}{\sigma} \sum_{j=1}^{\infty} (2j+1) e^{-j(j+1)\theta_R/T} \quad \text{avec} \quad \begin{array}{ll} \sigma = 2 & (\text{molécule symétrique}) \\ \sigma = 1 & (\text{molécule asymétrique}) \end{array}$$

$$\text{On en déduit : } F^{(\text{rot})}(T, N) = -NkT \ln \Sigma_0 \quad E^{(\text{rot})}(T, N) = Nk\theta_R \frac{\Sigma_1}{\Sigma_0}$$

$$S^{(\text{rot})}(T, N) = Nk \frac{\theta_R}{T} \frac{\Sigma_1}{\Sigma_0} + Nk \ln \Sigma_0 \quad C_V^{(\text{rot})}(T, N) = Nk \frac{\theta_R^2}{T^2} \left(\frac{\Sigma_2}{\Sigma_0} - \frac{\Sigma_1^2}{\Sigma_0^2} \right)$$

$$\text{avec} \quad \Sigma_n = \frac{1}{\sigma} \sum_{j=0}^{\infty} j^n (j+1)^n (2j+1) e^{-j(j+1)\theta_R/T} \quad (n = 0, 1, 2)$$

Ces sommes n'étant pas calculables analytiquement, on a recours à des méthodes numériques. On donne figure 4.4 l'allure de la courbe $C_V^{(\text{rot})}(T, N)$ obtenue.

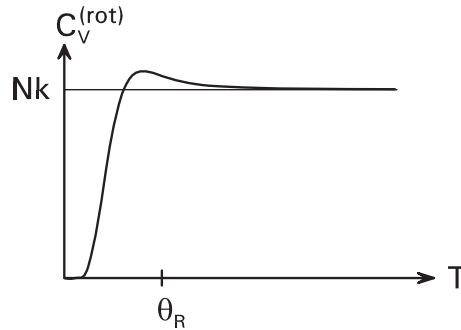


FIG. 4.4 – Capacité thermique : contribution de la rotation

À haute température, l'écart moyen entre deux niveaux d'énergie devient extrêmement faible par rapport à l'énergie moyenne. On peut alors remplacer l'indice j par une variable x continue et les sommes discrètes par des intégrales. On trouve :

$$q^{(\text{rot})}(T) \approx \frac{T}{\sigma \theta_R} \quad \Sigma_0 \approx \frac{1}{\sigma} \frac{T}{\theta_R} \quad \Sigma_1 \approx \frac{1}{\sigma} \frac{T^2}{\theta_R^2} \quad \Sigma_2 \approx \frac{2}{\sigma} \frac{T^3}{\theta_R^3}$$

$$E^{(\text{rot})}(T, N) \approx NkT \quad C_V^{(\text{rot})}(T, N) \approx Nk \quad S^{(\text{rot})}(T, N) \approx Nk + Nk \ln \frac{T}{\sigma \theta_R}$$

Remarque : en utilisant la définition de θ_R , $q^{(\text{rot})}(T)$ s'écrit :

$$q^{(\text{rot})}(T) \approx \frac{2\pi m kT}{h^2} \frac{4\pi r_0^2}{\sigma} = \frac{A}{\Lambda^2} \quad \text{avec } A = \frac{4\pi r_0^2}{\sigma}$$

C'est la fonction de partition d'une particule de masse m se déplaçant librement sur une sphère ou demi-sphère de rayon r_0 . Les résultats trouvés sont ceux d'un gaz parfait bidimensionnel. L'analogie avec le gaz parfait bidimensionnel permet d'estimer les vitesses moyennes de rotation. Les relations 3.7 donnent :

$$\bar{v} = \sqrt{\frac{\pi kT}{2m}} \Rightarrow \bar{\omega} = \frac{\bar{v}}{2\pi r_0} = \sqrt{\frac{kT}{8\pi m r_0^2}} = \frac{k}{h} \sqrt{\pi \theta_R T}$$

Dans le tableau ci-dessous, ces vitesses ont été calculées à 300 K ($\bar{v}$ en m/s et $\bar{\omega}$ en 10^{12} tours par seconde). La masse m utilisée est la masse réduite de la molécule.

moléc	θ_R	$\bar{v}$	$\bar{\omega}$	moléc	θ_R	$\bar{v}$	$\bar{\omega}$
H ₂	88,60 K	2798	6,02	N ₂	2,89 K	761	1,09
H D	66,45 K	2427	5,22	C O	2,78 K	750	1,07
D ₂	44,30 K	2009	4,26	N O	2,46 K	720	1,00
				O ₂	2,09 K	703	0,93
H Cl	15,30 K	2003	2,50	Cl ₂	0,172 K	336	0,27
H Br	12,30 K	1993	2,24	Br ₂	0,058 K	214	0,15
H I	9,50 K	2531	1,97	I ₂	0,027 K	183	0,11

Au vu de la courbe $C_V^{(\text{rot})}(T, N)$, le traitement classique est correct pour les températures supérieures à $2\theta_R$. Étant données les valeurs de θ_R ci-dessus, **le traitement classique est suffisant** dans la plupart des cas. Mais attention à l'hydrogène : le modèle classique est tout juste valable à l'ambiante pour H₂ ; pour les autres molécules hydrogénées, le traitement quantique n'est nécessaire qu'à très basse température.

4.3.4 Propriétés thermodynamiques

Le tableau ci-dessous rassemble les différents résultats relatifs aux molécules diatomiques dans les conditions les plus courantes :

- traitement classique des mouvements de translation et de rotation ;
- traitement quantique des mouvements de vibration ;
- pour les électrons, prise en compte uniquement du niveau fondamental (origine des énergies sur l'état fondamental nucléaire : $\varepsilon_n = 0$).

Remarques :

- Comme $pV = NkT$, le potentiel chimique s'écrit toujours :

$$\mu = \mu_T^0 + kT \ln \frac{p}{p_0} \quad \text{avec ici} \quad \mu_T^0 = \varepsilon_0 + kT \ln \left(\frac{2p_0 \sigma \theta_R \Lambda^3(T)}{\omega_n \omega_0 k T^2} \sinh \frac{\theta_V}{2T} \right)$$

- Pour certaines molécules (N O et O₂ par exemple), il faut prendre en compte le premier niveau électronique excité, et la dernière colonne du tableau est modifiée.

Il faut alors multiplier Q par $(1 + \omega e^{-\Delta\varepsilon/kT})^N$ et ajouter (on a posé $\Delta\varepsilon = \varepsilon_1 - \varepsilon_0$ et $\omega = \omega_1/\omega_0$) :

$$\begin{aligned}
 -kT \ln(1 + \omega e^{-\Delta\varepsilon/kT}) &\rightarrow \frac{F}{N} & \Delta\varepsilon \frac{\omega e^{-\Delta\varepsilon/kT}}{1 + \omega e^{-\Delta\varepsilon/kT}} &\rightarrow \frac{E}{N} \\
 -kT \ln(1 + \omega e^{-\Delta\varepsilon/kT}) &\rightarrow \mu & \frac{\Delta\varepsilon}{kT} \frac{\omega e^{-\Delta\varepsilon/kT}}{(1 + \omega e^{-\Delta\varepsilon/kT})^2} &\rightarrow \frac{C_V}{Nk} \\
 \ln(1 + \omega e^{-\Delta\varepsilon/kT}) + \frac{\Delta\varepsilon}{kT} \frac{\omega e^{-\Delta\varepsilon/kT}}{1 + \omega e^{-\Delta\varepsilon/kT}} &\rightarrow \frac{S}{Nk}
 \end{aligned}$$

		translation		rotation		vibrations ($u_v = \frac{\theta_v}{2T}$)		électrons et noyaux
Q	=	$\frac{1}{N!} \left(\frac{V}{\Lambda^3}\right)^N$	×	$\left(\frac{T}{\sigma\theta_R}\right)^N$	×	$\frac{1}{(2 \sinh u_v)^N}$	×	$(\omega_n \omega_0 e^{-\varepsilon_0/kT})^N$
$\frac{F}{N}$	=	$-kT + kT \ln \frac{N\Lambda^3}{V}$	+	$kT \ln \frac{\sigma\theta_R}{T}$	+	$kT \ln (2 \sinh u_v)$	+	$\varepsilon_0 - kT \ln(\omega_n \omega_0)$
$\frac{E}{N}$	=	$\frac{3}{2}kT$	+	kT	+	$k \frac{\theta_v}{2} \coth u_v$	+	ε_0
μ	=	$kT \ln \frac{N\Lambda^3}{V}$	+	$kT \ln \frac{\sigma\theta_R}{T}$	+	$kT \ln (2 \sinh u_v)$	+	$\varepsilon_0 - kT \ln(\omega_n \omega_0)$
$\frac{C_V}{Nk}$	=	$\frac{3}{2}$	+	1	+	$\left(\frac{u_v}{\sinh u_v}\right)^2$	+	0
$\frac{S}{Nk}$	=	$\frac{5}{2} + \ln \frac{N\Lambda^3}{V}$	+	$1 + \ln \frac{T}{\sigma\theta_R}$	+	$u_v \coth u_v - \ln (2 \sinh u_v)$	+	$\ln(\omega_n \omega_0)$

4.3.5 Molécules homonucléaires : effets quantiques

Le raisonnement est basé sur l'étude de la symétrie de la fonction d'onde de la molécule, et nous allons montrer que, aux températures inférieures à θ_R , les degrés de liberté nucléaires et rotationnels ne sont plus séparables. Pour H_2 et D_2 , ces effets d'origine purement quantique sont détectables à des températures faciles à atteindre (sous azote liquide par exemple). C'est pourquoi nous les avons choisis pour illustrer ce paragraphe.^[2]

² Des restrictions similaires mais plus complexes apparaissent également dans les molécules polyatomiques, particulièrement dans les composés isotopiques CH_4 , CDH_3 , CD_2H_2 , CD_3H , CD_4 .

On supposera que les conditions expérimentales (température et densité) sont telles que le mouvement de translation peut être traité classiquement, et telles que les molécules peuvent être considérées comme indépendantes.

Une molécule H_2 ou D_2 est constituée de deux noyaux identiques et de deux électrons. Chaque noyau est constitué d'un proton pour H_2 , et d'un proton et d'un neutron pour D_2 . Protons, neutrons et électrons sont des fermions. La fonction d'onde décrivant l'état d'une molécule est donc nécessairement antisymétrique vis-à-vis de la permutation des deux électrons, vis-à-vis de la permutation des deux protons, et éventuellement vis-à-vis de la permutation des deux neutrons.

C'est la symétrie vis-à-vis de la permutation des deux noyaux qui nous intéresse. C'est-à-dire, pour la molécule D_2 , la permutation simultanée des deux protons et des deux neutrons. On retiendra que :

- la fonction d'onde décrivant l'état de la molécule D_2 doit être symétrique vis-à-vis de la permutation des deux noyaux, car ils possèdent chacun un nombre pair de nucléons (ce sont deux bosons) ;
- la fonction d'onde décrivant l'état de la molécule H_2 doit être antisymétrique vis-à-vis de la permutation des deux noyaux, car ceux-ci possèdent chacun un nombre impair de nucléons (ce sont deux fermions).

Parmi tous les mouvements possibles de la molécule (translation, vibration, rotation), seule la rotation autorise la permutation des deux noyaux. Lorsqu'on écrit la fonction d'onde comme un produit $\Psi = \Psi^{(\text{transl})} \times \Psi^{(\text{vib})} \times \Psi^{(\text{élec})} \times \Psi^{(\text{rot,nucl})}$, on suppose que le produit $\Psi^{(\text{transl})} \times \Psi^{(\text{vib})} \times \Psi^{(\text{élec})}$, puisque non affectées par la permutation des deux noyaux, est symétrique vis-à-vis de cette permutation.

La fonction d'onde totale Ψ a donc la symétrie de $\Psi^{(\text{rot,nucl})}$; $\Psi^{(\text{rot,nucl})}$ doit donc être symétrique pour D_2 , antisymétrique pour H_2 vis-à-vis de la permutation. Plus généralement, $\Psi^{(\text{rot,nucl})}$ doit être symétrique si les noyaux sont des bosons (ici D^+), antisymétrique si ce sont des fermions (ici H^+).

Conclusion :

On peut toujours écrire $\Psi^{(\text{rot,nucl})} = \Psi^{(\text{nucl})} \times \Psi^{(\text{rot})}$ puis $q^{(\text{rot,nucl})} = q^{(\text{nucl})} \times q^{(\text{rot})}$, mais :

- si les noyaux sont des bosons (ici D^+), les fonctions d'onde $\Psi^{(\text{nucl})}$ et $\Psi^{(\text{rot})}$ doivent être, soit toutes deux symétriques, soit toutes deux antisymétriques vis-à-vis de la permutation des deux noyaux ; on doit donc écrire :

$$q_{D_2}^{(\text{rot,nucl})} = q_{D_2}^{(\text{nucl,sym})} \times q_{D_2}^{(\text{rot,sym})} + q_{D_2}^{(\text{nucl,antisym})} \times q_{D_2}^{(\text{rot,antisym})}$$

- si les noyaux sont des fermions (ici H^+), les fonctions $\Psi^{(\text{nucl})}$ et $\Psi^{(\text{rot})}$ doivent être, l'une symétrique, l'autre antisymétrique vis-à-vis de la permutation des deux noyaux, et donc :

$$q_{H_2}^{(\text{rot,nucl})} = q_{H_2}^{(\text{nucl,antisym})} \times q_{H_2}^{(\text{rot,sym})} + q_{H_2}^{(\text{nucl,sym})} \times q_{H_2}^{(\text{rot,antisym})}$$

Pour les noyaux, seuls le niveau fondamental est pris en compte. Si on place l'origine des énergies sur ce niveau, les fonctions de partition nucléaires sont réduites aux nombres de complexions. On supposera que la dégénérescence ω_n du premier niveau nucléaire est uniquement due au spin du noyau considéré.

Le noyau H^+ est un fermion de spin $1/2$ (c'est un proton), alors que le noyau D^+ est un boson de spin 1 (système de deux fermions différents de spin $1/2$). La

dégénérescence $\omega_n = 2s_n + 1$ est donc égale à 2 pour le noyau H^+ , et égale à 3 pour le noyau D^+ .

Le niveau fondamental du système constitué de deux noyaux identiques est ω_n^2 fois dégénéré. On sait (voir Ω_B et Ω_F au paragraphe 2.2.2 avec $N = 2$) que, parmi les ω_n^2 états, $\omega_n(\omega_n + 1)/2$ sont symétriques, et $\omega_n(\omega_n - 1)/2$ sont antisymétriques (pour D_2 : respectivement 6 et 3, pour H_2 : respectivement 3 et 1). On a donc :

$$q_{D_2}^{(\text{rot,nucl})} = 6 q_{D_2}^{(\text{rot,sym})} + 3 q_{D_2}^{(\text{rot,antisym})} \quad q_{H_2}^{(\text{rot,nucl})} = q_{H_2}^{(\text{rot,sym})} + 3 q_{H_2}^{(\text{rot,antisym})}$$

Pour la rotation, il faut aller un peu plus loin dans la théorie quantique du rotateur rigide.

On l'a vu, les niveaux d'énergie sont de la forme $\varepsilon_j = j(j+1) k\theta_R$ avec $\omega_j = 2j+1$. Ce qu'il faut savoir en plus ici, c'est que chaque fonction d'onde est symétrique si j est pair, et antisymétrique si j est impair. Ce qui donne :

$$q_{D_2}^{(\text{rot,nucl})} = 6 \sum_{j \text{ pair}} (2j+1) e^{-j(j+1)\theta_R^{D_2}/T} + 3 \sum_{j \text{ impair}} (2j+1) e^{-j(j+1)\theta_R^{D_2}/T}$$

$$q_{H_2}^{(\text{rot,nucl})} = \sum_{j \text{ pair}} (2j+1) e^{-j(j+1)\theta_R^{H_2}/T} + 3 \sum_{j \text{ impair}} (2j+1) e^{-j(j+1)\theta_R^{H_2}/T}$$

Le *para-hydrogène*, noté p- H_2 , est une molécule H_2 dans un état rotationnel pair. L'*ortho-hydrogène*, noté o- H_2 , est une molécule H_2 dans un état rotationnel impair.

À température infinie, tous les états j sont équiprobables : il y a trois fois plus d'ortho-hydrogène que de para-hydrogène.

À $T = 0$ K, l'état fondamental fait partie de la première somme, il n'y a plus en théorie (en pratique, voir ci-dessous) que du para-hydrogène.

Entre 10 K et 300 K, le mouvement de translation peut être traité classiquement, et les degrés de liberté vibrationnels et électroniques ne sont pas excités. Ce sont les conditions idéales : la courbe $C_V(T)$ fera apparaître des variations dues uniquement au mouvement de rotation des molécules. On donne figure 4.5 les résultats des calculs numériques comparés aux résultats expérimentaux.

- La courbe en pointillé est celle trouvée au paragraphe précédent.
- p- H_2 est la courbe qu'on obtiendrait avec du para-hydrogène pur.
- o- H_2 est la courbe qu'on obtiendrait avec de l'ortho-hydrogène pur.
- e- H_2 est la courbe d'équilibre théorique.
- m- H_2 est la courbe qu'on trouve expérimentalement.

La courbe théorique e- H_2 et la courbe expérimentale m- H_2 sont radicalement différentes. Il y a donc contradiction (apparente) entre théorie et expérience.

Denisson trouva l'explication en constatant que la courbe expérimentale correspond parfaitement à celle qu'on calcule en "gelant" la composition para-ortho haute température (25 %, 75 %), c'est-à-dire en supposant que la composition reste constante lorsqu'on abaisse la température. Il en déduisit que la réaction $o \rightarrow p$ est extrêmement lente, et qu'on a affaire à un mélange métastable de composition (25 %, 75 %) quelle que soit T . *Bonhæffer et Harteck* trouvèrent le bon catalyseur de la réaction $o \rightarrow p$ (charbon actif) et ont pu ainsi obtenir une courbe expérimentale en parfait accord avec la courbe théorique e- H_2 .

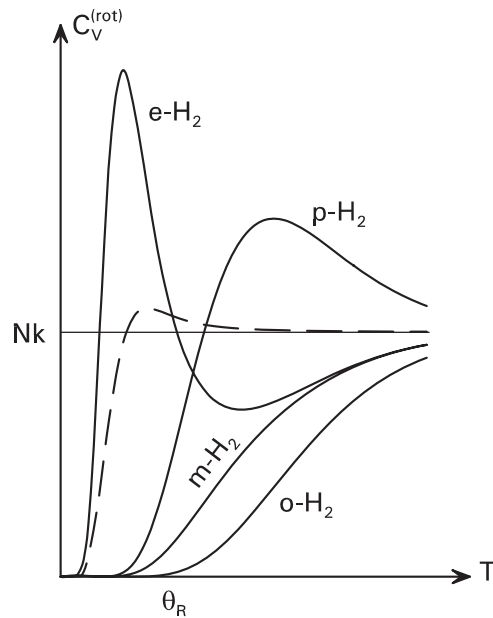


FIG. 4.5 – Capacité thermique : effets quantiques

Cette étude est essentielle pour le stockage de l'hydrogène liquide : la réaction $o \rightarrow p$ étant exothermique, il est impératif que les cuves ne contiennent que du para-hydrogène.

4.4 Les gaz parfaits de molécules polyatomiques

4.4.1 Identification des degrés de liberté

La molécule est maintenant constituée de n noyaux ($n \geq 3$). On a vu au paragraphe 4.2.2 qu'on pouvait étudier le mouvement des noyaux en considérant qu'ils étaient soumis à un champ de potentiel $U_0(R_{k=1 \text{ à } p})$ créé par le cortège électronique, p étant le nombre nécessaire et suffisant de distances internucléaires pour décrire la configuration de la molécule.

Question : quel est le nombre de distances internucléaires nécessaire et suffisant pour décrire la configuration de la molécule ?

On a au total $3n$ degrés de liberté. Pour connaître le nombre p , on retire à $3n$ le nombre de variables nécessaires pour définir la position et l'orientation de la molécule dans l'espace (on sait que U_0 est indépendant de cette position et de cette orientation).

La position de la molécule est définie par les trois coordonnées de son centre de masse. Ce sont les trois degrés de liberté de translation.

Quant à son orientation, elle est définie en général par trois variables angulaires : il faut deux angles pour définir l'orientation d'un axe quelconque de la molécule, et un troisième pour définir l'orientation de la molécule autour de cet axe. On a donc en général trois degrés de liberté de rotation. Ce nombre peut descendre à deux si la molécule est linéaire dans sa configuration d'équilibre. Bilan : $p = 3n - 6$ degrés de liberté de vibration (ou $3n - 5$ si la molécule est linéaire).

4.4.2 Degrés de liberté de rotation

On se place dans l'approximation du rotateur rigide libre, c'est-à-dire que les distances internucléaires sont supposées fixées à leurs valeurs d'équilibre.

Si la molécule est linéaire (CO_2 , C_2H_2 , N_2O , ...), la démarche est identique à celle utilisée pour la molécule diatomique. On remplace simplement mr_0^2 par le moment d'inertie I de la molécule par rapport à son centre de masse G. Les molécules polyatomiques étant plus lourdes, *le modèle classique est ici toujours suffisant* :

$$I = \sum_{i=1}^n m_i \rho_i^2 \quad \text{et} \quad \theta_R = \frac{h^2}{8\pi^2 I k}$$

$$q^{(\text{rot})}(T) = \frac{T}{\sigma \theta_R} \quad \text{avec} \quad \begin{cases} \sigma = 2 & (\text{molécule symétrique}) \\ \sigma = 1 & (\text{molécule asymétrique}) \end{cases}$$

Si la molécule n'est pas linéaire, le calcul de ses moments d'inertie principaux, notés (I_A, I_B, I_C) , est nécessaire pour pouvoir remplacer le mouvement de rotation à trois degrés de liberté par trois mouvements de rotation indépendants à un seul degré de liberté chacun. Le problème quantique est en général trop compliqué pour pouvoir calculer les niveaux d'énergie de rotation, d'autant plus que l'étude des symétries est encore plus complexe que celle exposée au paragraphe 4.3.5.

Heureusement, mis à part CH_4 en dessous de 80 K, le traitement classique est toujours suffisant. Il conduit à :

$$q^{(\text{rot})}(T) = \frac{\sqrt{\pi}}{\sigma} \sqrt{\frac{T}{\theta_A}} \sqrt{\frac{T}{\theta_B}} \sqrt{\frac{T}{\theta_C}}$$

$$\text{avec} \quad \theta_A = \frac{h^2}{8\pi^2 I_A k} \quad \theta_B = \frac{h^2}{8\pi^2 I_B k} \quad \theta_C = \frac{h^2}{8\pi^2 I_C k}$$

Le facteur de symétrie σ tient compte de l'indiscernabilité des noyaux identiques. Pour calculer σ , on compte le nombre d'orientations de la molécule conduisant, indiscernablement, à la même configuration.

Exemples : $\sigma = 12$ pour CH_4 (tétraèdre régulier), $\sigma = 12$ pour C_6H_6 (hexagone régulier), $\sigma = 4$ pour C_2H_4 (molécule plane ayant toutes les symétries du rectangle), $\sigma = 3$ pour NH_3 (pyramide aplatie à base triangulaire équilatérale), $\sigma = 2$ pour H_2O (molécule plane ne possédant qu'un seul axe de symétrie).

4.4.3 Degrés de liberté de vibration

Pour résoudre le problème mécanique, le raisonnement est tout à fait similaire à celui utilisé au paragraphe 2.5.3 pour décrire les vibrations dans un cristal.

