

PHYSIQUE STATISTIQUE : des processus élémentaires aux phénomènes collectifs

Christophe Texier

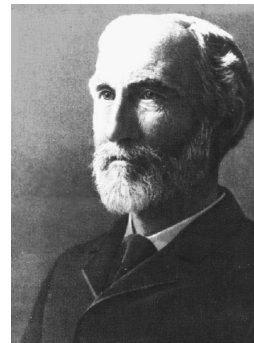
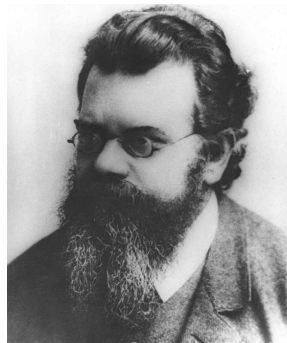
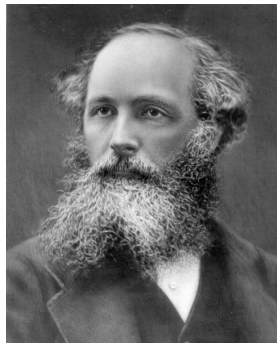
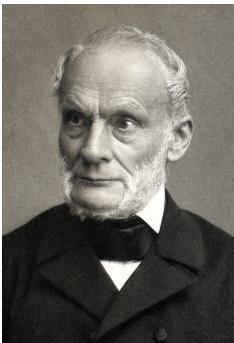
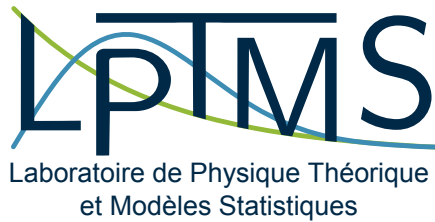


Table des matières

1	Le périmètre de la physique statistique	1
I	À quoi sert la physique statistique ?	1
II	Processus élémentaires et phénomènes collectifs	2
II.A	Échelle microscopique : l'hypothèse atomique	2
II.B	La dichotomie micro/macro	3
II.C	La physique statistique par rapport à la thermodynamique	3
II.D	La relation avec la thermodynamique comme guide	4
III	Fin du réductionnisme et émergence	5
III.A	La démarche de la physique statistique	5
III.B	Phénomènes collectifs et auto-organisation	5
III.C	Universalité	7
IV	Conseils bibliographiques	10
	Annexe 1.A : Repères historiques	11
2	L'approche probabiliste	15
I	Nécessité de l'approche probabiliste	15
I.A	Une nécessité pratique	16
I.B	Une nécessité de principe : chaos, perturbations extérieures et mécanique quantique	17
II	Quelques éléments de théorie des probabilités	17
II.A	Évènements et probabilités	17
II.B	Variable aléatoire discrète	18
II.C	Variable aléatoire continue – Distribution/densité de probabilité	18
II.D	Loi jointe – loi marginale	20
II.E	Moyenne, variance, moments, cumulants, corrélations,...	20
II.F	Fonctions génératrices	24
II.G	Exemples	25
II.H	Remarque : et si les moments n'existent pas ?	25
II.I	Théorème de la limite centrale	26
III	Microétats, macroétats, information et entropie statistique	28
III.A	Microétats, macroétats	28
III.B	Quantifier l'information manquante : entropie statistique	28
III.C	Propriétés de l'entropie	29
	Annexe 2.A : Fonction de grande déviation	32
	Annexe 2.B : Exercices	36
3	Postulat fondamental – Ensemble microcanonique	37
I	Première approche dans un cadre classique	39
I.A	Espace des phases	39
I.B	Pourquoi l'espace des phases ? Théorème de Liouville	39

I.C	Expérience (numérique) des sphères dures – Du problème de Sinai au postulat fondamental	41
I.D	Ergodicité	43
I.E	Énoncé du postulat fondamental	44
II	Description quantique : espace de Hilbert	47
II.A	Rappel des postulats de la mécanique quantique	47
II.B	Microétats et macroétats, moyennes	48
II.C	Dénombrer les microétats : densité d'états	49
II.D	Densité d'états pour des particules identiques (donc indiscernables)	53
III	Postulat fondamental	56
III.A	La distribution microcanonique	56
III.B	Entropie microcanonique	57
IV	Relâchement de contraintes	60
IV.A	Variables externes et variables internes	60
IV.B	Contact thermique entre deux systèmes – Température microcanonique	60
IV.C	Échange de volume – Pression microcanonique	64
IV.D	Échange de particules – Potentiel chimique microcanonique	66
IV.E	Cas général : distribution d'une variable interne quelconque	67
V	Irréversibilité	68
VI	Remarques sur le rôle de l'indiscernabilité	69
VI.A	Validité du traitement semi-classique de l'indiscernabilité – cas du gaz parfait monoatomique	69
VI.B	Entropie de mélange et paradoxe de Gibbs	70
Annexe 3.A	La méthode des multiplicateurs de Lagrange	74
Annexe 3.B	Rappel de mécanique quantique : postulat de symétrisation	75
Annexe 3.C	Exercices	77
Problème	Isothermes et isentropes des gaz parfaits	77
4	Ensemble canonique	79
I	Ensemble canonique	79
I.A	Distribution canonique	80
I.B	Propriétés de l'ensemble canonique	82
II	Relâchement d'une contrainte	86
II.A	Contexte physique	86
II.B	Relâchement de la contrainte	86
II.C	Énoncé du théorème de minimisation	87
II.D	Discussion	87
III	Ensembles microcanonique et canonique	88
III.A	Introduction	88
III.B	Analyse d'un exemple : le cristal paramagnétique	89
III.C	Analyse générale de l'équivalence entre ensembles	90
IV	Retrouver les « principes » de la thermodynamique	92
V	Formulation semi-classique	94
V.A	Distribution canonique classique et fonction de partition	94
V.B	Distribution de Maxwell des vitesses	95
V.C	Théorème d'équipartition de l'énergie	95
VI	Indiscernabilité et approximation de Maxwell-Boltzmann	95
VI.A	Indiscernabilité	96
VI.B	Approximation de Maxwell-Boltzmann	96
VI.C	Apparition des effets quantiques	98

Annexe 4.A : Transformations de Legendre et potentiels thermodynamiques	100
Annexe 4.B : Exercices	102
5 Description semi-classique des gaz	105
I Gaz parfait monoatomique	105
I.A Propriétés thermodynamiques	105
I.B Ordres de grandeur dans un gaz d'Argon	107
II Gaz parfaits moléculaires	107
II.A Aspects microscopiques : cas des molécules diatomiques	108
II.B Thermodynamique des gaz de molécules diatomiques hétéronucléaires . .	109
II.C Molécules diatomiques homonucléaires : cas de H_2	113
II.D Molécules polyatomiques	114
III Gaz réels	114
6 Thermodynamique des oscillateurs harmoniques	117
I Molécule diatomique	118
I.A Analyse classique	118
I.B Traitement quantique	118
II Vibrations d'un corps solide	120
II.A Aspects microscopiques	120
II.B Thermodynamique des vibrations	123
III Thermodynamique du rayonnement	128
III.A Description du problème du corps noir	128
III.B Modes propres du champ électromagnétique dans une cavité	128
III.C Thermodynamique du rayonnement	130
Annexe 6.A : Modes propres d'une chaîne de ressorts	136
Annexe 6.B : Corde vibrante : quantification « classique » <i>versus</i> « quantique » . . .	138
Annexe 6.C : Cristal – Approximation harmonique et indiscernabilité	140
7 Transitions de phase – Approximation de champ moyen	143
I Introduction	143
II Liquide-gaz	143
II.A Description du problème	144
II.B Approximation de champ moyen	144
II.C Analyse des isothermes de van der Waals – transition liquide-gaz	147
III Para-Ferro	152
III.A Introduction	152
III.B Traitement de champ moyen du modèle d'Ising Ferro	153
III.C Brisure spontanée de symétrie et paramètre d'ordre	161
III.D Comparaison avec la transition liquide-gaz	162
Annexe 7.A : Exercices	165
8 Ensemble grand canonique	167
I Ensemble grand canonique	167
I.A Introduction : Position du problème	167
I.B Ensemble grand canonique	168
I.C Propriétés de l'ensemble grand canonique	170
II Particules identiques sans interaction	170
II.A Motivation	170
II.B Description microscopique : facteurs d'occupation	174
II.C Décomposition de la grande fonction de partition	174

II.D	Occupation des états individuels	175
9	Statistique de Fermi-Dirac	181
I	Introduction	181
I.A	Contexte physique	181
I.B	Les électrons d'un métal	182
II	Gaz parfait de fermions libres complètement dégénéré ($T = 0$)	185
II.A	Formules grand canoniques	185
II.B	Interprétation canonique	185
III	Développements de basse température	187
III.A	Rappel des formules grand canoniques	187
III.B	Stratégie de l'analyse	188
III.C	Développements de Sommerfeld	188
III.D	Potentiel chimique canonique	189
III.E	Énergie et chaleur spécifique	190
III.F	Entropie	191
III.G	Équation d'état	192
10	Statistique de Bose-Einstein	195
I	Phonons et photons	195
I.A	Rappel : description canonique	195
I.B	Description grand canonique	196
I.C	Remarque sur le cas où $\mu = 0$: Gap ou masse ?	196
II	Gaz parfait de bosons libres	197
II.A	Gaz de bosons libres – Information microscopique	197
II.B	Analyse du potentiel chimique	197
II.C	Énergie et chaleur spécifique pour $T < T_{BE}$	202
II.D	Équation d'état	203
II.E	Entropie et chaleur latente	204
III	Réalisations expérimentales	204
III.A	Diverses réalisations	204
III.B	Condensation dans un piège harmonique	205
11	Aspects cinétiques	209
I	Théorie cinétique des gaz	209
12	La physique statistique : mode d'emploi	211
I	Quelques notions importantes	211
II	La « boîte à outils » de la physique statistique	211
A	Formulaire	217
B	Corrections	221
	Bibliographie	222
	Index	225

Chapitre 1

Le périmètre de la physique statistique

Plan du chapitre

I	À quoi sert la physique statistique ?	1
II	Processus élémentaires et phénomènes collectifs	2
II.A	Échelle microscopique : l'hypothèse atomique	2
II.B	La dichotomie micro/macro	3
II.C	La physique statistique par rapport à la thermodynamique	3
II.D	La relation avec la thermodynamique comme guide	4
III	Fin du réductionnisme et émergence	5
III.A	La démarche de la physique statistique	5
III.B	Phénomènes collectifs et auto-organisation	5
III.C	Universalité	7
IV	Conseils bibliographiques	10
	Annexe 1.A : Repères historiques	11

I À quoi sert la physique statistique ?

La physique statistique est l'une des plus importantes **théories** physiques « **cadre** », fondée (dans sa version classique) dans la seconde moitié du XIX^e siècle (Clausius, Maxwell, Boltzmann & Gibbs), puis fertilisée par le développement de la mécanique quantique au cours de la première moitié du XX^e siècle. Elle s'attelle à l'étude des systèmes à très grand nombre de degrés de liberté.

La physique statistique permet de déduire, par l'utilisation de méthodes probabilistes, les propriétés des systèmes aux échelles macroscopiques à partir des lois qui gouvernent les constituants élémentaires aux échelles microscopiques.

Cette proposition montre qu'il aurait mieux valu poser la question : « à quoi ne sert pas la physique statistique ? » ! On peut en effet en juger en examinant la petite liste (loin d'être exhaustive) d'exemples d'application de la physique statistique (dont nous étudierons certaines) :

1. Propriétés thermodynamiques des gaz et des fluides (fluides normaux, superfluides),...
2. Une grande partie de la physique des solides :

- vibrations des atomes des cristaux
 - dynamique des électrons
(métaux, semiconducteurs, semi-métaux, supraconductivité,...)
 - Défauts et alliages
 - Propriétés magnétiques
(paramagnétisme, ferromagnétisme, antiferromagnétisme, ferrimagnétisme,...)
 - etc
3. Thermodynamique du rayonnement dans une étoile.
 4. Stabilité des étoiles naines blanches (pression quantique des électrons) et des étoiles à neutrons.
 5. Matière molle :
 - polymères
 - cristaux liquides (phases nématique, smectique,...)
 - micelles
 - etc
 6. Physique des plasmas
 7. thermo-chimie
 8. biophysique (encore un sujet très vaste)

Les questions posées peuvent concerner aussi bien des situations à l'équilibre (thermodynamique d'un solide, vibrations,...) que des aspects hors équilibre (transport de la chaleur,...)

Plus récemment les méthodes de la physique statistique ont été appliquées dans des domaines plus éloignés de la physique :

9. étude des réseaux (réseaux de neurones, réseaux de communication,...)
10. informatique théorique : optimisation combinatoire (ex : problème du voyageur de commerce), problème de satisfaisabilité de contraintes,...
11. modèles socio-économiques : réseaux sociaux, propagation d'épidémie, modèle de vote, étude des marchés financiers, etc.
12. etc.

II Processus élémentaires et phénomènes collectifs

La diversité des exemples d'applications, qui sortent même du champ de la physique, nous conduit à redéfinir le micro et le macroscopique. La distinction entre « microscopique » et « macroscopique » ne doit pas être prise trop littéralement et n'est pas nécessairement corrélée à des échelles microscopiques ou macroscopiques au sens strict. Il s'agit plutôt d'une distinction entre des *processus élémentaires* et des *phénomènes collectifs*.

II.A Échelle microscopique : l'hypothèse atomique

Il semble une tautologie de souligner que la physique statistique postule l'existence d'une échelle « microscopique » où la physique est contrôlée par des processus élémentaires. Autrement dit il existe une échelle où la description de la matière perd le caractère *continu* perceptible aux grandes échelles qui nous sont familières, et acquiert un caractère *discret*. Historiquement, la physique statistique s'est développée pour construire une théorie cinétique des gaz, dans le contexte de laquelle on parlait au XIX^e siècle « d'hypothèse atomique ». Cette dernière est ancienne et remonte à l'antiquité grecque (Démocrite (460-370 av JC) ; « *ατομος* », indivisible). Bien que les développements de la chimie au XIX^e siècle en fournissaient de précieux indices, avec

la « loi des proportions définies », l'hypothèse atomique a rencontré une incroyable résistance dans la communauté des physiciens à la fin du XIX^e siècle, dont une partie étaient en faveur d'une description purement continue. On peut dire que ce n'est qu'à l'orée du XX^e siècle que l'hypothèse atomique fut validée définitivement dans une expérience de physique par Jean Perrin (en 1908) avec son étude du mouvement brownien fournissant estimation du nombre d'Avogadro.

Il est intéressant de chercher à comprendre la raison de la résistance d'une partie des physiciens du XIX^e : tout d'abord, comme nous l'avons dit, les preuves d'existence des atomes ne sont venues qu'au début du XX^e. Mais plus profondément, l'idée que la description la plus fondamentale ¹ (la physique statistique) repose sur une formulation de nature probabiliste était inconcevable (d'un point de vue philosophique) pour de nombreux scientifiques [3].

II.B La dichotomie micro/macro

L'existence d'une échelle microscopique étant acquise, nous aurons toutefois à clarifier un certain nombre de contradictions apparentes qui semblent opposer les deux échelles et que nous résumons dans le tableau :

Microscopique		Macroscopique
• discret • petit nombres de lois simples (postulats de la MQ, équations de Maxwell,...) • comportement aléatoire • équations réversibles	$\Rightarrow$ Physique statistique $\Rightarrow$	• continu • diversité, complexité, grand nombre de théories effectives • déterministe • irréversibilité

TABLE 1.1 : L'opposition micro-macro.

II.C La physique statistique par rapport à la thermodynamique

Comme il n'aura pas échappé au lecteur, le champ d'étude de la physique statistique recouvre partiellement celui de la thermodynamique, aussi il convient de préciser la relation entre les deux théories. Contrairement à la physique statistique, la thermodynamique ne cherche pas à fonder ses résultats à partir d'une description microscopique, mais se place directement à l'échelle macroscopique. La thermodynamique est donc une **théorie phénoménologique** par opposition à la physique statistique qui permet de construire des **théories microscopiques**. Cette dernière se fixe donc des objectifs beaucoup plus ambitieux que la première et possède une puissance prédictive nettement supérieure.

Exemple 1 : Gaz.— En supposant qu'il est justifié de décrire la dynamique des molécules d'un gaz par la mécanique newtonienne, la physique statistique permet de déduire les propriétés thermodynamiques de celui-ci : équation d'état, coefficients calorimétriques, etc.

Exemple 2 : Coefficients calorimétriques.— Les coefficients calorimétriques jouent un rôle très important en thermodynamique : capacités calorifiques ² $C_V = \left(\frac{\partial E}{\partial T}\right)_V$ et $C_p = \left(\frac{\partial H}{\partial T}\right)_p$, compressibilité $\kappa_T = -\frac{1}{V}\left(\frac{\partial V}{\partial p}\right)_T$, coefficient de dilatation $\alpha = \frac{1}{V}\left(\frac{\partial V}{\partial T}\right)_p$, etc). La thermodynamique

¹ « fondamental » dans le sens de « fondation » ou « fondement ».

² C_X caractérise la chaleur reçue pour un accroissement de température, lorsque la grandeur X est fixée : $dQ^{\text{rev}} = C_X dT + a dX$ ou encore, puisque $dQ^{\text{rev}} = T dS$, $C_X = T (\partial S / \partial T)_X$.
Si $X \rightarrow V$, on a $C_V = T (\partial S / \partial T)_V$. En utilisant $dE = T dS - p dV$ on déduit $C_V = (\partial E / \partial T)_V$.
Si $X \rightarrow p$, on a $C_p = T (\partial S / \partial T)_p$ et il faut plutôt commencer par effectuer une transformation de Legendre en introduisant l'enthalpie $H = E + pV$, d'où $dH = T dS + V dp$ et $C_p = (\partial H / \partial T)_p$.

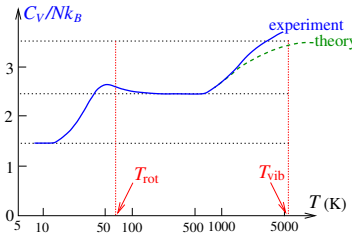


FIGURE 1.1 : **Chaleur spécifique du dihydrogène deutéré** (molécule HD formée d'un atome d'hydrogène 1_1H et de son isotope 2_1H (noté D). Données tirées du livre de R. Balian [3] (p. 378).

permet de trouver des relations entre ces coefficients, telles la relation de Mayer $C_p - C_V = T(\frac{\partial V}{\partial T})_p(\frac{\partial p}{\partial T})_V = TV\alpha^2/\kappa_T$, cependant ces contraintes ne permettent pas de déterminer tous ces coefficients et en laissent au moins un indéterminé, par exemple C_V . À l'inverse la physique statistique fournit les moyens de calculer (en principe) le potentiel thermodynamique qui encode toutes les propriétés thermodynamiques – l'équation d'état, l'ensemble des coefficients calorimétriques, etc – en fonction des paramètres microscopiques du modèle. Par exemple, nous serons en mesure de prédire l'allure des courbes à la structure aussi complexe que les courbes de chaleurs spécifiques de gaz diatomiques représentées sur la figure 1.1 (ce qui est hors de portée de la thermodynamique, mais sera à celle du lecteur dans dès le chapitre 5).

II.D La relation avec la thermodynamique comme guide

Une idée qui nous aidera dans les développements ultérieurs, par exemple pour définir des quantités comme la température, la pression, etc, est la correspondance avec la thermodynamique dans une certaine limite que nous préciserons.

Il est intéressant de rappeler brièvement les bases sur lesquelles se construit la thermodynamique, une théorie axiomatique dont nous pouvons ébaucher le squelette :

- Le postulat fondamental : l'existence d'une fonction d'état, caractérisant l'état d'équilibre macroscopique du système, l'énergie $E(S, V, N, \dots)$ fonction de l'entropie S , du volume V , du nombre de particules N (et éventuellement d'autres paramètres comme le champ magnétique, etc), et dont la différentielle (l'égalité fondamentale de la thermodynamique) ³

$$dE = \underbrace{T dS}_{dQ^{\text{rev}}} - \underbrace{p dV}_{dW^{\text{rev}}} + \mu dN + \dots \quad (1.1)$$

définit température, pression, potentiel chimique, etc.

La donnée de cette fonction de ces paramètres encode l'ensemble des propriétés thermodynamiques. En pratique, cette fonction n'est en général pas connue.

☞ La physique statistique nous permettra justement de calculer $E(S, V, N, \dots)$ en termes des paramètres microscopiques.

- Le premier principe de la thermodynamique (1847) postule l'équivalence entre travail et chaleur : la somme du travail et de la chaleur, deux quantités qui dépendent du « chemin » suivi lors d'une transformation, est égale à la variation de la fonction d'état, qui ne dépend que des points initial et final de la transformation : $\Delta E = Q + W$. Dans une version infinitésimale, $dE = dQ + dW$.

☞ Cette distinction entre travail et chaleur apparaîtra dans le cadre de la physique statistique comme une séparation entre différents termes de la variation d'énergie moyenne.

³ La notation dQ désigne une quantité infinitésimale qui n'est pas la différentielle d'une fonction.

- L'étude des machines thermiques, initiée par Sadi Carnot, conduisit Clausius à formuler le second principe (1954) : $\oint \frac{dQ}{T} \geq 0$ pour tout cycle de transformations (nul si le cycle est réversible $\oint \frac{dQ^{\text{rev}}}{T} = 0$). Par conséquent $dS = dQ^{\text{rev}}/T$ est la différentielle d'une fonction d'état $S(E, V, N, \dots)$ (fonction inverse de la précédente).
- ☞ Le second « principe » sera démontré dans le cadre de la physique statistique (dédit du théorème H de Boltzmann par exemple).

III Fin du réductionnisme et émergence

III.A La démarche de la physique statistique

Comme nous l'avons dit, la physique statistique permet de faire des prédictions en partant de l'information sur la dynamique aux échelles microscopiques. Elle n'est pas auto-suffisante et se présente donc comme une « *sur-couche* » de la description microscopique. Par conséquent la démarche procédera toujours en deux temps :

1. Caractérisation de la dynamique à l'échelle microscopique. Cette échelle de description peut être classique ou quantique (en fonction des besoins... et de notre courage).
2. Ces ingrédients microscopiques seront utilisés, avec des considérations probabilistes, pour construire la théorie physique décrivant la dynamique à l'échelle macroscopique.

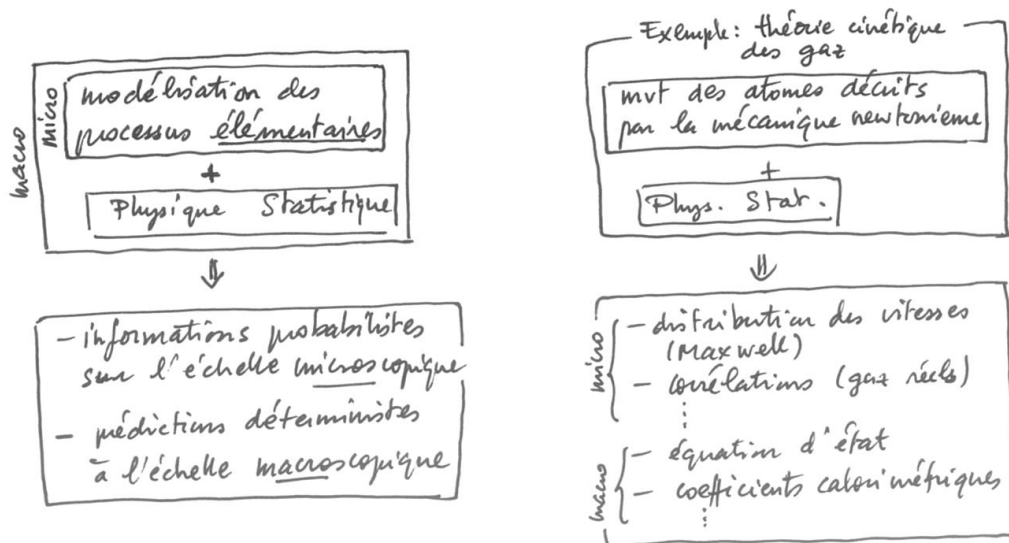


FIGURE 1.2 : La physique statistique est une sur-couche au dessus d'une description microscopique.

III.B Phénomènes collectifs et auto-organisation

Il faut maintenant dissiper un malentendu qui aurait pu émerger à la lecture des précédentes remarques. Puisque qu'au cœur d'un modèle développé dans le cadre de la physique statistique, nous trouvons des informations sur l'échelle microscopique, il pourrait sembler à première vue que là réside le niveau fondamental de la théorie. C'est la logique **réductionniste** (la « théorie du tout ») : caractérisons la dynamique à l'échelle la plus élémentaire, puis nous pourrions remonter à toutes les échelles supérieures. Pour caricaturer : le comportement de l'univers est contenu dans la dynamique des quarks !

Nous allons prendre l'exact contre-pied de cette logique, car attention :

Élémentaire $\neq$ Fondamental

La **physique statistique pulvérise le mythe réductionniste : chaque échelle requiert son propre niveau d'analyse**. La physique à l'échelle microscopique ne dit que très peu sur celle à l'échelle macroscopique. L'origine du découplage entre les échelles repose sur l'existence de puissants principes d'auto-organisation à l'échelle macroscopique, ce qui explique en grande partie la disparition de l'information sur l'échelle microscopique, et une certaine **universalité** de nombreuses propriétés à l'échelle macroscopique [1, 22].

Probablement une des meilleures illustrations de l'idée d'**émergence** et d'auto-organisation est fournie par l'étude des **transitions de phases**. Par exemple si nous partons d'une description microscopique d'un ensemble d'atomes formant un fluide, nous pouvons prendre en compte dans le modèle l'interaction inter-atomique. Pour des atomes simples, celle-ci est bien décrite par le potentiel de Lennard-Jones $u(r) = u_0 \left[\left(\frac{a}{r} \right)^{12} - \left(\frac{a}{r} \right)^6 \right]$. Nous serons en mesure de montrer plus tard que cette interaction conduit à la possible organisation du fluide en **deux phases**,⁴ liquide ou gazeuse, i.e. à des comportements très différents à l'échelle macroscopique, quand bien même aucun des ingrédients microscopiques n'a changé (figure 1.3) !