On écrit l'énergie potentielle et les équations du mouvement en fonction des déplacements des noyaux :

$$U = -\frac{1}{2} \sum_{\alpha=1}^n \sum_{\beta=1}^n \sum_{i=1}^3 \sum_{j=1}^3 K_{\alpha i}^{\beta j} x_{\alpha i} x_{\beta j} \quad \text{avec} \quad K_{\alpha i}^{\beta j} = - \left. \frac{\partial^2 U}{\partial x_{\alpha i} \partial x_{\beta j}} \right|_{\text{équil}} \quad (4.5)$$

$$m_{\alpha} \ddot{x}_{\alpha i} = F_{\alpha i} = - \frac{\partial U}{\partial x_{\alpha i}} = \sum_{\beta=1}^n \sum_{j=1}^3 K_{\alpha i}^{\beta j} x_{\beta j} \quad (\alpha = 1 \text{ à } n \text{ et } i = 1 \text{ à } 3) \quad (4.6)$$

Il est également possible, en utilisant l'invariance par translation, de considérer le système comme un ensemble de masses ponctuelles liées entre elles par des ressorts :

$$U = \frac{1}{4} \sum_{\alpha=1}^n \sum_{\beta \neq \alpha} C_{\alpha}^{\beta} \delta \ell_{\alpha\beta}^2 = \frac{1}{4} \sum_{\alpha=1}^n \sum_{\beta \neq \alpha} C_{\alpha}^{\beta} \left(\vec{e}_{\alpha}^{\beta} \cdot (\vec{r}_{\alpha} - \vec{r}_{\beta}) \right)^2 \quad (4.7)$$

On peut remarquer que l'énergie potentielle n'est fonction que des distances relatives des noyaux entre eux. Il n'est donc pas nécessaire de rechercher d'éventuelles conditions d'invariance par rotation.

Ce n'est plus le cas si la molécule est plane ou linéaire à l'équilibre : l'énergie potentielle 4.7 ne rend pas compte des déplacements perpendiculaires au plan ou la ligne de la molécule. Ces déplacements peuvent être de type translation ou rotation mais il peut y avoir également des modes de distortion. Pour rendre compte de ces distortions, on doit ajouter des termes de type $K_{\alpha i}^{\beta j} x_{\alpha i} x_{\beta j}$ (avec $x_{\alpha i}$ et $x_{\beta j}$ perpendiculaires au plan ou à la ligne de la molécule) et leur appliquer les conditions d'invariance par translation et rotation.

Les solutions particulières recherchées ici sont de la forme $x_{\alpha i} = u_{\alpha i} e^{-i\omega t}$. On remplace dans 4.6 et on obtient un système d'équations linéaires de la forme :

$$\omega^2 u_{\alpha i} - \sum_{\beta=1}^n \sum_{j=1}^3 \Delta_{\alpha i}^{\beta j} u_{\beta j} = 0 \quad (4.8)$$

En écrivant que le déterminant de ce système doit être nul, on obtient les $3n$ pulsations propres du système. Parmi ces $3n$ valeurs propres, six sont nulles (ou cinq si la molécule est linéaire). Elles correspondent aux déplacements de type translation et rotation. Les $p = 3n - 6$ ou $3n - 5$ non nulles définissent les modes propres de vibration.

Connaissant les p pulsations propres, on peut écrire la fonction de partition $q^{(\text{vib})}(T)$ comme le produit de fonctions de partition d'oscillateurs linéaires :

$$q^{(\text{vib})}(T) = \prod_{i=1}^p \frac{1}{2 \sinh u_{v,i}} \quad \text{avec} \quad u_{v,i} = \frac{\theta_{v,i}}{2T} \quad \text{et} \quad \theta_{v,i} = \frac{\hbar \omega_i}{k}$$

Quelques valeurs numériques :

- CO_2 ($p = 3n - 5 = 4$) : $\theta_{v,i} = 954 \text{ K}, 954 \text{ K}, 1890 \text{ K}, 3360 \text{ K}$;
- N_2O ($p = 3n - 5 = 4$) : $\theta_{v,i} = 850 \text{ K}, 850 \text{ K}, 1840 \text{ K}, 3200 \text{ K}$;
- BF_3 ($p = 3n - 6 = 6$) : $\theta_{v,i} = 631 \text{ K}, 631 \text{ K}, 995 \text{ K}, 1270 \text{ K}, 2070 \text{ K}, 2070 \text{ K}$.

4.4.4 Propriétés thermodynamiques

On donne les résultats relatifs aux conditions les plus courantes :

- traitement classique pour la translation et la rotation ;
- traitement quantique des mouvements de vibration ;
- prise en compte uniquement du niveau fondamental électronique.

Par rapport au tableau relatif aux molécules diatomiques (paragraphe 4.3.4), il faut modifier la colonne "vibration" (avec $p = 3n - 5$ ou $3n - 6$) :

$$\frac{F^{(\text{vib})}}{N} = kT \sum_{i=1}^p \ln (2 \sinh u_{v,i}) \quad \frac{E^{(\text{vib})}}{N} = k \sum_{i=1}^p \frac{\theta_{v,i}}{2} \coth u_{v,i}$$

$$\mu^{(\text{vib})} = kT \sum_{i=1}^p \ln (2 \sinh u_{v,i}) \quad \frac{C_V^{(\text{vib})}}{Nk} = \sum_{i=1}^p \left(\frac{u_{v,i}}{\sinh u_{v,i}} \right)^2$$

$$\frac{S^{(\text{vib})}}{Nk} = \sum_{i=1}^p \left(u_{v,i} \coth u_{v,i} - \ln (2 \sinh u_{v,i}) \right)$$

Pour les molécules non linéaires, on modifie également la colonne "rotation" :

$$\frac{F^{(\text{rot})}}{N} = kT \ln \left(\frac{\sigma}{\sqrt{\pi}} \sqrt{\frac{\theta_A}{T}} \sqrt{\frac{\theta_B}{T}} \sqrt{\frac{\theta_C}{T}} \right) \quad \frac{E^{(\text{rot})}}{N} = \frac{3}{2} kT \quad \frac{C_V^{(\text{rot})}}{Nk} = \frac{3}{2}$$

$$\mu^{(\text{rot})} = kT \ln \left(\frac{\sigma}{\sqrt{\pi}} \sqrt{\frac{\theta_A}{T}} \sqrt{\frac{\theta_B}{T}} \sqrt{\frac{\theta_C}{T}} \right) \quad \frac{S^{(\text{rot})}}{Nk} = \frac{3}{2} + \ln \left(\frac{\sigma}{\sqrt{\pi}} \sqrt{\frac{\theta_A}{T}} \sqrt{\frac{\theta_B}{T}} \sqrt{\frac{\theta_C}{T}} \right)$$

Remarque : certaines molécules (essentiellement hydrocarbures) possèdent un ou plusieurs degrés de liberté appelés mode de torsion. Dans le cas de l'éthane $\text{H}_3\text{C}-\text{CH}_3$ par exemple, les deux groupes méthyles peuvent tourner l'un par rapport à l'autre, sans déformation, autour de la liaison C-C. La courbe $C_V(T)$ montre que ce mouvement correspond à une oscillation à basse température de faible fréquence ($\theta_V \approx 350 \text{ K}$), et à une rotation interne libre à haute température. Aux températures intermédiaires, ce n'est ni une vibration ni une rotation libre : c'est toujours une rotation, mais elle est entravée par une énergie potentielle fonction d'une coordonnée angulaire décrivant l'orientation d'un groupe méthyle par rapport à l'autre. La rotation nécessite le passage de barrières de potentiels ($\approx 22 \text{ zJ}$) pour passer d'une configuration stable à une autre.

4.5 Équilibres chimiques entre gaz parfaits

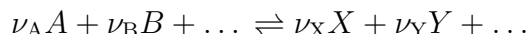
4.5.1 Thermodynamique des équilibres chimiques

Soit une réaction chimique : $\nu_A A + \nu_B B + \dots \longrightarrow \nu_X X + \nu_Y Y + \dots$

Les molécules A, B, ... sont les réactifs, et les molécules X, Y, ... sont les produits.

On dira que la réaction est totale si elle évolue jusqu'à disparition complète de l'un des réactifs, si on attend suffisamment longtemps. Si la composition du système correspond à la stœchiométrie, il y a en fin de réaction disparition complète de tous les réactifs.

Il existe des réactions qui, même après un temps très long, ne conduisent pas à la disparition complète des réactifs. Il s'établit un équilibre dont la composition en réactifs et produits est entièrement déterminée par les conditions d'expérience (la température et la pression). C'est ce qu'on appelle un **équilibre chimique** et on le note :



La composition d'équilibre est indépendante de la composition initiale.

Conséquence : si on modifie les conditions expérimentales, le système trouve une nouvelle composition d'équilibre ; si on revient aux anciennes conditions expérimentales, le système revient à l'ancien équilibre.

Soit par exemple un système constitué de M unités formulaires $\nu_A A + \nu_B B + \dots$. Pour ce système, *l'avancement de la réaction* est nul. Supposons qu'à un instant donné, m unités $\nu_A A + \nu_B B + \dots$ ce sont transformées en m unités $\nu_X X + \nu_Y Y + \dots$.

Pour des conditions expérimentales données, l'état du système, si la stœchiométrie est respectée, est défini entièrement par M et m .

À température et pression constantes par exemple, on peut écrire :

$$N_A = \nu_A(M - m) \quad N_B = \nu_B(M - m) \quad \dots \quad N_X = \nu_X m \quad N_Y = \nu_Y m \quad \dots$$

Et donc :

$$\begin{aligned} dG &= -S dT + V dp + \mu_A dN_A + \mu_B dN_B + \dots + \mu_X dN_X + \mu_Y dN_Y + \dots \\ &= -S dT + V dp + (\mu_A \nu_A + \mu_B \nu_B + \dots) dM \\ &\quad + (\mu_X \nu_X + \mu_Y \nu_Y + \dots - \mu_A \nu_A - \mu_B \nu_B - \dots) dm \end{aligned}$$

Lors d'un processus infinitésimal isotherme ($\delta T = 0$), isobare ($\delta p = 0$) dans une enceinte fermée ($\delta M = 0$), le deuxième principe dit que, à l'équilibre, l'enthalpie libre du système est minimale vis-à-vis des modifications infinitésimales de la composition. On doit donc avoir :

$$\delta G = (\mu_X \nu_X + \mu_Y \nu_Y + \dots - \mu_A \nu_A - \mu_B \nu_B - \dots) \delta m = 0$$

On suppose que durant ce processus infinitésimal, les potentiels chimiques n'ont pas varié. D'où la **condition d'équilibre** :

$$\Delta\mu = 0 \quad \text{avec} \quad \Delta\mu = \mu_X \nu_X + \mu_Y \nu_Y + \dots - \mu_A \nu_A - \mu_B \nu_B - \dots$$

Dans un **gaz parfait**, les potentiels chimiques s'écrivent sous la forme :

$$\begin{aligned} \mu_A &= \mu_{T(A)}^0 + kT \ln \frac{p_A}{p_0} & \mu_B &= \mu_{T(B)}^0 + kT \ln \frac{p_B}{p_0} & \dots \\ \mu_X &= \mu_{T(X)}^0 + kT \ln \frac{p_X}{p_0} & \mu_Y &= \mu_{T(Y)}^0 + kT \ln \frac{p_Y}{p_0} & \dots \end{aligned}$$

où $p_A, p_B \dots p_X, p_Y \dots$ sont les pressions partielles relatives à chacun des constituants. Les potentiels chimiques standards $\mu_{T(A)}^0$, etc sont donnés paragraphe 3.3.1.

En remplaçant dans la condition d'équilibre $\Delta\mu = 0$, on trouve :

$$\frac{(p_X/p_0)^{\nu_X} (p_Y/p_0)^{\nu_Y} \dots}{(p_A/p_0)^{\nu_A} (p_B/p_0)^{\nu_B} \dots} = \exp \frac{\nu_A \mu_{T(A)}^0 + \nu_B \mu_{T(B)}^0 + \dots - \nu_X \mu_{T(X)}^0 - \nu_Y \mu_{T(Y)}^0 - \dots}{kT}$$

Le terme de droite est une fonction de T caractéristique de l'équilibre étudié. C'est par définition la **constante d'équilibre** $K_P(T)$.

$$\text{Condition d'équilibre équivalente à } \Delta\mu = 0 : \frac{(p_X/p_0)^{\nu_X} (p_Y/p_0)^{\nu_Y} \dots}{(p_A/p_0)^{\nu_A} (p_B/p_0)^{\nu_B} \dots} = K_P(T)$$

D'autres constantes d'équilibre peuvent dans certaines conditions être plus pratiques d'emploi. Pour un gaz parfait par exemple, on peut, en utilisant les relations $p_X = \rho_X kT$, etc., travailler avec les concentrations volumiques (ici en molécules par unité de volume). On définit alors la constante d'équilibre :

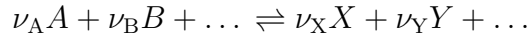
$$K_C(T) = K_P(T) \left(\frac{p_0}{kT} \right)^{\nu_X + \nu_Y + \dots - \nu_A - \nu_B - \dots} \quad \text{Equilibre : } \frac{\rho_X^{\nu_X} \rho_Y^{\nu_Y} \dots}{\rho_A^{\nu_A} \rho_B^{\nu_B} \dots} = K_C(T)$$

Attention : K_P est une grandeur sans dimension mais pas K_C (c'est un volume à la puissance $\nu_A + \nu_B + \dots - \nu_X - \nu_Y$).

4.5.2 Statistique des équilibres chimiques entre gaz parfaits

L'objectif est de calculer les constantes d'équilibre entre gaz parfaits en fonction de leurs propriétés microscopiques.

Soit un gaz parfait constitué de N_A molécules A, N_B molécules B ... en équilibre avec N_X molécules X, N_Y molécules Y ... selon la relation d'équilibre :



Le système est constitué de sous-systèmes macroscopiques, donc discernables. Sa fonction de partition peut s'écrire comme le produit de fonctions de partition :

$$Q(T, V, N_A, N_B \dots N_X, N_Y \dots) = \frac{q_A(T, V)^{N_A}}{N_A!} \frac{q_B(T, V)^{N_B}}{N_B!} \dots \frac{q_X(T, V)^{N_X}}{N_X!} \frac{q_Y(T, V)^{N_Y}}{N_Y!}$$

De l'expression des potentiels chimiques, on exprime N_A en fonction de q_A :

$$\mu_A = -kT \frac{\partial \ln Q}{\partial N_A} = kT \ln \frac{N_A}{q_A} \quad \Rightarrow \quad \rho_A = \frac{N_A}{V} = \frac{q_A}{V} e^{\mu_A/kT}$$

De même pour les autres constituants :

$$\rho_A = \frac{q_A}{V} e^{\mu_A/kT} \quad \rho_B = \frac{q_B}{V} e^{\mu_B/kT} \dots \quad \rho_X = \frac{q_X}{V} e^{\mu_X/kT} \quad \rho_Y = \frac{q_Y}{V} e^{\mu_Y/kT} \dots$$

Que l'on remplace dans $K_C(T)$:

$$K_C(T) = \frac{\rho_X^{\nu_X} \rho_Y^{\nu_Y} \dots}{\rho_A^{\nu_A} \rho_B^{\nu_B} \dots} = \frac{(q_X/V)^{\nu_X} (q_Y/V)^{\nu_Y} \dots}{(q_A/V)^{\nu_A} (q_B/V)^{\nu_B} \dots} e^{\Delta\mu/kT}$$

la condition d'équilibre $\Delta\mu = \mu_X \nu_X + \mu_Y \nu_Y + \dots - \mu_A \nu_A - \mu_B \nu_B - \dots = 0$ donne :

$$K_C(T) = \frac{(q_X/V)^{\nu_X} (q_Y/V)^{\nu_Y} \dots}{(q_A/V)^{\nu_A} (q_B/V)^{\nu_B} \dots}$$

4.5.3 Exemples d'équilibres chimiques

Exemple 1 : Ionisation de l'atome d'hydrogène

On considère ici l'équilibre : $\text{H} \rightleftharpoons \text{H}^+ + \text{e}^-$.

Le but est de calculer le taux d'ionisation, c'est-à-dire la proportion α de protons H^+ ou d'électrons e^- , en fonction de la température. Si partant de N atomes, n sont ionisés, le nombre total de particules vaut $N+n$, dont n sont des électrons, n sont des protons, et $N-n$ sont des atomes. Le taux d'ionisation est par définition $\alpha = n/N$. Les concentrations volumiques s'écrivent :

$$\rho_{\text{H}^+} = \rho_{\text{e}^-} = \frac{n}{V} = \alpha \rho_0 \quad \rho_{\text{H}} = \frac{N-n}{V} = (1-\alpha) \rho_0 \quad \text{avec} \quad \rho_0 = \frac{N}{V}$$

On en déduit $K_{\text{C}}(T) = \frac{\rho_{\text{H}^+} \rho_{\text{e}^-}}{\rho_{\text{H}}} = \rho_0 \frac{\alpha^2}{1-\alpha}$ et donc :

$$\alpha(T, \rho_0) = \frac{K_{\text{C}}(T)}{2\rho_0} \left(\sqrt{1 + \frac{4\rho_0}{K_{\text{C}}(T)}} - 1 \right) \quad \text{avec} \quad K_{\text{C}}(T) = \frac{(q_{\text{H}^+}/V)(q_{\text{e}^-}/V)}{(q_{\text{H}}/V)}$$

Chacune des fonctions de partition apparaissant dans $K_{\text{C}}(T)$ est le produit de deux fonctions de partition : l'une tenant compte des degrés de liberté de translation (V/Λ^3), l'autre tenant compte des degrés de liberté internes. Les degrés de liberté de translation seront traités classiquement, même pour l'électron (l'ionisation n'est détectable qu'au dessus de 10 000 K).

- Pour l'atome H : $\frac{q_{\text{H}}}{V} = 2\omega_p \text{e}^{-\varepsilon_0/kT} \left(\frac{2\pi(m_p + m_e)kT}{h^2} \right)^{3/2}$
- Pour le proton : $\frac{q_{\text{H}^+}}{V} = \omega_p \left(\frac{2\pi m_p kT}{h^2} \right)^{3/2}$
- pour l'électron : $\frac{q_{\text{e}^-}}{V} = 2 \left(\frac{2\pi m_e kT}{h^2} \right)^{3/2}$

m_p est la masse du proton, m_e est la masse de l'électron. Les facteurs 2 sont les contributions du spin de l'électron ($s = 1/2 \Rightarrow \omega_e = 2s + 1 = 2$). La fonction de partition interne du proton est plus complexe, mais elle disparaît dans l'expression de $K_{\text{C}}(T)$. Elle se réduit à une dégénérescence ω_p du niveau fondamental d'énergie nulle (choix d'une origine des énergies) qui seul est à prendre en compte. Pour l'atome H, seul le niveau fondamental électronique ($\varepsilon_0 \approx -2,180 \text{ aJ}$) est pris en compte. Les états excités sont considérés comme des étapes intermédiaires du processus de dissociation.

Ce qui donne finalement : $K_{\text{C}}(T) = \frac{\text{e}^{\varepsilon_0/kT}}{\Lambda^3(T)}$ avec $\Lambda(T) = \frac{h}{\sqrt{2\pi m kT}}$
 m est la masse réduite du système proton-électron : $m = \frac{m_p m_e}{m_p + m_e} \approx m_e$.^[3]

Et pour α : $\alpha(T, \rho_0) = \frac{\sqrt{1 + 4\rho_0 \Lambda^3(T) \text{e}^{-\varepsilon_0/kT}} - 1}{2\rho_0 \Lambda^3(T) \text{e}^{-\varepsilon_0/kT}}$

³ $m_e \approx 0,911 \cdot 10^{-30} \text{ kg} \ll m_p \approx 1,673 \cdot 10^{-27} \text{ kg}$

Comme le gaz est supposé parfait, on a $\rho_0 \Lambda^3(T) \ll 1$ pour tout T . Mais attention, ce n'est pas le cas de $\rho_0 \Lambda^3(T) e^{-\varepsilon_0/kT}$: ce n'est valable qu'à très haute température.

Ci-dessous les fractions $\alpha(T)$ d'atomes ionisés calculées à différentes températures pour une densité $\rho_0 = 0,01$ mol/L. On peut considérer que l'ionisation est quasiment absente en dessous de 10 000 K, et qu'elle est quasiment complète au-dessus de 50 000 K.

T (K)	0	5000	10000	15000	20000	25000	30000	40000	50000	∞
α	0,000	10^{-6}	0,007	0,131	0,472	0,785	0,921	0,984	0,995	1,000

Exemple 2 : Dissociation de la molécule d'hydrogène

On considère maintenant l'équilibre : $\text{H}_2 \rightleftharpoons 2 \text{H}$.

Pour observer ce phénomène, il faut également travailler à haute température, mais pas trop haute de façon à négliger l'ionisation ($T < 10000$ K). Dans cette gamme de température, la rotation de la molécule peut être traitée classiquement ($\theta_R = 85,4$ K), mais la vibration doit être traitée quantiquement ($\theta_V = 6210$ K). Pour les états électroniques, nous ne prendrons en compte que les niveaux fondamentaux de l'atome ($\varepsilon_1 = 0$, $\omega_1 = 2$) et de la molécule ($\varepsilon_2 = -0,7085$ aJ, $\omega_2 = 1$).

Comme dans l'exemple précédent, on considère N molécules dont n sont dissociées : le gaz est donc composé de $N - n$ molécules et de $2n$ atomes, ce qui fait en tout $N + n$ particules.

Le taux de dissociation est défini par $\alpha = \frac{n}{N}$.

Les concentrations volumiques s'écrivent :

$$\rho_H = \frac{2n}{V} = 2\alpha \rho_0 \quad \text{et} \quad \rho_{\text{H}_2} = \frac{N - n}{V} = (1 - \alpha) \rho_0 \quad \text{avec} \quad \rho_0 = \frac{N}{V}$$

Ce qui donne $K_C = \frac{\rho_H^2}{\rho_{\text{H}_2}} = \rho_0 \frac{4\alpha^2}{1 - \alpha}$ et donc :

$$\alpha = \frac{K_C(T)}{8\rho_0} \left(\sqrt{1 + \frac{16\rho_0}{K_C(T)}} - 1 \right) \quad \text{avec} \quad K_C(T) = \frac{(q_H/V)^2}{q_{\text{H}_2}/V}$$

Les fonctions de partition valent :

$$\frac{q_H}{V} = 2 \left(\frac{2\pi m_H kT}{h^2} \right)^{3/2} \quad \text{et} \quad \frac{q_{\text{H}_2}}{V} = \left(\frac{4\pi m_H kT}{h^2} \right)^{3/2} \frac{T}{2\theta_R} \frac{e^{-\varepsilon_2/kT}}{2 \sinh(\frac{\theta_V}{2T})}$$

On trouve numériquement, avec $\rho_0 = 0,01$ mol/L :

$$K_C(T) \approx 952,42 \sqrt{T} (e^{-51691/T} - e^{-57901/T})$$

Ce qui donne pour α les valeurs ci-dessous. La dissociation est quasi inexistante en dessous de 2000 K, et quasi complète au dessus de 10000 K.

T (K)	0	1000	2000	3000	4000	5000	10000	∞
α	0,0000	10^{-9}	0,0014	0,1034	0,6052	0,9297	0,9995	1,0000

Exemple 3 : Équilibre isotopique

On considère pour terminer l'équilibre : $\text{H}_2 + \text{D}_2 \rightleftharpoons 2\text{HD}$.

Pour la plupart des éléments, les effets isotopiques sont très difficiles à mettre en évidence. L'hydrogène est encore une fois particulier : la présence d'un neutron dans le deutéron (le noyau du deutérium D) rend sa masse deux fois plus importante que celle du proton, noyau de l'hydrogène habituel. On notera m la masse commune (à 0,2 % près) du proton et du neutron ($m \approx 1,674 \cdot 10^{-27}$ kg).

Pour l'hydrogène, les effets isotopiques sont détectables à l'ambiante. De façon à ce que ni dissociation, ni ionisation ne soient à envisager, on travaillera à des températures inférieures à 2000 K. De même, de façon à traiter la rotation classiquement, on travaillera à des températures supérieures à 150 K.

Pour le calcul de la constante d'équilibre, on sépare les diverses contributions :

$$K_C(T) = \frac{(q_{\text{HD}}/V)^2}{(q_{\text{HH}}/V)(q_{\text{DD}}/V)} = \frac{(q_{\text{HD}}^{(\text{transl})}/V)^2}{(q_{\text{HH}}^{(\text{transl})}/V)(q_{\text{DD}}^{(\text{transl})}/V)} \times \frac{(q_{\text{HD}}^{(\text{rot})})^2}{q_{\text{HH}}^{(\text{rot})} q_{\text{DD}}^{(\text{rot})}} \times \frac{(q_{\text{HD}}^{(\text{vib})})^2}{q_{\text{HH}}^{(\text{vib})} q_{\text{DD}}^{(\text{vib})}}$$

Les degrés de liberté électroniques et nucléaires ont été omis : ils sont identiques dans les deux membres de l'équation d'équilibre, puisque le niveau électronique fondamental (le seul pris en compte) est le même quelle que soit la molécule (les interactions sont électrostatiques : elles ne dépendent pas de la masse des particules concernées).

Les contributions des degrés de liberté de translation et de rotation sont indépendantes de la température :

Translation :

$$\frac{q_{\text{HH}}^{(\text{transl})}}{V} = \left(\frac{4\pi m kT}{h^2} \right)^2 \quad \frac{q_{\text{HD}}^{(\text{transl})}}{V} = \left(\frac{6\pi m kT}{h^2} \right)^2 \quad \frac{q_{\text{DD}}^{(\text{transl})}}{V} = \left(\frac{8\pi m kT}{h^2} \right)^2$$

Et donc :

$$K_C^{(\text{transl})} = \frac{(q_{\text{HD}}^{(\text{transl})}/V)^2}{(q_{\text{HH}}^{(\text{transl})}/V)(q_{\text{DD}}^{(\text{transl})}/V)} = \frac{27}{16\sqrt{2}} \approx 1,193$$

Rotation (pour les températures caractéristiques, voir paragraphe 4.3.3) :

$$q_{\text{HH}}^{(\text{rot})} = \frac{T}{2\theta_{\text{R}}^{\text{HH}}} = \frac{T}{177,20} \quad q_{\text{HD}}^{(\text{rot})} = \frac{T}{\theta_{\text{R}}^{\text{HD}}} = \frac{T}{66,45} \quad q_{\text{DD}}^{(\text{rot})} = \frac{T}{2\theta_{\text{R}}^{\text{DD}}} = \frac{T}{88,60}$$

Et donc :

$$K_C^{(\text{rot})} = \frac{(q_{\text{HD}}^{(\text{rot})}/V)^2}{q_{\text{HH}}^{(\text{rot})} q_{\text{DD}}^{(\text{rot})}} = \frac{32}{9} \approx 3,556$$

On notera que ces deux contributions sont supérieures à 1 : elles sont favorables à la formation de molécules HD. À noter également que sans les facteurs de symétrie, la contribution des degrés de liberté de rotation serait 4 fois moins importante.

On multiplie maintenant $K_C^{(\text{transl})} \times K_C^{(\text{rot})} = 3\sqrt{2}$ par la contribution des degrés de liberté de vibration (pour les températures caractéristiques, voir paragraphe 4.3.2) :

$$K_C(T) = 3\sqrt{2} \frac{\sinh \frac{\theta_V^{\text{HH}}}{2T} \sinh \frac{\theta_V^{\text{DD}}}{2T}}{\left(\sinh \frac{\theta_V^{\text{HD}}}{2T} \right)^2} \quad (4.9)$$

Sachant que les températures caractéristiques de vibration varient en $1/\sqrt{m}$, m étant la masse réduite de la molécule, on a :

$$\theta_V^{\text{HH}} = \frac{2}{\sqrt{3}} \theta_V^{\text{HD}} \quad \theta_V^{\text{DD}} = \sqrt{\frac{2}{3}} \theta_V^{\text{HD}} \quad \text{et donc} \quad K_C(T) \rightarrow 4 \text{ quand } T \rightarrow \infty$$

Soit $\alpha(T)$ le taux de molécules HD et ξ le taux de deutérium du mélange. On vérifiera que les taux de molécules H_2 et D_2 valent respectivement $1 - \xi - \frac{\alpha(T)}{2}$ et $\xi - \frac{\alpha(T)}{2}$. On a donc :

$$K_C(T) = \frac{\alpha^2(T)}{(1 - \xi - \frac{\alpha(T)}{2})(\xi - \frac{\alpha(T)}{2})} \quad (4.10)$$

Ce qui permet d'exprimer α en fonction de ξ et $K_C(T)$:

$$\alpha(T) = \frac{\sqrt{1 + 2\alpha_{(\text{al})}\Gamma(T)} - 1}{\Gamma(T)} \quad \text{avec} \quad \Gamma(T) = \frac{4 - K_C(T)}{K_C(T)} \quad \text{et} \quad \alpha_{(\text{al})} = 2\xi(1 - \xi)$$

Le facteur $\alpha_{(\text{al})} = 2\xi(1 - \xi)$ est le taux de molécules HD que l'on observerait si les mariages entre atomes se faisaient de façon aléatoire. On constate que c'est la limite de $\alpha(T)$ quand T tend vers l'infini.

On constate également que $\alpha(T)$ est inférieur à $\alpha_{(\text{al})}$ quelle que soit la température. Un exemple de plus qui montre que la nature a une prédilection pour la symétrie.

On peut traiter à titre d'exemple le cas $\xi = 1/2$ ($\alpha_{(\text{al})} = 1/2$). On trouve :

T (K)	0	10	20	50	100	200	500	1000	2000	$+\infty$
% HD	0,00	2,10	12,92	32,18	41,14	45,90	48,80	49,70	49,96	50,00
% H_2	50,00	48,95	43,54	33,91	29,43	27,05	25,60	25,15	25,02	25,00
% D_2	50,00	48,95	43,54	33,91	29,43	27,05	25,60	25,15	25,02	25,00

4.6 Réaction chimique dans un gaz parfait

4.6.1 Taux d'avancement d'une réaction chimique

Soit l'équilibre chimique $\nu_A \text{A} + \nu_B \text{B} \rightleftharpoons \nu_X \text{X} + \nu_Y \text{Y}$.

On supposera pour simplifier que la **stœchiométrie** est respectée : à chaque instant, il existe M et m tels que :

$$N_A = \nu_A(M - m) \quad N_B = \nu_B(M - m) \quad N_X = \nu_X m \quad N_Y = \nu_Y m$$

On suppose M fixé, m variant de 0 à M . Chaque fois que m augmente de δm unités, N_A et N_B diminuent de $\delta N_A = \nu_A \delta m$ et $\delta N_B = \nu_B \delta m$, alors que N_X et N_Y augmentent de $\delta N_X = \nu_X \delta m$ et $\delta N_Y = \nu_Y \delta m$.

Attention : en thermodynamique, les quantités de matière (n_A, n_B, n_X, n_Y, m) sont exprimées en moles ; à l'échelle macroscopique, ce sont des variables continues (elles varient par sauts de l'ordre de 10^{-24}). En thermodynamique, on écrit sans hésiter :

$$dm = -\frac{dn_A}{\nu_A} = -\frac{dn_B}{\nu_B} = \frac{dn_X}{\nu_X} = \frac{dn_Y}{\nu_Y} \text{ au lieu de } \delta m = -\frac{\delta n_A}{\nu_A} = -\frac{\delta n_B}{\nu_B} = \frac{\delta n_X}{\nu_X} = \frac{\delta n_Y}{\nu_Y}$$

En thermodynamique statistique, (N_A, N_B, N_X, N_Y, m) sont des nombres de molécules : les variations δm sont des entiers relatifs. C'est $\delta m/M$ qui est extrêmement petit. C'est pourquoi l'**avancement élémentaire** $\delta\xi$ est défini ici comme étant :

$$\delta\xi = \frac{\delta m}{M} = -\frac{\delta N_A}{M \nu_A} = -\frac{\delta N_B}{M \nu_B} = \frac{\delta N_X}{M \nu_X} = \frac{\delta N_Y}{M \nu_Y}$$

Comme δm est extrêmement petit devant M (une unité face au nombre d'Avogadro), on pourra assimiler la variation $\delta\xi$ à la différentielle $d\xi$.

Les définitions qui suivent font référence à un mélange initial caractérisé par $m_0 = m(t=0)$ et à un mélange final caractérisé par $m_\infty = m(t \rightarrow \infty)$. Elles ne sont valables que pour les mélanges à la stœchiométrie. Elles peuvent présenter des différences (d'unités) par rapport à celles utilisées par les chimistes. Toutes les grandeurs citées ci-dessous sont sans dimension.

– **L'avancement de la réaction** : $\xi(t) = \frac{m(t) - m_0}{M}$

Il est nul à $t = 0$, positif lorsque, par rapport au mélange initial, l'équilibre s'est déplacé vers la droite, et négatif dans le cas contraire.

– **Le taux absolu d'avancement** : $\alpha(t) = \frac{m(t)}{M}$

C'est un nombre toujours compris entre 0 et 1. On remarquera que la valeur initiale n'est pas nulle (sauf si $m_0 = 0$, et que la valeur finale n'est pas égale à 1 (sauf si $m_\infty = 0$). C'est une fonction croissante si les réactants de gauche disparaissent pour donner les réactants de droite, et décroissante dans le cas contraire.

– **Le taux relatif d'avancement** : $\alpha_R(t) = \frac{m(t) - m_0}{m_\infty - m_0}$

Il est nul à $t = 0$ et égal à 1 pour $t \rightarrow \infty$. C'est une fonction croissante dans tous les cas, quel que soit le sens de réaction.

Tableau récapitulatif.

	minimum absolu (que des réactants)	mélange initial	mélange final	maximum absolu (que des produits)
$\xi(t)$	$\xi_{\min} = -\frac{m_0}{M}$	$\xi_0 = 0$	$\xi_{\infty} = \frac{m_{\infty} - m_0}{M}$	$\xi_{\max} = \frac{M - m_0}{M}$
$\alpha(t)$	$\alpha_{\min} = 0$	$\alpha_0 = \frac{m_0}{M}$	$\alpha_{\infty} = \frac{m_{\infty}}{M}$	$\alpha_{\max} = 1$
$\alpha_R(t)$	$\alpha_R^{\min} = -\frac{m_0}{m_{\infty} - m_0}$	$\alpha_R^0 = 0$	$\alpha_R^{\infty} = 1$	$\alpha_R^{\max} = \frac{M - m_0}{m_{\infty} - m_0}$

4.6.2 Surfaces équipotentielles

On considère la réaction chimique $A + BC \rightarrow AB + C$. On suppose qu'à tout moment de la réaction, les trois atomes A, B, C sont alignés : l'énergie potentielle du système constitué des trois atomes n'est fonction que de deux distances R_{AB} et R_{BC} . La figure 4.6 de gauche représente un exemple de diagramme d'équipotentielles parmi les plus simples.

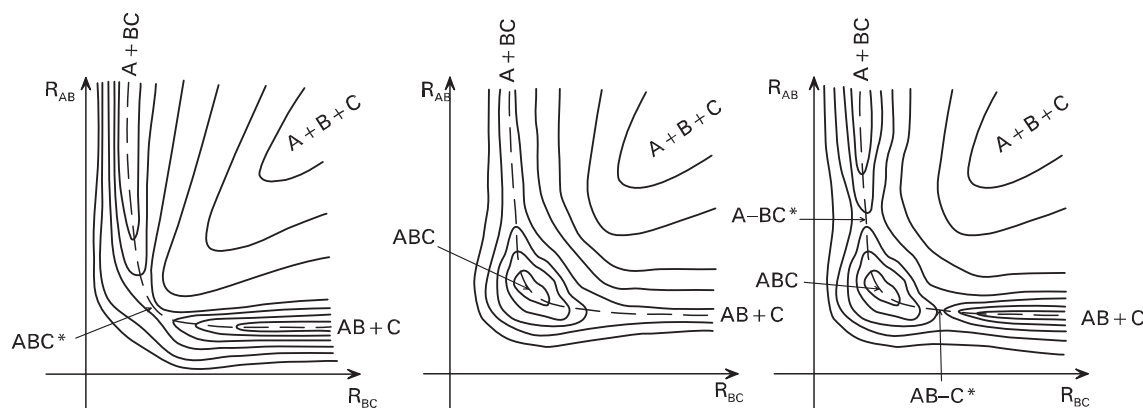


FIG. 4.6 – Surfaces équipotentielles : un col, un puits, un bassin

En haut à gauche et en bas à droite, on trouve deux vallées correspondant aux configurations stables $A + BC$ et $AB + C$. Celle de droite est la plus profonde ($AB + C$ est plus stable que $A + BC$) et la plus encaissée (le "ressort" $A - B$ est plus dur que le "ressort" $B - C$). En haut à droite, le haut plateau correspond aux trois atomes séparés. La position ABC^* est une configuration de col : le passage de la configuration $A + BC$ à la configuration $AB + C$ nécessite le franchissement d'une barrière de potentiel et la configuration ABC^* est la plus petite barrière qu'il soit possible de trouver. La configuration ABC^* n'est pas vraiment une molécule : on l'appelle **complexe activé**. C'est la configuration intermédiaire, instable et éphémère, présentant la plus haute énergie au cours de la réaction. La ligne pointillée est celle suivie au cours de la réaction. C'est le chemin que suivraient deux gouttes

d'eau lâchées au niveau du col, l'une vers la vallée $A + BC$, l'autre vers la vallée $AB + C$. On peut exprimer l'énergie potentielle au cours de la réaction en fonction d'une seule variable : l'abscisse curviligne τ le long de cette courbe. Cette abscisse est appelée **coordonnée de réaction**.

L'énergie d'activation est la hauteur de la barrière à traverser. Si U_1 est l'énergie du fond de vallée $A + BC$, et si U^* est l'énergie du complexe activé, l'énergie d'activation est l'énergie à fournir pour déclencher la réaction. Elle vaut $\Delta U^* = U^* - U_1$. **L'énergie motrice** est la différence $\Delta U_m = U_2 - U_1$ entre les fonds de vallée^[4]. Pour la réaction inverse, l'énergie d'activation et l'énergie motrice valent respectivement $\Delta U'^* = U^* - U_2$ et $\Delta U'_m = U_1 - U_2$.

Dans la théorie d'Eyring qui suit, on raisonnera dans le cas où le composé ABC est un complexe instable. Mais il faut savoir que d'autres profils d'équipotentiels existent. On peut citer, toujours pour les systèmes à deux degrés de liberté, deux cas où le composé ABC est une molécule stable. Par exemple :

- un puits de potentiel (figure 4.6 au milieu). Le produit devrait en principe être la molécule ABC : c'est la configuration de plus basse énergie. Mais, à température non nulle, il faut tenir compte de l'entropie : sa tendance à être maximale fera pencher la réaction, soit vers $A + BC$, soit vers $AB + C$.
- un bassin (figure 4.6 à droite). La molécule ABC est métastable (d'énergie supérieure à celles des produits initial et final mais sa valeur correspond à un minimum local de l'énergie potentielle). La réaction nécessite dans ce dernier cas le passage de deux barrières de potentiel (deux complexes activés $A - BC^*$ et $AB - C^*$).

4.6.3 Théorie d'Eyring

Il n'est pas question ici de faire un traitement général de la théorie d'Eyring. On choisit l'exemple (simple !) de la réaction $A + BC \rightarrow AB + C$. Son diagramme d'équipotentiels est supposé avoir l'allure représentée figure 4.6 à gauche. On rappelle les hypothèses permettant de tracer un tel diagramme :

- à chaque instant, les trois atomes A, B, C sont alignés : l'énergie potentielle du système est fonction de deux distances R_{AB} et R_{BC} ;
- l'énergie d'activation est due au franchissement d'un seul col (existence d'un complexe activé intermédiaire ABC^*) ;
- l'énergie motrice est négative (la vallée $AB + C$ est plus profonde que la vallée $A + BC$).

À l'état initial le mélange (stœchiométrique) est en équilibre à la température T_0 . Sa composition est entièrement définie par deux nombre M et m_0 : on y compte m_0 molécules AB, m_0 atomes C, $M - m_0$ atomes A et $M - m_0$ molécules BC.

À l'instant $t = 0$, on plonge le système dans un bain à la température T . Au cours du temps, la composition évolue vers une composition finale caractérisée par les deux nombres M et m_∞ .

Le but de la théorie d'Eyring est de calculer la vitesse de la réaction $m(t)$, autrement dit le nombre de passages de col par unité de temps.

⁴ En termes familiers, l'énergie motrice est ce qui donne envie à $A + BC$ de devenir $AB + C$, et l'énergie d'activation est ce qu'elle doit dépenser pour assouvir cette envie.

On relie tout d'abord m_0 et m_∞ aux fonctions de partition par l'intermédiaire des constantes d'équilibre. On a, avec les notations $q_A^0 = q_A(T_0)$, $q_A = q_A(T)$, etc :

$$\begin{aligned} K_C(T_0) &= \frac{q_{AB}^0 q_C^0}{q_A^0 q_{BC}^0} = \frac{m_0^2}{(M - m_0)^2} \Rightarrow \frac{m_0}{M} = \frac{\sqrt{K_C(T_0)}}{\sqrt{K_C(T_0)} + 1} \\ K_C(T) &= \frac{q_{AB} q_C}{q_A q_{BC}} = \frac{m_\infty^2}{(M - m_\infty)^2} \Rightarrow \frac{m_\infty}{M} = \frac{\sqrt{K_C(T)}}{\sqrt{K_C(T)} + 1} \end{aligned} \quad (4.11)$$

Supposons que la réaction se fait de gauche à droite. Durant la première phase de la réaction $A + BC \rightarrow ABC^*$ (montée du col), Eyring suppose que le produit de départ et le complexe activé sont en équilibre thermodynamique $A + BC \rightleftharpoons ABC^*$. C'est contradictoire avec le processus (une évolution ne peut pas être un équilibre) mais il se trouve que cette approximation donne de bons résultats.

Comparons le complexe activé ABC^* à une vulgaire molécule ABC . En considérant les trois atomes comme des masses ponctuelles (on ignore les degrés de liberté électroniques et nucléaires), la molécule ABC (linéaire) possède neuf degrés de liberté : trois degrés de liberté de translation, deux degrés de liberté de rotation et quatre degrés de liberté de vibration. Les degrés de liberté de translation et de rotation n'ont aucune influence sur le diagramme d'équipotentiels. Les seuls degrés de liberté qui nous intéressent ici sont les degrés de liberté de vibration. Après analyse en coordonnées normales (méthode NCA, voir paragraphe 4.4.3), on trouve quatre modes normaux : deux modes dits *transverses* qui décrivent les mouvements vibrationnels perpendiculaires à l'axe de la molécule, et deux modes dits *longitudinaux* qui décrivent les mouvements vibrationnels parallèles à l'axe de la molécule. Par hypothèse, les deux modes transverses sont ignorés : ils introduisent des torsions dans la molécule et nous avons supposé qu'elle restait linéaire à chaque instant. Il nous reste les deux modes longitudinaux de vibration qui correspondent sur le diagramme d'équipotentiels, d'une part à la coordonnée de réaction τ (localement tangente à la courbe de réaction), d'autre part à la coordonnée η normale à la courbe de réaction.

Ce qui différencie ABC^* de ABC , c'est que suivant l'axe tangent à la courbe de réaction, l'énergie potentielle présente un maximum. Suivant l'axe perpendiculaire à cette courbe, le profil est par contre habituel. La fréquence propre associée à la variable η est trouvée comme d'habitude réelle et positive. Par contre, dans le hamiltonien du système, le terme d'énergie potentielle correspondant à la variable τ porte un coefficient négatif, ce qui conduit par le calcul à une fréquence propre imaginaire pure. C'est ce qui caractérise d'ailleurs la coordonnée de réaction : si le complexe activé possède une géométrie plus quelconque, la coordonnée de réaction sera toujours la seule parmi les $3n$ degrés de liberté à posséder cette propriété.

En résumé, pour un complexe activé à n noyaux, parmi ses $3n$ degrés de liberté, $3n - 1$ peuvent être traités en considérant le complexe activé comme une molécule. Seule la coordonnée de réaction τ requiert un traitement particulier. Si ce degré de liberté était "normal", il ferait partie des degrés de liberté de vibration.

Les configurations intermédiaires entre molécule et complexe activé seront nommées tout simplement "complexes".

L'avancement de la réaction $\xi(t)$ est défini ici comme étant le nombre, rapporté à M de complexes ayant traversé le col à l'instant t . Autrement dit, $M \dot{\xi}(t) dt$ est le nombre de complexes traversant le col entre t et $t + dt$.

$\dot{\xi}(t) = \frac{d\xi}{dt}$ est par définition la **vitesse d'avancement** de la réaction.

Le nombre de complexes traversant le col pendant un intervalle de temps dt sont ceux pour lesquels la vitesse $\dot{\tau}$ est positive et pour lesquels la valeur de τ est comprise entre $-\dot{\tau} dt$ et 0 (on a placé l'origine de τ au niveau du col).

Pour le calcul de $\dot{\xi}$, on procède par étapes.

- Nombre de complexes tels que $-\dot{\tau} dt \leq \tau \leq 0$:

$$M \int_{-\dot{\tau} dt}^0 g(\tau) d\tau \approx M g^* \dot{\tau} dt \quad (4.12)$$

$M g(\tau) d\tau$ est le nombre de complexes pour lesquels la coordonnée de réaction est comprise entre τ et $\tau + d\tau$. L'approximation faite ici consiste à dire que, dt étant un infiniment petit, la densité $g(\tau)$ est constante et égale à g^* dans tout l'intervalle $[-\dot{\tau} dt, 0]$.

- Fraction de complexes de vitesse comprise entre $\dot{\tau}$ et $\dot{\tau} d\dot{\tau}$:

$$\frac{\exp\left(\frac{-m^* \dot{\tau}^2}{2kT}\right) d\dot{\tau}}{\int_{-\infty}^{+\infty} \exp\left(\frac{-m^* \dot{\tau}^2}{2kT}\right) d\dot{\tau}} = \sqrt{\frac{m^*}{2\pi kT}} \exp\left(\frac{-m^* \dot{\tau}^2}{2kT}\right) d\dot{\tau} \quad (4.13)$$

- Nombre de complexes traversant le col entre t et $t + dt$ (c'est par définition $M \dot{\xi}$) : on multiplie 4.12 par 4.13 et on intègre sur $\dot{\tau}$ de 0 à $+\infty$. On trouve :

$$M \dot{\xi} = M g^* dt \sqrt{\frac{m^*}{2\pi kT}} \int_0^{\infty} \exp\left(\frac{-m^* \dot{\tau}^2}{2kT}\right) \dot{\tau} d\dot{\tau} = M g^* dt \sqrt{\frac{kT}{2\pi m^*}}$$

On en déduit $\dot{\xi} = g^* \sqrt{\frac{kT}{2\pi m^*}}$. Il reste à calculer g^* .

$M g^*$ est le nombre de complexes activés par unité de longueur sur la courbe de réaction au niveau du col. On peut relier ce nombre aux fonctions de partition par l'intermédiaire de la constante de l'équilibre $A + BC \rightleftharpoons ABC^*$, à condition de définir une longueur $\delta\tau$ au voisinage du col où le potentiel $U(\tau)$ peut être considéré comme constant. On pourra alors écrire :

$$N^* = M g^* \delta\tau \quad \text{et} \quad \frac{N^*}{N_A N_{BC}} = \frac{q^*}{q_A q_{BC}} \quad \Rightarrow \quad g^* = \frac{1}{\delta\tau} \frac{q^*}{q_A q_{BC}} \frac{N_A N_{BC}}{M}$$

La valeur de $\delta\tau$ n'a aucune importance. Ce paramètre arbitraire disparaîtra en fin de calcul. Les fonctions de partition q_A et q_{BC} sont connues. Seul le calcul de q^* requiert un traitement particulier.

On suppose que tous les degrés de liberté sont séparables et on écrit $q^* = q_{\bar{\tau}}^* \times q_{\tau}^*$:

- $q_{\bar{\tau}}^*$ est la fonction de partition du complexe activé calculée sans tenir compte de la coordonnée de réaction τ . Elle se calcule de la manière habituelle.