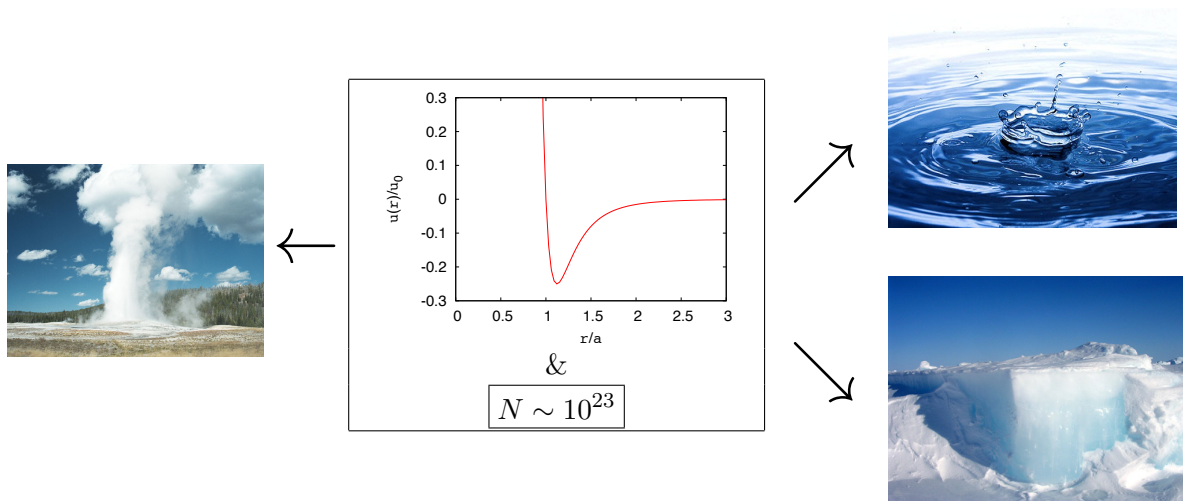


FIGURE 1.3 : **Émergence.**— *Les mêmes détails microscopiques (ici un modèle simple d'interaction à deux corps) peuvent conduire à des phénomènes collectifs très différents (lorsque $N \rightarrow \infty$) : des phases gazeuse, liquide ou solide (notons que le potentiel d'interaction entre molécules d'eau est plus complexe).*

À l'échelle macroscopique, il y a **émergence** de phénomènes **collectifs**, imprévisibles avec la seule compréhension de la dynamique à l'échelle microscopique.

Le niveau **fondamental** est donc dans le passage du microscopique au macroscopique, rendu possible grâce aux concepts transversaux de la physique statistique, i.e. conduisant à une vision unificatrice des sciences macroscopiques [3].

⁴ Le modèle prédit également une troisième phase solide, même si nous ne serons pas en mesure de le prouver.

III.C Universalité

Afin de souligner la puissance des principes d'auto-organisation en œuvre aux échelles macroscopiques, discutons deux exemples concrets illustrant la notion d'universalité, une idée centrale de la physique statistique des transitions de phases.

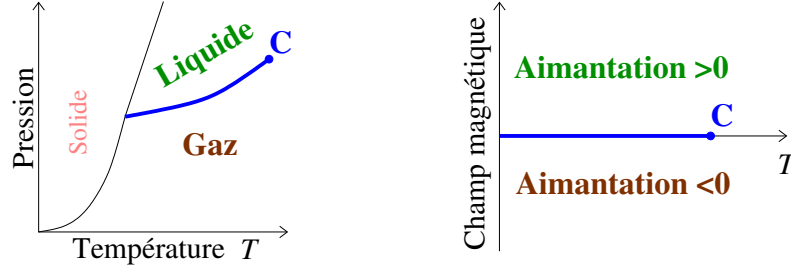


FIGURE 1.4 : Deux diagrammes de phases : liquide-gaz-solide et para-ferromagnétique. Le point critique est repéré par un C.

Le point critique de la transition liquide-gaz La phénoménologie du diagramme de phases d'un fluide simple est bien connue : on peut représenter un diagramme de phase $p - T$ où les lignes séparent les phases gazeuse, liquide et solide (figure 1.4). Ces lignes correspondent à des transitions de phases « du premier ordre », caractérisée par l'existence d'une chaleur latente (par exemple la conversion, à p et T fixés, d'une certaine quantité de liquide en gaz, a un coût énergétique dont toute personne sortant de sa douche a une perception claire). La transition liquide-gaz a toutefois une particularité qui est l'interruption de la ligne en un point appelé point de « l'opalescence critique ». En ce « point critique », noté C, la distinction entre liquide et gaz s'estompe et la chaleur latente de vaporisation s'annule. On dit que ce point correspond à une transition de phase « du second ordre » (nous expliquerons cette terminologie et reviendrons sur la physique des transitions de phases dans un chapitre ultérieur). Au point critique, la physique devient *invariante d'échelle* et de nombreuses quantités physiques présentent des comportements en loi de puissance. Par exemple si on mesure la chaleur spécifique, on observe qu'elle diverge au point critique comme $C(T) \sim |T - T_c|^{-\alpha}$ pour un exposant $\alpha = 0.111(1)$, mesuré sur un fluide de CO_2 (cf. tableau 1.2). L'écart à la densité critique (le « paramètre d'ordre ») de ce même fluide de CO_2 se comporte, sous la température critique, comme $\rho(T) - \rho_c \sim (T_c - T)^\beta$ pour $\beta = 0.324(2)$. La compressibilité du gaz (maintenant du xenon), $\kappa_T \stackrel{\text{def}}{=} \frac{1}{\rho} \left(\frac{\partial \rho}{\partial p} \right)_T$, présente également un comportement divergent au point critique $\kappa_T(T) \sim |T - T_c|^{-\gamma}$ pour $\gamma = 1.246(10)$. Le point remarquable est que si des études similaires sont menées pour d'autres fluides, on observe que la température critique T_c dépend du fluide (propriété non-universelle), alors que les exposants sont tout à fait comparables (propriétés universelles).

Transition paramagnétique-ferromagnétique en champ nul On considère maintenant la transition paramagnétique-ferromagnétique dans un aimant. En l'absence de champ magnétique et au dessus d'une température dite de Curie T_c , le matériau est dans une phase paramagnétique (pas d'aimantation). En dessous de la température de Curie on voit apparaître une aimantation spontanée $m(T)$ (segment épais de la figure 1.4). On étudie de même les comportements de différentes quantités au voisinage du point critique (la transition para-ferro en champ nul) pour un matériau FeF_2 : on obtient une chaleur spécifique divergente : $C(T) \sim |T - T_c|^{-\alpha}$ pour un exposant $\alpha = 0.11(3)$, une aimantation $m(T) \sim (T_c - T)^\beta$ pour $\beta = 0.325(2)$. Finalement on analyse la susceptibilité magnétique $\chi(T) \stackrel{\text{def}}{=} \left. \frac{\partial m}{\partial B} \right|_{B=0} \sim |T - T_c|^{-\gamma}$ pour $\gamma = 1.25(1)$ (cf.

Transition	liquide-gaz	para-ferro
Chaleur spécifique	$C(T) \sim t ^{-\alpha}$ CO ₂ $\alpha = 0.111(1)$	$C(T) \sim t ^{-\alpha}$ FeF ₂ $\alpha = 0.11(3)$
Paramètre d'ordre	densité $\rho(T) - \rho_c \sim (-t)^\beta$ CO ₂ $\beta = 0.324(2)$	aimantation $m(T, 0) \sim (-t)^\beta$ FeF ₂ $\beta = 0.325(2)$
Susceptibilité	Compressibilité $\kappa_T(T) \sim t ^{-\gamma}$ Xe $\gamma = 1.246(10)$	Susceptibilité $\chi(T) \sim t ^{-\gamma}$ FeF ₂ $\gamma = 1.25(1)$

TABLE 1.2 : **Exposants critiques des transitions liquide-gaz et para-ferromagnétique.** (*Transitions de phases de la classe d'universalité d'Ising 3D*). $t \stackrel{\text{def}}{=} (T - T_c)/T_c$ mesure l'écart au point critique. Données expérimentales tirées de H. W. Blöte, E. Luijten & J. R. Heringa, *J. Phys. A* **28**, 6289 (1995).

tableau 1.2). Là encore, si la température critique T_c varie d'un matériau à l'autre, on observe une universalité des exposants critiques.

Coarse-graining et universalité Il est extrêmement frappant que des phénomènes à première vue aussi différents que la transition liquide-gaz et la transition paramagnétique-ferromagnétique, présentent de telles similarités : toute une gamme d'exposants critiques identiques au pourcent près ! Des études numériques (cf. tableau 1.3) ont permis de prédire des valeurs des exposants très proches des valeurs mesurées expérimentalement (cf. tableau 1.2). Aujourd'hui l'origine de ces similarités est bien comprise : elle réside dans les propriétés d'universalité des phénomènes critiques et l'invariance d'échelle. Si la physique statistique fonde ses résultats sur une description microscopique, les détails de celle-ci sont gommés (coarse-graining) lorsqu'on s'intéresse à la physique de grande échelle, ce qui explique que certaines propriétés (comme les exposants critiques introduits plus haut) soient complètement *insensibles aux détails sous-jacents*. Les résultats universels sont obtenus pour une large classe de problèmes physiques appartenant tous à la même « classe d'universalité » (ici la « classe d'universalité d'Ising 3D ») : l'appartenance à une classe d'universalité dépend de propriétés très générales, comme la dimension d'espace, la nature du paramètre d'ordre (pour le modèle d'Ising, un scalaire) et les symétries du problème.

Quantité	exposant
$C(T) \sim t ^{-\alpha}$	$\alpha = 0.110(1)$
$m(T, 0) \sim (-t)^\beta$	$\beta = 0.3265(3)$
$\chi(T) \sim t ^{-\gamma}$	$\gamma = 1.2372(5)$
$m(T_c, B) \sim B^{1/\delta}$	$\delta = 4.789(5)$

TABLE 1.3 : **Exposants critiques du modèle d'Ising 3D (théorie).** $t \stackrel{\text{def}}{=} (T - T_c)/T_c$. Résultats théoriques obtenus pour le modèle d'Ising 3D, tirées de la revue : Andrea Pelissetto & Ettore Vicari, « Critical phenomena and renormalization-group theory », *Phys. Rep.* **368**, 549–727 (2002)

Dans la classe d'universalité d'Ising 3D, on trouve des problèmes physiques aussi divers que : la transition liquide-gaz et para-ferro, donc, les fluides binaires (mélanges), les alliages désordonnés, les systèmes ferroélectriques, etc.

☺ Les idées importantes

- La distinction entre processus élémentaires et phénomènes collectifs.

📖 Pour en savoir plus

- Pour ceux qui sont intéressés par les aspects historiques, je recommande le chapitre d'introduction de l'ouvrage de Roger Balian [3] et le paragraphe §3.4 sur l'histoire de la thermodynamique.

IV Conseils bibliographiques

A. Physique statistique

Les ouvrages les plus faciles d'accès :

- Un livre d'introduction : volume V du cours de Berkeley de F. Reif [31].
- Un livre au classicisme très français plutôt synthétique et plaisant : R. Castaing [7].
- Pédagogique et plutôt synthétique : C. & H. Ngô [27].
- L'excellent polycopié de cours d'A. Georges & M. Mézard [12] (difficile à se procurer ?)
- Très (trop ?) complet et pédagogique : B. Diu, C. Guthmann, D. Lederer & B. Roulet [11]. Ne pas se laisser effrayer par le volume (~ 1000 pages) : le cœur de l'ouvrage (la partie « cours ») ne fait que 235 pages, le reste étant une série de « compléments » et illustrations.

Autres ouvrages :

- Un ouvrage très agréable, avec une longue discussion sur les transitions de phases et les phénomènes critiques : R. K. Pathria & P. D. Beale [28].
- Les ouvrages récents de M. Kardar méritent très certainement d'être mentionnés [16, 15] (le premier tome correspond davantage à notre programme).
- Deux ouvrages intéressants sont ceux de K. Huang [13] et plus original, celui de S.-K. Ma [25].
- Pour rester du côté asiatique on peut mentionner le classique livre de R. Kubo *et alter* [35] et, sur des sujets plus avancés (aspects hors équilibre,...) [20].

Ouvrages de référence :

- Re-citons : B. Diu, C. Guthmann, D. Lederer & B. Roulet [11].
- Un livre classique très complet est celui de F. Reif [30].
- Un autre classique (plus difficile) : le « Landau & Lifchitz » [21]. Discussion détaillée des symétries.
- Très profond et complet, mais également plus difficile : R. Balian [3, 4].

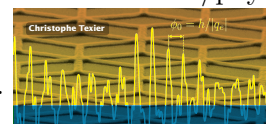
Exercices :

- L. Couture, C. Chahine & R. Zitoun [10]
- Récemment : H. Krivine & J. Treiner [19].

B. Réviser la mécanique quantique

- Un grand classique : C. Cohen-Tannoudji, B. Diu & F. Laloë [8].
- Deux excellents livres récents : J.-L. Basdevant & J. Dalibard [5] et M. Le Bellac. [23].
- Encore plus récent, assez condensé et avec une légère coloration « matière condensée/physique

mésoscopique », [34] (fait spécialement pour le magistère d'Orsay ☺).



Mécanique quantique

Cours et exercices corrigés

2^e édition

Licence 3
Master
CAPES
Agrégation

DUNOD

C. Physique du solide

- Un livre assez descriptif et agréable : C. Kittel [17].
- Un classique assez détaillé : N. Ashcroft & N. Mermin [2].

Annexe 1.A : Repères historiques

Je reprends l'introduction des livres de Balian et de Pathria [3, 28]. La physique statistique s'est développée dans la seconde moitié du XIX^e siècle. Les précurseurs sont Daniel Bernoulli (1700-1782), John Herapath (1790-1868) et James Prescott Joule (1818-1889) pour leurs tentatives de construction d'une *théorie cinétique des gaz*. Cependant les véritables pères de la physique statistiques semblent être Rudolf J. E. Clausius (1822-1888), James Clerk Maxwell (1831-1879) et surtout Ludwig Eduard Boltzmann (1844-1906) et Josiah Willard Gibbs (1839-1903). Quelques dates importantes sont :

- 1727 : travaux précurseurs de Bernoulli qui propose une interprétation cinétique de la pression, due aux chocs des molécules.
- 1820 : La théorie cinétique se développe avec les travaux de Herapath et Joule (1821).
- 1847 : 1er principe de la thermodynamique, i.e. équivalence travail-chaleur, établie par Joule.
- 1854 : 2ème principe de la thermodynamique. Clausius montre que $\oint \frac{dQ}{T} \geq 0$ pour tout cycle de transformations (nul s'il s'agit d'un cycle de transformations réversibles $\oint \frac{dQ^{\text{rev}}}{T} = 0$; Par conséquent $dS = dQ^{\text{rev}}/T$ est la différentielle d'une fonction d'état, appelée l'entropie).
- 1857 : Clausius retrouve la loi des gaz parfaits à partir de l'image microscopique de particules en mouvement à l'aide de concepts statistiques ; il discute l'effet des collisions (et introduit la notion de libre parcours moyen élastique), ce qui jouera un rôle important dans le travail de Boltzmann.
- 1860 : le mémoire de Maxwell sur la théorie cinétique des gaz dans lequel il dérive la célèbre distribution des vitesses.
- 1866 : Influencé par les idées de Maxwell, Boltzmann comprend qu'il faut introduire des concepts statistiques.
- 1867 : Paradoxe du démon de Maxwell.
- 1868-1871 : Boltzmann développe la théorie cinétique, généralisant les résultats de Maxwell. Il introduit le théorème d'équipartition.
- 1872 : Boltzmann propose son équation et démontre son célèbre théorème H , qui étend la notion d'entropie à une situation hors équilibre et donne une fondation microscopique au second principe de la thermodynamique. L'idée d'*irréversibilité* est introduite sur de solides bases (travaux violemment critiqués par Loschmidt, Zermelo). En particulier Boltzmann montre que la distribution de Maxwell-Boltzmann est la seule distribution stable sous l'effet des collisions, vers laquelle toute autre distribution converge. Boltzmann subit également des critiques des énergétistes (en faveur de la thermodynamique et d'une vision purement continue de la matière, Mach, Duhem, Ostwald).
- 1874-1876 : Paradoxe de l'irréversibilité (Thomson & Loschmidt) : comment des équations microscopiques réversibles peuvent donner naissance à l'irréversibilité macroscopique ? Cf. § V.
- 1875 : Paradoxe de Gibbs (problème de l'indiscernabilité, cf. § VI.B).
- 1900 : Planck propose la loi éponyme décrivant la thermodynamique du rayonnement (cf. § III.C). Il résout une difficulté majeure de la physique classique. Le problème est à l'intersection de plusieurs théories fondamentales : physique statistique, électrodynamique, et signe l'acte de naissance de la théorie quantique.
- 1902 : Gibbs propose la théorie des ensembles dans son ouvrage « *Elementary principles of*

statistical mechanics ».

- 1900-1927 : développement de la mécanique quantique.
- 1907 : Einstein développe une première théorie de la chaleur spécifique des solides (cf. § II.B).
- 1912 : Debye améliore la théorie d'Einstein et explique les comportements de basse température de la chaleur spécifique des solides $C_V^{\text{vib}} \propto T^3$ (cf. § II.B).
- 1912 : Smoluchowski lève le paradoxe du démon de Maxwell pour un démon automatique.
- 1912 : Otto Sackur (1880-1914) et Hugo Tetrode (1895-1931) proposent (indépendamment) une formule pour l'entropie du gaz monoatomique.⁵ La physique statistique classique permettait d'aboutir (en tenant du paradoxe de Gibbs) à $S(E, V, N) = Nk_B [\ln(V/N) + \frac{3}{2} \ln(E/N) + s_0]$ où s_0 est une constante indéterminée. En utilisant des arguments théoriques et des données expérimentales sur la vapeur de Mercure, ils introduisent la constante de Planck⁶ afin de discrétiser le problème du comptage des microétats accessibles, ce qui fixe la constante s_0 (cf. § III.B).
- Années 1910 : premières théories des milieux magnétiques : (expériences de Curie), Weiss, Langevin, Brillouin (plus tard).
- 1924 : Statistique de Bose-Einstein pour des particules indiscernables de nature bosonique (dont la fonction d'onde est symétrique sous l'échange de particules).
- 1925 : Einstein applique ces idées aux gaz atomiques et prédit la possible condensation de Bose-Einstein (qui sera proposée comme explication du phénomène de supraconductivité par London en 1938). Cf. § II et II.
- 1926 : Suivant les idées de Pauli (principe de Pauli), Fermi et Dirac proposent (indépendamment) la statistique de Fermi-Dirac pour des particules indiscernables de nature fermionique (dont la fonction d'onde est antisymétrique sous l'échange). Cf. § II.
- 1927 : Landau et von Neumann introduisent la notion de matrice densité, l'équivalent quantique de la notion de distribution dans l'espace des phases classique.
- 1928 : Sommerfeld produit sa théorie des électrons dans les métaux (théorie de Drude-Sommerfeld). Cf. § III.
- 1937 : Théorie phénoménologique de Landau des transitions de phases du second ordre.
- Fin des années 1930 : théorie de la superfluidité par Landau (qui lui vaudra le prix Nobel en 1962).
- 1939 : Théorème spin-statistique (Pauli, Fierz et Belinfante) : les propriétés de permutabilité de la fonction d'onde de particules identiques sont contrôlées par le spin de celles-ci (entier ou demi-entier).
- 1944 : Lars Onsager (prix Nobel de chimie 1968) obtient la solution exacte pour le modèle d'Ising bidimensionnel en champ magnétique nul (calcul de $Z(T, B = 0)$).
- 1948 : Théorie de l'information (Shannon & Weaver).
- 1950 : En interprétant le "paradoxe du démon de Maxwell" avec les idées de la théorie de

⁵On pourra consulter l'article récent de W. Grimus, « 100th anniversary of the Sackur-Tetrode equation », Ann. Phys. (Berlin) **512**(3), A32–A35 (2013) ; disponible sur arXiv:1112.3748.

⁶Jusque là la constante de Planck était apparue dans l'étude de la thermodynamique du rayonnement (Planck 1900, Einstein 1905), mais aussi dans la théorie d'Einstein de la chaleur spécifique des solides (Einstein, 1907).

l'information, Léon Brillouin montre comment lever le paradoxe pour un démon qui serait un être vivant et pas seulement un automate.

- Années 1950 : formulation du problème à N corps. Stimulera la recherche de nouveaux états de la matière.
- 1957 : théorie de la supraconductivité par John Bardeen, Leon Neil Cooper et John Robert Schrieffer (prix Nobel 1972).
- Années 1970 : théorie des phénomènes critiques.

Chapitre 2

L'approche probabiliste

Plan du chapitre

I	Nécessité de l'approche probabiliste	15
I.A	Une nécessité pratique	16
I.B	Une nécessité de principe : chaos, perturbations extérieures et mécanique quantique	17
II	Quelques éléments de théorie des probabilités	17
II.A	Évènements et probabilités	17
II.B	Variable aléatoire discrète	18
II.C	Variable aléatoire continue – Distribution/densité de probabilité	18
II.D	Loi jointe – loi marginale	20
II.E	Moyenne, variance, moments, cumulants, corrélations,...	20
II.F	Fonctions génératrices	24
II.G	Exemples	25
II.H	Remarque : et si les moments n'existent pas ?	25
II.I	Théorème de la limite centrale	26
III	Microétats, macroétats, information et entropie statistique	28
III.A	Microétats, macroétats	28
III.B	Quantifier l'information manquante : entropie statistique	28
III.C	Propriétés de l'entropie	29
	Annexe 2.A : Fonction de grande déviation	32
	Annexe 2.B : Exercices	36

I Nécessité de l'approche probabiliste

Comme nous l'avons dit, la physique statistique s'intéresse aux systèmes ayant un très grand nombre de degrés de liberté, typiquement de l'ordre du nombre d'Avogadro $N \sim \mathcal{N}_A$, où

$$\mathcal{N}_A = 6.022\,141\,29(27)\,10^{23}\,\text{mol}^{-1} \quad (2.1)$$

La physique statistique tourne à son avantage le fait que le nombre de constituants élémentaires soit extraordinairement grand, ce qui justifie d'abandonner l'idée de suivre l'évolution de chaque constituant élémentaire et de plutôt suivre une approche probabiliste.

I.A Une nécessité pratique

Si l'impossibilité pratique d'analyser la dynamique (classique) de tous les constituants d'un gaz était évidente à la fin du XIX^e siècle, examinons ce qu'il en est à l'heure de l'informatique. Considérons 1 cm³ de gaz parfait monoatomique dans les conditions normales de température et de pression ($T = 300 \text{ K} \simeq 27^\circ\text{C}$ et $p = 1 \text{ atm} = 10^5 \text{ Pa}$) et cherchons à caractériser l'état microscopique du système à un instant donné, entreprise dont les fruits semblent bien modestes !

Le nombre d'atomes dans le gaz est

$$N = \frac{10^{-3}}{24} \mathcal{N}_A \simeq 2.5 \cdot 10^{20} \text{ atomes} \quad (2.2)$$

Calculons le volume d'information que nous devons stocker pour caractériser à un instant t toutes les positions et les vitesses des N atomes.

Première question : combien de chiffres significatifs sont nécessaires pour caractériser les positions/vitesses des atomes ? La distance typique entre atomes est $\delta = n^{-1/3}$ où $n = N/V$ est la densité moyenne. Il faut au moins avoir la position d'un atome à une fraction $1/p$ de cette échelle, par rapport à la taille de la boîte $L = V^{1/3}$. Une coordonnée x d'un atome doit donc être repérée sur un segment $[0, L]$ ayant au moins pL/δ graduations ($p = 10$ veut dire qu'il y a dix graduations à l'échelle de δ). Il nous faut donc $\log_{10}(pLn^{1/3})$ chiffres significatifs par atome et par coordonnée (il y en a six : trois coordonnées d'espace et trois pour les vitesses). Donc l'information totale que nous devons stocker est, mesurée en octets (1 octet=8 bits, i.e. un chiffre entre 0 et 15),

$$\text{Mém} = N \times \underbrace{\overbrace{\log_{10}(pLn^{1/3}) \times 6}^{\substack{\text{\# de chiffres significatifs} \\ \text{par atome (base 10)}}}}_{\substack{\text{\# d'octets par atome}}} \times \underbrace{\overbrace{\frac{\ln 10}{\ln 16}}^{\substack{\text{conversion} \\ \text{en octets}}}} \quad (2.3)$$

soit

$$\text{Mém} = \frac{N \ln(p^3 N)}{2 \ln 2} \quad (2.4)$$

Application numérique (pour une précision $p = 3$ très modeste) : $\text{Mém} \approx 10^{21}$ octets, soit 10^{10} disques durs de 100 Go (la capacité standard aujourd'hui).

Conclusion, si chaque être humain sur terre (environ $7 \cdot 10^9$) possédait un disque dur de 100 Go, cela suffirait à peine à stocker l'information (de la mécanique classique) correspondant à une « photographie » du gaz contenu dans 1 cm³ !

D'un point de vue pratique, il est donc exclu de chercher à prédire la dynamique de l'ensemble des parties élémentaires d'un corps macroscopique. Non seulement l'objectif est **hors de portée** (les expériences numériques de dynamique moléculaire peuvent suivre aujourd'hui quelques milliers de particules sans problème ; la limite des capacités numériques atteint plutôt le million de particules) mais soulignons également qu'il n'a **pas d'intérêt**. En effet, le nombre d'informations « intéressantes » que l'on souhaite extraire est en général très faible : température, pression, densité, à la rigueur des distributions,... la connaissance du nombre faramineux de données spécifiant les positions et les vitesses des particules ne nous dit pas clairement comment remonter à ces informations intéressantes.

I.B Une nécessité de principe : chaos, perturbations extérieures et mécanique quantique

D'un point de vue plus fondamental remarquons que la capacité prédictive des lois microscopiques sur l'évolution est très faible pour plusieurs raisons.

1. Dans le cadre classique : dès que la dynamique d'un système à grand nombre de degrés de liberté est un peu complexe, ses propriétés sont essentiellement **chaotiques**. L'absence d'intégrabilité (i.e. la faiblesse du nombre de constantes du mouvement, comme l'énergie), conduit à une *extrême sensibilité* de la dynamique dans les *conditions initiales*. C'est-à-dire que des variations infimes sur l'état initial du système se répercutent extrêmement rapidement par des évolutions divergentes.
2. Pour insister encore sur la faiblesse prédictive des équations du mouvement pour un nombre macroscopique de constituants, remarquons qu'il est difficile en pratique de définir précisément l'hamiltonien du système. Tout système « isolé » réel sera soumis à des perturbations extérieures résiduelles perturbant sa dynamique.
3. La seconde raison fondamentale a trait à la nature intrinsèquement **probabiliste** de la description **quantique** des constituants élémentaires.

II Quelques éléments de théorie des probabilités

II.A Évènements et probabilités

Lorsque l'occurrence d'un phénomène présente un caractère stochastique, soit pour des raisons intrinsèques comme dans le cadre quantique, soit pour des raisons pratiques, parce que le résultat dépend d'un trop grand nombre de paramètres qu'on ne peut pas contrôler, il est plus efficace de raisonner en termes probabilistes. C'est par exemple le cas si l'on joue à pile ou face (pour la seconde raison) ou si l'on étudie la désintégration d'un noyau atomique instable (pour la première raison).

Pour formaliser le problème on introduit l'*espace des évènements* (ou *éventualités*) pouvant se produire lors d'une expérience *reproductible*, $\mathcal{E} = \{e_1, e_2, \dots, e_M\}$, que nous avons ici supposés dénombrables et en nombre fini (M). Chaque évènement représente un résultat possible lors de la réalisation de l'expérience.

Exemple : Pile ou face.— Dans le jeu de pile ou face, $M = 2$ et $\mathcal{E} = \{\text{pile}, \text{face}\}$.

Considérons N réalisations de l'expérience telles que l'évènement e_m se produise N_m fois (donc $\sum_{m=1}^M N_m = N$). Remarquons que ce nombre est aléatoire : si on réalise N autres fois l'expérience, e_m se produira N'_m fois, a priori différent de N_m . La probabilité de l'évènement e_m est

$$P_m \stackrel{\text{def}}{=} \lim_{N \rightarrow \infty} \frac{N_m}{N} \quad (2.5)$$

Par construction on a $0 \leq P_m \leq 1$ et les M probabilités satisfont la condition de normalisation

$$\boxed{\sum_{m=1}^M P_m = 1} \quad (\text{normalisation}). \quad (2.6)$$

II.B Variable aléatoire discrète

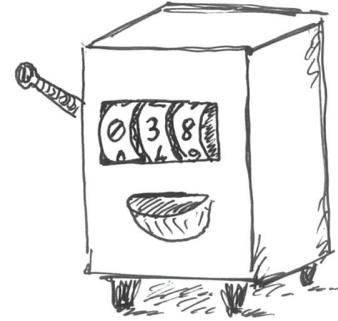
On considère une variable X prenant la valeur $X^{(m)}$ si l'évènement e_m est réalisé (X est entier, ou réel, ou complexe,...) : X est donc une *variable aléatoire*. Sa distribution est reliée aux probabilités $\{P_m\}$ comme

$$P(x) \stackrel{\text{def}}{=} \text{Proba}\{X = x\} = \sum_{m \text{ t.q. } X^{(m)}=x} P_m = \sum_{m=1}^M P_m \delta_{x, X^{(m)}}, \quad (2.7)$$

où le symbole de Kronecker $\delta_{i,j}$ (0 si les arguments sont distincts, 1 s'ils sont égaux) nous a permis de déplacer la contrainte ailleurs que sur la sommation. Notons que si les valeurs $X^{(m)}$ sont toutes distinctes, on a plus simplement $P(X^{(m)}) = P_m$.

Exemple : Machine à sous.— On joue avec une machine à sous qui peut sortir aléatoirement trois chiffres entre 0 et 9, i.e. $M = 10^3 = 1000$ résultats. Les 1000 évènements sont équiprobables : $P_1 = \dots = P_{1000} = 1/1000$ (sinon le casino triche).

- Les 10 évènements pour lesquels les trois chiffres sont identiques, notés e_1 (000) à e_{10} (999), font gagner $s_1 \stackrel{\text{def}}{=} S^{(1)} = \dots = S^{(10)} = 100 \text{ €}$.
- Les $10 \times 9 \times 3 = 270$ évènements pour lesquels deux chiffres sont égaux sur les trois, notés e_{11} (100), e_{12} (200), ..., e_{19} (900), e_{20} (010), ..., e_{280} (998), font gagner $s_2 \stackrel{\text{def}}{=} S^{(11)} = \dots = S^{(280)} = 10 \text{ €}$.
- Les 720 autres évènements ne font rien gagner $s_3 \stackrel{\text{def}}{=} S^{(281)} = \dots = S^{(1000)} = 0 \text{ €}$.



La somme S est donc une variable aléatoire qui peut prendre trois valeurs, $s_1 = 100 \text{ €}$ avec probabilité $P(s_1) = 1/100$, $s_2 = 10 \text{ €}$ avec probabilité $P(s_2) = 27/100$ et $s_3 = 0 \text{ €}$ avec probabilité $P(s_3) = 18/25$ (figure 2.1)

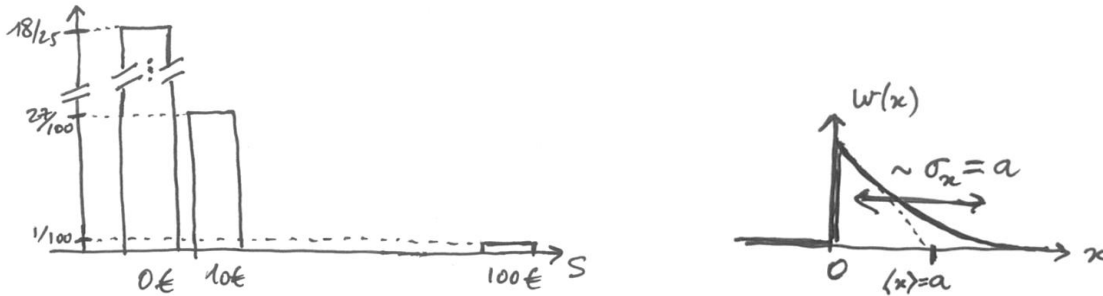


FIGURE 2.1 : **Distributions.** *distribution d'une variable aléatoire discrète (à gauche). À droite : distribution d'une variable continue (loi exponentielle).*

II.C Variable aléatoire continue – Distribution/densité de probabilité

Nous avons jusqu'à présent discuté le cas d'une *variable aléatoire discrète*, prenant des valeurs parmi un ensemble dénombrable, $\{0\text{€}, 10\text{€}, 100\text{€}\}$ dans l'exemple. Toutefois, dans de nombreux cas l'espace des évènements n'est pas un ensemble dénombrable et la variable aléatoire peut varier dans un certain domaine de $\mathbb{R}$ (ou de $\mathbb{C}$, ou de $\mathbb{R}^n$, etc) : on dit qu'il s'agit d'une *variable aléatoire continue*.

Exemple : Mouvement brownien.— Si on considère le mouvement brownien d'un grain de pollen à la surface de l'eau, sa position sera une variable aléatoire vectorielle $(x, y) \in \mathbb{R}^2$ distribuée dans le plan (selon une loi gaussienne, ainsi que l'a mesuré Jean Perrin en 1908), cf. Fig. 2.2.

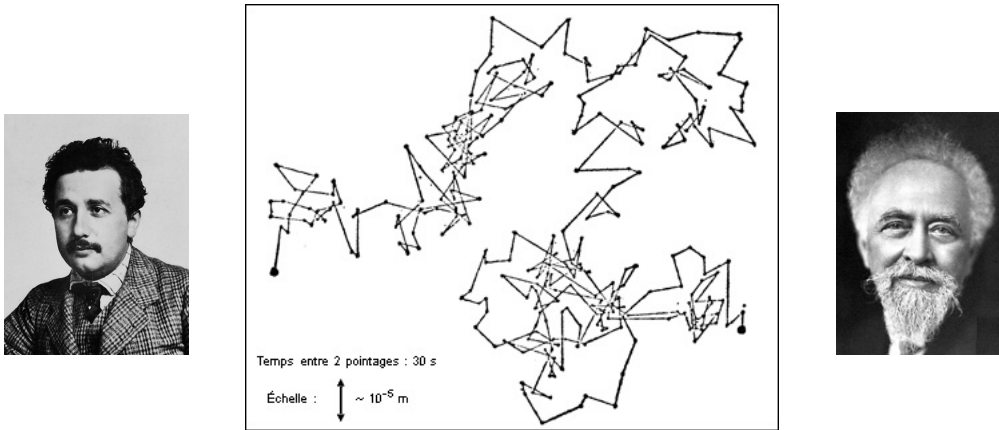


FIGURE 2.2 : Mouvement brownien d'un grain de pollen.

Nous avons déjà introduit la notion de distribution d'une variable aléatoire discrète : (2.7) est la probabilité pour que la variable aléatoire X prenne la valeur x parmi un ensemble discret de valeurs $\{x_1, x_2, \dots, x_V\}$ (notons que le nombre de valeurs distinctes est $V \leq M$ puisque la même valeur peut être réalisée pour plusieurs événements, comme on l'a vu sur notre exemple de machine à sous).

Si la variable est continue, on introduit la *densité de probabilité* $w(x)$ qui mesure la probabilité pour que la variable aléatoire X prenne la valeur x :

$$w(x) dx \stackrel{\text{def}}{=} \text{Proba}\{X \in [x, x + dx]\} . \quad (2.8)$$

La condition de normalisation est alors

$$\boxed{\int w(x) dx = 1} \quad (2.9)$$

où l'intégrale porte sur le domaine de définition de $w(x)$, le « support de la distribution ».

Remarque : changement de variable.— Considérons une variable (aléatoire) fonction de la variable aléatoire Y reliée à la variable X par $X = f(Y)$ où la fonction inverse f^{-1} est supposée bijective sur le support de la distribution de X . Question : étant donnée la distribution $w(x)$ de X , quelle est la distribution $p(y)$ de Y ?

$$p(y) dy = \text{Proba}\{X \in [f(y), f(y) + \underbrace{f'(y)dy}_{dx}]\} = w(f(y)) f'(y) dy \quad \Rightarrow p(y) = w(f(y)) f'(y) \quad (2.10)$$

Nous vérifions que cette transformation respecte la condition de normalisation $\int p(y) dy = \int w(x) dx = 1$; c'est un bon moyen de se rappeler que $\boxed{p(y) dy = w(x) dx}$.

Exemple : considérons une particule en une dimension d'énergie $E = \frac{1}{2m}p^2$. Connaissant la distribution de l'impulsion $f(p)$, nous obtenons la distribution de son énergie $p(E) = \sqrt{E/m} f(\sqrt{2mE})$.

II.D Loi jointe – loi marginale

Considérons le cas de n variables scalaires aléatoires $X_1, X_2, \dots, X_n$ (remarquons que nous pouvons regrouper ces variables scalaires dans une unique variable vectorielle $\vec{X}$). La distribution de ces variables est une fonction à plusieurs variables, appelée la « *loi de distribution jointe* » : $w(x_1, \dots, x_n)$. I.e.

$$w(x_1, \dots, x_n) dx_1 \cdots dx_n = \text{Proba}\{X_1 \in [x_1, x_1 + dx_1] \& \cdots \& X_n \in [x_n, x_n + dx_n]\}. \quad (2.11)$$

Nous rencontrerons de nombreuses situations où la loi jointe de plusieurs variables est connue mais où seule une de ces variables nous intéresse. La distribution de la variable X_1 , *sa loi marginale*, est obtenue en intégrant sur les autres variables :

$$w_1(x_1) dx_1 = dx_1 \int dx_2 \cdots \int dx_n w(x_1, \dots, x_n) \quad (2.12)$$

$$= \text{Proba}\{X_1 \in [x_1, x_1 + dx_1] \text{ sachant que les } n-1 \text{ autres variables peuvent valoir n'importe quoi}\}. \quad (2.13)$$

II.E Moyenne, variance, moments, cumulants, corrélations,...

Supposons que l'on réalise N expériences dont l'espace des événements est $\mathcal{E} = \{e_1, e_2, \dots, e_M\}$. Lors de chaque expérience on mesure la variable aléatoire X pouvant prendre V valeurs (dans l'exemple de la machine à sous : $X \rightarrow S$, $M = 1000$ et $V = 3$). L'expérience est réalisée N fois : se produisent les événements $e_{m_1}, \dots, e_{m_N}$, auxquels nous pouvons associer la suite de valeurs $X^{(m_1)}, \dots, X^{(m_N)}$.

Moyenne

Une information intéressante est donnée par la moyenne *statistique* sur un ensemble de N réalisations :

$$\bar{X}^{(N)} = \frac{1}{N} \sum_{i=1}^N X^{(m_i)} = \frac{1}{N} \sum_{m=1}^M N_m X^{(m)} \quad (2.14)$$

La première somme porte sur les N événements qui se sont effectivement réalisés (la suite pile, face, face, pile, etc), alors que la seconde porte sur l'ensemble des M événements (les deux possibilités pile ou face), ce pourquoi nous avons réintroduit le nombre d'occurrences N_m de l'événement e_m . Remarquons que si nous réalisons un autre ensemble de N expériences, nous obtenons une suite d'événements différente (et des N'_m distincts) et la moyenne statistique est également différente $\bar{X}^{(N)} \neq \bar{X}'^{(N)}$.

Nous définissons maintenant la moyenne *probabiliste* (aussi appelée « *espérance mathématique* », notée $\mathbb{E}(X)$) comme la limite de la moyenne statistique pour $N \rightarrow \infty$: $\langle X \rangle \stackrel{\text{def}}{=} \lim_{N \rightarrow \infty} \bar{X}^{(N)} = \lim_{N \rightarrow \infty} \bar{X}'^{(N)} = \dots$ (à la limite $N \rightarrow \infty$, les moyennes statistiques obtenues pour toute les suites distinctes d'événements coïncident, ce qui sera prouvé plus bas). Nous déduisons que la moyenne probabiliste peut s'exprimer en termes de la distribution des événements

$$\langle X \rangle = \sum_{m=1}^M P_m X^{(m)} \quad (2.15)$$

Remarquons que lorsqu'on s'intéresse aux propriétés de la variable aléatoire, il est plus naturel d'exprimer les moyenne à l'aide de la distribution de celle-ci :

$$\langle X \rangle = \sum_{v=1}^V P(x_v) x_v \quad (2.16)$$

où la somme porte sur les valeurs distinctes prises par la variable aléatoire. Par exemple, sur l'exemple de la machine à sous discuté ci-dessus, la variable S ne prend que trois valeurs, $s_v \in \{0 \text{ €}, 10 \text{ €}, 100 \text{ €}\}$. Sa distribution $P(s_v)$ a été donnée plus haut.

Dans le cas d'une variable continue distribuée selon la densité de probabilité $w(x)$, par analogie avec l'expression précédente, nous voyons que la moyenne prend la forme

$$\langle X \rangle = \int dx w(x) x. \quad (2.17)$$

Remarquons finalement que les formules (2.16) et (2.17) peuvent être généralisées pour calculer la moyenne de n'importe quelle fonction de la variable aléatoire,

$$\langle f(X) \rangle = \sum_{v=1}^V P(x_v) f(x_v) \quad (2.18)$$

pour une variable aléatoire discrète et

$$\langle f(X) \rangle = \int dx w(x) f(x) \quad (2.19)$$

pour une variable aléatoire continue.

Variance et écart-type

$\langle X \rangle$ nous dit en moyenne quelle valeur prendra la variable aléatoire si elle est mesurée un nombre $N \rightarrow \infty$ de fois. Une autre information intéressante serait de caractériser l'ordre de grandeur des écarts à la moyenne $X - \langle X \rangle$. Pour cela nous définissons la variance ¹

$$\text{Var}(X) \stackrel{\text{def}}{=} \langle (X - \langle X \rangle)^2 \rangle = \langle X^2 \rangle - \langle X \rangle^2 \quad (2.20)$$

où l'ordre de grandeur des *fluctuations* de la variable aléatoire autour de sa moyenne est caractérisé par l'*écart type* (qui a la dimension de X)

$$\sigma_X \stackrel{\text{def}}{=} \sqrt{\text{Var}(X)} \quad (2.21)$$

Ce qui veut dire que si on réalise une expérience, la variable prendra une valeur (en gros) $X^{(m)} \approx \langle X \rangle$ à $\pm \sigma_X$ près; notons que cette formulation approximative n'a d'intérêt que si $\sigma_X \ll |\langle X \rangle|$, sinon nous ne pouvons pas synthétiser aussi simplement l'information contenue dans la distribution.

Exemple : loi exponentielle Considérons par exemple la loi exponentielle $w(x) = \frac{1}{a} e^{-x/a}$ pour $x > 0$ (figure 2.1). Nous obtenons $\langle X \rangle = a$ et $\langle X^2 \rangle = 2a^2$, donc $\text{Var}(X) = a^2$ et $\sigma_X = a$.

Indépendance et corrélations

Considérons deux événements A et B . S'ils sont statistiquement indépendants (non corrélés) la probabilité d'occurrence des deux événements est donnée par le produit des probabilités

$$P(A \& B) = P(A) P(B) \quad (\text{définition de l'indépendance statistique}) \quad (2.22)$$

En revanche, cette propriété ne sera pas respectée si les deux événements sont corrélés, $P(A \& B) \neq P(A) P(B)$.

¹Nous pourrions introduire une autre quantité pour caractériser cette information, comme $\langle |X - \langle X \rangle| \rangle$. Nous allons voir plus loin pourquoi considérer cette quantité n'est pas très judicieux.

Exemple : Considérons une situation simple : les deux événements correspondent à $A = \{\text{Alice dans la salle C20}\}$ et $B = \{\text{Bob dans la salle C115}\}$. Chaque jour les étudiants se distribuent aléatoirement parmi les cinq groupes de TD. Trois situations :

1. Premier jour de la rentrée, Alice et Bob ne se connaissent pas, les étudiants se répartissent aléatoirement dans les salles de TD $\Rightarrow P(A \& B) = P(A) P(B)$.
2. Après un mois, des affinités se sont formées, Alice et Bob souhaitent être dans le même groupe de TD (corrélations sur leurs positions) $\Rightarrow P(A \& B) < P(A) P(B)$.
3. Après deux mois, premières brouilles, Alice et Bob sont fâchés et s'évitent (anti-corrélations sur leurs positions) $\Rightarrow P(A \& B) > P(A) P(B)$.

Variables indépendantes ou corrélées.— Dans la pratique il est courant de chercher à révéler l'existence de corrélations par l'étude de propriétés simples de variables aléatoires plutôt que par l'analyse de la distribution complète.

Considérons par exemple deux variables aléatoires X_A et X_B (par exemple les positions de Alice et Bob). Nous pouvons rencontrer trois situations (correspondant à trois types de distributions de ces deux variables) :

1. Les deux variables sont **indépendantes**, la moyenne du produit est :

$$\langle X_A X_B \rangle = \sum_{x, y} x y \underbrace{\text{Proba}\{X_A = x \& X_B = y\}}_{=\text{Proba}\{X_A=x\} \text{Proba}\{X_B=y\}} = \langle X_A \rangle \langle X_B \rangle \quad (2.23)$$

où la somme $\sum_x$ porte sur les différentes valeurs prises par les deux variables aléatoires (par exemple les positions des salles de TD dans lesquelles se trouvent Alice et Bob).

2. Les positions sont **corrélées** (Alice et Bob veulent rester ensemble) : une fluctuation de X_A autour de $\langle X_A \rangle$ sera corrélée à une fluctuation de X_B autour de $\langle X_B \rangle$ de même signe

$$\langle X_A X_B \rangle > \langle X_A \rangle \langle X_B \rangle . \quad (2.24)$$

3. Les positions sont **anti-corrélées** (Alice et Bob s'évitent) : une fluctuation de $X_A - \langle X_A \rangle > 0$ sera corrélée à une fluctuation $X_B - \langle X_B \rangle < 0$:

$$\langle X_A X_B \rangle < \langle X_A \rangle \langle X_B \rangle . \quad (2.25)$$

De manière générale, si l'on considère N variables aléatoires, il est commode d'introduire la matrice des covariances

$$C_{ij} \stackrel{\text{def}}{=} \langle X_i X_j \rangle - \langle X_i \rangle \langle X_j \rangle = \langle (X_i - \langle X_i \rangle)(X_j - \langle X_j \rangle) \rangle \quad (2.26)$$

afin de caractériser les corrélations entre elles :

$$C_{ij} \begin{cases} = 0 & X_i \& X_j \text{ indépendantes} \\ > 0 & X_i \& X_j \text{ corrélées} \\ < 0 & X_i \& X_j \text{ anti-corrélées} \end{cases}$$

Remarquons que $C_{ii} = \text{Var}(X_i)$. En outre on peut montrer que la matrice C_{ij} est symétrique définie positive, i.e. il existe une matrice orthogonale $\mathcal{O}$ telle que $C = \mathcal{O} \text{diag}(\sigma_1^2, \dots, \sigma_N^2) \mathcal{O}^T$. Autrement dit les nouvelles variables $Y_i = \sum_j \mathcal{O}_{ji} X_j$ sont des variables indépendantes de variances σ_i^2 .

Moments et cumulants

La notion de « *moment* » généralise la notion de moyenne et de valeur quadratique moyenne. Le n ième moment est par définition $\mu_n \stackrel{\text{def}}{=} \langle X^n \rangle$. Notons que la donnée de l'ensemble des moments est aussi informative que la donnée de la distribution (sauf distribution pathologique) : il s'agit de deux manières d'encoder l'information sur les propriétés statistiques de la variable aléatoire.

Si les fluctuations relatives de la variable sont faibles $\sigma_X \ll |\langle X \rangle|$, nous voyons que le second moment est essentiellement dominé par le premier moment : $\langle X^2 \rangle = \langle X \rangle^2 + \text{Var}(X) \simeq \langle X \rangle^2$. L'information intéressante (dans le sens « l'information apportée par $\langle X^2 \rangle$ qui n'est pas contenue dans $\langle X \rangle$ ») est donc masquée, ce qui permet de comprendre l'intérêt de la notion de variance dans ce cas. La variance est aussi appelée le *second cumulant* $c_2 \stackrel{\text{def}}{=} \text{Var}(X)$ de la distribution (la moyenne étant le premier $c_1 \stackrel{\text{def}}{=} \langle X \rangle$). On peut également généraliser la notion et introduire le cumulant d'ordre supérieur d'ordre n , noté c_n , caractérisant l'information statistique qui n'est pas contenue dans les cumulants précédents. Précisons ce point :

1. *Moyenne.* – La première information est donnée par la moyenne $c_1 \stackrel{\text{def}}{=} \mu_1 = \langle X \rangle$.
2. *Variance.* – Le second cumulant $c_2 \stackrel{\text{def}}{=} \mu_2 - (\mu_1)^2 = \text{Var}(X) = \langle \tilde{X}^2 \rangle$, où $\tilde{X} \stackrel{\text{def}}{=} X - \langle X \rangle$, caractérise la *largeur* de la distribution (cf. figure 2.3).

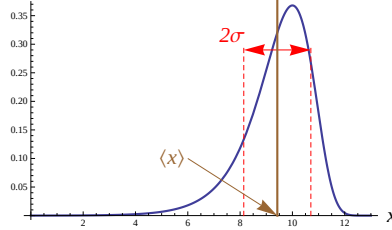


FIGURE 2.3 : L'écart type $\sigma = \sqrt{\text{Var}(X)}$ donne une mesure de la largeur de la distribution.

3. *Skewness.* – Le second cumulant indique l'ordre de grandeur des fluctuations autour de la moyenne, mais ne dit pas si ces fluctuations sont préférentiellement positives ou négatives. Nous introduisons le troisième cumulant $c_3 \stackrel{\text{def}}{=} \mu_3 - 3\mu_1\mu_2 + 2(\mu_1)^3 = \langle \tilde{X}^3 \rangle$ caractérisant l'*asymétrie* de la distribution autour de sa moyenne ($c_3 > 0$ signifie que les fluctuations positives sont plus probables). Le coefficient de dissymétrie (« skewness »), adimensionné, est défini comme $\gamma_1 = c_3/c_2^{3/2}$. Cf. figure 2.4.

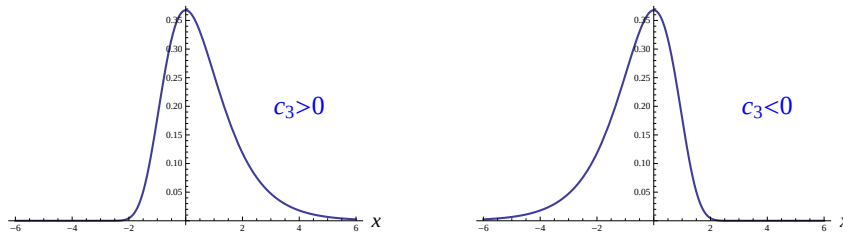


FIGURE 2.4 : À gauche : la loi de Gumbel $f(x) = \exp(-x - e^{-x})$ a pour premiers cumulants : $c_1 = \mathbf{C} = 0.577\dots$ (constante d'Euler), $c_2 = \pi^2/6$, $c_3 = 2\zeta(3)$ (fonction zeta de Riemann), $c_4 = \pi^4/15$, etc. Le signe de c_3 indique si les fluctuations positives ou négatives (relativement à $\langle X \rangle$) sont favorisées.

4. *Kurtosis.* – La distribution entièrement caractérisée par ses deux premiers cumulants est la distribution gaussienne pour laquelle $\langle \tilde{X}^4 \rangle = 3\langle \tilde{X}^2 \rangle^2$. Nous introduisons une quantité qui

caractérise la *déviaton à la gaussienne* : $c_4 \stackrel{\text{def}}{=} \mu_4 - 3(\mu_2)^2 + 12\mu_2(\mu_1)^2 - 4\mu_3\mu_1 - 6(\mu_1)^4 = \langle \tilde{X}^4 \rangle - 3\langle \tilde{X}^2 \rangle^2$ ($c_4 > 0$ veut dire que les grandes fluctuations sont plus favorables que dans le cas d'une distribution gaussienne). On définit un coefficient d'aplatissement (« kurtosis ») $\gamma_2 = c_4/c_2^2$, adimensionné. Cf. figure 2.5.

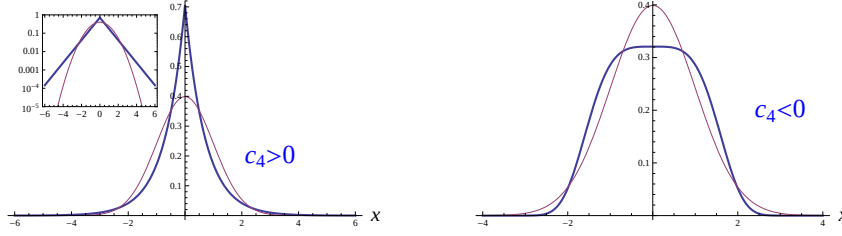


FIGURE 2.5 : La distribution symétrique $g_a(x) = \frac{a\sqrt{\Gamma(3/a)}}{2\Gamma(1/a)^{3/2}} \exp\left(-\left|\sqrt{\frac{\Gamma(3/a)}{\Gamma(1/a)}} x\right|^a\right)$, ici tracée pour $a = 1$ et $a = 4$, a pour premiers cumulants (non nuls) : $c_2 = 1$, $c_4 = \frac{\Gamma(1/a)\Gamma(5/a)}{\Gamma(3/a)^2} - 3$, etc. Le quatrième cumulant change de signe pour $a = 2$. La courbe rose est la distribution gaussienne ($a = 2$).