- q_τ^* est la fonction de partition relative à la coordonnée de réaction. On la calcule en utilisant les résultats du paragraphe 3.6.1 (équation 3.31) :

$$q_\tau^* = \frac{1}{h} \int_0^{\delta\tau} d\tau \int_{-\infty}^{+\infty} \exp\left(-\frac{p_\tau^2}{2m^* kT}\right) dp_\tau = \frac{\delta\tau}{h} \sqrt{2\pi m^* kT}$$

Ce qui donne finalement :

$$\dot{\xi} = K(T) \frac{N_A N_{BC}}{M} \quad \text{avec} \quad K(T) = \frac{kT}{h} \frac{q_\tau^*}{q_A q_{BC}} = \frac{kT}{h} \frac{q_0^*}{q_A^0 q_{BC}^0} e^{-\Delta U^*/kT}$$

Les fonctions de partition q_A^0 , q_{BC}^0 et q_0^* sont les fonctions de partitions prenant en comptes les degrés de liberté de translation, rotation et vibration, avec pour chacune d'elles l'origine des énergies placée sur son niveau électronique fondamental.

$K(T)$ est par définition la **constante d'avancement** de la réaction^[5]. C'est le nombre de complexes traversant le col par seconde, par atome A, par molécule BC.^[6]

Il faut maintenant envisager la réaction inverse. Les calculs sont tout à fait similaires aux précédents. Ils conduisent à :

$$\dot{\xi} = -K'(T) \frac{N_{AB} N_C}{M} \quad \text{avec} \quad K'(T) = \frac{kT}{h} \frac{q_\tau^*}{q_{AB} q_C} = \frac{kT}{h} \frac{q_0^*}{q_{AB}^0 q_C^0} e^{-\Delta U'^*/kT}$$

Les deux réactions ayant lieu simultanément, on doit prendre en compte les deux contributions. Avec $N_A = N_{BC} = M - m$ et $N_{AB} = N_C = m$, on doit résoudre :

$$\frac{dm}{dt} = M \dot{\xi} = K(T) (M - m)^2 - K'(T) m^2$$

La résolution de cette équation différentielle ne pose aucune difficulté. Elle conduit au taux absolu d'avancement $\alpha(t) = m(t)/M$ ($\alpha_0 = m_0/M$) :

$$\frac{\alpha(t) - \alpha_\infty}{\alpha(t) - \alpha_1} = \frac{\alpha_0 - \alpha_\infty}{\alpha_0 - \alpha_1} e^{-\omega t} \quad \text{avec} \quad \alpha_\infty = \frac{\sqrt{K(T)}}{\sqrt{K(T)} + \sqrt{K'(T)}} = \frac{\sqrt{K_C(T)}}{\sqrt{K_C(T)} + 1}$$

$$\alpha_1 = \frac{\sqrt{K(T)}}{\sqrt{K(T)} - \sqrt{K'(T)}} = \frac{\sqrt{K_C(T)}}{\sqrt{K_C(T)} - 1} \quad \text{et} \quad \omega = 2M \sqrt{K(T) K'(T)}$$

On constate que $\alpha(t)$ tend vers α_∞ quand t tend vers l'infini. C'est en excellent accord avec la théorie des équilibres chimiques (équation 4.11).

⁵ Si les surfaces équipotentiels sont plus complexes, on multiplie la constante d'avancement trouvée par un **facteur de transmission** κ . On peut traiter à titre d'exemple le cas du bassin : pour passer de $A + BC$ à $AB + C$, les trois atomes doivent passer une première barrière ΔU_1 (complexe activé $A-BC^*$), traverser un bassin (molécule métastable ABC), et enfin traverser une seconde barrière ΔU_2 (complexe activé $AB-C^*$). Les molécules ayant traversé la première barrière et se trouvant dans le bassin ont le choix entre passer la deuxième barrière (κ est la probabilité de ce choix) et retraverser la première.

⁶ son unité ici est la s^{-1} ; en thermodynamique chimique, la définition est légèrement différente : c'est la vitesse de réaction rapportée aux concentrations en mol/L.

4.6.4 Application à la diffusion de surface

Le modèle d'Eyring peut être appliqué à de nombreux problèmes physiques, chaque fois que le processus fait intervenir le passage d'une barrière de potentiel. On traite ici, pour des raisons de simplicité de représentation, le problème de la diffusion de surface.

Une position stable (ou site d'adsorption ou puits de potentiel) sera caractérisée par une énergie minimale U_0 (énergie de fond de puits) et par deux pulsations :

- une pulsation $\omega_\perp$ pour les vibrations perpendiculaires à la surface ;
- une pulsation $\omega_\parallel$ pour les vibrations parallèles à la surface (suivant x et y).

La fonction de partition d'un atome adsorbé en position stable s'écrit :

$$q_0 = e^{-U_0/kT} q_\perp q_\parallel^2 \quad \text{avec} \quad q_\perp = \frac{1}{2 \sinh \frac{\hbar \omega_\perp}{2kT}} \quad \text{et} \quad q_\parallel = \frac{1}{2 \sinh \frac{\hbar \omega_\parallel}{2kT}}$$

Soit un atome traversant un col dans une direction τ (x ou y). L'autre direction parallèle à la surface (y ou x) sera notée η . La variable τ joue le même rôle que la coordonnée de réaction du paragraphe précédent : dans cette direction, l'énergie potentielle au niveau du col présente un maximum (noté U^*). Suivant les directions η et z , les mouvements sont vibrationnels, mais les pulsations sont en principe différentes de celles mesurées au fond des puits. Elles seront notées ω_η^* et $\omega_\perp^*$.

Moyennant quoi, la fonction de partition d'un atome situé dans un intervalle $\delta\tau = \dot{\tau} \delta t$ au voisinage d'une position de col s'écrit :

$$q^*(\dot{\tau}) = e^{-U^*/kT} q_\perp^* q_\eta^* \frac{\dot{\tau} \delta t}{h} \sqrt{2\pi m kT} \quad \text{avec} \quad q_\perp^* = \frac{1}{2 \sinh \frac{\hbar \omega_\perp^*}{2kT}} \quad \text{et} \quad q_\eta^* = \frac{1}{2 \sinh \frac{\hbar \omega_\eta^*}{2kT}}$$

Cette fonction de partition est celle des atomes possédant la vitesse $\dot{\tau}$ et traversant le col dans l'intervalle de temps δt . On cherche la fonction de partition de tous les atomes (indépendamment de leur vitesse) traversant le col dans l'intervalle de temps δt .

Il faut donc multiplier $q^*(\dot{\tau})$ par $\sqrt{\frac{m}{2\pi kT}} e^{-m\dot{\tau}^2/2kT}$ et intégrer sur $\dot{\tau}$ de 0 à $+\infty$:

$$q^* = \sqrt{\frac{m}{2\pi kT}} \int_0^\infty q^*(\dot{\tau}) e^{-m\dot{\tau}^2/2kT} d\dot{\tau} = \frac{kT}{h} q_\perp^* q_\eta^* \delta t e^{-U^*/kT}$$

Le rapport entre q^* et q_0 est égal au rapport entre le nombre δN^* d'atomes en situation de passage de cols durant δt et le nombre N_0 d'atomes en position stable. Ce dernier étant très grand par rapport au premier, on l'assimile au nombre total d'atomes. Pour obtenir la fréquence de saut par atome, il suffit ensuite de diviser ce rapport par δt . On obtient :

$$\Gamma(T) = \frac{kT}{h} \frac{q_\perp^* q_\eta^*}{q_\perp q_\parallel^2} e^{-\Delta U^*/kT} \quad \text{avec} \quad \Delta U^* = U^* - U_0$$

Pour obtenir un ordre de grandeur, on assimile $q_\perp^*$ et q_η^* à $q_\perp$ et $q_\parallel$. C'est ce qu'on trouve d'ailleurs si on utilise l'énergie potentielle 4.3 du paragraphe 4.1.3. On obtient :

$$\Gamma(T) \sim \frac{kT}{h} 2 \sinh \frac{h\nu_\parallel}{2kT} e^{-\Delta U^*/kT} \sim \nu_\parallel e^{-\Delta U^*/kT}$$

Résultat physiquement correct. La fréquence $\nu_\parallel$ est une fréquence de tentatives de saut, et le facteur de Boltzmann $e^{-\Delta U^*/kT}$ est un taux de réussite.

4.7 Gaz parfait diélectrique sous champ électrique

4.7.1 Conditions d'équilibre sous champ de force extérieur

On considère un système Σ de particules toutes identiques (indépendantes ou non) que l'on soumet à un champ de forces : chaque particule est soumise, non seulement aux forces d'interactions exercées par ses congénères, mais aussi à une force $\vec{f}(\vec{r})$ dépendant uniquement de sa position dans le système. Elle ne dépend ni explicitement du temps, ni de la position des autres particules (autrement dit, cette force ne modifie pas les forces d'interaction entre particules). On supposera également que cette force dérive d'un potentiel $\Phi(\vec{r})$, autrement dit $\vec{f}(\vec{r}) = -\vec{\nabla}\Phi(\vec{r})$, et que ses variations en fonction de $\vec{r}$ sont inappréciables à l'échelle microscopique : au voisinage de la position $\vec{r}$, on pourra toujours définir un élément de volume δV suffisamment petit à l'échelle macroscopique pour pouvoir considérer que le potentiel y est uniforme, et suffisamment grand à l'échelle microscopique pour pouvoir appliquer un traitement statistique aux particules qu'il contient.

On choisit justement cet élément de volume δV comme système $\delta\Sigma$, et on notera δN le nombre de particules qu'il contient. Soit j un microétat accessible au micro-système $\delta\Sigma$ et $E_j^0(\delta V, \delta N)$ l'énergie correspondant à cet état lorsqu'il n'y a pas de forces extérieures. Lorsqu'on applique la force extérieure, cette énergie devient :

$$E_j(\delta V, \delta N, \vec{r}) = E_j^0(\delta V, \delta N) + \Phi(\vec{r}) \delta N$$

On peut exprimer la fonction de partition canonique $Q(T, \delta V, \delta N, \vec{r})$ de $\delta\Sigma$ en fonction de la fonction de partition sans champ $Q_0(T, \delta V, \delta N)$:

$$Q(T, \delta V, \delta N, \vec{r}) = Q_0(T, \delta V, \delta N) e^{-\Phi(\vec{r}) \delta N / kT}$$

La fonction de partition grand canonique s'écrit :

$$Z(T, \delta V, \mu, \vec{r}) = \sum_{\delta N=0}^{\infty} Q(T, \delta V, \delta N, \vec{r}) e^{\mu \delta N / kT} = \sum_{N=0}^{\infty} Q_0(T, \delta V, \delta N) e^{(\mu - \Phi(\vec{r})) \delta N / kT}$$

On définit le **potentiel chimique local** par : $\tilde{\mu}(\vec{r}) = \mu - \Phi(\vec{r})$

Il permet d'écrire le grand potentiel de la manière habituelle :

$$\Psi(T, \delta V, \tilde{\mu}(\vec{r})) = -kT \ln Z(T, \delta V, \tilde{\mu}(\vec{r}))$$

$$\text{avec } Z(T, \delta V, \tilde{\mu}(\vec{r})) = \sum_{\delta N=0}^{\infty} Q_0(T, \delta V, \delta N) e^{\tilde{\mu}(\vec{r}) \delta N / kT}$$

Les propriétés locales du système sous champ se calculent comme celles du système sans champ. Il suffit de remplacer dans les expressions le potentiel chimique par le potentiel chimique local :

$$p(\vec{r}) = - \left. \frac{\partial \Psi}{\partial \delta V} \right]_{T, \tilde{\mu}(\vec{r})} \quad \delta S(\vec{r}) = - \left. \frac{\partial \Psi}{\partial T} \right]_{\delta V, \tilde{\mu}(\vec{r})} \quad \delta N(\vec{r}) = - \left. \frac{\partial \Psi}{\partial \tilde{\mu}(\vec{r})} \right]_{T, \delta V}$$

Pour finir, on exprime l'équilibre entre les différents microsystemes en écrivant que le potentiel chimique est homogène dans tout le système : $\mu = \tilde{\mu}(\vec{r}) + \Phi(\vec{r}) = \text{cste}$. L'équilibre apparaît comme étant le résultat d'une compétition entre potentiel extérieur et potentiel chimique : le potentiel extérieur tend à concentrer les particules là où il est minimum, alors que le potentiel chimique tend à homogénéiser le système.

En exemple on peut traiter le cas d'un gaz parfait **dans le champ de pesanteur**. Sans forces extérieures, le potentiel chimique local s'écrit :

$$\mu(z) = \mu_T^0 + kT \ln \frac{p(z)}{p_0}$$

Sous l'effet de la pesanteur, chaque molécule acquiert une énergie potentielle $m g z$. L'équilibre global est assuré si :

$$\mu(z) + m g z = \text{cste} = \mu(0)$$

On en déduit :

$$\mu_T^0 + kT \ln \frac{p(z)}{p_0} + m g z = \mu_T^0 + kT \ln \frac{p(0)}{p_0} \quad \Rightarrow \quad p(z) = p(0) e^{-m g z / kT}$$

Un résultat qui nécessite en mécanique la résolution d'une équation différentielle.

Autre exemple : un gaz parfait **en contact avec une surface attractive**. La surface est supposée perpendiculaire à l'axe z . Le gaz, situé dans le demi-espace $z > 0$, est soumis de la part de la surface à une énergie potentielle $U(z)$ tendant vers 0 quand z tend vers $+\infty$, présentant un minimum (négatif) à une certaine distance de la surface (le rayon R de l'atome), et tendant vers $+\infty$ quand z tend vers 0.

La condition d'équilibre s'écrit ici :

$$\mu(z) + U(z) = \text{cste} = \mu(R) + U(R)$$

ou encore :

$$\rho(z) = \rho(R) e^{(U(R)-U(z))/kT}$$

Soit H la hauteur de l'enceinte, A l'aire de la surface et ρ_0 la densité moyenne du gaz, autrement dit la densité homogène qu'aurait le gaz en l'absence de la surface attractive. Connaissant le profil $U(z)$, on calcule $\rho(R) e^{U(R)/kT}$ en écrivant :

$$\int_0^H \rho(z) dz = N = \rho_0 A H \quad \Rightarrow \quad \rho(R) e^{U(R)/kT} = \frac{\rho_0 A H}{\int_0^H e^{-U(z)/kT} dz}$$

Pour calculer la densité superficielle $\sigma = N_{\text{ads}}/A$ de la phase adsorbée, on considère qu'un atome est adsorbé si la distance entre cet atome et la surface est inférieure à R :

$$N_{\text{ads}} = \rho(R) e^{U(R)/kT} \int_0^R e^{-U(z)/kT} dz \quad \Rightarrow \quad \sigma = \frac{N_{\text{ads}}}{A} = \rho_0 H \frac{\int_0^R e^{-U(z)/kT} dz}{\int_0^H e^{-U(z)/kT} dz}$$

4.7.2 Thermodynamique des milieux diélectriques

On considère un gaz constitué de deux types de particules les unes chargées positivement (cations), les autres négativement (anions). Le système global sera supposé électriquement neutre. Pour simplifier, nous considérerons un système constitué de N cations de charge $+q$ et de N anions de charge $-q$.

Le récipient est un parallélépipède rectangle inséré entre deux plaques, l'une chargée positivement et l'autre négativement (figure 4.7 à gauche). On suppose le gaz suffisamment dilué pour négliger les interactions électrostatiques entre particules par rapport aux interactions avec les plaques (le gaz est supposé parfait). La pesanteur n'intervient pas dans l'étude.

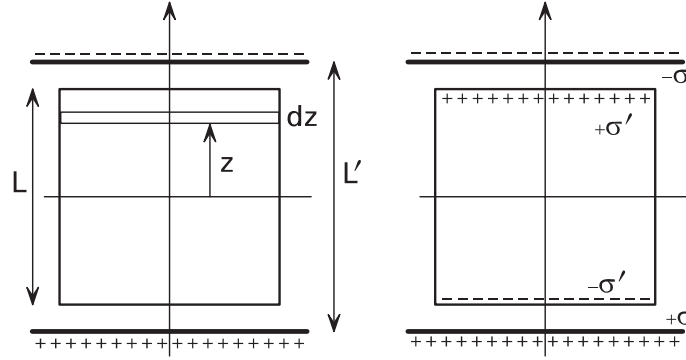


FIG. 4.7 – Gaz sous champ électrique : géométrie et répartition des charges fictives

On notera $+\sigma$ et $-\sigma$ les densités de charges de chacune des plaques. Pour éviter les effets de bord, on supposera les deux plaques de dimension infinie. Le champ électrique uniforme entre les deux plaques est dirigé du bas vers le haut sur la figure et vaut $\mathcal{E} = +\sigma/\epsilon_0$. Les cations y sont donc soumis à une force $+q\mathcal{E}$ (dirigée vers le haut), et les anions à une force $-q\mathcal{E}$ (dirigée vers le bas). Tout se passe comme si les particules baignaient dans deux champs d'énergie potentielle opposés :

$$U^+(z) = -q\mathcal{E}z \quad \text{pour les cations} \quad \text{et} \quad U^-(z) = +q\mathcal{E}z \quad \text{pour les anions}$$

On obtient, avec le raisonnement utilisé pour l'étude du champ de pesanteur, les densités $\rho^+(z)$ et $\rho^-(z)$ (en cations par m^3 et en anions par m^3) :

$$\rho^+(z) = \rho_0^+ e^{+q\mathcal{E}z/kT} \quad \text{et} \quad \rho^-(z) = \rho_0^- e^{-q\mathcal{E}z/kT}$$

Si on appelle $\rho_0 = N/V$ la densité commune en anions et en cations en l'absence de champ, on détermine les deux constantes en écrivant :

$$\frac{1}{L} \int_{-L/2}^{+L/2} \rho^+(z) dz = \frac{1}{L} \int_{-L/2}^{+L/2} \rho^-(z) dz = \rho_0 \quad \Rightarrow \quad \rho_0^+ = \rho_0^- = \rho_0 \frac{\frac{q\mathcal{E}L}{2kT}}{\sinh \frac{q\mathcal{E}L}{2kT}}$$

Les centres de gravité des charges positives et négatives se trouvent respectivement aux cotes z^+ et z^- que l'on calcule grâce aux relations :

$$z^+ = -z^- = \frac{\int_{-L/2}^{+L/2} z \rho^+(z) dz}{\int_{-L/2}^{+L/2} \rho^+(z) dz} = -\frac{\int_{-L/2}^{+L/2} z \rho^-(z) dz}{\int_{-L/2}^{+L/2} \rho^-(z) dz} = \frac{kT}{q\mathcal{E}} \left(\frac{\frac{q\mathcal{E}L}{2kT}}{\tanh \frac{q\mathcal{E}L}{2kT}} - 1 \right)$$

On peut remarquer que le modèle n'est pas satisfaisant à basse température : lorsque T tend vers 0, z^+ tend vers $+L/2$ et z^- vers $-L/2$. Cela signifie que les charges sont agglutinées sur les parois supérieure et inférieure et donc que l'enceinte est entièrement vide. Ce que nous n'avons pas pris en compte, c'est que la répartition non homogène des charges crée un champ électrique induit venant se superposer (s'opposer) au champ appliqué. Ce champ induit est impossible à calculer puisqu'il nécessite la connaissance de la répartition des charges, c'est-à-dire du résultat. Le milieu est dit diélectrique : le champ électrique effectif présent dans le milieu est différent du champ appliqué. À partir de maintenant, $\mathcal{E}$ désignera le champ effectif. Le champ appliqué $+\sigma/\varepsilon_0$ sera noté D . **La constante diélectrique** ε du milieu est définie par la relation $D = \varepsilon \mathcal{E}$.

Soit $D' = \mathcal{E} - D$ le champ induit. Tout comme à D est associée une densité de charge σ , on associe à D' une charge fictive σ' : le champ D' est égal à celui que provoquerait la présence de densités de charge $+\sigma'$ et $-\sigma'$ respectivement sur les parois supérieure et inférieure du récipient (figure 4.7 à droite).

On écrira donc : $D' = -\frac{\sigma'}{\varepsilon_0}$ (il s'oppose au champ appliqué).

En thermodynamique, D est une variable intensive. Tout comme p et T , il peut être contrôlé de l'extérieur (par l'intermédiaire de σ). Le champ D' par contre est incontrôlable : il dépend de la réaction du milieu à l'action de D . Le développement qui suit a pour but de déterminer la variable extensive associée à D .

Le champ électrique total dans le milieu vaut : $\mathcal{E} = D + D' = D - \frac{\sigma'}{\varepsilon_0}$.

La variable extensive associée à D est celle dont le produit avec D donne l'énergie emmagasinée par le système sous l'effet de D . On la notera M : c'est le **moment dipolaire total** du système.

Calculons le travail fourni au système lorsque σ passe à $\sigma + d\sigma$. Si A est l'aire des plaques en regard avec l'enceinte, ce travail est utilisé pour transporter une quantité de charge $A d\sigma$ de la plaque supérieure (chargée négativement) vers la plaque inférieure (chargée positivement). À ce déplacement s'oppose un champ $\mathcal{E}$ sur une longueur L , et un champ D sur une longueur $L' - L$. D'où le travail total fourni : $dW_{\text{tot}} = \mathcal{E} L A d\sigma + D (L' - L) A d\sigma$. Si l'espace entre les deux plaques était vide, ce travail vaudrait $dW_0 = D L' A d\sigma$. La contribution du système vaut donc $dW = dW_{\text{tot}} - dW_0$. On trouve, avec $V = L \times A$:

$$dW = (\mathcal{E} - D) V d\sigma = D' V d\sigma = -\sigma' V \frac{d\sigma}{\varepsilon_0} = -\sigma' V dD$$

La densité de charge σ' est habituellement notée P : c'est la **polarisation**. La grandeur extensive M recherchée vaut au signe près $P V$. On a donc, avec les notation du paragraphe 1.4.8, une contribution $-M dD$ à $d\Phi_M$ et par conséquent une contribution $+D dM$ à dE :

$$\textbf{Équation de Gibbs : } dE = T dS - p dV + D dM + \mu dN$$

4.7.3 Molécules dipolaires dans un champ électrique

La diélectricité à l'échelle microscopique peut avoir différentes origines. Dans tous les cas, le champ électrique appliqué provoque un déplacement des charges, de sorte que le centre de gravité des charges positives ne coïncide plus avec le centre de gravité des charges négatives. Le système devient porteur d'un moment dipolaire induit porté par la droite liant les deux centres de gravité. On en a vu un exemple au paragraphe précédent. En voici un autre.

Dans le système étudié ici, il n'est plus question d'anions et de cations. On considère un système d'atomes ou de **molécules électriquement neutres**. Le moment dipolaire est dû à une déformation des nuages électroniques : les noyaux sont attirés par la plaque chargée négativement, et les nuages électroniques par la plaque chargée positivement. Chaque molécule devient porteuse d'un moment dipolaire, et le moment dipolaire total est la somme des moments dipolaires individuels.

Certaines molécules, comme H_2O ou HCl , portent un **moment dipolaire permanent**. Même en l'absence de champ, les deux centres de gravité ne coïncident pas. Lorsqu'on applique le champ, d'une part un moment dipolaire induit vient s'ajouter au permanent (déformation des nuages électroniques), d'autre part le moment dipolaire résultant (permanent et induit) a tendance à s'aligner à la direction du champ. Il y a donc perturbation des degrés de liberté de rotation de la molécule. Les degrés de liberté de vibration sont en toute rigueur également perturbés (modification des niveaux d'énergie électroniques) mais nous n'en tiendront pas compte dans l'étude. Les degrés de liberté de translation par contre ne sont pas du tout perturbés par la présence du champ : les molécules restent électriquement neutres et ne présentent aucune attirance vers l'une ou l'autre des plaques chargées.

On considère un gaz parfait monoconstituant constitué de N molécules indépendantes que l'on plonge dans un champ électrique uniforme orienté suivant un axe Oz .