Il ne semble pas évident de poursuivre cette construction afin de définir des coefficients c_5 , c_6 , etc. Nous expliquons dans le paragraphe suivant comme définir systématiquement les cumulants.

II.F Fonctions génératrices

L'intérêt de l'analyse des moments (ou des cumulants) est de proposer un moyen de caractériser les propriétés statistique, autre que l'analyse directe de la distribution $w(x)$, souvent plus intuitif comme l'a montré la discussion sur la moyenne et l'écart type. Dans les problèmes concrets il s'avère bien souvent commode de calculer la *fonction génératrice des moments* :

$$g(p) \stackrel{\text{def}}{=} \langle e^{-pX} \rangle = \int dx w(x) e^{-px} \quad (2.27)$$

qui représente la transformée de Laplace de la distribution (nous considérons le cas d'une variable aléatoire continue ; nous aurions aussi bien pu discuter le cas discret). Notons que la condition de normalisation prend la forme $g(0) = 1$. Le paramètre p pourra être complexe pour assurer la convergence de l'intégrale : $g(ik)$ est alors la transformée de Fourier de la distribution. L'intérêt de cette notion est qu'il s'agit souvent de la quantité aisée à calculer, la physique statistique nous en donnera de multiples illustrations. La fonction génératrice ne porte pas de sens « physique » mais son intérêt est qu'étant connue, elle permet d'obtenir directement les moments $\mu_n \stackrel{\text{def}}{=} \langle X^n \rangle$ par dérivation puisque

$$g(p) = \sum_{n=0}^{\infty} \frac{(-p)^n}{n!} \mu_n \quad \Rightarrow \quad \mu_n = (-1)^n g^{(n)}(0). \quad (2.28)$$

De même nous introduisons la *fonction génératrice des cumulants* :

$$\omega(p) \stackrel{\text{def}}{=} \sum_{n=1}^{\infty} \frac{(-p)^n}{n!} c_n \quad \Rightarrow \quad c_n = (-1)^n \omega^{(n)}(0). \quad (2.29)$$

On prendra maintenant la relation entre les deux fonctions génératrices

$$\omega(p) = \ln g(p) \quad (2.30)$$

comme une **définition des cumulants**. En identifiant les deux développements en puissance de p , on peut vérifier que l'on retrouve les définitions des quatre premiers cumulants données plus haut. Par exemple (à l'ordre 2) $-c_1p + \frac{1}{2}c_2p^2 + \mathcal{O}(p^3) = \ln[1 - \mu_1p + \frac{1}{2}\mu_2p^2 + \mathcal{O}(p^3)]$, d'où, en développant le logarithme, $-c_1p + \frac{1}{2}c_2p^2 + \mathcal{O}(p^3) = -\mu_1p + \frac{1}{2}(\mu_2 - \mu_1^2)p^2 + \mathcal{O}(p^3)$. QED.

II.G Exemples

Loi binomiale.— Considérons N évènements, chacun associé à deux éventualités possibles (par exemple pile ou face). L'évènement « pile » se produit avec probabilité p et l'évènement face avec probabilité $1-p$. La probabilité pour que n évènements pile se produisent est : $\Pi_N(n) = C_N^n p^n (1-p)^{N-n}$ (attention, ici N est un paramètre fixé, tandis que n est la variable aléatoire). Il est commode de prendre la définition suivante pour la fonction génératrice : $g_N(s) \stackrel{\text{def}}{=} \langle s^n \rangle$ (si on écrit $s = e^{-p}$, cela correspond à la définition donnée ci-dessus). On obtient sans difficulté : $g_N(s) = [1 + p(s-1)]^N$. On peut trouver les deux premiers cumulants en utilisant $\langle n \rangle = g'_N(1)$ et $\langle n(n-1) \rangle = g''_N(1)$, d'où $\langle n \rangle = Np$ et $\text{Var}(n) = Np(1-p)$.

Loi de Poisson.— Considérons des évènements *indépendants*. La probabilité pour qu'un évènement se produise pendant un intervalle de temps infinitésimal dt est indépendante du temps, donnée par λdt . Par exemple il peut s'agir du passage des électrons arrachés à la cathode d'une lampe à vide (Schottky, 1918) ; dans ce cas $q_e \lambda$ est le courant électrique moyen. La probabilité pour que n évènements se produisent pendant un temps T (avec $\nu = \lambda T$) est donnée par la loi de Poisson $P_\nu(n) = \frac{\nu^n}{n!} e^{-\nu}$. Les fonctions génératrices des moments et des cumulants sont respectivement $g_\nu(p) = \exp[\nu(e^p - 1)]$ et $\omega_\nu(p) = \nu(e^p - 1)$, ce qui montre que les cumulants sont tous égaux : $c_k = \nu$. En particulier $\langle n \rangle = \text{Var}(n) = \nu$.

Distribution gaussienne.— Considérons une variable continue X distribuée selon la loi (à deux paramètres) : $p_{a,\sigma}(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp -\frac{(x-a)^2}{2\sigma^2}$. Nous calculons sans difficulté la fonction génératrice $g_{a,\sigma}(p) = \exp[-ap + \frac{1}{2}\sigma^2 p^2]$. La fonction génératrice des cumulants est quadratique $\omega_{a,\sigma}(p) = -ap + \frac{1}{2}\sigma^2 p^2$, ce qui montre que la loi gaussienne est la seule loi dont seuls les deux premiers cumulants $c_1 = \langle X \rangle = a$ et $c_2 = \text{Var}(X) = \sigma^2$ sont non nuls.

II.H Remarque : et si les moments n'existent pas ?

En écrivant (2.28), nous avons implicitement admis l'existence des moments μ_n . On peut toutefois imaginer des situations où les moments deviennent infinis : il suffit que la distribution décroisse à l'infini en loi de puissance. Par exemple si $w(x) \sim |x|^{-\mu-1}$ pour $|x| \rightarrow \infty$, le moment μ_n est fini si $n < \mu$. L'existence de moments divergents se manifeste dans la fonction génératrice par un comportement non analytique.

Exemples

Loi de Cauchy.— Considérons par exemple la loi $w(x) = \frac{a/\pi}{a^2+x^2}$ (l'exposant est $\mu = 1$, donc tous les moments sont divergents, $\mu_n = \infty \forall n$). Nous pouvons calculer la fonction génératrice $\tilde{g}(k) \stackrel{\text{def}}{=} \langle e^{-ikX} \rangle = e^{-a|k|}$. Cette fonction est $\mathcal{C}_0(\mathbb{R})$, i.e. continue, mais non dérivable en $k = 0$, ce qui signale l'absence d'un développement de Taylor (2.28), i.e. la divergence de tous les moments $\langle X^{2n} \rangle = \infty$.

Un autre exemple.— Étudions maintenant la loi $w(x) = \frac{4/\pi}{4+x^4}$; en utilisant le théorème des résidus, on vérifiera que $\tilde{g}(k) = \langle e^{-ikX} \rangle = [\cos k + \sin |k|] e^{-|k|} \in \mathcal{C}_2(\mathbb{R})$ (en effet $\tilde{g}(k) = 1 - k^2 + \frac{2}{3}|k|^3 + \mathcal{O}(k^4)$) ce qui indique que seuls les deux premiers moments sont finis : $\langle X \rangle = 0$, $\langle X^2 \rangle = 2$ et $\langle X^{2n} \rangle = \infty$ pour $n > 1$.

II.I Théorème de la limite centrale

Un problème tout à fait générique est le suivant : considérons un ensemble de N variables aléatoires $X_1, \dots, X_N$

- *statistiquement indépendantes*, non corrélées (hypothèse 1).
- *Identiques*, i.e. toutes distribuées selon la même loi de probabilité que nous notons $p(x)$ (hypothèse 2).
- Nous ne faisons aucune hypothèse sur la forme de la distribution $p(x)$, sinon que ses deux premiers moments existent $\langle X_n \rangle < \infty$ et $\langle X_n^2 \rangle < \infty$ (hypothèse 3).

Nous souhaitons déterminer la distribution de la somme de ces variables aléatoires

$$S = \sum_{n=1}^N X_n \quad (2.31)$$

(qui est elle même une variable aléatoire).

Illustrations : Nous pouvons donner deux illustrations concrètes :

- Nous avons introduit plus haut la moyenne statistique $\bar{X}^{(N)} = \frac{1}{N} \sum_{n=1}^N X_n$, qui coïncide en effet à la variable S , à un facteur $1/N$ près. Une question intéressante est de caractériser les fluctuations de la moyenne statistique (pour contrôler l'erreur qui serait faite en pratique par un statisticien).
- Si nous considérons un gaz d'atomes ou de molécules, son énergie est dominée par l'énergie cinétique (c'est presque la définition du gaz, l'énergie potentielle caractérisant l'interaction entre atomes/molécules est négligeable devant l'énergie cinétique, $E_c \gg E_{\text{int}}$). Une question pourrait être de caractériser la distribution de l'énergie du gaz atomique $E_c = \sum_{i=1}^N \frac{1}{2m} \vec{p}_i^2$.

Détermination de la distribution de S : Afin de déterminer la distribution $P_N(S)$ de la variable S , la manière la plus simple de procéder est de faire appel à la notion de fonction génératrice. Considérons la fonction génératrice de la somme $G_N(p) \stackrel{\text{def}}{=} \langle \exp(-pS) \rangle = \langle \exp(-p \sum_{n=1}^N X_n) \rangle$. En utilisant l'indépendance statistique des variables (hyp.1), nous faisons apparaître le produit des fonctions génératrices $G_N(p) = \prod_{n=1}^N \langle e^{-pX_n} \rangle$. Le fait que les variables soient identiquement distribuées (hyp.2) nous permet de relier simplement la fonction génératrice des variables à celle de leur somme :

$$G_N(p) = [g(p)]^N. \quad (2.32)$$

Cette relation nous suggère d'écrire la fonction génératrice sous une forme exponentielle $g(p) = \exp \omega(p)$, i.e. de faire apparaître la fonction génératrice des *cumulants*. Nous notons $\Omega_N(p) \stackrel{\text{def}}{=} \sum_{k=1}^{\infty} \frac{(-p)^k}{k!} C_k = \ln G_N(p)$ la fonction génératrice des cumulants de S . Nous obtenons

$$\Omega_N(p) = N \omega(p) = N \sum_{k=1}^{\infty} \frac{(-p)^k}{k!} c_k \quad \text{i.e.} \quad \boxed{C_k = N c_k} \quad (2.33)$$

Pour les premiers :

$$\langle S \rangle = N \langle X \rangle \quad (\text{évident}), \quad (2.34)$$

$$\text{Var}(S) = N \text{Var}(X) \quad (\text{déjà moins trivial}), \quad (2.35)$$

etc.

En particulier la seconde relation montre que les écarts types (i.e. l'ordre de grandeur des fluctuations) sont reliés comme

$$\boxed{\sigma_S = \sqrt{N} \sigma_X} \quad (2.36)$$

Les fluctuations de la somme de N variables indépendantes se comportent typiquement² comme $\delta S \sim \sigma_S \propto \sqrt{N}$. Nous justifions maintenant que dans la limite $N \rightarrow \infty$ nous pouvons négliger les autres cumulants de la variables S . La raison est la suivante : tous les cumulants de S croissent linéairement avec N : $C_k \propto N$, or ce dernier contribue aux fluctuations de S comme $\delta_k S \stackrel{\text{def}}{=} (C_k)^{1/k} \propto N^{1/k}$. Cette contribution décroît plus vite que celle du second cumulant :

$$(C_k)^{1/k} / (C_2)^{1/2} \propto N^{-(k-2)/(2k)} \xrightarrow{N \rightarrow \infty} 0$$

(cette relation peut être comprise par analyse dimensionnelle puisque $[C_k] = [S]^k$). Finalement cela justifie de négliger tous les cumulants C_k avec $k > 2$:

$$G_N(p) \simeq e^{-Nc_1 p + \frac{1}{2} N c_2 p^2} \text{ pour } N \rightarrow \infty \quad (2.37)$$

i.e.

$$\boxed{P_N(S) \simeq \frac{1}{\sqrt{2\pi N \text{Var}(X)}} \exp \left\{ -\frac{(S - N \langle X \rangle)^2}{2N \text{Var}(X)} \right\}} \quad \text{pour } N \gg 1. \quad (2.38)$$

Théorème de la limite centrale : Soit $S = \sum_{n=1}^N X_n$ la somme de N variables aléatoires identiques et indépendantes dont les deux premiers moments sont finis, $\langle X \rangle < \infty$ et $\langle X^2 \rangle < \infty$.

Dans la limite $N \rightarrow \infty$, la distribution de S tend vers une loi **gaussienne** centrée sur $\langle S \rangle = N \langle X \rangle$ et d'écart type (i.e. de largeur) $\sigma_S = \sqrt{N} \sigma_X$.

Le théorème offre une illustration assez simple de la notion d'**émergence** introduite dans le premier chapitre : quelle que soit la distribution $p(x)$, pourvu que ses deux premiers moments soient finis, dans la limite $N \rightarrow \infty$ la distribution $P_N(S)$ tend vers une loi gaussienne. L'émergence de cette loi *universelle* n'est pas contenue dans les détails « microscopiques » de départ, la loi $p(x)$, mais dans le passage du microscopique au macroscopique (la limite $N \rightarrow \infty$).

Ce théorème sera d'une importance cruciale pour la physique statistique où nous aurons très souvent à considérer des grandeurs physiques (macroscopiques) somme d'un très grand nombre ($N \sim \mathcal{N}_A$) d'autres grandeurs (microscopiques) ; par exemple l'énergie d'un gaz, somme des énergies des particules qui le constituent.

Remarque : Une approche plus élégante consiste à « rescaler » la variable S en introduisant

$$Y \stackrel{\text{def}}{=} \frac{S - N \langle X \rangle}{\sqrt{N \text{Var}(X)}}. \quad (2.39)$$

Nous calculons la fonction génératrice de Y :

$$\gamma_N(p) \stackrel{\text{def}}{=} \langle e^{-pY} \rangle = \left\langle \exp -p \frac{\sum_n (X_n - c_1)}{\sqrt{N c_2}} \right\rangle \quad (2.40)$$

² Un point subtil : l'écart-type $\sigma_S \propto \sqrt{N}$ caractérise des fluctuations *typiques* (i.e. les plus probables), ce qui n'interdit pas l'existence de fluctuations atypiques (très peu probables). L'étude de ces dernières est dénommée l'étude des « grandes déviations » (cf. annexe).

où nous avons utilisé la notation $c_1 = \langle X \rangle$ et $c_2 = \text{Var}(X)$. Introduisons $\tilde{g}(p) = g(p) e^{c_1 p}$, la fonction génératrice de $X - \langle X \rangle$. Nous obtenons

$$\gamma_N(p) = \left[\tilde{g} \left(\frac{p}{\sqrt{N c_2}} \right) \right]^N = \exp N \left[\frac{1}{2} c_2 \left(\frac{p}{\sqrt{N c_2}} \right)^2 - \frac{1}{6} c_3 \left(\frac{p}{\sqrt{N c_2}} \right)^3 + \dots \right] \quad (2.41)$$

$$= \exp \left[\frac{1}{2} p^2 + \mathcal{O} \left(\frac{p^3}{\sqrt{N}} \right) \right] \quad (2.42)$$

Dans la limite $N \rightarrow \infty$ nous retrouvons la fonction génératrice $\gamma_\infty(p) = e^{\frac{1}{2} p^2}$ correspondant à la distribution gaussienne normalisée $p(y) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2} y^2}$. QED.

III Microétats, macroétats, information et entropie statistique

Dans la section précédente nous avons introduit les notions (le « vocabulaire ») de base qui nous serviront pour mettre en œuvre l'approche probabiliste dans le cadre de la physique statistique. Nous cherchons maintenant un principe nous guidant afin de construire la distribution de probabilité (l'ensemble des P_m) associée à un problème physique. Une notion centrale pour remplir cet objectif est celle d'*entropie statistique*.

III.A Microétats, macroétats

Nous introduisons deux notions essentielles pour la suite. Pour cela nous considérons l'exemple précis d'un gaz de particules (exemple ô combien important historiquement). Le gaz est décrit dans le cadre de la mécanique classique.

- **Microétat (état pur).**— Si tous les paramètres microscopiques définissant l'état du système (toutes les positions et les vitesses des particules) sont connus, nous disons que le système se trouve dans un « *microétat* » ou « *état pur* ».

Dans cette section nous notons $\{e_m\}$ les microétats (on a repris la notation de la section II, i.e. un microétat est ici vu comme une éventualité de l'état du système parmi un ensemble d'états possibles) :

$$e_m \leftrightarrow \text{microétat} .$$

- **Macroétat (mélange statistique).**— Si nous ne savons pas précisément dans quel microétat le système se trouve, nous pouvons en revanche caractériser avec quelle probabilité il occupe les différents microétats. La donnée des probabilités d'occupation $\{P_m\}$ des microétats définit un « *macroétat* » ou « *mélange statistique* » (de microétats)

$$\text{distribution } \{P_m\} \leftrightarrow \text{macroétat} .$$

Dans le cas du gaz, le macroétat est spécifié en précisant la densité de probabilité pour que les particules se trouvent en différentes positions de l'espace avec certaines vitesses.

III.B Quantifier l'information manquante : entropie statistique

La distinction entre les notions de microétat et de macroétat fait référence à l'*information* dont nous disposons sur le système : soit toute l'information nécessaire permettant de spécifier l'état microscopique (soit pile, soit face), soit l'information partielle spécifiant avec quelles probabilités sont occupés les microétats (l'ensemble des probabilités $\{P_{\text{pile}}, P_{\text{face}}\}$).

Il est donc essentiel de disposer d'un « outil » permettant une mesure de l'information, ou de la perte d'information liée à notre absence de connaissance de l'occupation précise des microétats.

La grandeur permettant de mesurer le *manque d'information* lié à l'occupation aléatoire des microétats est « l'entropie statistique » (ou entropie de Shannon, ou de Gibbs) :

$$S(\{P_m\}) \stackrel{\text{def}}{=} -k_B \sum_m P_m \ln P_m \quad (2.43)$$

où la constante k_B est arbitraire et correspond au « quantum d'entropie » (i.e. le quantum d'information manquante) ; nous la choisirons égale à la constante de Boltzmann, le rapport de la constante des gaz parfaits et du nombre d'Avogadro $k_B = R/\mathcal{N}_A$, afin d'assurer la correspondance avec l'entropie de la thermodynamique : ³

$$k_B = 1.380\,6505(24) \times 10^{-23} \text{ J.K}^{-1} \quad (2.44)$$

III.C Propriétés de l'entropie

D'où vient la définition (2.43) de l'entropie ? La définition (2.43) a été proposée en recherchant une fonction obéissant à un certain nombre de propriétés que nous énonçons maintenant. L'entropie de von Neumann (2.43) est une fonction des $\{P_m\}$ qui obéit à toutes ces propriétés « raisonnables ». ⁴ On peut même montrer que l'entropie de von Neumann (2.43) est la *seule* fonction qui permet de retrouver une entropie *extensive* avec la taille du système (on pourra consulter l'article pédagogique [14]).

Minimum L'entropie d'un microétat est minimum. Dans ce cas il ne manque aucune information pour spécifier l'état microscopique.

$$S_{\min} = S(\text{microétat}) = 0 \quad (2.45)$$

Très facile à vérifier puisque les probabilités associées au microétat e_ℓ sont $P_m = \delta_{m,\ell}$.

Maximum Le cas où nous avons le moins d'information sur l'occupation des M microétats est lorsque ceux-ci sont **équiprobables**, $P_m = 1/M \forall m$. Dans ce cas

$$S_{\max} = S(1/M, \dots, 1/M) = k_B \ln M \quad (2.46)$$

(pour une démonstration, cf. annexe 3.A 74).

Croissance L'entropie maximale croît avec le nombre de microétats M .

Évènements impossibles L'existence de microétats inaccessibles (i.e. d'évènements impossibles) ne change pas l'entropie :

$$S(P_1, \dots, P_M, 0, \dots, 0) = S(P_1, \dots, P_M) \quad (2.47)$$

Symétrie Les microétats jouent des rôles symétriques du point de vue de la mesure de l'information, i.e. $S(P_1, \dots, P_M)$ est une fonction symétrique sous l'échange de ses arguments.

³ En théorie de l'information, l'information est mesurée en bit. Le manque d'information maximum sur un bit étant $\ln 2$ on choisit plutôt $k_B \rightarrow 1/\ln 2$ dans ce contexte.

⁴ Un autre choix possible est l'entropie de Rényi $S_\alpha(\{P_m\}) \stackrel{\text{def}}{=} \frac{k_B}{1-\alpha} \ln(\sum_m P_m^\alpha)$, qui tend vers l'entropie de von Neumann dans la limite $\alpha \rightarrow 1$. Remarquons $S_\alpha^{\max} = k_B \ln M$, comme pour l'entropie de von Neumann, cependant elle ne vérifie pas toutes les propriétés énoncées plus bas.

Additivité Nous pourrions séparer arbitrairement l'ensemble des microétats $\mathcal{E} = \{e_1, \dots, e_M\}$ en deux sous ensembles $A = \{e_1, \dots, e_m\}$ et $B = \{e_{m+1}, \dots, e_M\}$. La probabilité de se trouver dans un des microétats de A est $q_A = \sum_{n \in A} P_n$. Les probabilités $\{P_1/q_A, \dots, P_m/q_A\}$ définissent donc un ensemble de probabilités normalisées associées aux microétats du sous ensemble. On vérifie la propriété d'additivité

$$S(P_1, \dots, P_M) = \overbrace{S(q_A, q_B)}^{\text{choix entre } A \text{ et } B} + \underbrace{q_A S(P_1/q_A, \dots, P_m/q_A)}_{\text{entropie du sous ensemble } A} + \underbrace{q_B S(P_{m+1}/q_B, \dots, P_M/q_B)}_{\text{entropie du sous ensemble } B} \quad (2.48)$$

L'équation (2.48), que nous pouvons écrire plus schématiquement $S(A \oplus B) = S(q_A, q_B) + q_A S(A) + q_B S(B)$, exprime comment additionner les entropies des deux sous ensembles A et B , le premier terme $S(q_A, q_B)$ correspondant à l'information manquante quant au choix entre A et B .

Sous-additivité Considérons des microétats spécifiés par la donnée de deux informations. Par exemple si nous disposons de deux sous systèmes, l'un dans un microétat parmi $A \in \{e_1, \dots, e_N\}$ et l'autre parmi $B \in \{e'_1, \dots, e'_M\}$. L'état du système complet est spécifié par la donnée de l'état de chaque sous système : $e_{n,m} = e_n \otimes e'_m$ (j'utilise la notation de la mécanique quantique puisque le point que nous discutons correspond précisément à la notion de produit tensoriel). Le microétat $e_{n,m}$ est réalisé avec probabilité $p_{n,m}$. Nous introduisons la probabilité

$$p_n^A = \sum_{m \in B} p_{n,m} \quad \text{resp.} \quad p_m^B = \sum_{n \in A} p_{n,m} \quad (2.49)$$

pour que le système A (resp. B) se trouve dans le microétat e_n (resp. e'_m), indépendamment de toute information sur l'état du système de B (resp. A). Remarquons qu'en général on aura $p_{n,m} \neq p_n^A p_m^B$, et c'est seulement lorsque les deux systèmes sont *décorrélés* que $p_{n,m} = p_n^A p_m^B$. Nous prouvons facilement l'égalité

$$S(\{p_{n,m}\}) = S(\{p_n^A\}) + S(\{p_m^B\}) - \underbrace{k_B \sum_{n,m} p_{n,m} \ln \left(\frac{p_{n,m}}{p_n^A p_m^B} \right)}_{\text{corrélations } \mathcal{C}(A,B) \geq 0}, \quad (2.50)$$

ce que nous écrivons plus schématiquement $S(A \otimes B) = S(A) + S(B) - \mathcal{C}(A, B)$. En remarquant que le terme de corrélation est *positif*,^{5,6} nous obtenons l'inégalité

$$S(\{p_{n,m}\}) \leq S(\{p_n^A\}) + S(\{p_m^B\}) \quad (2.52)$$

Si les systèmes sont indépendants, $p_{n,m} = p_n^A p_m^B$, alors les entropies s'ajoutent simplement et l'égalité est réalisée : il y a alors autant d'information à les considérer comme un tout, $e_n \otimes e'_m$, ou séparément, e_n et e'_m . En revanche si les systèmes sont corrélés, les considérer comme un tout contient une information supplémentaire (sur les corrélations entre A et B) par rapport à la situation où ils sont séparés : $A \otimes B$ contient plus d'information (son entropie est plus petite) que A et B .

⁵ Nous avons notée la corrélation $\mathcal{C}(A, B)$: ce terme porte une « information » (d'où le signe $-$), i.e. convention inverse de l'entropie (manque d'information).

⁶ Montrons que $\mathcal{C}(A, B) \geq 0$: nous utilisons l'inégalité de convexité $\ln(x) \leq x - 1$, donc $\ln(x) \geq 1 - 1/x$:

$$\mathcal{C}(A, B) \stackrel{\text{def}}{=} k_B \sum_{n,m} p_{n,m} \ln \left(\frac{p_{n,m}}{p_n^A p_m^B} \right) \geq k_B \sum_{n,m} p_{n,m} \left(1 - \frac{p_n^A p_m^B}{p_{n,m}} \right) = k_B \sum_{n,m} (p_{n,m} - p_n^A p_m^B) = 0. \quad (2.51)$$

Remarquons que lorsque $A \otimes B$ décrit un état pur, $S(A \otimes B) = 0$, les corrélations $\mathcal{C}(A, B)$, si elles existent, sont de nature purement quantique et on parle plutôt d'entropie « d'intrication ».