Les conditions de travail (T, V, D, N constants) conduisent tout naturellement à travailler dans l'ensemble statistique de fonction caractéristique $\Phi_M(T, V, D, N)$:

$$\Phi_M(T, V, D, N) = E - TS - DM = F - DM = -kT \ln \Delta_M(T, V, D, N)$$

$$\text{avec } \Delta_M(T, V, D, N) = \sum_M Q(T, V, M, N) e^{+DM/kT}$$

$$\text{et } d\Phi_M = -SdT - p dV - M dD + \mu dN$$

Que l'on travaille à M ou à D constant, les molécules sont indépendantes. Dans la limite classique (voir paragraphe 2.2.2), on a avec les notations du paragraphe 1.4.8 :

$$- \text{ à } M \text{ constant : } Q(T, V, M, N) = \frac{q(T, V, M)^N}{N!}$$

$$- \text{ à } D \text{ constant : } \Delta_M(T, V, D, N) = \frac{\delta_M(T, V, D)^N}{N!}$$

Le plus simple est de travailler à D constant. Le problème revient à calculer la fonction de partition $\delta_M(T, V, D)$ relative à une molécule. Deux phénomènes interviennent :

- Le moment dipolaire permanent de la molécule a tendance à s'aligner avec le champ électrique appliqué $\vec{D}$. L'énergie emmagasinée par un dipôle de moment dipolaire $\vec{\mu}_0$ dans un champ $\vec{D}$ vaut (θ est l'angle entre $\vec{\mu}_0$ et $\vec{D}$) :

$$-\vec{\mu}_0 \cdot \vec{D} = -\mu_0 D \cos \theta$$

- Un moment dipolaire induit $\vec{\mu}_i$ apparaît. En champ faible, celui-ci est proportionnel à $\vec{D}$: $\mu_i = \alpha \vec{D}$. Le coefficient α est une propriété de la molécule appelée *polarisabilité*. Le travail nécessaire pour créer ce moment induit vaut :

$$\int_0^D -\mu_i(D) dD = \int_0^D -\alpha D dD = -\frac{\alpha D^2}{2}$$

La somme des deux contributions est appelée énergie potentielle rotationnelle :

$$U_R(\vec{\mu}, \vec{D}) = U_R(\theta, D) = -\mu_0 D \cos \theta - \frac{\alpha D^2}{2}$$

Pour calculer $\delta_M(T, V, D)$, on somme sur tous les vecteurs $\vec{\mu} = \vec{\mu}_0 + \vec{\mu}_i$ possibles :

$$\delta_M(T, V, D) = \sum_{\vec{\mu}} q(T, V) e^{-U_R(\vec{\mu}, \vec{D})/kT}$$

Cette somme est en fait une intégrale : toutes les orientations de $\vec{\mu}$ sont permises :

$$\delta_M(T, V, D) = \int_0^{2\pi} d\varphi \int_0^\pi q(T, V) \exp \frac{\mu_0 D \cos \theta + \frac{1}{2} \alpha D^2}{kT} \sin \theta d\theta$$

On trouve, en remarquant que $\delta_M^0(T, V) = \delta_M(T, V, 0)$ est égal à $4\pi q(T, V)$:

$$\delta_M(T, V, D) = \delta_M^0(T, V) \frac{\sinh y}{y} e^{\alpha D^2/2kT} \quad \text{avec} \quad y = \frac{\mu_0 D}{kT}$$

$$\text{et donc} \quad \Delta_M(T, V, D, N) = \frac{\delta_M^0(T, V)^N}{N!} \left(\frac{\sinh y}{y} \right)^N e^{N \alpha D^2/2kT}$$

$$\text{puis} \quad \Phi_M(T, V, D, N) = \Phi_M^0(T, V, N) - N \frac{\alpha D^2}{2} - NkT \ln \frac{\sinh y}{y}$$

$$\text{avec} \quad \Phi_M^0(T, V, N) = -kT \ln \frac{\delta_M^0(T, V)^N}{N!}$$

On en déduit par dérivations toutes les propriétés thermodynamiques du système. Nous nous limiterons au calcul du moment dipolaire, de l'entropie et de la capacité thermique à V et D constants. La fonction $\Phi_M^0(T, V, N)$ sera supposée connue. Ses contributions à l'entropie, à l'énergie interne et à la capacité thermique seront notées S_0 , E_0 , C_V^0 . Ce sont les deux autres termes qui nous intéressent.

Moment dipolaire.

De $d\Phi_M = -S dT - p dV - M dD + \mu dN$ on déduit :

$$M = - \left. \frac{\partial \Phi_M}{\partial D} \right]_{T,V,N} = N(\alpha D + \mu_0 \mathcal{L}(y)) \quad \text{avec} \quad \mathcal{L}(y) = \coth y - \frac{1}{y}$$

$\mathcal{L}$ est la fonction de Langevin. Son graphe est représenté figure 4.8. On en déduit la permittivité ε du milieu :

$$D = \mathcal{E} + \frac{P}{\varepsilon_0} = \frac{D}{\varepsilon} + \frac{M}{\varepsilon_0 V} = D \left(\frac{1}{\varepsilon} + \frac{N\alpha}{\varepsilon_0 V} + \frac{N\mu_0^2 \mathcal{L}(y)}{\varepsilon_0 V y} \right)$$

$$\Rightarrow \quad \frac{1}{\varepsilon} = 1 - \frac{\rho}{\varepsilon_0} \left(\alpha + \mu_0^2 \frac{\mathcal{L}(y)}{y} \right) \quad \text{avec} \quad \rho = \frac{N}{V} \quad \text{et} \quad y = \frac{\mu_0 D}{kT}$$

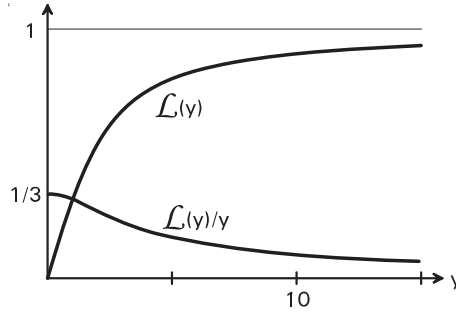


FIG. 4.8 – Fonction de Langevin $\mathcal{L}(y) = \coth y - \frac{1}{y}$

Étudions les deux cas extrêmes.

- *En champ fort ou à basse température*, $\mathcal{L}(y)$ tend vers 1. On obtient alors $\bar{\mu} \approx \alpha D + \mu_0$. Tous les moments sont alignés à $\vec{D}$. La rotation est bloquée.
- *Sous faible champ ou à haute température*, y est petit. On peut dans ces conditions utiliser les développements limités :

$$\frac{\sinh y}{y} \approx 1 + \frac{y^2}{6} \quad \text{et} \quad \mathcal{L}(y) \approx \frac{y}{3} - \frac{y^3}{45}$$

Le moment dipolaire moyen par molécule devient :

$$\bar{\mu} = \alpha D + \mu_0 \mathcal{L}(y) \approx \alpha D + \frac{\mu_0^2 D}{3kT}$$

Le premier terme est le moment dipolaire induit. On peut vérifier que le deuxième correspond à une moyenne de la composante parallèle à $\vec{D}$ du moment dipolaire permanent $\vec{\mu}_0$, dans une situation où l'orientation de celui-ci est aléatoire :

$$\frac{\int_0^\pi (\mu_0 \cos \theta) e^{\mu_0 D \cos \theta / kT} \sin \theta d\theta}{\int_0^\pi e^{\mu_0 D \cos \theta / kT} \sin \theta d\theta} = \frac{\mu_0^2 D}{3kT}$$

Conclusion : à haute température ou sous faible champ, les rotateurs sont libres ; les orientations des moments dipolaires permanents sont distribués aléatoirement.

Entropie et capacité thermique.

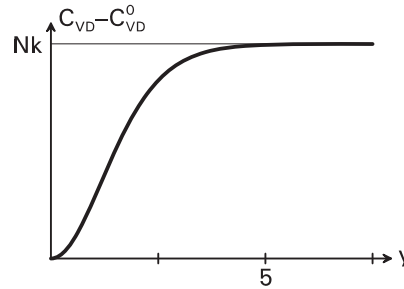


FIG. 4.9 – Milieux diélectriques : capacité thermique

L'entropie est obtenue par dérivation par rapport à T :

$$S = - \left. \frac{\partial \Phi_M}{\partial T} \right]_{V,D,N} = S_0 - Nk \left(y \mathcal{L}(y) - \ln \frac{\sinh y}{y} \right)$$

On vérifie facilement que S est toujours inférieure à S_0 : le champ électrique limite la liberté de mouvement de rotation des dipôles, et il y a diminution d'entropie.

Pour la capacité thermique C_{VD} , il faut exprimer le potentiel thermodynamique $E - D M$ en fonction de (T, V, D, N) .^[7] On trouve (figure 4.9) :

$$E - D M = \Phi_M + T S = \Phi_M^0 + T S_0 - \frac{N\alpha D^2}{2} - N\mu_0 D \mathcal{L}(y)$$

$$C_{VD} = - \left. \frac{\partial(E - DM)}{\partial T} \right]_{V,D,N} = C_{VD}^0 + Nk \left(1 - \frac{y^2}{\sinh^2 y} \right)$$

⁷ Pour C_{VM} , on exprime E en fonction de (T, V, M, N) ; pour C_{pM} , on exprime $E + pV$ en fonction de (T, p, M, N) ; pour C_{pD} , on exprime $H = E + pV - DM$ en fonction de (T, p, D, N) .

Chapitre 5

Les gaz parfaits quantiques

5.1 Les différents cas pratiques

On a vu (paragraphe 3.3.1) que le modèle haute température est valable si $\rho \Lambda^3 \ll 1$, avec $\Lambda = h/\sqrt{2\pi m kT}$. Cette condition est vérifiée dans la plupart des cas pratiques. On étudie ici deux systèmes de particules matérielles pour lesquels le modèle haute température n'est plus utilisable.

L'hélium est particulier pour deux raisons :

- c'est un atome de *faible masse* ($^3\text{He} : A = 3$, $^4\text{He} : A = 4$) ;
- c'est un gaz rare : *les interactions entre atomes sont extrêmement faibles*.

Les interactions sont si faibles que l'on peut considérer le gaz comme quasiment parfait jusqu'à l'état liquide. Les densités valent $\rho \approx 1,60 \cdot 10^{28}$ atomes par m^3 pour ^3He , et $\rho \approx 2,26 \cdot 10^{28}$ atomes par m^3 pour ^4He . Moyennant quoi, la condition $\rho \Lambda^3(T) \ll 1$ s'écrit $T \gg 6 \text{ K}$ pour les deux isotopes (voir équation 3.2). Sachant que les températures de liquéfaction valent 3,2 K (^3He) et 4,2 K (^4He), le modèle classique décrit au paragraphe 3.3.1 ne convient plus à basse température, les effets quantiques devenant de plus en plus importants au fur et à mesure que la température décroît. À noter que, bien que faibles, les interactions classiques à l'état liquide existent. Les résultats ne pourront donc être que qualitatifs.

Autre propriété à prendre en compte : le spin. L'hélium 3 est un atome de spin $1/2$ (c'est donc un **fermion**) alors que l'hélium 4 est un atome de spin nul (c'est donc un **boson**). A basse température, les deux isotopes auront des comportements radicalement différents.

L'électron : dans les métaux dits normaux, le système constitué des électrons de conduction peut être considéré comme un gaz parfait. C'est le modèle du gaz d'électrons libres de Fermi. Les principales caractéristiques de ce gaz :

- *très faible masse* de la particule ($m \approx 10^{-30} \text{ kg}$), spin $1/2$;
- *haute densité du gaz* ($\rho \approx 3 \cdot 10^{28} \text{ e}^-/\text{m}^3$).

La condition $\rho \Lambda^3(T) \ll 1$ s'écrit ici $T \gg 50\,000 \text{ K}$. Jusqu'à la fusion du métal, la température sera considérée comme très basse : le gaz d'électrons est dit dégénéré.

Il existe également des gaz purement quantiques, constitués de particules immatérielles (particules de masse nulle comme les photons ou de phonons). Ce type de gaz, au traitement particulier, a fait l'objet des paragraphes 2.4 et 2.5.

5.2 L'hélium à haute température

Ce paragraphe est consacré aux gaz légers (comme l'hélium) à (relativement) haute température. Le gaz parfait sera tridimensionnel et décrit dans l'approximation du continuum. Les variables choisies comme fixées sont (T, V, N) , en fonction desquelles on exprimera les autres grandeurs (E, S, p, μ, C_V) . On rappelle les équations de base :

$$N = (2s + 1) 2\pi V \left(\frac{2m}{h^2} \right)^{3/2} \int_0^\infty \frac{\sqrt{\varepsilon} d\varepsilon}{e^{(\varepsilon-\mu)/kT} + \sigma} \quad (5.1)$$

$$E = \frac{3}{2} p V = (2s + 1) 2\pi V \left(\frac{2m}{h^2} \right)^{3/2} \int_0^\infty \frac{\varepsilon \sqrt{\varepsilon} d\varepsilon}{e^{(\varepsilon-\mu)/kT} + \sigma} \quad (5.2)$$

On ignore les degrés de liberté électroniques et nucléaires. Ils n'ont aucune influence dans le déroulement des calculs. On se reportera au paragraphe 4.1.1 pour savoir ce qu'il faut ajouter à μ et S pour en tenir compte.

5.2.1 Calcul du potentiel chimique

Dans l'étude du gaz parfait classique (paragraphe 3.3.1), nous avons tout simplement retiré σ du dénominateur. L'approximation adoptée ici (développement au premier ordre) est moins violente. On peut bien entendu étendre le raisonnement à des développements d'ordre plus élevé, mais le premier ordre suffit à distinguer le comportement des bosons de celui de fermions.

De l'expression ?? de N on doit calculer le **potentiel chimique** μ , de façon à pouvoir exprimer E en fonction de (T, V, N) .

On pose $\rho = \frac{N}{V}$, $x = \frac{\varepsilon}{kT}$ et $\gamma = -\frac{\mu}{kT}$.

En remplaçant dans l'expression 5.1, on trouve que γ doit vérifier :

$$\frac{2}{\sqrt{\pi}} \int_0^\infty \frac{\sqrt{x} dx}{e^{x+\gamma} + \sigma} = \varphi \quad \text{avec} \quad \varphi = \frac{\rho \Lambda^3(T)}{2s + 1} \quad \text{et} \quad \Lambda(T) = \frac{h}{\sqrt{2\pi m k T}} \quad (5.3)$$

φ est une grandeur sans dimension qui définit entièrement le gaz (la masse et le spin d'une particule et la densité et la température du gaz). Elle prend des valeurs très grandes à basse température ou à haute densité, et très faibles à haute température ou à basse densité. Dans un gaz parfait classique, on a $\varphi = \lambda = e^{-\gamma}$ rigoureusement.

À haute température, $\gamma = -\mu/kT$ est très grand. On a donc $e^{-(x+\gamma)} = \lambda e^{-x}$ petit pour tout x . On réécrit la relation 5.3 en faisant apparaître ce terme :

$$I(\lambda) = \varphi \quad \text{avec} \quad I(\lambda) = \frac{2\lambda}{\sqrt{\pi}} \int_0^\infty \frac{\sqrt{x} e^{-x} dx}{1 + \sigma \lambda e^{-x}} \quad (5.4)$$

On remplace maintenant $\frac{1}{1 + \sigma \lambda e^{-x}}$ par $\sum_{j=0}^\infty (-\sigma \lambda e^{-x})^j$.

On trouve, sachant $\frac{2}{\sqrt{\pi}} \int_0^\infty \sqrt{u} e^{-u} du = 1 \Rightarrow \frac{2}{\sqrt{\pi}} \int_0^\infty \sqrt{x} e^{-(j+1)x} dx = (j+1)^{-3/2}$:

$$I(\lambda) = \lambda \sum_{j=0}^\infty \frac{(-\sigma \lambda)^j}{(j+1)^{3/2}} \Rightarrow \varphi = I(\lambda) = \lambda - \frac{\sigma}{2\sqrt{2}} \lambda^2 + \frac{1}{3\sqrt{3}} \lambda^3 - \frac{\sigma}{4\sqrt{4}} \lambda^4 + \dots$$

φ et λ étant deux infiniment petits du même ordre, on accède à λ par inversion de série^[1]. On obtient :

$$\lambda = \varphi + \frac{\sigma}{2\sqrt{2}} \varphi^2 + \left(\frac{1}{4} - \frac{1}{3\sqrt{3}}\right) \varphi^3 - \frac{\sigma}{2\sqrt{2}} \left(\frac{1}{2\sqrt{2}} + \frac{5}{3\sqrt{3}} - \frac{5}{4\sqrt{4}}\right) \varphi^4 + \dots$$

Pour ce qui nous concerne, les deux premiers termes suffisent :

$$\varphi \approx \lambda - \frac{\sigma}{2\sqrt{2}} \lambda^2 \quad \lambda \approx \varphi + \frac{\sigma}{2\sqrt{2}} \varphi^2$$

$$\mu = kT \ln \lambda \approx kT \ln \varphi + kT \ln \left(1 + \frac{\sigma}{2\sqrt{2}} \varphi\right) \approx kT \ln \varphi + kT \frac{\sigma}{2\sqrt{2}} \varphi$$

5.2.2 Les autres grandeurs thermodynamiques

On réexprime d'abord E (5.2) en posant $\lambda = e^{\mu/kT}$, $x = \frac{\varepsilon}{kT}$ et $\varphi = \frac{\rho \Lambda^3}{2s+1}$:

$$E = \frac{NkT}{\varphi} J(\lambda) \quad \text{avec} \quad J(\lambda) = \frac{2\lambda}{\sqrt{\pi}} \int_0^\infty \frac{x^{3/2} e^{-x} dx}{1 + \sigma \lambda e^{-x}}$$

Avec $\frac{2}{\sqrt{\pi}} \int_0^\infty u^{3/2} e^{-u} du = 3/2$, on trouve par un calcul similaire à celui de $I(\lambda)$:

$$J(\lambda) = \frac{3}{2} \lambda \sum_{j=0}^\infty \frac{(-\sigma \lambda)^j}{(j+1)^{5/2}} = \frac{3}{2} \left(\lambda - \frac{\sigma}{4\sqrt{2}} \lambda^2 + \frac{1}{9\sqrt{3}} \lambda^3 - \frac{\sigma}{16\sqrt{4}} \lambda^4 + \dots \right)$$

Et on remplace λ par $\varphi + \frac{\sigma}{2\sqrt{2}} \varphi^2$. Le calcul au premier ordre conduit à :

$$E \approx \frac{3}{2} NkT \left(1 + \frac{\sigma}{4\sqrt{2}} \varphi \right)$$

On calcule enfin la capacité thermique C_V en dérivant E par rapport à T (attention en dérivant : φ dépend de T), la pression avec la relation $pV = \frac{2}{3}E$, et l'entropie grâce à $TS = E + pV - N\mu$; Ci-dessous le bilan complet de l'étude :

$$\begin{aligned} \lambda &\approx \varphi + \frac{\sigma}{2\sqrt{2}} \varphi^2 & \frac{\mu}{kT} &\approx \ln \varphi + \frac{\sigma \varphi}{2\sqrt{2}} \\ \frac{E}{NkT} &\approx \frac{3}{2} + \frac{3\sigma \varphi}{8\sqrt{2}} & \frac{pV}{NkT} &\approx 1 + \frac{\sigma \varphi}{4\sqrt{2}} \\ \frac{C_V}{Nk} &\approx \frac{3}{2} - \frac{3\sigma \varphi}{16\sqrt{2}} & \frac{S}{Nk} &\approx \left(\frac{5}{2} - \ln \varphi\right) + \frac{\sigma \varphi}{8\sqrt{2}} \end{aligned} \quad (5.5)$$

Dans chacune des six expressions, le premier terme (celui ne contenant pas σ) correspond au gaz parfait classique où les effets quantiques sont ignorés. Les termes contenant σ représentent les effets quantiques. L'expression de E montre que les effets quantiques font augmenter l'énergie des fermions ($\sigma = +1$) et diminuer celle des

¹ $y = x + ax^2 + bx^3 + cx^4 + \dots \Rightarrow x = y - ay^2 - (b - 2a^2)y^3 - (c - 5ab + 5a^3)y^4 - \dots$

bosons ($\sigma = -1$). On parle de *répulsion quantique pour les fermions*, et d'*attraction quantique pour les bosons*. Mais il ne faut pas pousser trop loin l'analogie avec les interactions classiques : celles-ci provoquent toujours une diminution d'entropie (qu'il y ait répulsion ou attraction entre particules), alors que la répulsion quantique chez les fermions provoque une augmentation d'entropie.

Pour quantifier ces interactions, on peut calculer la fonction de partition canonique à partir de l'énergie libre comme au paragraphe 3.3.2. On trouve ici :

$$F(T, V, N) = -kT \ln \left(\frac{1}{N!} \left(\frac{(2s+1)V}{\Lambda^3(T)} \right)^N e^{-N\varphi/4\sqrt{2}} \right)$$

$$\text{et donc } Q(T, V, N) = \frac{(2s+1)^N}{N! \Lambda^{3N}} V^N e^{-N\sigma\varphi/4\sqrt{2}}$$

Le terme $(2s+1)^N$ est ce que nous avons appelé Q_{quant} au paragraphe 3.6.2. On en déduit que $V^N e^{-N\sigma\varphi/4\sqrt{2}}$ est l'intégrale de configuration classique du système^[2] (voir équation 3.33). Soit U l'énergie potentielle classique censée représenter les interactions quantiques :

$$Z_N = \int \cdots \int dq_1 \dots dq_{3N} e^{-U/kT} = V^N e^{-N\sigma\varphi/4\sqrt{2}}$$

En remplaçant U par $N\bar{u}$, sachant que l'énergie moyenne d'interaction $\bar{u}$ par atome est indépendant de la position dans l'enceinte, et sachant que $dq_1 \dots dq_{3N}$ s'écrit en coordonnées cartésiennes $dx_1 dy_1 dz_1 \dots dx_N dy_N dz_N$ on trouve :

$$e^{-N\sigma\varphi/4\sqrt{2}} = e^{-N\bar{u}/kT} \left(\frac{1}{V} \int_V dx_1 dy_1 dz_1 \right) \times \dots \times \left(\frac{1}{V} \int_V dx_N dy_N dz_N \right) = e^{-N\bar{u}/kT}$$

On en déduit l'énergie moyenne d'interaction (pour $\bar{v}$, voir équation 3.5) :

$$\bar{u} = kT \frac{\sigma\varphi}{4\sqrt{2}} = \frac{\sigma}{4\sqrt{2}} \left(\frac{h^2}{2\pi m} \right)^{3/2} \frac{\rho}{(2s+1)\sqrt{kT}} = \frac{N}{2s+1} \times \frac{\sigma h^3}{4\pi^2 \sqrt{2} V m^2 \bar{v}}$$

Résultat tout à fait correct :

- $\bar{u}$ est du signe de σ : répulsion pour les fermions, attraction pour les bosons ;
- $|\bar{u}|$ augmente quand ρ augmente ou quand T diminue ;
- pour une particule dans un état de spin donné, $\frac{N}{2s+1}$ est le nombre de particules dans le même état de spin : les interactions ne concernent que deux particules dans le même état de spin ; c'est également ce qu'on trouve dans le traitement quantique de Hartree-Fock des interactions entre électrons ;

On peut également remarquer que $\bar{u}$ est inversement proportionnel à la vitesse moyenne $\bar{v}$: plus les particules sont rapides, plus les interactions sont faibles.

² plus exactement, c'est l'intégrale de configuration classique équivalente : on cherche ici à quoi ressemblent classiquement les interactions quantiques.