Illustration de la propriété de sous-additivité : corrélations entre deux spins.— Nous considérons deux particules A et B de spin $1/2$. Chaque particule peut être dans deux états de spin $\in \{|\uparrow\rangle, |\downarrow\rangle\}$. Nous considérons deux situations pour lesquelles la probabilité de trouver chaque spin dans l'un ou l'autre état est $p_\uparrow = p_\downarrow = 1/2$. L'entropie associée à cette distribution est $S(A) = S(B) = k_B \ln 2$

- Dans un premier temps nous considérons la situation où les états de spin accessibles sont non corrélés et étudions le « macroétat » $A \otimes B$ décrivant la situation où les **quatre** états de l'espace de Hilbert $|\uparrow\rangle_A \otimes |\uparrow\rangle_B$, $|\uparrow\rangle_A \otimes |\downarrow\rangle_B$, $|\downarrow\rangle_A \otimes |\uparrow\rangle_B$ et $|\downarrow\rangle_A \otimes |\downarrow\rangle_B$, sont occupés de manière équiprobable. Par commodité, les probabilités sont regroupées dans une matrice

$$p_{\sigma,\sigma'} : \frac{1}{4} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \quad (2.53)$$

Ce qui montre que $p_{\sigma,\sigma'}$ se factorise comme $p_{\sigma,\sigma'} = p_\sigma^A p_{\sigma'}^B$ avec $p_\sigma^{A,B} = 1/2$. La probabilité qu'un spin soit dans l'état $|\uparrow\rangle$ ou $|\downarrow\rangle$ est indépendante de l'état de l'autre spin. L'entropie des deux spins non corrélés est $S(A \otimes B) = S(A) + S(B) = k_B \ln 4$.

- Nous considérons maintenant la situation où les états des deux spins sont corrélés. Le « macroétat » $\widetilde{A \otimes B}$ décrit l'occupation équiprobable des **deux** états ⁷ $|\uparrow\rangle_A \otimes |\downarrow\rangle_B$ et $|\downarrow\rangle_A \otimes |\uparrow\rangle_B$. Les probabilités sont maintenant

$$\tilde{p}_{\sigma,\sigma'} : \frac{1}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad (2.54)$$

On vérifie que la probabilité d'occupation d'un état de spin est la même que dans le cas précédent : $\tilde{p}_\sigma^A = \sum_{\sigma'} \tilde{p}_{\sigma,\sigma'} = 1/2$ pour $\sigma = \uparrow$ et $\downarrow$. Les probabilités d'occupation des états à un spin sont inchangées par rapport au cas précédent. Cependant si l'on connaît l'état d'un spin, cela donne l'information sur l'état du second spin, ce qui n'était pas le cas avec le premier macroétat. Le macroétat $\{\tilde{p}_{\sigma,\sigma'}\}$ contient donc plus d'information que $\{p_{\sigma,\sigma'}\}$. Son entropie est inférieure : $S(\widetilde{A \otimes B}) = k_B \ln 2 \leq S(A \otimes B) = S(A) + S(B)$.

On trouvera une présentation d'autres propriétés de l'entropie dans les références données ci-dessous. En particulier les théorèmes de Shannon sur les problèmes de communication et de traitement de données.

☺ Les idées importantes

- Quelques notions de probabilité : distribution, comment calculer des moyennes, variance, écart-type.
- Si l'introduction de la notion de fonction génératrice est apparue trop technique, nous la rencontrerons à nouveau dans des contextes plus physiques.
- Garder en tête le théorème de la limite centrale nous aidera à comprendre bien des choses plus tard.
- **Important** : microétat/macroétat. La définition de l'entropie statistique.

🔍 Pour en savoir plus

- Sur l'entropie : Chapitre III du livre de R. Balian [3]. Article de E. T. Jaynes [14].

⁷ **Attention** à ne pas confondre la **superposition quantique** (cohérente) entre les deux états quantiques, $\frac{1}{\sqrt{2}}[|\uparrow\rangle_A \otimes |\downarrow\rangle_B \pm |\downarrow\rangle_A \otimes |\uparrow\rangle_B]$, et le **mélange statistique** $\{p_{\uparrow\downarrow} = \frac{1}{2}, p_{\downarrow\uparrow} = \frac{1}{2}\}$. Pour mieux comprendre la nuance entre les deux situations, on peut utiliser le formalisme de la *matrice densité* ([3], complément I.H de [11],...) qui permet de décrire les deux types d'états sur un pied d'égalité (cf. note de bas de page n°9 page 49).

Annexe 2.A : Fonction de grande déviation

2.A.1. Préliminaires : méthode du col et approximation de la phase stationnaire

a) Méthode du col

Il arrive très souvent que l'on ait à considérer des intégrales de la forme

$$I(\lambda) = \int_a^b dx e^{-\lambda \varphi(x)} \quad (2.55)$$

où $\lambda \in \mathbb{R}_+$ est un « grand » paramètre et $\varphi(x)$ une fonction réelle. Dans ce cas il est possible de trouver une bonne approximation de l'intégrale à l'aide la méthode du col, dont nous exposons l'idée lorsque la fonction $\varphi(x)$ possède un unique minimum (absolu) en $x_* \in]a, b[$, « loin » des bords (on suppose de plus que $\varphi''(x_*) > 0$, non nul). Dans la limite $\lambda \rightarrow \infty$ l'intégrale est dominée par les contributions du voisinage du minimum :

$$I(\lambda) = \int_a^b dx e^{-\lambda [\varphi(x_*) + \frac{1}{2} \varphi''(x_*) (x-x_*)^2 + \mathcal{O}((x-x_*)^3)]} \quad (2.56)$$

$$= \frac{e^{-\lambda \varphi(x_*)}}{\sqrt{\lambda \varphi''(x_*)}} \underbrace{\int_{-\sqrt{\lambda \varphi''(x_*)}(x_*-a)}^{+\sqrt{\lambda \varphi''(x_*)}(b-x_*)} du e^{-\frac{1}{2} u^2 + \mathcal{O}(u^3 \lambda^{-1/2})}}_{\xrightarrow{\lambda \rightarrow \infty} \int_{-\infty}^{+\infty} du e^{-\frac{1}{2} u^2} = \sqrt{2\pi}} \quad (2.57)$$

Finalement, on obtient :

$$I(\lambda) \underset{\lambda \rightarrow +\infty}{\simeq} \sqrt{\frac{2\pi}{\lambda \varphi''(x_*)}} e^{-\lambda \varphi(x_*)} \quad (2.58)$$

(les restrictions venant des hypothèses énoncées précédemment sont faciles à lever ; on pourra obtenir des généralisations de cette formule lorsque toutes ces hypothèses ne sont pas satisfaites, par exemple $a = x_*$, etc).

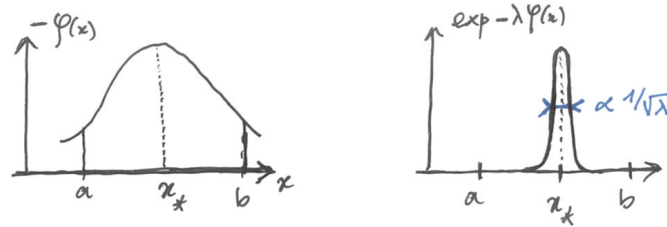


FIGURE 2.6 : Méthode du col.

b) Approximation de la phase stationnaire

La méthode du col peut-être étendue dans le plan complexe. Considérons l'intégrale

$$I(i\lambda) = \int_a^b dx e^{-i\lambda \varphi(x)} \quad (2.59)$$

où $\lambda \in \mathbb{R}$ est un « grand » paramètre et $\varphi(x)$ une fonction réelle. La phase de l'exponentielle oscille alors très rapidement et l'on s'attend à $\int dx e^{-i\lambda \varphi(x)} \sim 0$. Cependant l'intégrale est dominée par les extrema de $\varphi(x)$ qui correspondent à des points où la phase est « stationnaire ».

L'analogie de (2.58) est

$$I(i\lambda) \underset{\lambda \rightarrow \pm\infty}{\simeq} \sqrt{\frac{2\pi}{i\lambda\varphi''(x_*)}} e^{-i\lambda\varphi(x_*)} \quad (2.60)$$

lorsque la fonction $\varphi(x)$ possède un unique extremum (maximum ou minimum) dans l'intervalle.

2.A.2. Théorème de la limite centrale *versus* grandes déviations

Nous revenons sur l'analyse très importante du problème de la distribution de N variables aléatoires $\{X_1, \dots, X_N\}$ indépendantes et identiquement distribuées (i.i.d.) par une loi $p(x)$. Nous avons montré que la somme de ces variables,

$$S = \sum_{n=1}^N X_n, \quad (2.61)$$

est distribuée par une loi gaussienne centrée sur $\langle S \rangle = N \langle X \rangle$ et de variance $\text{Var}(S) = N \text{Var}(X)$, i.e. caractérisant des fluctuations d'ordre $\delta S \sim \sqrt{N}$. Ce résultat est très général : la forme de la distribution de S est **universelle**, i.e. est indépendante de la forme particulière de la loi $p(x)$, pourvu que ses deux premiers moments $\langle X \rangle$ et $\langle X^2 \rangle$ soient finis. Ce résultat ignore toutefois que la distribution de S présente des queues caractérisant des *événements rares*, sur des échelles $S \gtrsim N$, qui n'ont pas le caractère universel de la forme de la loi pour $\delta S \sim \sqrt{N}$. Nous allons discuter ce point que nous illustrerons sur un cas simple.

a) Fonction de grande déviation

Afin de caractériser commodément les propriétés statistiques des variables $\{X_n\}$, nous introduisons la fonction génératrice des cumulants (cf. chapitre 2)

$$\omega(ik) \stackrel{\text{def}}{=} \ln \left\langle e^{-ikX} \right\rangle = \sum_{n=1}^{\infty} \frac{c_n}{n!} (-ik)^n \quad (2.62)$$

où c_n est le cumulant d'ordre n (on rappelle que $c_1 = \langle X \rangle$ et $c_2 = \text{Var}(X)$).

Nous pouvons donc écrire la distribution de la somme sous la forme

$$P_N(S = Ny) = \underbrace{(p * \dots * p)}_{N \text{ fois}}(S) = \int_{\mathbb{R}} \frac{dk}{2\pi} e^{N[iky + \omega(ik)]}. \quad (2.63)$$

Dans la limite $N \rightarrow \infty$, nous nous attendons à ce que l'intégrale soit génériquement dominée par un unique point col k_* dans le plan complexe, qui correspond au point stationnaire de la fonction dans l'exponentielle $\frac{d}{dk} [iky + \omega(ik)] = 0$ ou

$$\omega'(ik) + y = 0 \quad \text{dont la solution est notée} \quad k = k_* \quad (2.64)$$

(k_* est fonction de y). La méthode du col (étendue dans le plan complexe) conduit à

$$\int_{\mathbb{R}} \frac{dk}{2\pi} e^{N[iky + \omega(ik)]} \underset{N \rightarrow \infty}{\simeq} \frac{1}{\sqrt{-2\pi N \omega''(k_*)}} e^{N[ik_* y + \omega(ik_*)]} \quad (2.65)$$

(Attention : si l'équation (2.64) admet plusieurs solutions, les déformations du contour d'intégration dans le plan complexe permettant de visiter ces solutions ne sont pas toutes licites ; nous en verrons un exemple plus bas). Finalement la distribution de S prend la forme d'échelle

$$P_N(S) \underset{N \rightarrow \infty}{\simeq} \exp \left\{ -N \Phi \left(\frac{S}{N} \right) \right\} \quad \text{où} \quad \Phi(y) = -ik_* y - \omega(ik_*) \quad (2.66)$$

est appelée la *fonction de grande déviation*. Remarquons que le théorème de la limite centrale doit être compris comme une limite particulière de ce résultat. On peut en effet le retrouver de la manière suivante : dans la limite $y \rightarrow 0$, la solution de (2.64) est obtenue en utilisant le développement de la fonction génératrice $\omega(ik) \simeq -\frac{1}{2}c_2k^2$ (nous supposons $c_1 = 0$ pour simplifier). La solution prend donc la forme $k_* \simeq iy/c_2$. D'après (2.66) on obtient donc

$$\Phi(y) \underset{y \rightarrow 0}{\simeq} \frac{y^2}{2c_2}. \quad (2.67)$$

La fonction de grande déviation présente donc un comportement universel, quadratique, à $y \rightarrow 0$ (théorème de la limite centrale),⁸ mais également un comportement non universel, i.e. dépendant de $p(x)$ via $\omega(ik)$, pour $|y| \gtrsim 1$, caractérisant les grandes déviations.

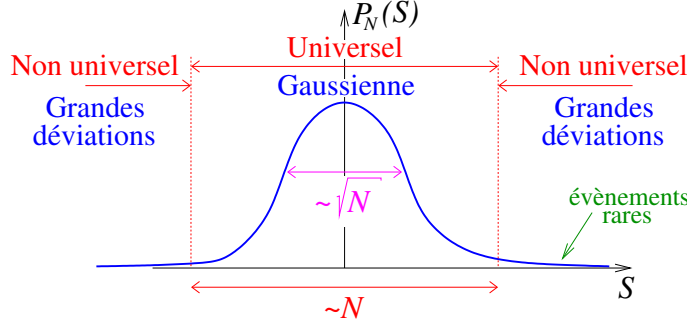


FIGURE 2.7 : Allure schématique de la distribution de S . Au centre la distribution prend la forme gaussienne universelle (indépendante de $p(x)$) alors que les grandes déviations caractérisant les événements rares dépendent des détails de la loi.

b) Application : loi exponentielle symétrique

À titre d'illustration nous considérons une loi exponentielle

$$p(x) = \frac{1}{2} e^{-|x|}. \quad (2.68)$$

La fonction génératrice (des moments) correspondante est $g(k) = \langle e^{-ikX} \rangle = \frac{1}{1+k^2}$ et la fonction génératrice des cumulants :

$$\omega(ik) = -\ln(1+k^2). \quad (2.69)$$

L'équation (2.64) déterminant le point dominant l'intégrale prend donc la forme

$$k^2 + \frac{2i}{y}k + 1 = 0 \quad \text{dont les solutions sont} \quad k_{\pm} = i \left(\pm \sqrt{\frac{1}{y^2} + 1} - \frac{1}{y} \right) \quad (2.70)$$

Pour savoir quelle solution de l'équation algébrique (2.64) correspond au point col, rappelons nous que l'intégrale contient un terme e^{Niky} . Pour visiter le point col k_* , le contour d'intégration de l'intégrale (2.63) doit être déformé dans le plan complexe de la variable k . Le terme exponentiel de l'intégrand impose de déformer le contour par le haut si $y > 0$ puisque $|e^{Niky}| = e^{-Ny \operatorname{Im}(k)}$ (cf. figure 2.8). Si $y > 0$ le point col dominant l'intégrale est $k_* = k_+$ (et $k_* = k_-$ si $y < 0$). Pour $y > 0$, nous obtenons $\Phi(y) = -ik_+y - \omega(ik_+)$, i.e.

$$\Phi(y) = \ln \left(2 \frac{\sqrt{y^2 + 1} - 1}{y^2} \right) + \sqrt{y^2 + 1} - 1, \quad (2.71)$$

⁸ On comprend mieux la terminologie « théorème de la limite centrale » puisque celui-ci caractérise le « centre » de la distribution $P_N(S)$.

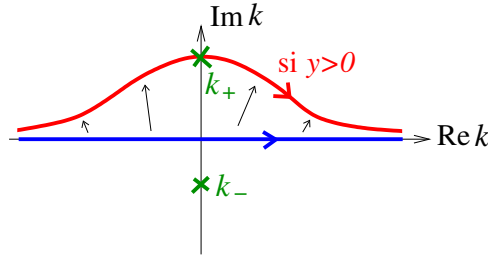


FIGURE 2.8 : Déformation du contour d'intégration de l'intégrale (2.63) dans le plan complexe permettant de visiter le point col en k_+ .

(finalement cette expression est aussi valable pour $y < 0$ puisque la loi est symétrique). La fonction est tracée sur la figure 2.9. Nous vérifions bien que $\Phi(y) \underset{y \rightarrow 0}{\simeq} \frac{y^2}{4}$, qui est l'expression du théorème de la limite centrale puisque $c_2 = 2$, i.e.

$$P_N(S) \underset{N \rightarrow \infty}{\simeq} \frac{1}{\sqrt{4\pi N}} e^{-\frac{1}{4N} S^2} \quad \text{pour} \quad |S| \ll N. \quad (2.72)$$

D'autre part le comportement $\Phi(y) \simeq |y| - \ln |y|$ de la fonction de grande déviation pour $|y| \gg 1$ montre que la distribution de la somme présente le comportement

$$P_N(S) \underset{N \rightarrow \infty}{\sim} |S|^N e^{-|S|} \quad \text{pour} \quad |S| \gg N. \quad (2.73)$$

Ce comportement, qui est donc assez différent de la décroissance gaussienne, est en effet typique de la convolution de N lois exponentielles.

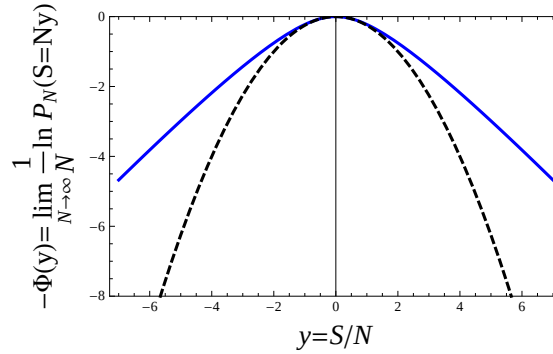


FIGURE 2.9 : Fonction de grandes déviations (2.71) ; comparaison avec la parabole (i.e. la distribution gaussienne).

En conclusion, retenons que si l'on s'intéresse à la statistique des événements rares (grandes déviations), ceux-ci ne sont pas caractérisés par le théorème de la limite centrale qui ne décrit que les déviations « normales » (sur l'exemple discuté, la forme approchée de la distribution donnée par le théorème de la limite centrale sous estime beaucoup le résultat correct).

Annexe 2.B : Exercices

✎ **Exercice 2.1 Séance de travaux pratiques :** Au cours d'une séance de travaux pratiques, des étudiants répètent N fois la mesure d'une grandeur physique X . Chaque étudiant obtient N données $X_1, \dots, X_N$ et donne une analyse statistique : il déduit une moyenne statistique $m = \frac{1}{N} \sum_{n=1}^N X_n$ et un écart-type $\Sigma = [\frac{1}{N} \sum_{n=1}^N (X_n - m)^2]^{1/2}$. Ces deux grandeurs sont elles-mêmes aléatoires et diffèrent d'un étudiant à un autre.

a) En supposant que $N \gg 1$, donner les formes approchées des distributions de m et Σ (exprimer la moyenne et la variance de ces deux lois en fonction de N et des moments de X_n).

b) Discuter la limite $N \rightarrow \infty$ de la distribution de m .

✎ **Exercice 2.2 : Entropie d'une distribution.**— Considérons N variables x_i distribuées dans un intervalle $I = [a, b]$. Soit $\rho(x)$ la distribution de ces variables :

$$\rho(x) = \frac{1}{N} \sum_{i=1}^N \delta(x - x_i), \quad (2.74)$$

i.e. $\rho(x)dx$ est la fraction des variables se trouvant dans l'intervalle $[x, x + dx]$.

La donnée de la distribution contient moins d'information sur la position des variables que la données des positions $\{x_i\}$. En particulier on perd une information sur l'ordonnancement des variables (si $x_1 < x_2$ ou $x_1 > x_2$ par exemple). On souhaite analyser l'entropie associée à cette perte d'information.

a) Supposons que l'intervalle I est découpé en M intervalles de largeurs égales $\epsilon = (b - a)/M$: $I = I_1 \cup I_2 \cup \dots \cup I_M$. On note n_α le nombre de variables dans l'intervalle I_α . Pour un choix particulier d'une séquence d'occupations $\{n_\alpha\}$, combien y-a-t'il de façons $\Omega[\{n_\alpha\}]$ de distribuer les N variables dans les M intervalles ? Trouver une forme approchée de l'entropie $S = \ln \Omega$ dans la limite où $n_\alpha \gg 1$.

b) Soit $X_\alpha = a + (\alpha - 1/2)\epsilon$, avec $\alpha \in \{1, \dots, M - 1\}$ la position de l'intervalle I_α . En écrivant que la densité est $\epsilon \rho(X_\alpha) \approx n_\alpha/N$, exprimer l'entropie de la distribution $S[\rho]$.

Réponse :

a) $\Omega = \frac{N!}{n_1! n_2! \dots n_M!}$, d'où $\ln \Omega \simeq N \ln N - \sum_\alpha n_\alpha \ln n_\alpha = - \sum_\alpha n_\alpha \ln(n_\alpha/N)$.

b) L'entropie associée à la distribution est donc $S[\rho] = \ln \Omega \approx - \sum_\alpha n_\alpha \ln(n_\alpha/N) = -N \sum_\alpha \epsilon \rho(X_\alpha) \ln [\epsilon \rho(X_\alpha)]$. Si $\rho(x)$ varie lentement à l'échelle de ϵ , on remplace la somme par une intégrale :

$$S[\rho] = -N \int dx \rho(x) \ln [\epsilon \rho(x)] \quad (2.75)$$

Chapitre 3

Postulat fondamental – Ensemble microcanonique

Plan du chapitre

I	Première approche dans un cadre classique	39
I.A	Espace des phases	39
I.B	Pourquoi l'espace des phases ? Théorème de Liouville	39
I.C	Expérience (numérique) des sphères dures – Du problème de Sinai au postulat fondamental	41
I.D	Ergodicité	43
I.E	Énoncé du postulat fondamental	44
II	Description quantique : espace de Hilbert	47
II.A	Rappel des postulats de la mécanique quantique	47
II.B	Microétats et macroétats, moyennes	48
II.C	Dénombrer les microétats : densité d'états	49
II.D	Densité d'états pour des particules identiques (donc indiscernables) . . .	53
III	Postulat fondamental	56
III.A	La distribution microcanonique	56
III.B	Entropie microcanonique	57
IV	Relâchement de contraintes	60
IV.A	Variables externes et variables internes	60
IV.B	Contact thermique entre deux systèmes – Température microcanonique	60
IV.C	Échange de volume – Pression microcanonique	64
IV.D	Échange de particules – Potentiel chimique microcanonique	66
IV.E	Cas général : distribution d'une variable interne quelconque	67
V	Irréversibilité	68
VI	Remarques sur le rôle de l'indiscernabilité	69
VI.A	Validité du traitement semi-classique de l'indiscernabilité – cas du gaz parfait monoatomique	69
VI.B	Entropie de mélange et paradoxe de Gibbs	70
Annexe 3.A : La méthode des multiplicateurs de Lagrange		74
Annexe 3.B : Rappel de mécanique quantique : postulat de symétrisation		75
Annexe 3.C : Exercices		77
Problème Isothermes et isentropes des gaz parfaits		77

Dans ce chapitre nous discutons la physique statistique des systèmes **isolés à l'équilibre** macroscopique. Nous introduisons le postulat fondamental, le socle de la physique statistique d'équilibre.

La situation étudiée : système isolé à l'équilibre macroscopique

Dans ce chapitre nous donnons une description des systèmes isolés, i.e. dont l'état macroscopique est caractérisé par un certain nombre de paramètres fixés comme : l'énergie du système E , le nombre de particules N , le volume V , et éventuellement d'autres quantités.

L'essentiel du cours sera consacré à la **physique statistique des systèmes à l'équilibre**. Considérons un exemple précis pour introduire la notion d'équilibre macroscopique.

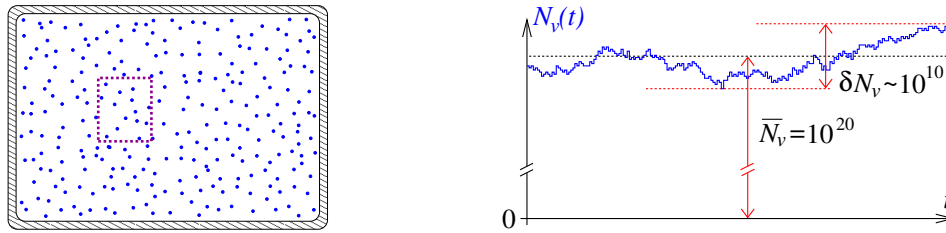


FIGURE 3.1 : **Équilibre macroscopique et fluctuations.**— On mesure le nombre de particules $N_v(t)$ contenues dans un petit volume d'un gaz. Pour un volume $v \sim 1 \text{ cm}^3$ de gaz dans les conditions de température et de pression normales, nous avons vu que $\overline{N}_v \sim 10^{20}$.

Nous isolons par la pensée un volume v dans un gaz (figure 3.1), ce qui constitue « le système ». Le nombre de particules dans le volume, que nous notons $N_v(t)$, est une quantité qui fluctue au cours du temps : des particules entrent et sortent du volume en permanence, traduisant l'**agitation thermique** dans le gaz. Parce que la densité est uniforme dans le volume, le nombre de particules vaut *en moyenne*

$$\overline{N}_v(t)^{(T,t_0)} = \frac{1}{T} \int_{t_0}^{t_0+T} dt N_v(t) \simeq n v \quad \text{pour } T \text{ "grand"}, \quad (3.1)$$

où $n = N/V$ est la densité moyenne dans l'enceinte (moyennée spatialement). On pourrait également chercher à caractériser d'autres propriétés du système : les fluctuations de $N_v(t)$, l'énergie des particules, etc. Si l'ensemble est isolé et ne subit aucune perturbation extérieure, on peut supposer que la moyenne $\overline{N}_v(t)^{(T,t_0)}$ est *indépendante de t_0* , à condition que T soit « assez grand ». Dans ces conditions toutes les grandeurs macroscopiques sont stationnaires. En outre il n'y a globalement aucun flux d'une quelconque quantité physique (flux de particules, flux d'énergie, flux d'entropie, etc) : le flux des particules entrant dans le volume v est exactement compensé, en moyenne, par les particules sortant du volume. Le petit volume est dans un état d'équilibre macroscopique, bien qu'à l'échelle microscopique les particules soient animées d'un mouvement incessant.

Retenons que ¹

l'équilibre macroscopique requiert

- (i) la *stationnarité* des quantités macroscopiques, et
- (ii) *l'absence de flux*.

¹Stationnaire=indépendant du temps. Équilibre=stationnaire et tous les flux nuls.