5.3 L'hélium 3 à basse température et l'électron

On pose maintenant $\sigma = +1$:

$$N = \int_0^\infty g(\varepsilon) f_{\text{FD}}(\varepsilon, T, \mu) d\varepsilon \quad E = \frac{3}{2} p V = \int_0^\infty \varepsilon g(\varepsilon) f_{\text{FD}}(\varepsilon, T, \mu) d\varepsilon \quad (5.6)$$

$$\text{avec } g(\varepsilon) = (2s+1) 2\pi V \left(\frac{2m}{h^2} \right)^{3/2} \sqrt{\varepsilon} \quad \text{et} \quad f_{\text{FD}}(\varepsilon, T, \mu) = \frac{1}{e^{(\varepsilon-\mu)/kT} + 1}$$

5.3.1 L'état fondamental

L'état fondamental est celui qu'on observerait au zéro absolu. Au zéro absolu, $f_{\text{FD}}(\varepsilon, 0, \mu)$ est une marche : elle vaut 1 pour tout ε inférieur à μ , et 0 pour tout ε supérieur à μ . Le potentiel chimique $\mu_{(T=0\text{K})}$ est le plus haut niveau d'énergie occupé. On l'appelle le niveau de Fermi et on le note ε_F :

$$f_{\text{FD}}^0(\varepsilon) = f_{\text{FD}}(\varepsilon, 0, \mu) = \begin{cases} 1 & \text{si } \varepsilon < \varepsilon_F \\ 0 & \text{si } \varepsilon > \varepsilon_F \end{cases}$$

On en déduit qu'au zéro absolu on a :

$$N = \int_0^{\varepsilon_F} g(\varepsilon) d\varepsilon \quad \text{et} \quad E = \frac{3}{2} p V = \int_0^{\varepsilon_F} \varepsilon g(\varepsilon) d\varepsilon$$

La première expression permet de calculer ε_F en fonction de $\rho = \frac{N}{V}$:

$$\varepsilon_F = \frac{h^2}{2m} \left(\frac{3\rho}{4\pi(2s+1)} \right)^{2/3} = \frac{\hbar^2}{2m} \left(\frac{6\pi^2\rho}{2s+1} \right)^{2/3}$$

On peut exprimer en fonction de ε_F la densité d'états, ainsi que l'énergie interne et la pression au zéro absolu :

$$g(\varepsilon) = \frac{3}{2} \frac{N}{\varepsilon_F} \sqrt{\frac{\varepsilon}{\varepsilon_F}} = g(\varepsilon_F) \sqrt{\frac{\varepsilon}{\varepsilon_F}} \quad \text{avec} \quad g(\varepsilon_F) = \frac{3}{2} \frac{N}{\varepsilon_F}$$

$$E_0 = N \bar{\varepsilon}_0 = \frac{3}{5} N \varepsilon_F \quad p_0 = \frac{2}{5} \rho \varepsilon_F$$

5.3.2 Le calcul de Sommerfeld

On réécrit les équations 5.6 en remplaçant $(2s+1) 2\pi V \left(\frac{2m}{h^2} \right)^{3/2}$ par $\frac{3}{2} \frac{N}{\varepsilon_F}$:

$$\int_0^\infty \frac{\sqrt{\varepsilon} d\varepsilon}{e^{(\varepsilon-\mu)/kT} + 1} = \frac{\varepsilon_F^{3/2}}{3/2} \quad \text{et} \quad \bar{\varepsilon} = \frac{E}{N} = \frac{3/2}{\varepsilon_F^{3/2}} \int_0^\infty \frac{\varepsilon^{3/2} d\varepsilon}{e^{(\varepsilon-\mu)/kT} + 1}$$

On est ainsi amené à calculer des intégrales du type :

$$\Psi(T, \mu) = \int_0^\infty \frac{\phi(\varepsilon) d\varepsilon}{e^{(\varepsilon-\mu)/kT} + 1} \quad \text{avec} \quad \phi(\varepsilon) = \sqrt{\varepsilon} \quad \text{ou} \quad \phi(\varepsilon) = \varepsilon^{3/2}$$

Le calcul qui suit, dû à Sommerfeld, permet de trouver les termes correctifs de $\Psi(T, \mu)$ par rapport à $\Psi(0, \mu) = \int_0^\mu \phi(\varepsilon) d\varepsilon$ et de calculer toutes les grandeurs thermodynamiques à basse température.

Pour $T \neq 0 K$, les termes correctifs s'écrivent :

$$\delta\Psi(T, \mu) = \Psi(T, \mu) - \Psi(0, \mu) = \int_0^\infty \phi(\varepsilon) \delta f_{\text{FD}}(\varepsilon, T, \mu) d\varepsilon$$

$$\text{avec } \delta f_{\text{FD}}(\varepsilon, T, \mu) = \begin{cases} f_{\text{FD}}(\varepsilon, T, \mu) - 1 & \text{si } \varepsilon < \mu \\ f_{\text{FD}}(\varepsilon, T, \mu) & \text{si } \varepsilon > \mu \end{cases}$$

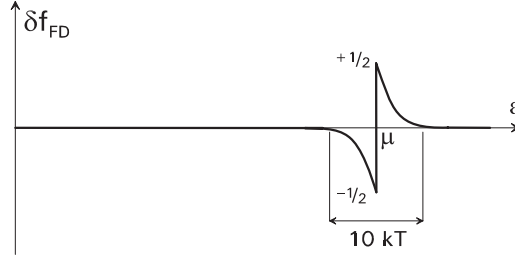


FIG. 5.1 – La fonction $\delta f_{\text{FD}}(\varepsilon, T, \mu)$

Cette fonction δf_{FD} (figure 5.1) a deux particularités :

- son graphe en fonction de ε possède un centre de symétrie : le point $(\mu, 0)$;
- elle prend des valeurs appréciables uniquement dans un domaine de quelques kT centré sur $\varepsilon = \mu$.

On peut donc sans commettre d'erreur appréciable remplacer la borne inférieure de l'intégrale par $-\infty$, et appliquer à la fonction ϕ un développement de Taylor au voisinage de μ .

On pose tout d'abord $x = (\varepsilon - \mu)/kT$: $f_{\text{FD}}(\varepsilon, T, \mu)$ et $\delta f_{\text{FD}}(\varepsilon, T, \mu)$ deviennent des fonctions à une seule variable $f(x)$ et $\delta f(x)$. On remarquera que la fonction δf est impaire :

$$f(x) = \frac{1}{e^x + 1} \quad \text{et} \quad \delta f(x) = \begin{cases} f(x) - 1 = \frac{-1}{e^{-x} + 1} & \text{si } x < 0 \\ f(x) = \frac{1}{e^x + 1} & \text{si } x > 0 \end{cases}$$

Les valeurs de x pour lesquelles $\delta f(x)$ est appréciable appartiennent à un domaine large de quelques unités centré sur 0.

La fonction $\delta\Psi$ s'écrit maintenant :

$$\delta\Psi(T, \mu) = kT \int_{-\infty}^0 \phi(\mu + kT x') \frac{-dx'}{e^{-x'} + 1} + kT \int_0^{+\infty} \phi(\mu + kT x) \frac{dx}{e^x + 1}$$

Ce qui donne, après changement de variable $x' = -x$:

$$\delta\Psi(T, \mu) = kT \int_0^{+\infty} [\phi(\mu + kT x) - \phi(\mu - kT x)] \frac{dx}{e^x + 1}$$

Si on se contente du premier ordre, le crochet vaut tout simplement $2kT x \phi'(\mu)$. Si ce n'est pas suffisant,^[3] on développe le crochet en puissances de $kT x$:

$$\phi(\mu + kT x) - \phi(\mu - kT x) = 2 \sum_{j=1}^{\infty} \frac{(kT x)^{2j-1}}{(2j-1)!} \phi^{(2j-1)}(\mu)$$

On en déduit :

$$\delta\Psi(T, \mu) = 2 \sum_{j=1}^{\infty} \left(\frac{(kT)^{2j-1}}{(2j-1)!} \phi^{(2j-1)}(\mu) \int_0^{+\infty} \frac{x^{2j-1} dx}{e^x + 1} \right)$$

Les intégrales restantes sont connues. Les trois premières valent :

$$\int_0^{+\infty} \frac{x dx}{e^x + 1} = \frac{\pi^2}{12} \quad \int_0^{+\infty} \frac{x^3 dx}{e^x + 1} = \frac{7\pi^4}{120} \quad \int_0^{+\infty} \frac{x^5 dx}{e^x + 1} = \frac{31\pi^6}{252}$$

Nous nous contenterons des deux premiers termes :

$$\delta\Psi(T, \mu) = \frac{\pi^2}{6} (kT)^2 \phi'(\mu) + \frac{7\pi^4}{360} (kT)^4 \phi'''(\mu)$$

On peut maintenant revenir à nos deux équations :

$$\int_0^{\infty} \frac{\sqrt{\varepsilon} d\varepsilon}{e^{(\varepsilon-\mu)/kT} + 1} = \frac{\varepsilon_F^{3/2}}{3/2} \quad \text{et} \quad \bar{\varepsilon} = \frac{E}{N} = \frac{3/2}{\varepsilon_F^{3/2}} \int_0^{\infty} \frac{\varepsilon^{3/2} d\varepsilon}{e^{(\varepsilon-\mu)/kT} + 1}$$

Pour la première équation, on applique ce qui précède à $\phi(\varepsilon) = \sqrt{\varepsilon}$:

$$\Psi(0, \mu) = \int_0^{\mu} \sqrt{\varepsilon} d\varepsilon = \frac{\mu^{3/2}}{3/2} \quad \delta\Psi(T, \mu) = \frac{\pi^2}{12} \frac{(kT)^2}{\sqrt{\mu}} + \frac{7\pi^4}{960} \frac{(kT)^4}{\mu^{5/2}} + \dots$$

$$\varepsilon_F^{3/2} = \frac{2}{3} (\Psi(0, \mu) + \delta\Psi(T, \mu)) = \mu^{3/2} \left(1 + \frac{\pi^2}{8} \left(\frac{kT}{\mu} \right)^2 + \frac{7\pi^4}{640} \left(\frac{kT}{\mu} \right)^4 + \dots \right)$$

On élève la série à la puissance $-2/3$, et on multiplie le tout par kT . On écrit ainsi l'infiniment petit $\frac{kT}{\varepsilon_F}$ en fonction de l'infiniment petit $\frac{kT}{\mu}$:

$$\frac{kT}{\varepsilon_F} = \frac{kT}{\mu} \left(1 - \frac{\pi^2}{12} \left(\frac{kT}{\mu} \right)^2 + \frac{\pi^4}{720} \left(\frac{kT}{\mu} \right)^4 + \dots \right)$$

On inverse la série (voir note au bas de la page 167) :

$$\frac{kT}{\mu} = \frac{kT}{\varepsilon_F} \left(1 + \frac{\pi^2}{12} \left(\frac{kT}{\varepsilon_F} \right)^2 + \frac{7\pi^4}{360} \left(\frac{kT}{\varepsilon_F} \right)^4 + \dots \right) \quad (5.7)$$

Et enfin, on divise par kT et on élève à la puissance -1 :

$$\mu = \varepsilon_F \left(1 - \frac{\pi^2}{12} \left(\frac{kT}{\varepsilon_F} \right)^2 - \frac{\pi^4}{80} \left(\frac{kT}{\varepsilon_F} \right)^4 + \dots \right) \quad (5.8)$$

³ c'est malheureusement le cas, afin de distinguer la capacité thermique et l'entropie

Pour la deuxième équation, on a $\phi(\varepsilon) = \varepsilon^{3/2}$:

$$\Psi(0, \mu) = \int_0^\mu \varepsilon^{3/2} d\varepsilon = \frac{\mu^{5/2}}{5/2} \quad \delta\Psi(T, \mu) = \frac{\pi^2}{4} (kT)^2 \sqrt{\mu} - \frac{7\pi^4}{960} \frac{(kT)^4}{\mu^{3/2}} + \dots$$

$$\bar{\varepsilon} = \frac{2/3}{\varepsilon_F^{3/2}} (\Psi(0, \mu) + \delta\Psi(T, \mu)) = \frac{4}{15} \frac{\mu^{5/2}}{\varepsilon_F^{3/2}} \left(1 + \frac{5\pi^2}{8} \left(\frac{kT}{\mu} \right)^2 - \frac{7\pi^4}{384} \left(\frac{kT}{\mu} \right)^4 + \dots \right)$$

En utilisant 5.7 et 5.8, on exprime E en fonction de (T, N, ε_F) :

$$E = N \bar{\varepsilon} = \frac{3}{5} N \varepsilon_F \left(1 + \frac{5\pi^2}{12} \left(\frac{kT}{\varepsilon_F} \right)^2 - \frac{\pi^4}{16} \left(\frac{kT}{\varepsilon_F} \right)^4 + \dots \right)$$

On calcule la capacité thermique C_V en dérivant E par rapport à T (sachant que ε_F n'est fonction que de N et V), la pression avec la relation $pV = \frac{2E}{3}$, et l'entropie grâce à $TS = E + pV - N\mu$:

$$p = \frac{2}{5} \rho \varepsilon_F \left(1 + \frac{5\pi^2}{12} \left(\frac{kT}{\varepsilon_F} \right)^2 - \frac{\pi^4}{16} \left(\frac{kT}{\varepsilon_F} \right)^4 + \dots \right)$$

$$C_V = Nk \frac{\pi^2}{2} \frac{kT}{\varepsilon_F} \left(1 - \frac{3\pi^2}{10} \left(\frac{kT}{\varepsilon_F} \right)^2 + \dots \right) \quad S = Nk \frac{\pi^2}{2} \frac{kT}{\varepsilon_F} \left(1 - \frac{\pi^2}{10} \left(\frac{kT}{\varepsilon_F} \right)^2 + \dots \right)$$

Au premier ordre, la loi $C_V(T)$ est linéaire : c'est ce qu'on constate également expérimentalement. La pression ne s'annule pas lorsque T tend vers 0. La pression résiduelle p_0 est appelée pression quantique. Elle est due au principe d'exclusion de Pauli : les fermions ne pouvant pas tous se mettre sur le niveau fondamental (vitesse nulle), les chocs contre les parois existent même au zéro absolu.

5.3.3 Conditions de validité

Le modèle **basse température** de Sommerfeld exige $kT \ll \varepsilon_F$, ou encore :

$$T \ll T_F \quad \text{avec} \quad T_F = \left(\frac{3}{4\pi} \right)^{2/3} \frac{h^2}{2mk} \left(\frac{\rho}{2s+1} \right)^{2/3} \approx 0,385 \frac{h^2}{2mk} \left(\frac{\rho}{2s+1} \right)^{2/3}$$

T_F est appelée température de Fermi. Plus la densité du gaz est faible, plus la température doit être faible également. Le système dans ces conditions constitue ce qu'on appelle un gaz de fermions dégénéré dans lequel les effets quantiques sont importants.

Le modèle **haute température** est valable lorsque $\frac{\rho \Lambda^3(T)}{2s+1} \ll 1$. Ce qui donne :

$$T \gg T_C \quad \text{avec} \quad T_C = \frac{1}{\pi} \frac{h^2}{2mk} \left(\frac{\rho}{2s+1} \right)^{2/3} \approx 0,318 \frac{h^2}{2mk} \left(\frac{\rho}{2s+1} \right)^{2/3}$$

Les effets quantiques deviennent insignifiants, et le gaz devient classique.

T_F et T_C ont même ordre de grandeur. Pour savoir si tel ou tel modèle est valable, **les deux tests sont équivalents.**

On peut, soit comparer la température de travail à la température de Fermi (fonction de la densité du gaz), soit comparer la distance moyenne entre particules (de l'ordre de $\rho^{-1/3}$) à la longueur thermique de de Broglie (fonction de la température).

Pour les électrons dans un métal ($T_F \approx 40\,000\text{ K}$), on peut à toute température ne garder dans les expressions que le premier terme. On en déduit, et c'est vérifié expérimentalement, que le potentiel chimique et la pression sont indépendants de T , et que la capacité thermique et l'entropie suivent une loi linéaire en T :

$$\mu = \varepsilon_F = \frac{h^2}{2m} \left(\frac{3\rho}{8\pi} \right)^{2/3} = \frac{\hbar^2}{2m} (3\pi^2 \rho)^{2/3} \quad (5.9)$$

$$E = N \bar{\varepsilon} = \frac{3}{5} N \varepsilon_F \quad p = \frac{2}{5} \rho \varepsilon_F \quad C_V = S = Nk \frac{\pi^2}{2} \frac{kT}{\varepsilon_F}$$

Remarque : Un cas plus exotique est celui des naines blanches. Ce sont des étoiles qui ont épuisé tout le combustible qui alimente les réactions nucléaires. Dans une étoile habituelle, la pression engendrée par ces réactions contrebalance les forces gravitationnelles, et la densité se stabilise à une valeur proche de celle de la terre. Dans une naine blanche, c'est la pression quantique du gaz d'électrons qui contrebalance les forces gravitationnelles. La température qui y règne est de l'ordre de 10^7 K . On pourrait penser qu'on a affaire à un gaz classique. Mais les mesures montrent que masse volumique et densité du gaz d'électrons atteignent des valeurs énormes ($\sim 5\text{ tonnes par cm}^3$ et $10^{36}\text{ e}^-/\text{m}^3$), et ils ne sont pas seuls ...), ce qui donne $T_F \sim 5 \cdot 10^9\text{ K}$, et par conséquent, à 10^7 K , un gaz d'électrons totalement dégénéré.

Pour l'hélium 3 ($T_F \approx 4,9\text{ K}$ à l'état liquide), les formules valables pour l'électron ne le sont ici qu'aux températures inférieure à $0,5\text{ K}$. Au delà, les développements limités sont nécessaires, avec de plus en plus de termes au fur et à mesure que la température augmente. De même avec le modèle haute température (paragraphe 5.2) où il faut de plus en plus de termes dans les développements au fur et à mesure que la température décroît. Aux températures intermédiaires, on peut constater que les deux modèles se rejoignent parfaitement^[4]. Les courbes obtenues se trouvent au paragraphe 5.4 en regard de celles trouvées pour l'hélium 4.

5.3.4 L'effet Pomerantchuk

L'hélium, en raison de sa faible masse atomique, présente la particularité de n'exister à l'état solide que sous forte pression (^4He : 25 bars, ^3He : 29 bars). En effet, le critère de Lindemann, bien vérifié expérimentalement, estime que la fusion d'un solide a lieu lorsque l'extension spatiale des vibrations atomiques devient supérieure au dixième de la distance interatomique. Or, l'extension spatiale varie de façon inversement proportionnelle à la racine carrée de $m\omega$. L'atome d'hélium étant très léger, cette extension spatiale est, aux pressions habituelles, comparable aux distances interatomiques : même à l'état fondamental, c'est-à-dire au zéro absolu,

⁴ à condition d'éviter les transitions gaz—liquide et liquide—solide. L'état solide ne peut être obtenu qu'au delà de 29 bars. Pour éviter la transition gaz—liquide, il faut se placer à des pressions supérieures à celle du point critique (1,15 bars). Le système est alors homogène à toute température.

l'extension spatiale des oscillateurs reste très supérieure au dixième de la distance interatomique : la solidification est donc impossible. Lorsqu'on augmente fortement la pression, la distance interatomique ne diminue que très peu. Par compte la pulsation ω augmente fortement : l'extension spatiale diminue et la solidification peut avoir lieu en dessous d'une certaine température.

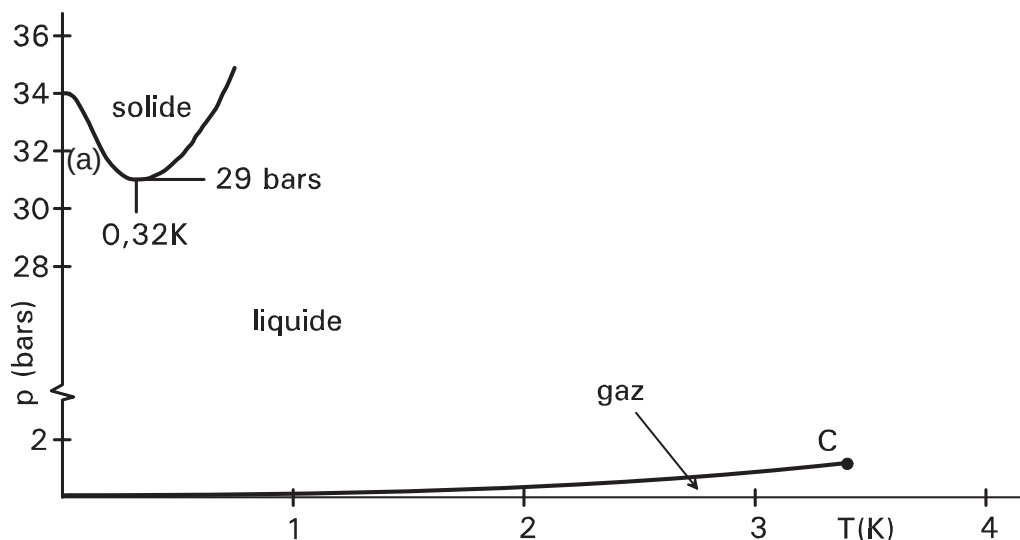


FIG. 5.2 – Diagramme p - T de l'hélium 3

Le diagramme p - T de l'hélium 3 représenté figure 5.2 fait apparaître une autre particularité de ce fluide : une courbe de solidification décroissante de 0 à 0,32 K. C'est un cas unique. Si, porté à 1 K et à 32 bars par exemple, on lui retire de la chaleur, sa température décroît jusqu'à environ 0,7 K où il passe de l'état liquide à l'état solide. Une fois la dernière goutte de liquide solidifiée, sa température décroît à nouveau jusqu'à 0,2 K, où il subit la transformation inverse : il repasse à l'état liquide.

Conclusion : dans la poche notée (a) sur le diagramme, l'entropie du liquide est plus faible que celle du solide. C'est l'effet Pomerantchuk. Ce phénomène qui semble anormal s'explique très simplement, du moins si l'on reste qualitatif.

D'abord, rappelons que :

- l'hélium 3 est un fermion de spin $1/2$;
- à l'état liquide, les atomes sont indiscernables : ils sont tous identiques et occupent en commun la même région de l'espace ;
- à l'état solide, ils occupent les sites d'un réseau : ils deviennent discernables.

Comparons l'enthalpie libre $G = E - TS + pV$ du système pour ces deux états à une pression p et une température T données.

- L'énergie interne est plus faible à l'état solide qu'à l'état liquide.
- L'entropie ne prenant en compte que l'état orbital tend vers 0 quand T tend vers zéro aussi bien à l'état liquide qu'à l'état solide. On pourra négliger son effet dans notre étude qualitative.
- La différence de volume entre la phase liquide et la phase solide est également négligeable.

C'est l'entropie de configuration des spins qui change radicalement lorsqu'on passe d'une phase à l'autre.

A l'état liquide, les N atomes (fermions) du système ont de plus en plus tendance à occuper les N états de plus basse énergie au fur et à mesure que la température diminue. Pour atteindre cet état fondamental, il est nécessaire que la moitié des fermions soient dans un état de spin donné, et l'autre moitié dans l'autre (de façon à pouvoir occuper chaque état orbital par deux fermions). Mais ces fermions étant indiscernables à l'état liquide, l'entropie de configuration des spins est nulle.

À l'état solide, les atomes sont localisés sur les sites d'un réseau cristallin. Ils deviennent discernables et peuvent donc tous occuper le même état vibratoire, ce qui a pour effet de faire chuter l'énergie interne. Par contre, l'entropie augmente fortement : en l'absence de champ magnétique, ils peuvent posséder de façon équiprobable les deux états de spin disponibles. Il y a pour le système 2^N possibilités, et donc une contribution à l'entropie valant $Nk \ln 2$ qui ne tend pas vers 0 quand T tend vers 0.

Ce fait n'est pas en contradiction avec le troisième principe : à température beaucoup plus basse (estimée à quelques dizaines de nanokelvins pour l'hélium), il faut tenir compte des interactions entre spins causant une mise en ordre de ceux-ci (et donc un champ magnétique induit), ne laissant au système qu'un seul état de spin possible. Il en résulte une entropie nulle au zéro absolu.

On peut se demander pourquoi on n'observe pas ce phénomène avec les autres fermions, en particulier pour les autres gaz rares qui ont une énergie de cohésion à l'état condensé équivalente à celle de l'hélium. L'entropie $S_L(T)$ à l'état liquide varie en T/T_F : rien que pour le néon, elle est, à température équivalente, sept fois plus forte que celle de l'hélium. Partant du zéro absolu, $S_L(T)$ devient supérieure à l'entropie S_C de l'état solide à température beaucoup plus basse, inférieure à la température de mise en ordre des spins qui annule S_C .

L'effet Pomerantchuk est à la base d'une technique permettant d'atteindre le millikelvin. On amène l'hélium 3 à la pression ordinaire à la température de 0,2 K (en pompant la vapeur au dessus du liquide^[5]). On comprime à T constant jusqu'à la courbe de solidification (32 bars). On l'isole et on comprime à nouveau. La solidification ne peut s'opérer puisque ce faisant le système doit libérer de la chaleur (il est isolé). Son point représentatif sur le diagramme p - T doit donc suivre la courbe de solidification. Toute autre substance au cours de cette étape verrait sa température augmenter. Pour l'hélium 3 c'est l'inverse : sa température diminue. On peut ainsi descendre aux alentours du millikelvin.