I Première approche dans un cadre classique

I.A Espace des phases

Considérons un système de N particules en mécanique classique. L'état microscopique est caractérisé par la donnée des $3N$ composantes des positions des particules, ² notées $\{q_i\}_{i=1,\dots,3N}$, et par les $3N$ composantes de ses impulsions, ³ $\{p_i\}_{i=1,\dots,3N}$. L'**espace des phases** est l'espace de dimension $6N$ des coordonnées $\{q_i, p_i\}_{i=1,\dots,3N}$. La donnée de la position dans l'espace des phases caractérise entièrement l'évolution du système, contrôlée par les *équations de Hamilton*

$$\dot{q}_i(t) = \frac{\partial H}{\partial p_i}(\{q_i(t), p_i(t)\}) \quad (3.2)$$

$$\dot{p}_i(t) = -\frac{\partial H}{\partial q_i}(\{q_i(t), p_i(t)\}) \quad (3.3)$$

où $H(\{q_i, p_i\})$ est la fonction de Hamilton (l'énergie sur une trajectoire physique). Deux trajectoires physiques sont représentées par deux lignes qui ne s'intersectent pas. Un microétat (classique) correspond donc à un point dans l'espace des phases.

microétat = un point de l'espace des phases

La notion de macroétat correspond à la donnée d'une distribution de probabilité d'occupation des microétats. Puisque ces derniers forment un espace continu, on a

macroétat = densité de probabilité $\rho(\{q_i, p_i\})$ dans l'espace des phases

où $\rho(\{q_i, p_i\}) \prod_i dq_i dp_i$ est la probabilité pour que l'état du système soit dans un volume $\prod_i dq_i dp_i$ autour de $(\{q_i, p_i\})$.

Q : Comment trouver $\rho(\{q_i, p_i\})$?

I.B Pourquoi l'espace des phases ? Théorème de Liouville

Pourquoi l'espace des phases, i.e. des couples de variables canoniquement conjuguées, joue-t-il un rôle si important ? Pourquoi ne pas considérer d'autres fonctions des q_i et p_i ? Ce point est éclairci par l'énoncé du théorème de Liouville, que nous discutons maintenant.

Considérons l'évolution pendant un intervalle de temps « petit » δt :

$$q_i(t + \delta t) \simeq q_i(t) + \dot{q}_i(t) \delta t \quad (3.4)$$

$$p_i(t + \delta t) \simeq p_i(t) + \dot{p}_i(t) \delta t \quad (3.5)$$

D'après les équations de Hamilton, nous pouvons interpréter ces équations comme une transformation en général non linéaire des coordonnées :

$$\begin{pmatrix} \vdots \\ q_i \\ p_i \\ \vdots \end{pmatrix} \equiv \begin{pmatrix} \vdots \\ q_i(t) \\ p_i(t) \\ \vdots \end{pmatrix} \longrightarrow \begin{pmatrix} \vdots \\ \tilde{q}_i \\ \tilde{p}_i \\ \vdots \end{pmatrix} \equiv \begin{pmatrix} \vdots \\ q_i(t + \delta t) \\ p_i(t + \delta t) \\ \vdots \end{pmatrix} = \begin{pmatrix} \vdots \\ q_i + \frac{\partial H}{\partial p_i}(\{q_i, p_i\}) \delta t \\ p_i - \frac{\partial H}{\partial q_i}(\{q_i, p_i\}) \delta t \\ \vdots \end{pmatrix} \quad (3.6)$$

² Il n'est en fait pas nécessaire que les $3N$ degrés de liberté soient des degrés de liberté de translation, comme c'est implicite ici. On reviendra plus tard sur ce point.

³ l'impulsion est le moment canoniquement conjugué, i.e. obtenu à partir du lagrangien $p_i \stackrel{\text{def}}{=} \frac{\partial L}{\partial \dot{q}_i}$.

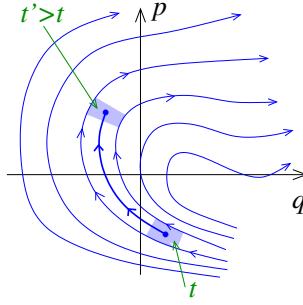


FIGURE 3.2 : **Théorème de Liouville.** On a représenté des trajectoires physiques dans l'espace des phases. L'évolution temporelle conserve le volume dans l'espace des phases, i.e. les deux volumes hachurés, représentant un ensemble de microétats à deux instants de leur évolution, sont égaux.

Le Jacobien de cette transformation est donné par

$$J = \det(1 + K \delta t) \quad \text{où} \quad K \stackrel{\text{def}}{=} \begin{pmatrix} \vdots & \vdots & \vdots & \vdots \\ \dots & \frac{\partial^2 H}{\partial q_i \partial p_i} & \frac{\partial^2 H}{\partial p_i^2} & \dots \\ \dots & -\frac{\partial^2 H}{\partial q_i^2} & -\frac{\partial^2 H}{\partial q_i \partial p_i} & \dots \\ \vdots & \vdots & \vdots & \vdots \end{pmatrix} \quad (3.7)$$

En remarquant que ⁴ $\det(1 + K \delta t) \simeq 1 + \text{Tr}(K) \delta t$, nous obtenons finalement

$$J \simeq 1 + \sum_i \left(\frac{\partial^2 H}{\partial q_i \partial p_i} - \frac{\partial^2 H}{\partial q_i \partial p_i} \right) \delta t = 1. \quad (3.8)$$

Autrement dit (théorème de Liouville ⁵)

$$\boxed{\prod_i d\tilde{q}_i d\tilde{p}_i = \prod_i dq_i dp_i} \quad (3.9)$$

i.e. si l'on considère un domaine dans l'espace des phases, il évoluera au cours du temps de telle sorte que son *volume soit conservé* (figure 3.2).

Remarquons qu'il est important d'avoir identifié la mesure conservée lors de l'évolution temporelle, ainsi toute la dépendance temporelle peut-être reportée dans la densité de probabilité $\rho(\dots; t)$.

⁴ On peut prouver cette relation à partir de la définition du déterminant en identifiant les termes d'ordre 1 en δt . On peut également écrire : $\det(1 + K \delta t) = \exp \ln \det(1 + K \delta t) = \exp \text{Tr} \ln(1 + K \delta t) \simeq \exp \text{Tr}(K \delta t) \simeq 1 + \text{Tr}(K \delta t)$.

⁵ *Stricto sensu* la formulation du théorème de Liouville est un peu différente, même si elle porte le même sens : on introduit la densité de probabilité $\rho(\{q_i, p_i\}; t)$ dans l'espace des phases. Localement, sa variation au cours du temps est donnée par $\partial \rho / \partial t$. Considérons un volume $\prod_i dq_i dp_i$ autour d'une trajectoire physique $(\{q_i(t), p_i(t)\})$ solution des équations du mouvement. Ce volume se meut au cours du temps. La variation de la densité, dans le volume suivant le flot défini par les équations du mouvement, est donnée par $\frac{d\rho}{dt} = \sum_i (\dot{q}_i \frac{\partial \rho}{\partial q_i} + \dot{p}_i \frac{\partial \rho}{\partial p_i}) + \frac{\partial \rho}{\partial t}$, ce que l'on peut écrire comme $\frac{d\rho}{dt} = \{\rho, H\} + \frac{\partial \rho}{\partial t}$, où $\{A, B\} \stackrel{\text{def}}{=} \sum_i (\frac{\partial A}{\partial q_i} \frac{\partial B}{\partial p_i} - \frac{\partial A}{\partial p_i} \frac{\partial B}{\partial q_i})$ est un crochet de Poisson (cf. cours de mécanique analytique). Le **théorème de Liouville** est $\frac{d\rho}{dt} = 0$. Si on note $\vec{V} = (\dots, \dot{q}_i, \dot{p}_i, \dots)$ le champ de "vitesse" dans l'espace des phases, on remarque l'analogie avec la mécanique des fluides, où $(\rho, \rho \vec{V})$ définit la densité de probabilité et son courant associé. Le théorème de Liouville prend la forme de l'équation de conservation $\frac{\partial \rho}{\partial t} + \text{div}(\rho \vec{V}) = 0$, où la divergence implique les dérivées partielles par rapport à l'ensemble des coordonnées de l'espace des phases.

Exemple : mouvement sur une sphère.— Nous considérons une particule de masse m se déplaçant sur la surface d'une sphère de rayon R . Sa position est repérée par un couple d'angles (θ, φ) , où θ est l'angle entre l'axe vertical et le vecteur position et φ l'angle azimutal (le problème est équivalent à l'analyse du mouvement de rotation d'un segment, une molécule diatomique rigide, dans l'espace tridimensionnel). Si le mouvement est libre, le Lagrangien est donné par l'énergie cinétique de rotation $L = \frac{1}{2}(\dot{\theta}^2 + \sin^2 \theta \dot{\varphi}^2)$, où $I = mR^2$ est un moment d'inertie. Nous obtenons les deux moments canoniquement conjugués $p_\theta = I\dot{\theta}$ et $p_\varphi = I \sin^2 \theta \dot{\varphi}$ en terme desquels nous exprimons l'hamiltonien

$$H = \frac{1}{2I} \left(p_\theta^2 + \frac{p_\varphi^2}{\sin^2 \theta} \right)$$

La mesure d'intégration dans l'espace des phases conservée par l'évolution temporelle est donc $d\theta dp_\theta d\varphi dp_\varphi$.

I.C Expérience (numérique) des sphères dures – Du problème de Sinai au postulat fondamental

En introduisant la notion de « macroétat », nous avons suggéré qu'il sera plus efficace d'étudier les propriétés d'un système à l'aide de l'information probabiliste contenue dans la densité de probabilité $\rho(\{q_i, p_i\})$. Il faut maintenant nous donner les moyens de déterminer cette distribution. Dans le cours, nous proposerons diverses manières de « construire » différentes distributions, ce qui correspondra aux différents « ensembles » de la physique statistique. Dans ce chapitre nous nous sommes placés dans la situation la plus épurée, cas d'un système isolé, afin d'introduire un postulat fondamental. Dans quelle direction chercher pour faire émerger le postulat fondamental? Pour cela nous nous plaçons dans une situation particulière dont l'analyse va nous permettre, par *induction*, d'énoncer un principe qui transcendera cette expérience spécifique. Nous nous plaçons ainsi dans une position analogue à celle de Galilée étudiant le mouvement de divers objets au fond de la cale d'un bateau, observations qui le conduiront à proposer le principe d'inertie de la mécanique classique. Nous choisissons ici une expérience numérique donnant une description simplifiée d'un gaz de particules (considérer le cas d'un gaz est naturel : nous verrons que celui-ci est le système modèle sur lequel nous appliquerons systématiquement les concepts de la physique statistique).

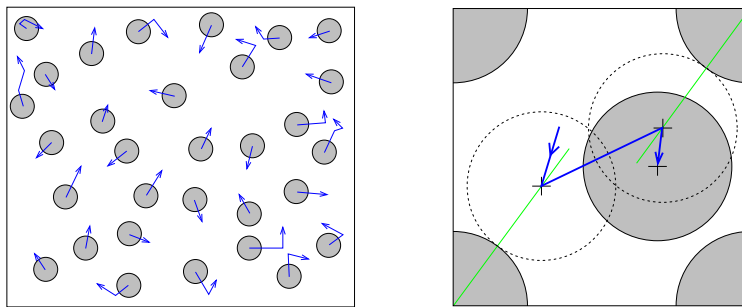


FIGURE 3.3 : **Le gaz de sphères dures et le problème de Sinai.** À gauche : nous pouvons grossièrement modéliser un gaz atomique comme un gaz de boules impénétrables (des « sphères dures ») de rayon fini, s'entrechoquant et frappant les parois. On a indiqué quelques mouvements ultérieurs des sphères dures. À droite : pour simplifier encore l'étude de la dynamique des sphères dures, nous isolons par la pensée une sphère, placée dans un environnement de quatre autres sphères supposées immobiles aux coins d'un carré.

Une description simplifiée d'un gaz atomique est de modéliser les atomes comme des petites

boules impénétrables (des « sphères dures ») de rayon fini R . On prend ainsi en compte la partie fortement répulsive de l'interaction entre atomes. Les sphères s'entrechoquent en permanence et frappent les parois du volume qui les contient. Un des intérêts de ce modèle simplifié est de n'introduire aucune échelle d'énergie (puisque le potentiel « inter-atomique » est soit nul soit infini) : seul importe l'analyse des possibles configurations spatiales, ce qui offre un avantage certain pour visualiser la dynamique du système en profitant du découplage position-impulsion (figure 3.3). Ce problème reste néanmoins extrêmement riche et difficile. Aussi nous allons encore le simplifier en nous plaçant tout d'abord dans la *situation bidimensionnelle*, mais surtout en nous focalisant sur une unique sphère, dont nous supposons l'environnement gelé. Afin de « forcer le destin » de cette sphère et induire de fréquentes collisions, nous la plaçons dans un carré de côté égal à $2R$, aux quatre coins duquel sont placées quatre sphères immobiles (partie droite de la figure 3.3). Nous espérons ainsi rendre compte de l'histoire d'une sphère dure au sein du gaz et de l'effet des collisions avec les autres particules. Bien que nous ayons drastiquement simplifié le problème du gaz de sphères dures, nous allons pouvoir maintenant tirer des conclusions simples à l'aide d'un petit programme simulant le mouvement de la sphère.

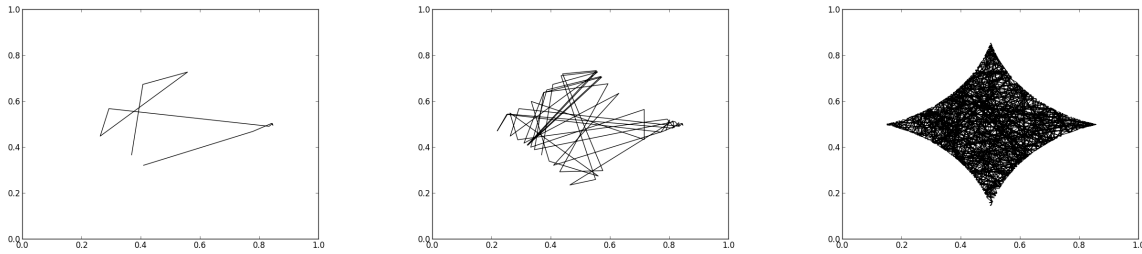


FIGURE 3.4 : **Ergodicité dans le problème de Sinai.** *Trajectoire du centre de la sphère dure après 10, 50 et 500 chocs.* (Remerciements à Alberto Rosso pour le programme python).

Nous lançons la sphère dans une direction quelconque et suivons le mouvement de son centre. La première partie de la figure 3.4 montre la trajectoire obtenue après 10 chocs, la figure du centre après 50 chocs, et la figure de droite après 500 chocs. Nous constatons que la trajectoire semble avoir exploré tout l'espace que les contraintes lui permettent de visiter. Autrement dit, *la distribution de probabilité de trouver la sphère en un endroit est non nulle partout où les contraintes stériques l'autorisent*. Si nous introduisons $\rho(\vec{r})$ la densité de probabilité pour trouver la sphère en $\vec{r}$, le résultat de la simulation suggère que cette fonction est non nulle partout, cependant nous ne connaissons pas encore la valeur de la distribution elle-même.

Un autre petit programme nous fournit toutefois cette information : tout au long de l'évolution, ce nouveau programme mesure la fraction du temps $d\tau(x, y)$ passé dans un petit élément de surface $dx dy$. Le résultat est représenté sur la figure 3.5 et montre que la particule passe autant de temps partout où elle a le droit de se trouver (aux fluctuations numériques près, qui diminueraient si la trajectoire de la sphère était analysée sur un temps T plus long). On peut donc écrire $\frac{d\tau(x, y)}{T} = \frac{dx dy}{S_{\text{autor}}}$, où S_{autor} est la surface totale visitée, où encore que **la densité de probabilité est un uniforme** : $\rho(\vec{r}) = 1/S_{\text{autor}}$ si $\vec{r} \in \text{domaine permis}$.

En conclusion, pour ce problème ultra-simplifié d'une sphère dans un espace bidimensionnel, en interaction avec quatre sphères immobiles, **l'espace des phases accessible est visité uniformément**. Nous allons donc *postuler* que cette assertion reste vraie pour des systèmes plus complexes (gaz de sphères dures ou gaz plus réalistes), avec un nombre macroscopique de degrés de liberté. La justification est que nous pouvons espérer que la complexité et les interactions vont favoriser une exploration de l'espace des phases accessible.

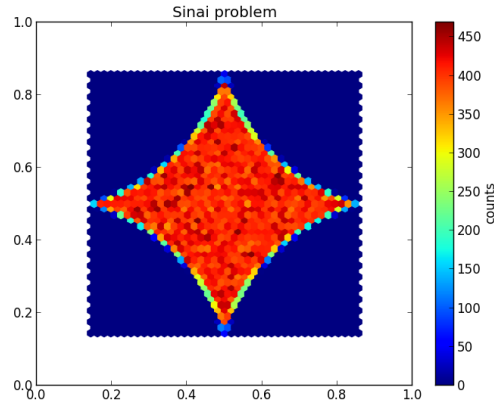


FIGURE 3.5 : **Ergodicité dans le problème de Sinai.** La densité de probabilité de présence est uniforme dans la région de l'espace des phases que la sphère dure est autorisée à explorer. (Remerciements à Alberto Rosso).

📖 **Pour en savoir plus sur le gaz de sphères dures :** l'ouvrage (spécialisé) de Werner Krauth [18].

I.D Ergodicité

Nous pouvons maintenant formaliser les conclusions de l'étude de l'expérience numérique.

Moyenne temporelle (sur une réalisation)

Considérons une grandeur physique, fonction des coordonnées dans l'espace des phases $A(\{q_i, p_i\})$. Cette fonction pourrait être l'énergie, ou le nombre de particules $N_v(t)$ dans un sous volume du gaz (figure 3.1). Cette grandeur physique est mesurée pendant d'une expérience : le résultat de cette mesure est *a priori* une certaine fonction du temps (qui dépend des conditions initiales). Typiquement l'observable est mesurée sur un temps expérimental T très grand devant les échelles caractérisant la dynamique microscopique et, plutôt qu'une valeur instantanée, on mesure en général une moyenne temporelle :

$$\overline{A}^{(T)} = \frac{1}{T} \int_0^T dt A(\{q_i(t), p_i(t)\}) , \quad (3.10)$$

où $\{q_i(t), p_i(t)\}$ sont les solutions des équations du mouvement.

Moyenne probabiliste (moyenne d'ensemble)

Le langage probabiliste que nous avons introduit plus haut nous permet de définir un autre type de moyenne. Un macroétat est défini par la donnée d'une distribution $\rho(\{q_i, p_i\})$ dans l'espace des phases caractérisant les probabilités d'occupation des microétats. La moyenne (probabiliste) de la grandeur physique A , lorsque le système est dans le macroétat, est donnée par

$$\langle A \rangle \stackrel{\text{def}}{=} \int \prod_i dq_i dp_i \rho(\{q_i, p_i\}) A(\{q_i, p_i\}) . \quad (3.11)$$

Hypothèse ergodique

D'un point de vue théorique, les calculs des deux moyennes (3.10) et (3.11) requièrent des informations très différentes : la première demande l'ensemble de l'information microscopique, i.e. la connaissance de la solution des équations du mouvement (3.2,3.3), une masse d'information en pratique phénoménale. La seconde moyenne requière une information probabiliste sur l'occupation des microétats, dont nous montrerons qu'elle découlera de règles simples.

Si le temps d'observation T est très « long », permettant au système d'explorer une grande partie de son espace des phases, nous pouvons supposer que la moyenne (3.10) tend vers une limite lorsque $T \rightarrow \infty$ (cf. figure 3.1). Bien que cela soit difficile à prouver rigoureusement, la raison sous jacente est très proche de la discussion autour du théorème de la limite centrale et de la distinction entre moyenne statistique et moyenne probabiliste (cf. justification ci-dessous).

Cette hypothèse nous permet de **remplacer la moyenne temporelle** par une **moyenne d'ensemble** (moyenne probabiliste) :

$$\lim_{T \rightarrow \infty} \overline{A}^{(T)} \stackrel{\text{hyp. ergod.}}{=} \langle A \rangle \quad (3.12)$$

Cette hypothèse centrale de la physique statistique, l'**hypothèse ergodique**, correspond à *remplacer la moyenne temporelle sur un unique système, par une moyenne sur un ensemble (virtuel) de systèmes* ayant tous les mêmes propriétés à l'échelle macroscopique.

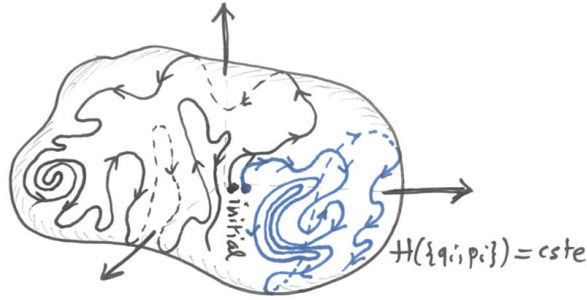


FIGURE 3.6 : **Ergodicité.** Deux trajectoires initialement très proches évoluent de manière très différentes sur la couche d'iso-énergie (attention, l'espace des phases a en principe une dimension paire !). Il a été prouvé rigoureusement que la distance entre les deux trajectoires croît comme $\exp \gamma t$ (où γ est un exposant de Lyapunov).

I.E Énoncé du postulat fondamental

Reformulons l'idée maîtresse que nous venons d'extraire de notre expérience numérique dans un cadre plus général. La question est de discuter la distribution $\rho(\{q_i, p_i\})$ dans l'espace des phases. Une première indication est donnée par le théorème de Liouville, tel qu'il est formulé dans la note de bas de page n°5 page 40, caractérisant l'évolution temporelle de la densité de probabilité dans l'espace des phases $\frac{\partial \rho}{\partial t} = -\{\rho, H\}$. Nous cherchons une mesure stationnaire, i.e. le crochet de Poisson est nul $\{\rho, H\} = \sum_i \left(\frac{\partial \rho}{\partial q_i} \frac{\partial H}{\partial p_i} - \frac{\partial \rho}{\partial p_i} \frac{\partial H}{\partial q_i} \right) = 0$, ce qui nous permet de déduire que la mesure dans l'espace des phases est une fonction de l'hamiltonien :

$$\text{densité stationnaire} \Rightarrow \rho(\{q_i, p_i\}) = \text{fct} \left(H(\{q_i, p_i\}) \right). \quad (3.13)$$

Nous déterminons maintenant la fonction, à l'aide d'arguments physiques. Si le système est isolé, en l'absence d'autre constante du mouvement, la seule contrainte est celle de la conservation de

l'énergie et la trajectoire dans l'espace des phases de dimension $6N$ évolue sur une « surface » (de dimension $6N - 1$) d'iso-énergie, i.e. le sous espace de l'espace des phases tel que

$$E = H(\{q_i, p_i\}) . \quad (3.14)$$

On « espère » qu'après un temps T « assez long », la trajectoire ait exploré *uniformément* la surface d'iso-énergie (figure 3.6). Cela revient à postuler que la mesure est **uniforme dans l'espace des phases accessible** :

$$\rho^*(\{q_i, p_i\}) = \frac{1}{\Sigma(E)} \delta[E - H(\{q_i, p_i\})] \quad \text{où} \quad \Sigma(E) = \int \prod_i dq_i dp_i \delta[E - H(\{q_i, p_i\})] . \quad (3.15)$$

La présence de $\Sigma(E)$, « l'aire » de la « surface » iso-énergie, assure la normalisation. $\rho^*(\{q_i, p_i\})$ est appelée la **distribution microcanonique**.

Le théorème de Liouville nous montre que la mesure uniforme est en effet stable sous l'évolution temporelle : si $\prod_i dq_i dp_i$ est invariante lors de l'évolution, alors $\rho^*(\{q_i, p_i\}) = \text{cste}$ est une solution stationnaire. Le point subtil est qu'il n'est pas évident de démontrer que toute mesure converge vers la mesure invariante.

Remarquons qu'il sera souvent plus agréable (et plus physique ?) de considérer que l'énergie est fixée à une incertitude expérimentale près δE et d'écrire :

$$\rho^*(\{q_i, p_i\}) = \begin{cases} \frac{1}{\mathcal{V}(E, \delta E)} & \text{si } E \leq H(\{q_i, p_i\}) \leq E + \delta E \\ 0 & \text{sinon} \end{cases} \quad \text{où } \mathcal{V}(E, \delta E) = \Sigma(E) \delta E \quad (3.16)$$

est le volume de l'espace des phases délimité par les deux contraintes $E \leq H(\{q_i, p_i\}) \leq E + \delta E$. Cette écriture rendra plus aisée la discussion de la correspondance quantique → classique. L'équation (3.16) est l'écriture du postulat fondamental de la physique statistique, dans le cadre classique.

Exemple : oscillateur harmonique.— Petit exercice académique mais instructif : nous étudions la distribution microcanonique pour un oscillateur harmonique unidimensionnel $H(q, p) = \frac{1}{2}(p^2 + q^2)$ (pour simplifier nous posons $m = \omega = 1$). Si nous notons E l'énergie de l'oscillateur, la « surface » iso-énergie est le cercle $2E = p^2 + q^2$. La distribution microcanonique correspond à l'occupation équiprobable de tous les points du cercle. Calculons la normalisation de la distribution :

$$\Sigma(E) = \int dq dp \underbrace{\delta\left(E - \frac{1}{2}(p^2 + q^2)\right)}_{\frac{1}{|p|} [\delta(p - \sqrt{2E - q^2}) + \delta(p + \sqrt{2E - q^2})]} = \int_{-\sqrt{2E}}^{+\sqrt{2E}} \frac{2dq}{\sqrt{2E - q^2}} = 2 \int_0^1 \frac{dt}{\sqrt{t}\sqrt{1-t}} \quad (3.17)$$

$$= 2 B(1/2, 1/2) = 2\pi \quad (3.18)$$

où nous avons utilisé (A.16). La dernière intégrale est reliée à la fonction Beta d'Euler (A.4). Finalement :

$$\rho^*(q, p) = \frac{1}{\pi} \delta(2E - p^2 - q^2) . \quad (3.19)$$

Nous déduisons la probabilité de présence au point q :

$$w(q) = \int dp \rho^*(q, p) = \frac{1}{\pi \sqrt{2E - q^2}} \quad \text{pour } |q| \leq \sqrt{2E} \quad (3.20)$$

(le calcul utilise à nouveau (A.16)). La distribution dans l'espace physique nous permet de calculer des moyennes de fonctions de la position. Par exemple :

$$\frac{1}{2} \langle q^2 \rangle = \int dq w(q) \frac{1}{2} q^2 = \frac{1}{2} E \quad (3.21)$$

(l'intégrale est à nouveau donnée par (A.4)). Finalement nous reconnaissons un résultat bien connu : la moyenne *temporelle* de l'énergie potentielle de l'oscillateur harmonique, sur une période d'oscillation, est la moitié de l'énergie mécanique $\overline{E_{\text{pot}}(t)} = \frac{1}{2}q(t)^2 = \frac{1}{2}E$. Sur cet exemple trivial, nous avons vérifié l'hypothèse ergodique d'équivalence entre moyenne d'ensemble et moyenne temporelle.

Remarques

Le postulat fondamental déduit d'un principe d'entropie maximale.— En supposant la distribution d'équilibre uniforme sur la surface d'iso-énergie, i.e. que tous les microétats permis par la contrainte (énergie fixée) sont équiprobables, nous avons choisi la distribution ayant l'*entropie maximale*. La démarche est analogue avec l'analyse du problème de pile ou face : on ne sait avancer aucune prédiction sur l'état final de la pièce tournoyant en l'air, on postule donc que le « macroétat » est celui ayant l'entropie la plus grande (i.e. associé avec le minimum d'information), la distribution uniforme $P_{\text{pile}} = P_{\text{face}} = 1/2$. De même pour le système macroscopique isolé : en l'absence de toute contrainte autre que la conservation de l'énergie, nous postulons que le macroétat effectivement réalisé est celui ayant la plus grande entropie, i.e. que tel tous les microétats d'énergie E soient équiprobables.

La terminologie : les ensembles de la physique statistique.— Nous avons expliqué l'origine de la dénomination « ensemble microcanonique » : cette terminologie fait référence à un ensemble virtuel d'une infinité de systèmes équivalents sur lequel la moyenne probabiliste est réalisée ; les systèmes virtuels ont tous les mêmes propriétés à l'échelle macroscopique, compatibles avec les contraintes imposées (système isolé, donc $E, V, N, \dots$ fixés). À la fin du XIX^e, Gibbs a compris comment introduire d'autres « ensembles » correspondant à d'autres manières de construire des macroétats en fixant des contraintes différentes (par exemple, au chapitre suivant nous remplacerons la contrainte de conservation de l'énergie par une contrainte sur la température).

Hypothèse ergodique et théorème de la limite centrale.— Essayons de comprendre comment nous pouvons justifier le passage de la moyenne temporelle (3.10) à la moyenne d'ensemble (terminologie de la physique statistique pour désigner la moyenne probabiliste). Pour cela nous discutons l'existence de quelques temps caractéristiques contrôlant la dynamique du système macroscopique (par exemple un gaz).

- Tout d'abord nous pouvons introduire un temps microscopique τ_{micro} ; par exemple le temps de collision entre particules, $\tau_{\text{coll}} \sim 1$ ns dans un gaz dans les conditions normales de température et de pression, le temps typique entre deux collisions successives subies par un atome (i.e. pendant τ_{coll} , les $N \sim 10^{23}$ atomes subissent chacun typiquement une collision).
- Nous pouvons introduire un temps de corrélation τ_{corr}^A caractérisant la décroissance de la fonction de corrélation de l'observable en question, i.e. la largeur temporelle de la fonction $C_A(t - t') = \langle A(t)A(t') \rangle - \langle A(t) \rangle \langle A(t') \rangle$.
- Enfin nous introduisons un temps ergodique τ_{erg} (temps de Poincaré), qui est le temps nécessaire à la trajectoire pour explorer l'espace des phases accessible. À cause de la très grande dimension de celui-ci ($6N - 1$), ce temps est faramineux !

Typiquement on attend la hiérarchie $\tau_{\text{micro}} \ll \tau_{\text{corr}}^A \ll \tau_{\text{erg}}$.

Nous admettons que le temps « expérimental » T dans (3.10) est très grand devant le temps τ_{corr}^A et découpons l'intégrale en $\mathcal{N}_T$ morceaux associés à des intervalles de temps $\Delta T = T/\mathcal{N}_T \gtrsim \tau_{\text{corr}}^A$:

$$\overline{A}^{(T)} = \frac{1}{\mathcal{N}_T} \left[\int_0^{\Delta T} \frac{dt}{\Delta T} A(t) + \int_{\Delta T}^{2\Delta T} \frac{dt}{\Delta T} A(t) + \cdots + \int_{(\mathcal{N}_T-1)\Delta T}^{\mathcal{N}_T\Delta T} \frac{dt}{\Delta T} A(t) \right]. \quad (3.22)$$

Puisque $\Delta T \gtrsim \tau_{\text{corr}}^A$, les morceaux d'intégrales sont statistiquement indépendants. Le théorème de la limite centrale nous montre que $\overline{A}^{(T)}$ est une quantité fluctuante dont les fluctuations sont d'ordre $1/\sqrt{\mathcal{N}_T}$, i.e.

$$\overline{A}^{(T)} = \langle A \rangle + \underbrace{\mathcal{O}(1/\sqrt{T})}_{\text{terme aléatoire}} \quad (3.23)$$

Notons que notre argument n'a réglé qu'une partie du problème puisque nous n'avons pas prouvé que les morceaux d'intégrales sont des échantillons représentatifs d'une moyenne d'ensemble, i.e. que leur distribution correspond à la distribution uniforme sur la couche d'iso-énergie.

La recherche de résultats rigoureux sur la question de l'ergodicité fait l'objet d'un domaine appelé « théorie ergodique ». Toutefois assez peu de résultats exacts ont été prouvés (on peut toutefois citer la preuve de l'ergodicité pour le gaz classique de sphères dures par Ya. Sinai en 1963).

Pour terminer, remarquons que, si la question de l'ergodicité repose *en principe* sur la complexité de la dynamique du système macroscopique, les faibles perturbations extérieures (isoler un système n'est qu'une vue de l'esprit) peuvent jouer *en pratique* un rôle important.

II Description quantique : espace de Hilbert

II.A Rappel des postulats de la mécanique quantique

La mécanique quantique est le cadre définissant les lois de la physique aux échelles élémentaires. Rappelons les quelques règles qui régissent la dynamique quantique (les postulats) ; je renvoie à [5, 8, 24, 34] pour plus de détails.

- *États quantiques.* – L'état quantique d'un système est décrit par un vecteur d'état $|\psi\rangle$, élément d'un espace vectoriel des états, l'espace de Hilbert $\mathcal{H}$, construit sur le corps des nombres complexes. L'espace de Hilbert est muni d'un produit scalaire satisfaisant $\langle\psi|\phi\rangle = \langle\psi|\phi\rangle^*$.
- *Interprétation probabiliste.* – Le produit scalaire s'interprète comme l'amplitude de probabilité de transition (instantanée) entre deux états $\mathcal{A}(\psi \leftarrow \phi) = \langle\psi|\phi\rangle$, i.e. la probabilité de transition ⁶ est $\text{Proba}(\psi \leftarrow \phi) = |\langle\psi|\phi\rangle|^2$.
- *Observables et postulats de mesure.* – Les observables (grandeurs physiques) sont représentées par des opérateurs linéaires agissant dans l'espace de Hilbert. Une observable A est caractérisée par son spectre de valeurs propres et vecteurs propres $\{a_n, |\varphi_n\rangle\}$. Si le système est dans un état $|\psi\rangle$, le résultat d'une mesure de A est **une** des valeurs propres a_n avec probabilité $\text{Proba}(\varphi_n \leftarrow \psi) = |\langle\varphi_n|\psi\rangle|^2$. L'état du système après mesure est $|\varphi_n\rangle$ (*réduction du paquet d'ondes*). Le processus de mesure s'accompagne une évolution *stochastique et irréversible* de l'état quantique.

⁶ Autrement dit : $\text{Proba}(\psi \leftarrow \phi) = |\langle\psi|\phi\rangle|^2$ est la probabilité pour observer le système dans l'état $|\psi\rangle$, sachant qu'il se trouve dans l'état $|\phi\rangle$.

- *Évolution temporelle.* – L'évolution temporelle du vecteur d'état est contrôlée par l'équation de Schrödinger $i\hbar \frac{d}{dt} |\psi(t)\rangle = H |\psi(t)\rangle$ où H est l'opérateur représentant l'énergie (l'hamiltonien). C'est donc une évolution *déterministe et réversible* (d'un objet de nature probabiliste).
- *Permutabilité des particules identiques.* – La description de l'état quantique en termes d'amplitudes de probabilité conduit à postuler l'indiscernabilité des particules identiques, i.e. que les états à plusieurs particules possèdent une symétrie sous la permutation de deux particules. Notons $\psi(\vec{r}) = \langle \vec{r} | \psi \rangle$ la fonction d'onde, i.e. $\psi(\vec{r})$ est l'amplitude de probabilité de trouver la particule en $\vec{r}$. L'état à N particules est fonction de N coordonnées $\psi(\vec{r}_1, \dots, \vec{r}_i, \dots, \vec{r}_j, \dots, \vec{r}_N)$. Les particules sont classées en deux groupes :
 Les **bosons** (photon,...) dont l'état quantique est **invariant** sous la permutation. Pour deux bosons : $\psi(\vec{r}_1, \vec{r}_2) = \psi(\vec{r}_2, \vec{r}_1)$.
 Les **fermions** (électron, proton, neutron,...) dont l'état quantique est **antisymétrique** sous la permutation. Pour deux fermions : $\psi(\vec{r}_1, \vec{r}_2) = -\psi(\vec{r}_2, \vec{r}_1)$.
 Le théorème spin-statistique montre que les propriétés de permutabilité sont liées au *spin* (moment cinétique intrinsèque) : les bosons ont un spin entier et les fermions un spin demi-entier.

Puisque nous aurons à considérer un nombre macroscopique de particules identiques, le dernier point (postulat de symétrisation) jouera un rôle important. ⁷

II.B Microétats et macroétats, moyennes

Le vecteur d'état contient toute l'information quantique sur le système (une information de nature probabiliste) et correspond donc à la notion de microétat (*état pur*) :

$$\boxed{\text{microétat} = \text{un vecteur d'état } |\psi\rangle}$$

Si on ne connaît pas exactement dans quel état quantique se trouve le système mais savons qu'un état $|\psi_m\rangle$ est occupé avec probabilité P_m , on parle de *mélange statistique*. L'ensemble de ces vecteurs $\{|\psi_m\rangle\}$ doit former une base orthonormée (pour que les événements s'excluent mutuellement),

$$\boxed{\text{macroétat} = \text{ensemble des probabilités d'occupation } \{P_m\} \text{ des vecteurs } \{|\psi_m\rangle\}}$$

Nous devons donc prendre garde à ce qu'il y a une double incertitude sur l'occupation d'un état : quantique et statistique. ⁸ Ceci est bien illustré en écrivant les règles de projection pour un état pur et un mélange statistique :

$$\text{Proba}(\phi \leftarrow \text{état pur } \psi_m) = |\langle \phi | \psi_m \rangle|^2 \quad (3.24)$$

$$\text{Proba}(\phi \leftarrow \text{mélange statistique } \{P_m\}) = \sum_m P_m |\langle \phi | \psi_m \rangle|^2 \quad (3.25)$$

⁷ En principe les particules ne sont pas uniquement caractérisées par leur état dans l'espace physique mais possèdent souvent des degrés de liberté interne, comme le *spin* (son moment cinétique intrinsèque). Dans ce cas on doit également introduire une « coordonnée de spin » que nous notons σ (la projection d'une des composantes du moment cinétique intrinsèque. La fonction d'onde dépend également de cette coordonnée. Si on regroupe les coordonnées $\xi \equiv (\vec{r}, \sigma)$ alors on écrira en toute généralité pour deux particules $\psi(\xi_1, \xi_2) = \pm \psi(\xi_2, \xi_1)$ (+ pour des bosons et – pour des fermions).

⁸ **Attention** : dorénavant nous abolirons la distinction entre « probabiliste » et « statistique » (car le deuxième adjectif est d'usage courant en physique pour désigner la première notion).

P_m caractérise notre ignorance sur l'état quantique occupé par le système; $|\langle \phi | \psi_m \rangle|^2$ traduit l'incertitude pour que, l'état initial état supposé parfaitement connu, le système soit observé dans l'état final $|\phi\rangle$.

Si nous considérons une grandeur physique A , nous aurons à considérer sa moyenne associée au macroétat. D'après le postulat de mesure, si le système est dans l'état $|\psi_m\rangle$, la moyenne *quantique* de l'observable est $\langle A \rangle_{\psi_m} = \langle \psi_m | A | \psi_m \rangle$ (moyenne du résultat des mesures répétées, toujours à partir du même état $|\psi_m\rangle$). Ces moyennes quantiques doivent être pondérées par les probabilités d'occupation des états (moyenne *statistique*) :

$$\underbrace{\langle A \rangle_{\text{mélange}}}_{\text{statistique}} = \overbrace{\sum_m P_m \langle \psi_m | A | \psi_m \rangle}^{\text{moyenne statistique}} \quad (3.26)$$

Exemple : spin 1/2.— Considérons un spin 1/2 dont une base de l'espace de Hilbert est $\{|\uparrow\rangle, |\downarrow\rangle\}$, vecteurs propres de la composante du spin selon Oz , représentée par la matrice de Pauli

$$S_z = \frac{\hbar}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (3.27)$$

Nous étudions la moyenne de la composante du spin S_x :

$$S_x = \frac{\hbar}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \quad (3.28)$$

Un état pur.— Tout d'abord nous étudions la situation où le spin est dans l'état pur $|\psi\rangle = \frac{1}{\sqrt{2}}[|\uparrow\rangle + |\downarrow\rangle]$ (superposition quantique des deux vecteurs de base). Les probabilités pour se trouver dans l'un ou l'autre des états propres de S_z sont $\text{Proba}(\uparrow \leftarrow \psi) = \text{Proba}(\downarrow \leftarrow \psi) = 1/2$. Nous vérifions sans peine que $\langle S_x \rangle_\psi = \langle \psi | S_x | \psi \rangle = \frac{\hbar}{2}$.

Mélange statistique.— Nous considérons maintenant le mélange statistique décrit par les probabilités $P_\uparrow = P_\downarrow = 1/2$ (qui coïncident avec les probabilités quantiques $\text{Proba}(\uparrow \leftarrow \psi)$ et $\text{Proba}(\downarrow \leftarrow \psi)$ pour l'état pur précédent). Cependant la moyenne de la composante est $\langle S_x \rangle_{\text{mélange}} = P_\uparrow \langle \uparrow | S_x | \uparrow \rangle + P_\downarrow \langle \downarrow | S_x | \downarrow \rangle = 0$.

 **Pour en savoir plus :** Les deux situations peuvent être décrites sur un pied d'égalité grâce à la notion d'opérateur densité.⁹ Cf. [3] ou [11].

II.C Dénombrer les microétats : densité d'états

Définition et motivation

Nous avons vu que le postulat fondamental (le cœur de la physique statistique) fait intervenir le volume de l'espace des phases (classique) accessible. Dans la version quantique, la

⁹ **Opérateur densité ρ .**— Afin de décrire le mélange statistique, on introduit l'opérateur densité $\rho \stackrel{\text{def}}{=} \sum_m P_m |\psi_m\rangle \langle \psi_m|$, où $\{|\psi_m\rangle\}$ est une base orthonormée. La condition de normalisation des probabilités prend la forme $\text{tr}\{\rho\} = 1$ et la moyenne $\langle A \rangle = \text{tr}\{\rho A\}$. L'entropie de von Neumann (2.43) s'écrit $S = -k_B \text{tr}\{\rho \ln \rho\}$.

 **Exercice 3.1 :** On considère un spin 1/2. Calculer les entropies associées à (i) $\rho_m = \frac{1}{2}(|\uparrow\rangle \langle \uparrow| + |\downarrow\rangle \langle \downarrow|)$ puis (ii) $\rho_p = \frac{1}{2}[|\uparrow\rangle \langle \uparrow| + |\uparrow\rangle \langle \downarrow| + |\downarrow\rangle \langle \uparrow| + |\downarrow\rangle \langle \downarrow|]$ (indication : diagonaliser ρ dans chaque cas). Interpréter les résultats. Comprenez-vous pourquoi les termes non diagonaux de ρ_p sont appelés « cohérences » ? Calculer $\text{tr}\{\rho_m S_x\}$ puis $\text{tr}\{\rho_p S_x\}$.

problématique correspondante nous amènera à un calcul de *nombre d'états quantiques accessibles*, i.e. au problème de dénombrement des états quantiques ayant une certaine énergie. La notion de densité d'états va nous fournir le langage approprié pour étudier cette question.

Nous verrons que la densité d'états encodera bien souvent toute l'information microscopique suffisante pour la construction d'un modèle dans le cadre de la physique statistique.

Définitions Dans tout problème spectral, comme la caractérisation du spectre $\{E_n\}$ de l'hamiltonien de Schrödinger, une information intéressante est fournie par la *densité spectrale*, i.e. la distribution des valeurs propres :

$$\rho(E) \stackrel{\text{def}}{=} \sum_n \delta(E - E_n) \quad (3.29)$$

où la somme porte sur les **états** quantiques (ce qui se confond avec une somme sur les niveaux d'énergies uniquement en l'absence de dégénérescences ; en présence de dégénérescences on peut relier la somme sur les *états* à la somme sur les *niveaux d'énergie* en introduisant les dégénérescences $d_n : \sum_n \dots = \sum_{E_n} d_n \dots$). Cette quantité mesure le nombre d'états dans un intervalle :

$$\boxed{\rho(E) dE = \# \text{ d'états dans } [E, E + dE]} \quad (3.30)$$

En général cette fonction est extrêmement irrégulière : pour un spectre discret il s'agit d'une somme de pics δ (ce n'est même pas une fonction, mais une *distribution singulière*). Pour éviter de manipuler une distribution singulière, on préfère introduire la *densité d'états intégrée*

$$\Phi(E) \stackrel{\text{def}}{=} \int_{-\infty}^E dE' \rho(E') = \sum_n \theta_H(E - E_n) = \# \text{ d'états d'énergies inférieures à } E, \quad (3.31)$$

qui s'avèrera plus simple à analyser. $\Phi(E)$ présente des discontinuités lorsque E coïncide avec une valeur propre de l'énergie.

Si le spectre devient très dense (i.e. tend vers un continuum), la nature discrète de la distribution des énergies s'estompe. Autrement dit, sur des échelles d'énergie grandes devant l'écart entre niveaux, on peut considérer que la fonction $\Phi(E)$ est une fonction régulière de l'énergie (figure 3.7).

Terme de Weyl – Exemple de l'oscillateur harmonique 1D Considérons par exemple un oscillateur harmonique unidimensionnel dont le spectre est

$$E_n = \hbar\omega \left(n + \frac{1}{2} \right) \quad n \in \mathbb{N}. \quad (3.32)$$

Dans ce cas la densité d'états intégrée $\Phi(E)$ est une fonction en marches d'escalier régulières (Fig. 3.7). Elle est bien approximée par le terme lisse $\Phi_{\text{Weyl}}(E) = E/\hbar\omega$ (la droite qui passe au milieu des marches). On peut ainsi déduire de cette fonction le terme lisse de la densité d'états (la densité d'états moyennée à l'échelle de l'écart entre niveaux) $\rho_{\text{Weyl}}(E) = \Phi'_{\text{Weyl}}(E) = 1/\hbar\omega$, ce qui nous dit qu'il y a en moyenne un état dans un intervalle de largeur $\hbar\omega$ (alors que la densité d'états est un peigne de Dirac régulier). Le terme lisse de la densité d'états est appelé « *terme de Weyl* ».

Motivation L'intérêt de ces notions réside dans le fait que nous aurons souvent à considérer des sommes d'une fonction de l'énergie de la forme

$$\sum_n f(E_n) = \int dE \rho(E) f(E). \quad (3.33)$$

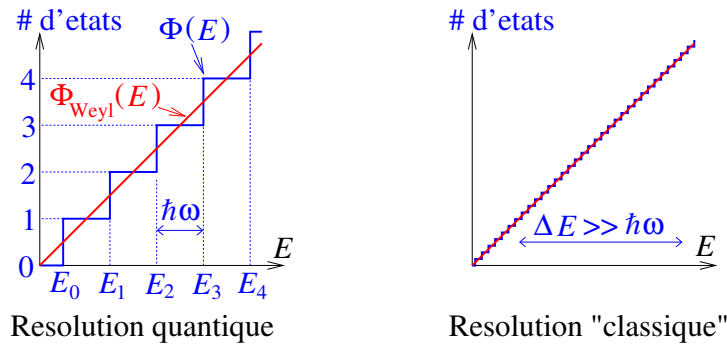


FIGURE 3.7 : **Densité d'états intégrée de l'oscillateur harmonique.** Comparaison de la densité d'états intégrée $\Phi(E)$ avec son approximation semi-classique $\Phi_{\text{Weyl}}(E)$. « Vue de près » (gauche) et « vue de loin » (droite).

Souvent, $f(x)$ sera une fonction qui varie lentement à l'échelle des écarts entre niveaux ; nous pourrons plus simplement écrire (approximation semi-classique)

$$\sum_n f(E_n) \simeq \int dE \rho_{\text{Weyl}}(E) f(E), \quad (3.34)$$

où l'intégrale réalise la moyenne locale sur une échelle plus grande que les écarts entre niveaux :

$$\rho_{\text{Weyl}}(E) \simeq \int_E^{E+\Delta E} \frac{dE'}{\Delta E} \rho(E') \quad \text{avec } \Delta E \gg E_{n+1} - E_n, \quad (3.35)$$

mais encore petite devant des échelles macroscopiques (plus précisément devant l'échelle caractéristique sur laquelle varie la fonction $f(E)$).

Insistons : le calcul de la somme $\sum_n f(E_n)$ requiert la connaissance précise de l'ensemble des valeurs propres de l'énergie, alors que l'intégrale $\int dE \rho_{\text{Weyl}}(E) f(E)$ ne nécessite que leur répartition moyenne sur l'axe réel. Il sera donc en général beaucoup plus facile de calculer l'intégrale (3.34) que la somme (3.33).

En outre, alors que la connaissance du spectre est un problème quantique (difficile), nous allons voir que la détermination de la densité moyenne $\rho_{\text{Weyl}}(E)$ peut être formulée en termes purement classiques (facile).

Densité d'états à l'approximation semi-classique (ou plutôt « semi-quantique »)

Nous discutons la densité d'états pour des particules libres, et montrons comment le calcul de cette quantité, qui caractérise la distribution des énergies des états quantiques, peut être formulé en termes classiques dans une certaine limite.

Particule libre La démonstration générale de la formule semi-classique pour $\rho_{\text{Weyl}}(E)$ fait appel à des outils que nous n'introduisons pas dans ce cours. Nous considérons plutôt l'exemple précis (et important) des particules libres afin d'établir la formulation semi-classique du calcul de la densité d'états intégrée, puis nous étendrons (sans démonstration) le résultat au cas général. Une particule de masse m se meut dans un espace de dimension d . Elle est confinée dans une boîte de forme hypercubique de volume $V = L^d$. Nous supposons des conditions aux limites périodiques, ce qui simplifiera l'analyse (on peut montrer que le terme de Weyl est indépendant de la nature des conditions aux bords). Les états propres du problème sont les ondes planes

$$\psi_{\vec{k}}(\vec{r}) = \frac{1}{\sqrt{V}} e^{i\vec{k} \cdot \vec{r}} \quad \text{où} \quad \vec{k} = \frac{2\pi}{L} (n_1, \dots, n_d) \quad (3.36)$$

avec $n_i \in \mathbb{Z}$ (la quantification du vecteur d'onde est imposée par les conditions aux limites périodiques). La densité d'états intégrée est

$$\Phi(E) = \sum_{\vec{k}} \theta_H \left(E - \frac{\hbar^2 \vec{k}^2}{2m} \right) \quad (3.37)$$

où la somme porte sur tous les vecteurs d'onde quantifiés selon (3.36) (i.e. $\sum_{\vec{k}} \equiv \sum_{n_1 \in \mathbb{Z}} \cdots \sum_{n_d \in \mathbb{Z}}$). $\theta_H(x)$ est la fonction de Heaviside. Autrement dit (3.37) compte tous les $\vec{k}$ (quantifiés) à l'intérieur d'une hypersphère de rayon $\sqrt{2mE}/\hbar$. Si le volume est « grand » (i.e. $E \gg \hbar^2/mL^2$), nous pouvons estimer que $\Phi(E)$ est donné par le volume de l'hypersphère de rayon $\sqrt{2mE}/\hbar$, divisé par le volume $(2\pi/L)^d$ occupé par un état quantique dans l'espace des vecteurs d'onde $\vec{k}$ (figure 3.8). Cela revient à faire une approximation continue et remplacer la somme par une intégrale :

$$\sum_{\vec{k}} \rightarrow \frac{V}{(2\pi)^d} \int d\vec{k} \quad (3.38)$$

Finalement nous substituons l'intégrale sur le vecteur d'onde par une intégrale sur l'impulsion $\vec{p} = \hbar\vec{k}$, et le volume par une intégrale sur les positions :

$$\Phi_{\text{Weyl}}(E) = \frac{1}{h^d} \int_V d\vec{r} \int d\vec{p} \theta_H \left(E - \frac{\vec{p}^2}{2m} \right) \quad (3.39)$$

L'intérêt de ce dernier changement de variable est d'avoir re-formulé le calcul de $\Phi_{\text{Weyl}}(E)$, une fonction comptant les énergies propres (problème quantique), en termes de grandeurs purement classiques. Nous interprétons donc ce résultat comme

$$\Phi_{\text{Weyl}}(E) = \frac{1}{h^d} (\text{volume occupé par les microétats de l'espace des phases d'énergies } < E) \quad (3.40)$$

h^d , où $h = 2\pi\hbar$ est la constante de Planck, s'interprète comme le volume qu'occuperait un état *quantique* dans l'espace des phases *classique*. d doit ici être interprété comme le **nombre de degrés de liberté**, i.e. la moitié de la dimension de l'espace des phases. Notons que ce résultat est réminiscent de la relation de Heisenberg $\delta q \delta p \sim h$.

Nous pouvons terminer le calcul dans le cas d'une particule libre : $\Phi_{\text{Weyl}}(E)$ est donné par le volume de l'hypersphère de rayon $\sqrt{2mE}$ dans un espace de dimension d : ¹⁰

$$\Phi_{\text{Weyl}}(E) = \frac{V}{\Gamma(\frac{d}{2} + 1)} \left(\frac{mE}{2\pi\hbar^2} \right)^{d/2}. \quad (3.41)$$

Cas général Nous admettons que la règle semi-classique (3.40) s'applique pour un hamiltonien général $H(\{q_i, p_i\})$:

$$\Phi_{\text{Weyl}}(E) = \int \prod_{i=1}^d \frac{dq_i dp_i}{h} \theta_H(E - H(\{q_i, p_i\})) \quad (3.42)$$

En pratique, l'approximation semi-classique est justifiée lorsque le spectre des valeurs propres de l'énergie est extrêmement dense, i.e. lorsqu'on n'est pas sensible aux effets de la quantification de l'énergie, ce qui est toujours le cas pour un volume macroscopique.