⁵ C'est une méthode classique ; au dessus du liquide se forme une vapeur dont la pression à saturation est fonction croissante de la température ; si on pompe cette vapeur, la pression diminue, et donc la température également ; on peut ainsi ajuster la température de façon continue depuis le point critique (5,2 K pour l'hélium 4 ordinaire, 3,2 K pour l'hélium 3) jusqu'au point triple du fluide ; l'hélium n'ayant pas de point triple, on peut en principe atteindre des températures quasi-nulles, mais la température minimale accessible dépend de la puissance de la pompe qui doit extraire la vapeur plus vite qu'elle ne se reforme ; avec l'hélium ordinaire, il n'est pas possible de descendre en-dessous de 0,7 K (sa pression de vapeur saturante décroît extrêmement rapidement avec la température) ; avec l'hélium 3, il est possible d'atteindre 0,2 K (à 0,5 K par exemple, sa pression de vapeur saturante est 10 000 fois plus importante que celle de l'hélium 4).

5.3.5 Le réfrigérateur à dilution d'hélium 3 dans l'hélium 4

Contrairement à la technique utilisant l'effet Pomerantchuk, le réfrigérateur à dilution d'hélium 3 dans l'hélium 4 (mis au point en 1960) permet une production continue de froid jusqu'à environ 2 millikelvins. En voici le principe.

Lorsqu'on mélange de l'hélium 3 et de l'hélium 4, on obtient un fluide qui se liquéfie, à pression ordinaire, à une température comprise entre 3,2 et 4,2 K dépendant de la concentration. Dans cette gamme de température, les deux isotopes sont miscibles en toutes proportions. Lorsqu'on abaisse la température, on constate qu'à 0,87 K, le liquide se sépare en deux phases, l'une pauvre en hélium 3 au fond du récipient, l'autre riche en hélium 3 plus légère qui surnage. Au fur et à mesure que la température s'abaisse, la phase riche supérieure s'enrichit et la phase pauvre inférieure s'appauvrit. Quand T tend vers 0, la phase supérieure tend vers de l'hélium 3 pur, alors que la phase inférieure ne tend pas vers de l'hélium 4 pur : il y restera toujours au moins 6,4 % d'atomes d'hélium 3.

Dans la phase inférieure, pauvre en hélium 3, du fait de la superfluidité de l'hélium 4 dans cette gamme de température (voir paragraphe 5.4.5), les atomes d'hélium 3 ont le comportement d'un gaz (l'habituelle pression doit simplement être remplacée par une pression osmotique). L'hélium 3 étant un fermion, cette pression ne s'annule pas au zéro absolu (la pression résiduelle vaut environ 15 millibars). Il est donc envisageable d'extraire ce que nous appellerons le quasi-gaz et de créer du froid même très près du zéro absolu.

À très basse température, le quasi-gaz d'hélium 3 de la phase inférieure est en équilibre avec l'hélium 3 à peu près pur de la phase supérieure : tout se passe comme dans un équilibre liquide - gaz habituel, à ceci près qu'ici la phase liquide flotte au dessus de la phase gazeuse. Un échange constant d'atomes se produit à travers la surface séparant les deux phases : "l'évaporation" (vers le bas) s'accompagne d'une absorption de chaleur, alors que la "condensation" (vers le haut) s'accompagne d'un dégagement de chaleur.

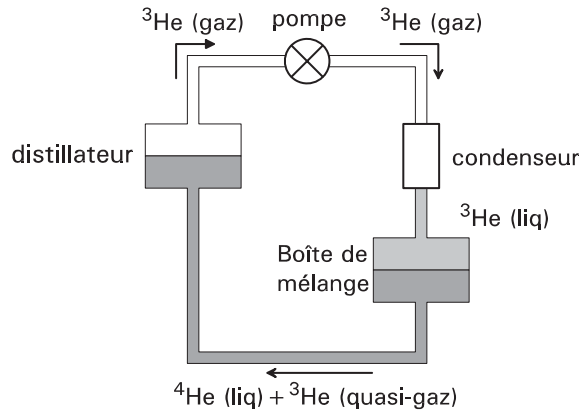


FIG. 5.3 – Schéma de principe de la pompe à dilution

Le principe est donc le même que celui d'un réfrigérateur à évaporation - condensation classique. La figure 5.3 représente le schéma de l'installation. L'hélium 3 sortant de la pompe est injecté sous pression à l'intérieur du condenseur (un bain d'hélium 4 à 1 K), où il est liquéfié et refroidi. Il est ensuite acheminé vers la boîte

de mélange où il est mis en contact avec le liquide inférieur contenant le quasi-gaz d'hélium 3. Grâce au pompage, celui-ci se trouve appauvri en hélium 3 dans le distillateur (1 % dans le distillateur pour 6,4 % dans la boîte de mélange). Il s'établit ainsi un flux d'hélium 3 de la boîte de mélange vers le distillateur, ce qui provoque "l'évaporation" (vers le bas) attendue dans la boîte de mélange.

5.4 L'hélium 4 à basse température

On pose maintenant $\sigma = -1$ et $s = 0$:

$$N = \int_0^\infty g(\varepsilon) f_{\text{BE}}(\varepsilon, T, \mu) d\varepsilon \quad E = \frac{3}{2} p V = \int_0^\infty \varepsilon g(\varepsilon) f_{\text{BE}}(\varepsilon, T, \mu) d\varepsilon \quad (5.10)$$

$$\text{avec } g(\varepsilon) = 2\pi V \left(\frac{2m}{h^2} \right)^{3/2} \sqrt{\varepsilon} \quad \text{et} \quad f_{\text{BE}}(\varepsilon, T, \mu) = \frac{1}{e^{(\varepsilon-\mu)/kT} - 1}$$

5.4.1 Température de Bose, volume atomique de Bose

Même principe qu'avec les fermions : de l'expression de N on tire le potentiel chimique. Mais ici, un problème se pose.

En posant $x = \frac{\varepsilon}{kT}$ et $\gamma = -\frac{\mu}{kT}$, on doit trouver γ tel que :^[6]

$$I(\gamma) = \varphi \quad \text{avec} \quad I(\gamma) = \frac{2}{\sqrt{\pi}} \int_0^\infty \frac{\sqrt{x} dx}{e^{x+\gamma} - 1} \quad \text{et} \quad \varphi = \rho \Lambda^3(T) = \frac{\Lambda^3(T)}{v}$$

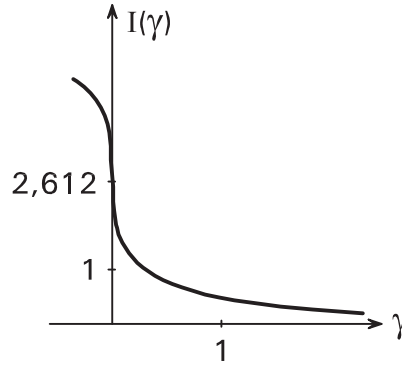


FIG. 5.4 – Graphe de la fonction $I(\gamma)$

On sait que le potentiel chimique est toujours inférieur à l'énergie du niveau fondamental. Étant donné que nous avons placé l'origine des énergies sur celui-ci, on doit donc avoir $\mu < 0$ et par conséquent $\gamma > 0$. Vu l'allure du graphe $I(\gamma)$ (figure 5.4), on ne pourra trouver une valeur positive à γ que si :

$$\varphi = \frac{\Lambda^3(T)}{v} = \frac{h^3}{v (2\pi m kT)^{3/2}} \leq 2,612$$

⁶ Dans ce paragraphe, les résultats seront exprimés, non pas en fonction de la densité $\rho = N/V$ mais en fonction du volume atomique $v = V/N$.

On peut envisager deux types d'expérience :

- travailler à **volume atomique constant** et faire varier la température pour tracer des isochores sur le diagramme p - T ; la condition $\varphi \leq 2,612$ devient :

$$T \geq T_B(v) \quad \text{avec} \quad T_B(v) = \frac{h^2}{2\pi m kT (2,612 v)^{2/3}} \quad \text{ou encore}$$

$$\Lambda(T) \leq \Lambda_B(v) \quad \text{avec} \quad \Lambda_B(v) = (2,612 v)^{1/3} \quad \text{et} \quad \Lambda(T) = \frac{h}{\sqrt{2\pi m kT}}$$

- travailler à **température constante** et faire varier le volume pour tracer des isothermes sur le diagramme p - v ; la condition $\varphi < 2,612$ s'écrit alors :

$$v \geq v_B(T) \quad \text{avec} \quad v_B(T) = \frac{\Lambda^3(T)}{2,612} = \frac{h^3}{2,612 (2\pi m kT)^{3/2}}$$

T_B est la **température de Bose**, v_B est le **volume atomique de Bose**.

Dans ce qui suit, nous raisonnerons à volume atomique constant en faisant varier la température. Le volume atomique de Bose n'aura son utilité qu'au moment du tracé du diagramme p - v .

Les démonstrations seront effectuées en conservant φ dans les expressions. On remplacera celui-ci :

- par $2,612 \frac{\Lambda^3(T)}{\Lambda_B^3(v)} = 2,612 \left(\frac{T_B(v)}{T} \right)^{3/2}$ lorsqu'on dérivera par rapport à T ,
- par $2,612 \frac{v_B(T)}{v}$ lorsqu'on dérivera par rapport à V ou N .

La température de Bose sépare deux domaines de température dans lesquels le système possède des propriétés totalement différentes. Pour l'hélium 4 à l'état liquide, on trouve $T_B = 3,20$ K.

Les températures très supérieures à T_B ont été traitées au paragraphe 5.2.2. Les résultats que nous y avons trouvés (équations 5.5, avec ici $\sigma = -1$ et $s = 0$) ont été obtenus à partir de développements limités du premier ordre. En poussant de plus en plus loin les développements limités, on peut en principe utiliser ce modèle jusqu'à des températures très voisines de T_B . Mais il existe un modèle mieux adapté aux températures légèrement supérieures à T_B .

5.4.2 Températures légèrement supérieures à T_B

Aux températures proches de T_B , γ tend vers 0 par valeurs supérieures, et $I(\gamma)$ tend vers $I(0) = 2,612$ par valeurs inférieures. Un calcul au premier ordre plutôt fastidieux donne :

$$\varphi = I(\gamma) = \frac{2}{\sqrt{\pi}} \int_0^\infty \frac{\sqrt{x} dx}{e^{x+\gamma} - 1} \approx 2,612 - \sqrt{4\pi} \gamma \quad \Rightarrow \quad \gamma = -\frac{\mu}{kT} \approx \frac{(2,612 - \varphi)^2}{4\pi}$$

Pour le calcul de l'énergie interne E , on repart de :

$$E = 2\pi V \left(\frac{2m}{h^2} \right)^{3/2} \int_0^\infty \frac{\varepsilon^{3/2} d\varepsilon}{e^{(\varepsilon-\mu)/kT} - 1} = \frac{NkT}{\varphi} J(\gamma) \quad \text{avec} \quad J(\gamma) = \frac{2}{\sqrt{\pi}} \int_0^\infty \frac{x^{3/2} dx}{e^{x+\gamma} - 1}$$

On pose $\delta J(\gamma) = J(\gamma) - J(0) \approx J(\gamma) - 2,0115$. Au premier ordre, on a :

$$\delta J(\gamma) \approx \gamma \left. \frac{dJ}{d\gamma} \right|_{\gamma=0} = \gamma \frac{2}{\sqrt{\pi}} \int_0^\infty x^{3/2} \frac{-e^x}{(e^x - 1)^2} dx = \gamma \frac{2}{\sqrt{\pi}} \int_0^\infty x^{3/2} \frac{d}{dx} \left(\frac{1}{e^x - 1} \right) dx$$

Après intégration par partie, on obtient :

$$\delta J(\gamma) \approx -\frac{3}{2} \gamma \frac{2}{\sqrt{\pi}} \int_0^\infty \frac{\sqrt{x} dx}{e^x - 1} \approx -3,918 \gamma \text{ et donc } E = NkT \frac{2,0115 - 3,918 \gamma}{\varphi}$$

On remplace γ par $\frac{(2,612 - \varphi)^2}{4\pi}$ et on trouve :

$$E \approx NkT \frac{2,0115 - 0,3118 (2,612 - \varphi)^2}{\varphi} \approx 1,6288 - \frac{0,1158}{\varphi} - 0,3118 \varphi$$

On remplace maintenant φ par $2,612 \left(\frac{T_B(v)}{T} \right)^{3/2}$:

$$E = \frac{3}{2} pV \approx NkT \left(1,6288 - 0,0443 \left(\frac{T}{T_B} \right)^{3/2} - 0,8144 \left(\frac{T_B}{T} \right)^{3/2} \right)$$

On calcule C_V en dérivant E par rapport à T et S grâce à $TS = E + pV - N\mu$:

$$C_V \approx Nk \left(1,6288 - 0,1107 \left(\frac{T}{T_B} \right)^{3/2} + 0,4072 \left(\frac{T_B}{T} \right)^{3/2} \right)$$

$$S \approx Nk \left(3,2575 - 0,0738 \left(\frac{T}{T_B} \right)^{3/2} - 1,0858 \left(\frac{T_B}{T} \right)^{3/2} + 0,5429 \left(\frac{T_B}{T} \right)^3 \right)$$

5.4.3 Températures inférieures à T_B

Partant de $T > T_B$, lorsqu'on diminue la température, le potentiel chimique (négatif) augmente jusqu'à devenir nul à $T = T_B$. La fonction $I(\gamma)$ atteint alors son maximum 2,612. Dès que T devient inférieure à T_B , φ devient supérieur à 2,612, et l'équation $I(\gamma) = \varphi$ n'a plus de solution pour γ .

On sait que, lorsque T diminue, les atomes ont une tendance de plus en plus marquée à se placer dans leur état fondamental d'énergie nulle. On peut calculer le potentiel chimique en fonction du nombre $N_{(\text{fond})}$ d'atomes dans l'état fondamental :

$$N_{(\text{fond})} = \frac{1}{e^{-\mu/kT} - 1} \quad \Rightarrow \quad \frac{\mu}{kT} = -\ln \left(1 + \frac{1}{N_{(\text{fond})}} \right)$$

À haute température, $\mu/kT \rightarrow -\infty$ et donc $N_{(\text{fond})}$ est quasi nul. Plus la température diminue, plus le potentiel chimique s'approche de 0, et plus $N_{(\text{fond})}$ augmente. L'approximation du continuum ne peut pas rendre compte de cette population puisque *la densité d'états est nulle pour $\varepsilon = 0$* . Tant que T est supérieure à T_B , $N_{(\text{fond})}$ est négligeable : le fait de les oublier n'affecte pas le résultat. En dessous d'une certaine température (très proche de T_B), le nombre d'atomes à l'état fondamental

devient macroscopique (on sait que $N_{(\text{fond})}$ doit atteindre la valeur $N_{(\text{fond})} = N$ au zéro absolu). Dès que $N_{(\text{fond})}$ devient macroscopique, le potentiel chimique devient, en valeur absolue, parfaitement négligeable devant kT (figure 5.5) :

$$\frac{\mu}{kT} = -\ln\left(1 + \frac{1}{N_{(\text{fond})}}\right) \approx -\frac{1}{N_{(\text{fond})}} \approx 0$$

À noter qu'il n'est pas nécessaire que $N_{(\text{fond})} = N$ soit d'ordre macroscopique pour obtenir ce résultat. Un millionième de millionième de mole suffit ($\sim 10^{12}$ atomes). Ce nombre est atteint à une température extrêmement voisine de T_B .

Conclusion : **Le potentiel chimique reste rigoureusement nul entre 0 et T_B .**

C'est là une caractéristique de particules, comme les photons, qui ont la faculté d'apparaître ou disparaître à volonté. Nous avons ici affaire à des particules matérielles. Bien entendu, elles ne disparaissent pas physiquement, mais elles disparaissent du calcul : les atomes dans l'état fondamental ne participent plus à l'énergie. Ils peuvent toujours subir des chocs de la part des autres atomes encore en mouvement, et donc passer dans un état excité, mais lorsqu'on abaisse la température la probabilité qu'un tel événement arrive est extrêmement faible par rapport à celle qu'ont les autres atomes à passer dans l'état fondamental. Ce phénomène est connu sous le nom de **condensation de Bose**. Mais attention, ce n'est pas une condensation habituelle :

- Elle s'effectue, non pas dans l'espace habituel, mais dans l'espace réciproque (espace des vitesses). Il n'y a pas deux phases distinctes. Les atomes dans l'état fondamental sont disséminés parmi les autres atomes. Le système reste toujours homogène, contrairement à ce qui se passe lors d'une transformation gaz–liquide, où les densités des phases liquide et gazeuse sont différentes.
- Habituellement, lorsqu'on retire de la chaleur à un gaz, la température s'abaisse jusqu'à ce qu'une première goutte de liquide apparaisse. La température s'arrête alors de baisser, et ce n'est que lorsque la transformation gaz–liquide est complète que la température peut à nouveau s'abaisser. La condensation de Bose, par contre, a lieu de façon continue de T_B à 0 K, sans interruption de la baisse de température. Nous en reparlerons au paragraphe 5.4.4.

À noter que la situation est très inhabituelle : un ensemble macroscopique de particules est *dans un état quantique pur*. Les fluctuations de $N_{(\text{fond})} = N$ sont donc extrêmement faibles, d'où des effets importants à l'échelle macroscopique. L'utilisation de la densité d'états reste correcte pour les états excités. Seul l'état fondamental est singulier. Si on note $N_{(\text{exc})}$ le nombre d'atomes excités, on a donc (avec $\mu = 0$) :

$$N_{(\text{exc})} = 2\pi V \left(\frac{2m}{h^2}\right)^{3/2} \int_0^\infty \frac{\sqrt{\varepsilon} d\varepsilon}{e^{\varepsilon/kT} - 1} = \frac{N}{\varphi} \frac{2}{\sqrt{\pi}} \int_0^\infty \frac{\sqrt{x} dx}{e^x - 1} = N \frac{2,612}{\varphi}$$

et donc (figure 5.5 à droite) :

$$N_{(\text{fond})} = N - N_{(\text{exc})} = N \left(1 - \frac{2,612}{\varphi}\right) = N \left(1 - \left(\frac{T}{T_B}\right)^{3/2}\right)$$

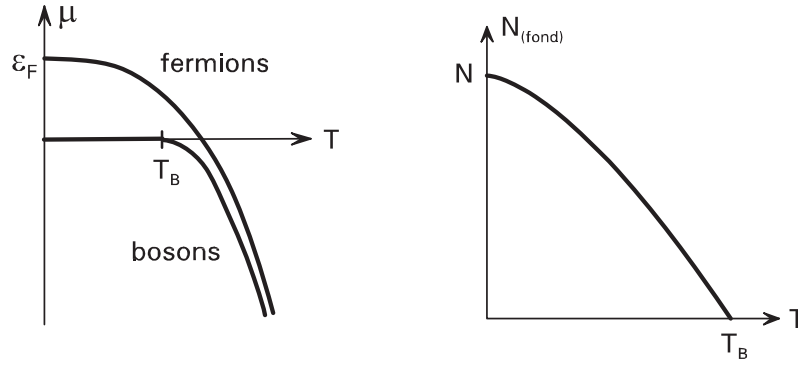


FIG. 5.5 – Potentiel chimique et nombre d'atomes "condensés"

Calculons maintenant l'énergie interne, la capacité thermique, la pression et l'entropie. Aux températures inférieures à T_B , le potentiel chimique est nul et seuls les atomes excités participent à l'énergie :

$$E = 2\pi V \left(\frac{2m}{h^2} \right)^{3/2} \int_0^\infty \frac{\varepsilon^{3/2} d\varepsilon}{e^{\varepsilon/kT} - 1} = \frac{VkT}{\Lambda^3} \frac{2}{\sqrt{\pi}} \int_0^\infty \frac{x^{3/2} dx}{e^x - 1} = 2,0115 \frac{VkT}{\Lambda^3} = 2,0115 \frac{NkT}{\Lambda^3/v}$$

Ce qui donne :

$$E = 2,0115 \frac{NkT}{\varphi} = 0,7701 NkT \left(\frac{T}{T_B} \right)^{3/2} = 0,7701 N_{\text{exc}} kT$$

$$C_V = \left. \frac{dE}{dT} \right]_{V,N} = 5,02875 \frac{Nk}{\varphi} = 1,9252 Nk \left(\frac{T}{T_B} \right)^{3/2} = 1,9252 N_{\text{exc}} k$$

$$S = \frac{E + pV}{T} = 3,3525 \frac{Nk}{\varphi} = 1,2835 Nk \left(\frac{T}{T_B} \right)^{3/2} = 1,2835 N_{\text{exc}} k$$

$$pV = \frac{2E}{3} = 1,3410 \frac{NkT}{\varphi} = 0,5134 NkT \left(\frac{T}{T_B} \right)^{3/2} = 0,5134 N_{\text{exc}} kT$$

On remarquera le comportement à peu près "normal" des atomes excités. La forte décroissance de E et de C_V lorsque T décroît est principalement due à la forte augmentation du nombre d'atomes condensés qui ne participent pas à l'énergie. On remarquera également que la pression ne dépend que de T . Elle sera donc constante sur les isothermes que nous tracerons sur le diagramme $p-v$.

En résumé, aux températures inférieures à T_B , il y a coexistence entre atomes condensés et atomes excités. Les atomes condensés ne participent pas à l'énergie totale du système. Dans nos hypothèses de départ, cette énergie est purement cinétique. Les atomes condensés sont donc immobiles et de ce fait ne participent pas à la pression (ils n'exercent aucune force sur les parois). Ils sont d'autre part dans un état quantique pur, non dégénéré : l'entropie du sous-système constitué des atomes condensés (indiscernables) est nulle. Le potentiel chimique étant nul, les atomes condensés sont virtuellement absents du récipient.

5.4.4 La courbe $C_V(T)$ et les diagrammes $p-v$ et $p-T$

Les expressions de la **capacité thermique** $C_V(T)$ trouvées dans les trois domaines de température permettent de tracer le graphe représenté sur la figure 5.6 à droite. Celui-ci présente un pic à la température de transition $T = T_B$. Le calcul de T_B donne, on l'a vu $T_B = 3,20$ K. Expérimentalement, ce pic est bien observé, mais à 2,17 K. Cet écart est à attribuer aux interactions entre atomes que nous avons négligées dans l'étude théorique.

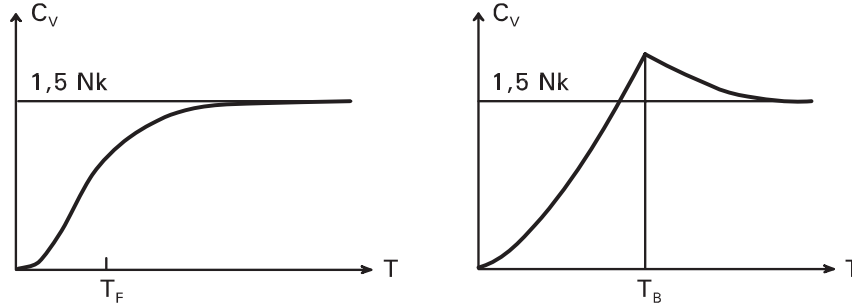


FIG. 5.6 – Capacité thermique : à gauche les fermions, à droite les bosons

Les unités habituelles des **diagrammes $p-v$ et $p-T$** sont le cm^3/mol pour le volume et le bar pour la pression. On réécrit donc tous les résultats utiles dans ce système d'unités.

– les grandeurs caractéristiques :

$$T_B(v) = \frac{28,63}{v^{2/3}} \quad v_B(T) = \frac{153,2}{T^{3/2}} \quad \varphi = \frac{400,2}{v T^{3/2}} \quad \frac{kT}{v} = 82,04 \frac{T}{v}$$

– la pression dans les trois domaines de température :

$$T \gg T_B \text{ ou } v \gg v_B \quad p \simeq 82,04 \frac{T}{v} \left(1 - \frac{70,74}{v T^{3/2}} \right)$$

$$T \gtrsim T_B \text{ ou } v \gtrsim v_B \quad p \simeq \frac{T}{v} \left(89,10 - \frac{6825}{v T^{3/2}} - 0,01585 v T^{3/2} \right)$$

$$T < T_B \text{ ou } v < v_B \quad p = 0,2749 T^{5/2}$$

Sur le diagramme $p-v$, on cherche d'abord l'équation de la courbe de transition séparant les deux domaines.

Dans le domaine $v < v_B$, la pression est indépendante de v : il y a donc sur le diagramme $p-v$ un plateau se terminant pour T donnée au point de coordonnées :

$$v_B(T) = \frac{153,2}{T^{3/2}} \quad \text{et} \quad p_B(T) = 0,2749 T^{5/2}$$

On en déduit l'équation de la courbe de transition : $p_B = \frac{1206}{v_B^{5/3}}$.

La condensation de Bose a lieu dans le domaine situé sous cette courbe.

Le diagramme p - v représenté figure 5.7 à gauche rassemble tous ces résultats. La courbe de transition (en pointillé) est l'ensemble des points (p_B, v_B) . Apparaissent également (en traits épais) les deux isothermes 3 K et 4 K. A droite on trouve le diagramme p - v d'un système "classique" (dans le sens où les effets quantiques sont indétectables) présentant une transition gaz-liquide.

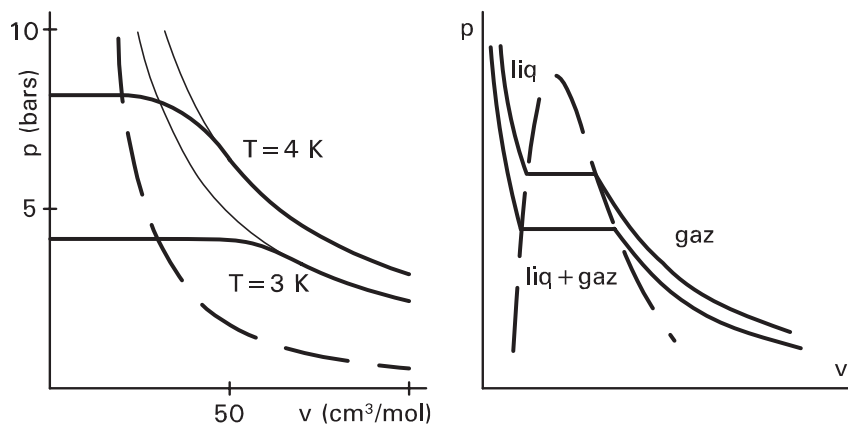


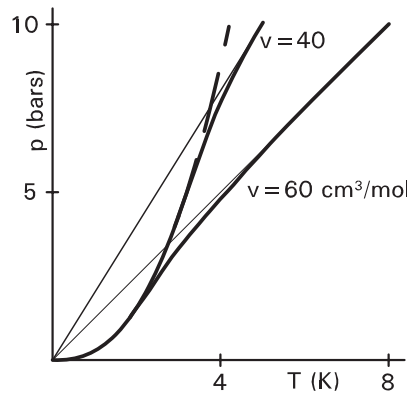
FIG. 5.7 – Diagrammes p - v : bosons à gauche, système classique à droite

Sur le diagramme p - v classique, les deux phases coexistent sous la courbe de saturation (en pointillé). A T constant (on suit une isotherme en trait plein), lorsque partant d'une très grande valeur on diminue le volume, la pression augmente jusqu'à ce que le point représentatif de l'état du système entre en contact avec la courbe de saturation. La première goutte de liquide apparaît alors, et la pression demeure constante (palier de liquéfaction) tant que la transformation n'est pas terminée. Celle-ci s'opère à pression constante, et s'achève à l'extrémité gauche du palier. Si on continue à réduire le volume, la pression augmente à nouveau de façon asymptotique : la valeur $v = 0$ n'est jamais atteinte.

Sur le diagramme p - v de Bose, lorsque le point représentatif de l'état du système est sous la courbe de transition, il y a également coexistence entre la phase condensée et la phase excitée. Mais ici, il n'y a pas séparation physique des deux phases : les atomes "condensés" sont disséminés dans le système. Dans la pratique, il ne peut y avoir transformation complète : on ne peut pas atteindre la valeur $v = 0$. Mais en théorie, c'est possible puisque les atomes sont considérés comme ponctuels et sans interaction.

Sur le **diagramme** p - T (figure 5.8), Les isochores (en traits épais) ont pour équations $p = kT/v$ à haute température et rejoignent tangentiellement la courbe de transition (en pointillé). Celle-ci, d'équation $p_B = 0,2749 T_B^{5/2}$, a la même allure que celle d'un corps classique ne présentant qu'une seule transition.

C'est dans l'utilisation du diagramme que réside la différence. Dans les deux cas, si par exemple on abaisse T à p constant, le point représentatif se déplace horizontalement jusqu'à atteindre la courbe de transition ou de liquéfaction. Habituellement, lorsque le point atteint la courbe de liquéfaction, la température s'arrête de baisser. Si on continue à retirer de la chaleur, le point représentatif stationne sur la courbe tant que la transformation gaz – liquide n'est pas complète. Lorsque le système est entièrement liquide, la température peut alors baisser à nouveau.

FIG. 5.8 – Diagrammes p - T d'un système de bosons

Sur le diagramme p - T de Bose, le point représentatif ne peut jamais se trouver au-dessus de la courbe de transition. Une fois la courbe de transition atteinte, le retrait de chaleur permet uniquement aux atomes de passer à l'état fondamental, ce qui a pour effet de diminuer la pression. Il est alors impossible de maintenir la pression constante. La température ne peut baisser que si la pression diminue également, et le point représentatif suit la courbe de transition sans jamais la traverser.

5.4.5 L'hélium 4 et la superfluidité

Mise en évidence expérimentale

À la température de 2,17 K (variant peu avec la pression), la capacité thermique (figure 5.9) et le coefficient de dilatation de l'hélium 4 présentent une discontinuité. C'est la signature d'une transition du second ordre (sans chaleur latente et sans coexistence entre les deux phases). On lui donna le nom de transition λ , du fait de la forme de la courbe $C_V(T)$. On baptisa hélium I la phase haute température et hélium II la phase basse température.

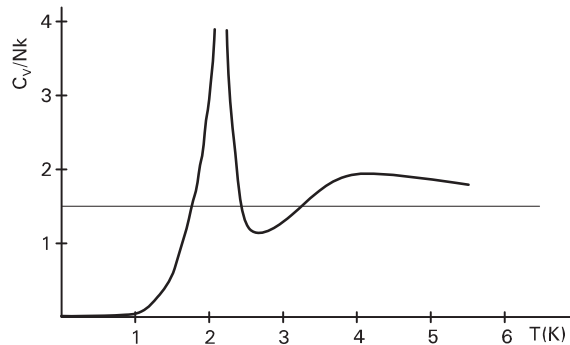


FIG. 5.9 – Capacité thermique de l'hélium 4

L'hélium I présente les propriétés habituelles des liquides. L'hélium II possède quant à elle des propriétés étonnantes n'appartenant à aucune autre substance. Mis à part cette transition, le diagramme p - T de l'hélium 4 (figure 5.10) a une allure voisine de celui de l'hélium 3.

On peut remarquer que la courbe de solidification est quasiment plate de 0 à 1 K : il existe en fait un minimum à peine marqué vers 0,8 K, et donc un très faible effet Pomerantchuk.

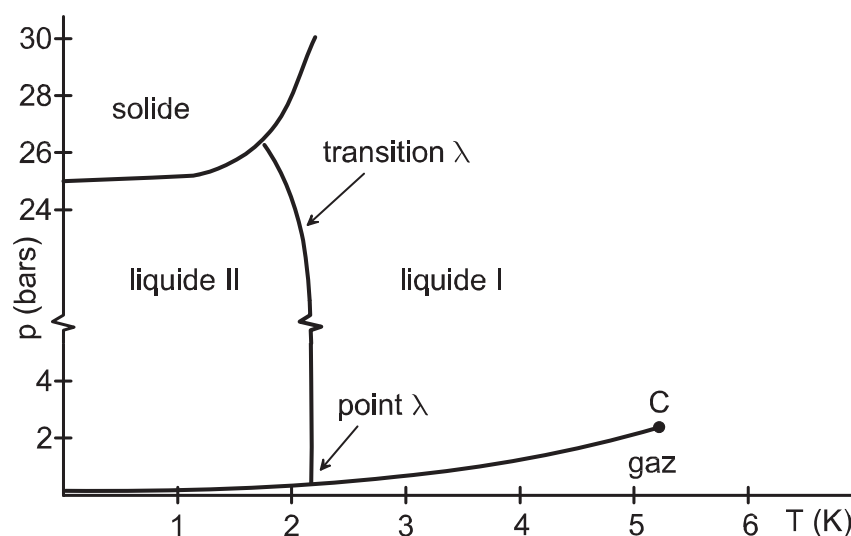


FIG. 5.10 – Diagramme p - T de l'hélium 4

Voici énumérés de façon chronologique les phénomènes "bizarres" apparaissant à la transition :

- Du point critique C au point λ , comme tout autre liquide en contact avec sa vapeur, l'hélium bout : il se vaporise en premier au fond du récipient où la température est plus élevée, d'où formation de bulles qui viennent éclater à la surface. En 1927, on s'aperçoit que, au point λ ^[7], le liquide s'arrête brusquement de bouillir. Interprétation (1936) : sa conductibilité thermique devient si élevée (2000 fois plus que celle du cuivre à 300 K), que plus aucun gradient de température n'est observé, et la vaporisation n'a lieu qu'en surface. L'expérience montre qu'il est impossible en dessous de cette température de créer des différences de température notables dans le liquide. On dit de lui que c'est un supraconducteur de chaleur.
- La viscosité de l'hélium II n'obéit pas aux lois classiques des fluides visqueux. Elle a une valeur normale dans certaines mesures (l'hélium II se comporte tout-à-fait normalement lorsqu'il s'écoule autour d'un objet), elle est nulle dans d'autres.

On constate d'abord (1935) que la viscosité de l'hélium II est inférieure à celle de l'hélium I. C'est déjà inhabituel : en général, la viscosité augmente quand la température diminue.

En 1937, le terme de "superfluide" apparaît : la viscosité de l'hélium II n'est pas mesurable. Il est capable de s'écouler entre deux disques polis et pressés l'un contre l'autre, dans des capillaires très fins ou dans des matières poreuses sans qu'il soit nécessaire de le comprimer : avec les moyens de l'époque, on en

⁷ Cette dénomination n'existait pas à l'époque : elle date de 1930, année du premier enregistrement de la courbe $C_V(T)$.

déduisit que la viscosité de l'hélium II était au moins 1000 fois plus faible que celle de l'hélium I.

On découvre ensuite qu'il grimpe aux parois du récipient : une éprouvette plongée partiellement dans un bain d'hélium II se remplit ou se vide lentement jusqu'à ce que les niveaux s'égalisent. Le transfert se fait par l'intermédiaire d'un film d'environ 30 nanomètres d'épaisseur qui recouvre toute surface en contact avec l'hélium. Ce film se déplace sans frottement à une vitesse d'environ 30 cm/s. On l'explique facilement : les interactions entre atomes sont si faibles qu'elles sont toujours inférieures aux interactions atome - paroi ; le liquide grimpe aux parois de façon à augmenter la surface de contact avec celle-ci. Cette ascension, dans tout autre fluide, serait bloquée par la viscosité.

- En 1938, on met en évidence l'effet thermomécanique, et F. London propose de relier toutes ces propriétés à la condensation de Bose, idée développée par L. Tisza qui, la même année, jette les bases du modèle à deux fluides.

Autre phénomène spectaculaire mis en évidence plus tard : lancé dans un anneau poreux, l'hélium superfluide ne subit à température constante aucun ralentissement observable en fonction du temps. C'est un cas de courant permanent tout comme celui des électrons dans un supraconducteur.

Le modèle à deux fluides

Selon le type d'expérience, l'hélium II se comporte, tantôt comme un fluide normal, tantôt comme un superfluide. C'est ce qui conduisit à considérer l'hélium II comme un mélange homogène de deux fluides, l'un "normal", l'autre superfluide. La condensation de Bose met également en évidence deux types d'atomes : les atomes condensés et les atomes excités. L'analogie entre les deux modèles est évidente. Mais il y a tout de même des différences. La théorie de Bose concerne les bosons indépendants, alors que l'hélium 4 en dessous de la transition est liquide. Les interactions classiques, bien que faibles, existent, avec pour principales conséquences :

- des températures de transition différentes (3,2 K en théorie, 2,17 K dans la pratique) ;
- à température nulle, le superfluide est constitué de 10 % des atomes seulement, alors que dans le modèle de Bose la phase condensée réunit tous les atomes.

En outre, on peut montrer que le condensat de Bose n'est pas superfluide en calculant la vitesse critique au delà de laquelle le liquide perd sa propriété de superfluidité. Considérons pour le montrer un objet de masse M se déplaçant dans un fluide. La viscosité représente un échange d'énergie et d'impulsion entre l'objet et le fluide au cours du déplacement. Si ces échanges sont nuls ou s'ils ne peuvent pas se produire, on dira que le milieu est superfluide.

Dans un système de particules indépendantes, chaque échange se fait entre l'objet et UNE particule. Si v_0 et v_1 sont les vitesses avant et après l'échange (le choc), et si $\delta\varepsilon$ et $\delta\vec{p}$ sont l'énergie et l'impulsion transmises à la particule, on écrit la conservation de l'énergie et de l'impulsion :

$$\frac{1}{2}Mv_0^2 = \frac{1}{2}Mv_1^2 + \delta\varepsilon \quad \text{et} \quad M\vec{v}_0 = M\vec{v}_1 + \delta\vec{p} \quad \Rightarrow \quad \vec{v}_0 \cdot \delta\vec{p} = \delta\varepsilon + \frac{\delta p^2}{2M} > \delta\varepsilon$$

Dans le condensat de Bose, l'énergie des atomes est nulle : lors d'un choc, elle ne peut qu'augmenter. On en déduit que $\vec{v}_0 \cdot \delta\vec{p}$ est positif, et donc, en utilisant $\vec{v}_0 \cdot \delta\vec{p} < v_0 \delta p$, on en déduit que :

$$\text{l'échange ne peut avoir lieu que si } v > v_C \text{ avec } v_C = \frac{\delta\varepsilon}{\delta p}$$

L'état superfluide ne peut exister qu'en dessous de cette vitesse critique. Un superfluide lancé à une vitesse supérieure se comporte comme un fluide normal. Pour une particule quantique de masse m confinée dans une boîte cubique d'arête L , l'énergie ne peut varier que par sauts de $\frac{h^2}{2mL^2}$ (voir paragraphe 3.2.1), ce qui donne pour l'impulsion des variations par sauts de $\frac{h}{L}$.

La vitesse critique v_C du condensat de Bose devient :

$$v_C = \frac{\frac{h^2}{2mL^2}}{\frac{h}{L}} = \frac{h}{2mL} \approx 0,5 \mu\text{m/s pour } L = 10 \text{ cm}$$

Le condensat de Bose ne peut donc être superfluide qu'à des vitesses extrêmement faibles. Or, la vitesse critique dans l'hélium est mesurée à plusieurs millimètres par seconde, c'est-à-dire environ 1000 fois plus que celle du condensat de Bose.

Seules les interactions peuvent être mises en cause : l'énergie et l'impulsion ne peuvent être transmises à un atome seul, mais **au système entier**. Le modèle qui suit (largement utilisé en physique du solide) décrit très bien le phénomène. Dans un système de particules interactives, on ne peut plus parler d'états quantiques individuels : le hamiltonien du système n'est plus à variables séparables, et on ne peut plus attribuer une fonction d'onde à chaque particule. Les fonctions d'onde sont macroscopiques : elles concernent le système entier. Lorsque les interactions sont faibles, la résolution de l'équation aux valeurs propres conduit à une description du système où les atomes sont animés de mouvements de translation individuels (comme s'ils étaient indépendants), et où les interactions engendrent un mouvement de vibration collective. Dans un superfluide, les mouvements de translation étant inexistantes, on a affaire à un oscillateur harmonique à $3N$ dimensions qui peut être décomposé en $3N$ oscillateurs harmoniques linéaires indépendants. À chaque oscillateur élémentaire est associé un vecteur d'onde $\vec{k}$ et une pulsation ω , liés entre eux, en milieu isotrope, par une certaine relation $\omega(k)$ (que l'on notera ω_k) dépendant de la nature des interactions, et appelée loi de dispersion. L'énergie vibrationnelle liée à un mode de vibration (c'est-à-dire à une valeur de k donnée) est quantifiée :

$$\varepsilon_k(n_k) = \hbar\omega_k(n_k + 1/2) \quad \text{avec } n_k \text{ entier positif ou nul}$$

La liste des n_k pour toutes les valeurs de k possibles définit l'état quantique du système vibratoire. On dit en physique quantique que le système contient, pour toutes les valeurs de k , n_k pseudo-particules, appelées phonons, dans le mode k . Expérimentalement, on trouve, pour les faibles valeurs de $k = 2\pi/\lambda$ (grandes longueurs d'onde ou ondes sonores), que la loi de dispersion $\omega(k)$ est linéaire. On parle alors de phonons acoustiques. La pente de la courbe est par définition la vitesse de groupe v_G de l'onde associée aux phonons, c'est-à-dire la vitesse du son dans le

milieu : $\omega(k) = v_G k$. À une certaine température, et en l'absence de perturbations extérieures, il s'établit dans le système un régime stationnaire correspondant à une combinaison linéaire (fonction de T) de tous les modes élémentaires. En termes de phonons, l'équilibre correspond à une certaine population moyenne (fonction de T) en phonons dans les différents modes. Lorsqu'on perturbe localement le système, ici par l'intermédiaire d'un objet qui se déplace, l'apport d'énergie et d'impulsion est transmise au gaz de phonons. Ces apports ne peuvent avoir lieu que s'ils peuvent être dissipés, c'est-à-dire s'ils sont suffisants pour créer au moins un phonon dans un mode n_k donné. Le liquide est alors visqueux. L'apport d'énergie doit au moins valoir $\delta\varepsilon = \hbar\omega_k = \hbar v_G k$. L'impulsion échangée dans le même temps vaut $\delta p = \hbar k$. Ce qui donne une vitesse critique :

$$v_C = \frac{\delta\varepsilon}{\delta p} = \frac{\hbar v_G k}{\hbar k} = v_G$$

La vitesse critique correspond donc à la vitesse de propagation des ondes sonores dans le fluide. Tant que cette vitesse n'est pas atteinte, la propagation ne peut avoir lieu et le milieu est superfluide. De la même façon, si on entraîne le superfluide à une certaine vitesse, l'état superfluide disparaît si cette vitesse est supérieure à la vitesse critique. Ou encore, si on met en rotation un récipient contenant de l'hélium superfluide, celui-ci reste immobile à vitesse lente (sa surface reste plane et il n'est pas entraîné par les parois en rotation). Lorsque la vitesse de rotation dépasse la vitesse critique, le liquide reprend les propriétés d'un fluide normal : il est partiellement entraîné par les parois du récipient en rotation et la surface libre devient parabolique comme dans tout autre liquide.

Compte tenu des résultats expérimentaux, la vitesse critique est estimée à 60 m/s. C'est beaucoup trop par rapport aux quelques millimètres par seconde que l'on mesure. La présence de tourbillons observés au cours des différentes expériences donne une explication : l'émission de tourbillons représente un mécanisme de dissipation plus efficace que la création de phonons. L'expérience montre qu'un écoulement superfluide lancé à grande vitesse ralentit en perdant des quantités d'énergie quantifiées à chaque émission d'un tourbillon. On parle de tourbillons quantiques.

En résumé, le modèle de Bose permet de comprendre qualitativement certains phénomènes observés dans l'hélium 4 à basse température, mais pas quantitativement. Si on compare les deux courbes 5.6 (à droite) et 5.9 représentant $C_V(T)$, on constate, d'une part que le pic à la température de transition est extrêmement plus prononcé sur la courbe expérimentale, d'autre part des comportements différents à très basse température (en $T^{3/2}$ pour la première, en T^3 pour la deuxième). C'est d'ailleurs cette loi en T^3 , caractéristique des vibrations dans un réseau cristallin, qui a suggéré que l'on pouvait décrire les interactions par la présence d'un gaz de phonons dans le fluide. La courbe de Bose décrit mal les excitations thermiques du fluide car celles-ci sont principalement liées aux interactions entre atomes : ce sont elles qui sont responsables de la propagation quasi instantanée de la chaleur dans le fluide. Le modèle à deux fluides, l'un constitué d'atomes superfluides, l'autre de phonons, permet de décrire plus précisément le système : plutôt que de décrire le système comme un mélange de deux fluides (l'un constitué des atomes "condensés",

l'autre des atomes "normaux" ou excités), on considère maintenant que la composante superfluide est un condensat constitué de tous les atomes d'hélium, et que la composante "normale" est un gaz de phonons.

Ce modèle est tout à fait correct jusqu'à environ 0,5 K : l'allure de la courbe $C_V(T)$, l'effet thermomécanique^[8], le calcul de la viscosité en fonction de la température, ainsi que l'interprétation de l'existence du second son^[9]. On doit simplement remplacer les anciennes densités en fluide normal et en superfluide par la densité du condensat et la densité du gaz de phonons, cette dernière étant très délicate à calculer, étant donné que les phonons sont des pseudo-particules de masse nulle. Quoiqu'il en soit, les calculs montrent que la densité du gaz de phonons suit une loi en T^4 et que la densité globale (condensat + gaz de phonons) reste constante en température.

Lorsqu'on augmente la température au delà de 0,5 K, on constate que ce modèle devient insuffisant. Le modèle de Debye qui décrit les vibrations dans un réseau cristallin prévoit qu'à haute température, la courbe $C_V(T)$ quitte la loi en T^3 "vers le bas". Pour l'hélium 4, au delà de 0,5 K, la courbe quitte la loi en T^3 "vers le haut". **Landau** explique le phénomène en introduisant un deuxième gaz de pseudo-particules (des "rotons") aux propriétés différentes de celles des phonons. Sans entrer dans les détails, ces rotons, quasiment inexcitables au voisinage du zéro absolu, prendraient le relais des phonons au delà de 0,5 K.

Si on augmente encore la température, le modèle de Landau est à son tour mis en échec : au delà de 1,8 K, aucun modèle théorique à ce jour ne permet d'interpréter l'allure quasi asymptotique de la courbe $C_V(T)$.

⁸ À chaque température correspond un certain taux de superfluide. Inversement, à chaque concentration correspond une température et une seule. Si on force l'écoulement du superfluide seul à travers un filtre poreux reliant deux réservoirs, on change les deux concentrations et il apparaît une différence de température entre les deux réservoirs. Inversement, si on chauffe d'un côté du filtre, un courant superfluide apparaît et une différence de pression s'établit. Ce phénomène est connu sous le nom *d'effet fontaine* à cause de l'expérience suivante. Un vase est rempli d'hélium porté à une température voisine de 1 K. Une canne dont l'extrémité inférieure est remplie d'un matériau poreux est plongée dans le vase. En chauffant suffisamment, on peut créer un jet d'hélium superfluide à l'extrémité supérieure de la canne.

⁹ Le second son, que l'on peut décrire comme une onde de compression dans le gaz de phonons, n'est pas propre à l'hélium superfluide. On peut l'observer également dans les cristaux très purs à basse température. Le son habituel, appelé premier son, correspond dans un cristal à une vibration en phase des atomes. C'est, dans la limite des milieux continus, une variation sinusoïdale de densité qui se propage dans le matériau. Il en est de même dans l'hélium superfluide : les deux fluides vibrent en phase et une oscillation de densité se propage. Dans l'hélium superfluide, on observe un autre type d'oscillation où les deux fluides vibrent en opposition de phase. C'est ce qu'on appelle le second son qui correspond à la propagation d'une onde de *taux de superfluide*, à densité constante. Ce taux étant directement lié à la température, on parle de propagation d'une onde de température.