¹⁰ Le volume de l'hypersphère de rayon 1 est $\mathcal{V}_d = \pi^{d/2}/\Gamma(\frac{d}{2} + 1)$.

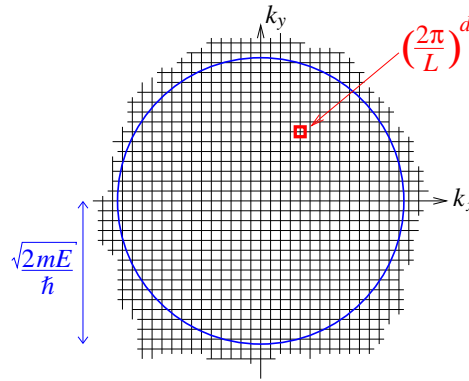


FIGURE 3.8 : Dans l'espace des vecteurs d'onde $\vec{k}$, chaque état quantique d'une particule libre est représenté par un nœud d'un réseau hypercubique de pas $2\pi/L$. Chaque état quantique occupe donc un volume $(2\pi/L)^d$ dans l'espace des $\vec{k} = \vec{p}/\hbar$. La densité d'états intégrée $\Phi(E)$ correspond au nombre de nœuds du réseau compris à l'intérieur de l'hypersphère de rayon $\sqrt{2mE}/\hbar$.

L'intérêt des équations (3.40) et (3.42) est de relier une quantité *quantique*, la densité d'états intégrée, à une propriété purement géométrique de l'espace des phases *classique* (le calcul d'un volume).

Exemple : oscillateur harmonique.— Appliquons la formule précédente au cas de l'oscillateur harmonique unidimensionnel : $H(q, p) = \frac{1}{2m}p^2 + \frac{1}{2}m\omega^2 q^2$. Nous obtenons

$$\Phi_{\text{Weyl}}(E) = \int \frac{dq dp}{h} \theta_H(E - H(q, p)) = \frac{\sqrt{2m}}{h} \underbrace{\sqrt{\frac{2}{m\omega^2}} \int dQ dP \theta_H(E - P^2 - Q^2)}_{\text{surface disque} = \pi E} = \frac{1}{\hbar\omega} E. \quad (3.43)$$

Nous avons retrouvé le résultat discuté au début du paragraphe (figure 3.7).

II.D Densité d'états pour des particules identiques (donc indiscernables)

Indiscernabilité et postulat de symétrisation.— Examinons les conséquences du dernier des postulats de la mécanique quantique concernant les propriétés de permutabilité des états à plusieurs particules identiques. Le postulat de symétrisation a des conséquences extrêmement importantes sur la *structure des états quantiques* et de l'espace de Hilbert, dont il affecte notamment la dimension. Pour le problème qui nous intéressera essentiellement, le *comptage des états quantiques* à travers la densité d'états, son effet est toutefois plus simple à discuter.

Un exemple très simple : deux particules dans un puits harmonique.— Pour clarifier ce point, commençons par un exemple très simple : deux particules sans interaction se déplaçant sur une ligne et soumises à un confinement harmonique. C'est un problème séparable, le spectre des énergies de l'hamiltonien à deux particules est simplement donné par l'addition des spectres à une particule : les énergies sont repérées par deux nombres quantiques (un par particule) $E_{n,m} = \hbar\omega(n + m + 1)$. Si l'on souhaite compter les états, calculer $\Phi(E)$, il nous faut dénombrer les *états*, i.e. caractériser les dégénérescences des *niveaux d'énergie*. Examinons trois cas.

- *Deux particules discernables.*— À chaque énergie $E_{n,m}$ sont associés deux états $|n\rangle \otimes |m\rangle$ et $|m\rangle \otimes |n\rangle$ (suivant que la particule 1 occupe $|n\rangle$ et la 2 occupe $|m\rangle$, ou l'inverse).

L'énergie $E_{n,n}$ est associée à un unique état $|n\rangle \otimes |n\rangle$. Une représentation graphique de ces états quantiques est donc l'ensemble des couples $(n, m) \in \mathbb{N}^2$ dans le quart de plan (figure 3.9). La densité d'états intégrée est donnée par

$$\Phi_{\text{disc}}(E) = \underbrace{\sum_{n=0}^{\infty} \sum_{m=0}^{\infty}}_{\text{sommés indep.}} \theta_H(E - E_{n,m}) \quad (3.44)$$

où la sommation sur les états quantiques est faite à l'aide de deux sommes indépendantes sur les deux nombres quantiques. Le calcul explicite de $\Phi_{\text{disc}}(E)$ peut se faire comme suit : $\Phi_{\text{disc}}(E)$ s'interprète comme le nombre de points du diagramme sous une ligne $n + m = E/(\hbar\omega) - 1$. Dans la limite $E/(\hbar\omega) \gg 1$, le nombre de microétats $\Phi_{\text{disc}}(E) \approx \frac{1}{2} \left(\frac{E}{\hbar\omega} \right)^2$ est la surface du triangle.

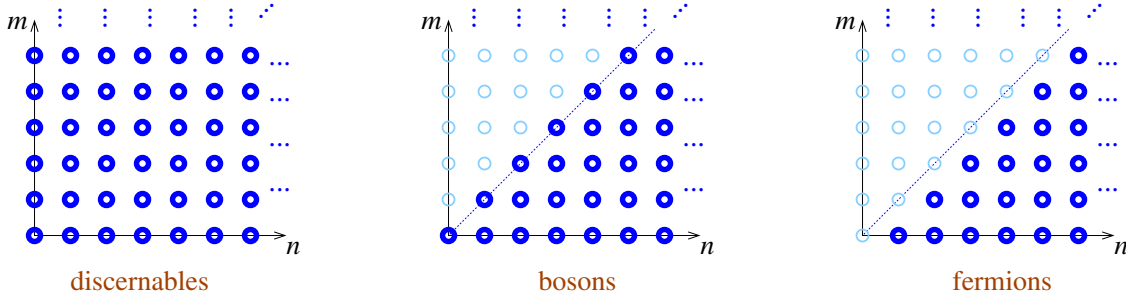


FIGURE 3.9 : Représentation graphique des états quantiques pour deux particules dans un puits harmonique. Les états sont indiqués en trait gras.

- *Deux bosons.* — L'état physique est invariant par permutation des particules, i.e. $|\text{particule 1 dans } n; \text{ particule 2 dans } m\rangle = |\text{particule 1 dans } m; \text{ particule 2 dans } n\rangle$. Un tel état peut être fabriqué à l'aide des états à un corps : $|n, m\rangle_{\text{bosons}} = \frac{1}{\sqrt{2}} [|n\rangle \otimes |m\rangle + |m\rangle \otimes |n\rangle]$. Cet état est invariant par échange des nombres quantiques : $|n, m\rangle = |m, n\rangle$, i.e. les niveaux d'énergie sont non dégénérés. Pour compter les états, il ne faut sommer qu'une fois des couples d'entiers distincts. La représentation graphique des états $|n, m\rangle_{\text{bosons}}$ montre qu'ils ne couvrent qu'un huitième du plan (figure 3.9). La densité d'états intégrée prend alors la forme

$$\Phi_{\text{bosons}}(E) = \underbrace{\sum_{m=0}^{\infty} \sum_{n=m}^{\infty}}_{\text{sommés contraintes}} \theta_H(E - E_{n,m}) \quad (3.45)$$

Le postulat de symétrisation affecte le comptage des microétats : les sommations sur les nombres quantiques ne sont plus indépendantes pour tenir compte du fait que $|n, m\rangle_{\text{bosons}}$ et $|m, n\rangle_{\text{bosons}}$ désignent le même état quantique.

- *Deux fermions.* — Les états quantiques à deux fermions sont antisymétriques sous la permutation, i.e. $|n, m\rangle_{\text{fermions}} = \frac{1}{\sqrt{2}} [|n\rangle \otimes |m\rangle - |m\rangle \otimes |n\rangle]$. Cette fois encore nous devons sommer une unique fois sur les couples d'entiers distincts. Par rapport au cas des bosons, nous devons exclure les états tels que $n = m$ (principe de Pauli). La densité d'états pour les deux fermions est donc

$$\Phi_{\text{fermions}}(E) = \underbrace{\sum_{m=0}^{\infty} \sum_{n=m+1}^{\infty}}_{\text{sommés contraintes}} \theta_H(E - E_{n,m}). \quad (3.46)$$

Lorsque $E/(\hbar\omega) \gg 1$, et puisque les états pour deux bosons et deux fermions ne couvrent que le huitième du plan, nous pouvons écrire très grossièrement que

$$\Phi_{\text{bosons}}(E) \simeq \Phi_{\text{fermions}}(E) \simeq \frac{1}{2!} \Phi_{\text{disc}}(E). \quad (3.47)$$

Cas du gaz parfait de particules libres.— Considérons le cas d'un ensemble de N particules libres identiques se mouvant dans un espace tridimensionnel. Si les particules sont **sans interaction**, les états propres de l'hamiltonien sont des produits tensoriels d'ondes planes $|\vec{k}_1\rangle \otimes \cdots \otimes |\vec{k}_N\rangle$ (états non symétrisés) d'énergie $E_{\vec{k}_1, \dots, \vec{k}_N} = \sum_{i=1}^N \frac{\hbar^2 \vec{k}_i^2}{2m}$. Les états physiques satisfaisant le postulat de symétrisation sont symétriques sous l'échange des particules : ils sont obtenus en appliquant un opérateur de symétrisation sur l'état non symétrisé. Alors que l'état non symétrisé $|\Psi_{\text{non sym}}\rangle = |\vec{k}_1\rangle \otimes \cdots \otimes |\vec{k}_N\rangle$ est repéré par la suite *ordonnée* des nombres quantiques (quelle particule est dans quel état), l'état symétrisé $|\Psi_{\text{sym}}\rangle = |\vec{k}_1, \dots, \vec{k}_N\rangle$ est repéré par la donnée de la suite *non ordonnée* des nombres quantiques (quel état est occupé par n'importe laquelle des N particules). Autrement dit **il ne faut pas sommer indépendamment sur les N vecteurs d'ondes, mais tenir compte des permutations des nombres quantiques correspondant à des états quantiques équivalents** : par exemple si tous les nombres quantiques sont distincts, $N!$ permutations des indices correspondent au même état quantique. Dans le cas des bosons, on a

$$\Phi_{\text{bosons}}(E) = \sum_{\vec{k}_1, \dots, \vec{k}_N} \frac{n_{\vec{k}_1}! \cdots n_{\vec{k}_N}!}{N!} \theta_H \left(E - \sum_{i=1}^N \epsilon_{\vec{k}_i} \right) \quad (3.48)$$

où $n_{\vec{k}_i}$ est le nombre de particules occupant l'état individuel $|\vec{k}_i\rangle$, et $\epsilon_{\vec{k}_i} = \frac{\hbar^2 \vec{k}_i^2}{2m}$. Les facteurs $n_{\vec{k}_i}!$ tiennent compte de l'éventualité que plusieurs vecteurs d'onde coïncident lors de la sommation (si deux vecteurs coïncident, le nombre de permutations équivalentes est en effet inférieur à $N!$). Pour des fermions, on doit exclure de la somme les termes où plusieurs particules occupent le même état individuel (principe de Pauli) :

$$\Phi_{\text{fermions}}(E) = \frac{1}{N!} \sum_{\substack{\vec{k}_1, \dots, \vec{k}_N \\ \text{tous différents}}} \theta_H \left(E - \sum_{i=1}^N \epsilon_{\vec{k}_i} \right). \quad (3.49)$$

La difficulté du calcul réside dans le fait que les sommations sur les vecteurs d'onde ne sont pas indépendantes puisqu'elles sont affectées par des facteurs combinatoires non triviaux pour les bosons, et sont contraintes par le principe de Pauli pour les fermions.

Facteurs d'occupation.— Nous introduisons une formulation qui se révélera très utile pour discuter les statistiques quantiques, vers la fin du cours. L'état quantique respectant le postulat de symétrisation est caractérisé par la donnée des états individuels occupés (sans préciser l'identité des particules occupant ces états). Autrement dit l'état quantique symétrisé $|\Psi_{\text{sym}}\rangle = |\{n_{\vec{k}}\}\rangle$ peut être spécifié par la donnée de l'ensemble des *facteurs d'occupations* des états quantiques individuels, où $n_{\vec{k}}$ est le nombre de particules dans l'état $|\vec{k}\rangle$. La densité d'états intégrée du problème à N particules s'écrit en ces termes :

$$\Phi_{\text{indisc}}(E) = \sum_{\substack{\{n_{\vec{k}}\} \text{ tq.} \\ \sum_{\vec{k}} n_{\vec{k}} = N}} \theta_H \left(E - \sum_{\vec{k}} n_{\vec{k}} \epsilon_{\vec{k}} \right) \quad \begin{cases} n_{\vec{k}} \in \mathbb{N} & (\text{bosons}) \\ n_{\vec{k}} \in \{0, 1\} & (\text{fermions}) \end{cases}. \quad (3.50)$$

Contrairement aux expressions (3.48) et (3.49), la somme $\sum_{\vec{k}}$ ne porte plus sur les particules, ou plutôt sur les états effectivement occupés par des particules, mais sur l'ensemble des états individuels (en nombre infini en général). La somme sur les états à N particules (i.e. sur l'ensemble des facteurs d'occupation) est clairement contrainte pour que la somme des occupations coïncide avec le nombre de particules. Pour des bosons $n_{\vec{k}} \in \mathbb{N}$ alors que pour des fermions $n_{\vec{k}} \in \{0, 1\}$.

Limite diluée.— Lorsque l'on considère la limite d'un fluide dilué (haute température/faible densité), la somme sur les états est dominée par des termes tels que les occupations des états individuels soient petites. En effet, lorsque l'énergie du système est « grande », le nombre de manières de distribuer cette énergie parmi les états individuels est en général faramineux : le nombre d'états accessibles est beaucoup plus grand que le nombre de particules. L'occupation d'un état $n_{\vec{k}_i}$ vaut la plupart du temps 0, rarement 1 et est encore plus rarement supérieur à 1. On peut alors négliger les contraintes sur les nombres quantiques lors des sommations dans (3.48) et (3.49). Nous obtenons alors un résultat identique pour les bosons et les fermions

$$\Phi_{\text{indisc}}(E) \simeq \frac{1}{N!} \sum_{\vec{k}_1, \dots, \vec{k}_N} \theta_H \left(E - \sum_{i=1}^N \frac{\hbar^2 \vec{k}_i^2}{2m} \right) \quad (3.51)$$

Autrement dit, si nous introduisons $\Phi_{\text{indisc}}(E) \simeq \Phi_{\text{bosons}}(E) \simeq \Phi_{\text{fermions}}(E)$, on a

$$\boxed{\Phi_{\text{indisc}}(E) \simeq \frac{1}{N!} \Phi_{\text{disc}}(E)} \quad (3.52)$$

où dans l'approximation semi-classique, $\Phi_{\text{disc}}(E)$ est donné par (3.42). L'approximation faite ici est équivalente à « l'approximation de Maxwell-Boltzmann » qui sera discutée plus tard (les expressions (3.48, 3.49, 3.50, 3.52) deviendront plus claires aux chapitres 9 et 10).

Dans le cas des particules libres, en remarquant que la sommation sur les N vecteurs tridimensionnels est équivalente à la sommation sur un unique vecteur de dimension $d = 3N$, nous pouvons écrire $\Phi_{\text{disc}}(E)$ sous la forme (3.37), et donc

$$\Phi_{\text{indisc}}^{\text{libres}}(E) \simeq \frac{1}{N!} \frac{V^N}{\Gamma(\frac{3N}{2} + 1)} \left(\frac{mE}{2\pi\hbar^2} \right)^{3N/2} \stackrel{\text{Stirling}}{\simeq} \frac{e^{5N/2}}{\pi\sqrt{6}N} \left(\frac{V}{N} \right)^N \left(\frac{mE}{3\pi\hbar^2 N} \right)^{3N/2}, \quad (3.53)$$

où $V = L^3$ est maintenant le volume dans l'espace tridimensionnel.

III Le postulat fondamental de la physique statistique

Nous donnons une formulation générale du postulat fondamental de la physique statistique.

Notations Notons (ℓ) les microétats du système. Dans le cas classique, (ℓ) représente un point $\{q_i, p_i\}$ de l'espace des phases. Dans le cas quantique, il désigne un état quantique $|\ell\rangle$ de l'espace de Hilbert.

III.A La distribution microcanonique

Dans le cas classique, l'hypothèse ergodique nous a conduit à l'idée que la mesure (la probabilité d'occupation des microétats) est uniforme sur la « surface » de l'espace des phases d'énergie E . Dans le cas quantique la dynamique complexe et les perturbations permettent au système d'explorer tous les états quantiques d'énergie fixée.

Postulat fondamental : tous les microétats accessibles d'un système isolé et à l'équilibre (macroscopique) sont équiprobables.

Nous introduisons $\Omega(E)$, le nombre d'états accessibles (le nombre de microétats dont l'énergie est E , à une incertitude δE près). Il s'exprime en termes de la densité d'états

$$\Omega(E) = \rho(E) \delta E \quad (3.54)$$

La probabilité pour que le système soit dans le microétat (ℓ) d'énergie E_ℓ est :

$$P_\ell^* = \begin{cases} \frac{1}{\Omega(E)} & \text{si } E \leq E_\ell \leq E + \delta E \\ 0 & \text{sinon} \end{cases} \quad (3.55)$$

Qu'on appelle la *distribution microcanonique*.

Remarque n°1 : Si nous choisissons $\delta E = 0$, le nombre d'états accessibles $\Omega(E)$ se confond avec la notion de dégénérescence d'un niveau d'énergie. Si $\delta E \neq 0$, le nombre $\Omega(E)$ est la somme des dégénérescences des niveaux dans $[E, E + \delta E]$.

Remarque n°2 : Dans le cas classique, (3.55) est une re-écriture de (3.16) (la correspondance précise est assurée en remarquant que le nombre de microétats accessibles est $\Omega(E) = \mathcal{V}(E, \delta E)/h^d$ où d est le nombre de degrés de liberté ; notons qu'au niveau de (3.16), l'effet de l'indiscernabilité n'avait pas encore été discuté, ce qui explique l'absence du facteur $1/N!$ dans cette dernière expression).

Remarque n°3 : Une démonstration du postulat ? Nous avons déjà remarqué dans la section I que la distribution microcanonique peut-être déduite d'un « principe d'entropie maximale », compte tenu des contraintes en présence (on parle parfois de « principe de désordre maximal »). On peut rendre l'argument assez précis : nous montrons dans l'annexe page 74 comment poser le problème de maximisation de l'entropie sous contrainte : nous prouvons que, compte tenu de la contrainte sur l'énergie, la distribution microcanonique maximise l'entropie statistique (2.43).

III.B Entropie microcanonique

D'après la définition (2.43), nous déduisons que l'entropie (l'information manquante) de la distribution canonique est

$$S^*(E) = k_B \ln \Omega(E) \quad (3.56)$$

Nous allons voir que l'entropie microcanonique, en fonction des paramètres macroscopiques du problème, volume V , nombre de particules N , etc, $S^*(E, V, N, \dots)$, est la fonction fondamentale permettant d'obtenir les *propriétés thermodynamiques* du système.¹¹ S^* jouera le rôle de *fonction génératrice des propriétés thermodynamiques*. Cette observation souligne le caractère remarquable de la relation fondamentale (3.56) qui relie $\Omega(E)$, une fonction qui dénombre des microétats, à des propriétés mesurables (équation d'état, chaleur spécifique,...).

¹¹ Il s'agit du postulat fondamental de la thermodynamique : cf. § II.

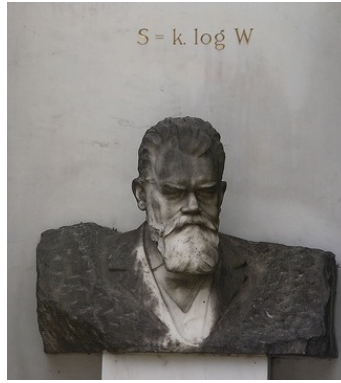


FIGURE 3.10 : La tombe de Ludwig Boltzmann (1844-1906) à Vienne.

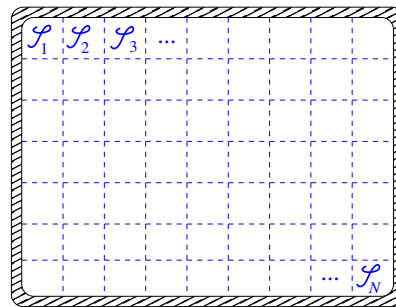
Extensivité de l'entropie

Variabes extensives et intensives Les grandeurs thermodynamiques peuvent être classées selon deux catégories : les variables extensives et les variables intensives. Considérons un système isolé de N particules, d'énergie E , occupant un volume V . Si le système est dupliqué, les variables « extensives » le sont également $N \rightarrow 2N$, $E \rightarrow 2E$, $V \rightarrow 2V$, etc alors que les variables « intensives » sont inchangées (densité $n = N/V$, température T , pression p , etc).

Propriétés d'extensivité de l'entropie pour des interactions à courte portée Nous considérons un système physique au sein duquel les interactions entre constituants sont à courte portée, ce qui sera toujours le cas dans les exemples que nous étudierons.¹² Nous séparons le système $\mathcal{S}$ en N sous-systèmes $\mathcal{S}_1, \dots, \mathcal{S}_N$. L'hypothèse d'interactions à courte portée nous permet d'affirmer que l'entropie du système global est la somme des entropies des sous-systèmes :

$$S(\mathcal{S}_1 \otimes \dots \otimes \mathcal{S}_N) \approx S(\mathcal{S}_1) + \dots + S(\mathcal{S}_N) \quad (3.57)$$

où nous avons utilisé la propriété de sous-additivité de l'entropie (2.52), ou plutôt (2.50); nous avons remarqué que l'égalité est réalisée lorsque les deux systèmes sont non corrélés, or seuls les sous-systèmes voisins sont faiblement corrélés si nous supposons que les interactions sont à courte portée.

FIGURE 3.11 : Un système $\mathcal{S}$ est subdivisé en N sous-systèmes $\mathcal{S}_1, \dots, \mathcal{S}_N$ (avec de faibles interactions mutuelles).

L'entropie est donc une grandeur *extensive*, i.e. qui croît linéairement avec la taille du système : $S \propto N$. Puisque l'énergie est également une grandeur extensive, nous concluons que

¹² Un contre exemple : les systèmes de particules en interaction gravitationnelle.

l'entropie est une fonction de la forme :

$$S(E, N) = N \times \text{fct} \left(\frac{E}{N} \right) \quad (3.58)$$

Si nous réintroduisons la dépendance dans le volume, également extensif, nous obtenons la structure

$$S(E, N, V) = N s \left(\frac{E}{N}, \frac{V}{N} \right) \quad (3.59)$$

où $s(E, V)$ est une entropie par particule. Bien que l'état d'un fluide simple soit caractérisé par trois grandeurs thermodynamiques, E , V et N , l'extensivité contraint la « fonction entropie », qui fait intervenir une fonction de *deux* variables, et non trois.

Nous déduisons que le nombre de microétats accessibles croît exponentiellement avec N (qui est macroscopique)

$$\Omega(E, N) \sim \exp \left[N \text{fct} \left(\frac{E}{N} \right) \right]. \quad (3.60)$$

C'est donc un nombre faramineux en général (cf. section I.B sur les ordres de grandeur) !

Exemple : entropie du gaz parfait monoatomique (formule de Sackur Tetrode)

Nous considérons un gaz parfait isolé de N atomes contenu dans une boîte de volume V . Nous avons calculé plus haut la densité d'états d'un gaz de particules libres indiscernables (3.53), dont nous pouvons déduire la densité d'états et le nombre de microétats accessibles $\Omega(E) = \rho(E)\delta E = \Phi'(E)\delta E$, puis l'entropie microcanonique (3.56). Remarquons toutefois qu'il est possible en pratique de s'épargner la dérivation par rapport à l'énergie : puisque la densité d'états est une loi de puissance $\Phi(E) \propto E^{3N/2}$ on a $\rho(E) = \frac{3N}{2E} \Phi(E) \propto E^{\frac{3N}{2}-1}$, d'où :

$$S^* = k_B \ln \left[\Phi(E) \frac{3N\delta E}{2E} \right] = k_B \left[\underbrace{\ln \Phi(E)}_{\text{extensif } \mathcal{O}(N)} + \underbrace{\ln \left(\frac{3N\delta E}{2E} \right)}_{\text{intensif } \mathcal{O}(N^0)} \right]. \quad (3.61)$$

Si N est un nombre de l'ordre de $\mathcal{N}_A$ on peut simplement négliger le second terme et écrire :

$$S^* \simeq k_B \ln [\Phi(E)] \quad \text{pour } N \rightarrow \infty \quad (3.62)$$

Autrement dit, pour le calcul de l'entropie à la limite thermodynamique $N \rightarrow \infty$, on peut prendre indifféremment comme argument du logarithme, soit le volume des états accessibles (couche d'épaisseur δE autour de l'énergie E dans l'espace des phases), soit le volume des états d'énergies inférieures à E . La croissance de $\Phi(E)$ avec E est tellement formidable, que sa valeur est dominée par l'effet de bord : on peut injecter dans le logarithme, soit $\Phi'(E)$, soit $\Phi(E)$.

En utilisant (3.53), on arrive à la *formule de Sackur-Tetrode* (cf. commentaires historiques de l'annexe page 11)

$$S^*(E, V, N) = N k_B \left[\frac{5}{2} + \ln \left(\frac{V}{N} \left[\frac{mE}{3\pi\hbar^2 N} \right]^{3/2} \right) \right] \quad (3.63)$$

On peut vérifier qu'elle présente les bonnes propriétés d'extensivité discutées plus haut. Notons toutefois qu'en passant de (3.62) à (3.63), nous avons négligé un terme $-k_B \ln(\pi\sqrt{6}N) \ll Nk_B$, i.e. au total, de la formule générale (3.56) au résultat (3.63), un terme $k_B \ln \left(\frac{1}{\pi} \sqrt{\frac{3}{2}} \frac{\delta E}{E} \right)$. Ce terme est non extensif, mais complètement négligeable, d'ordre $\mathcal{O}(\ln N)$.

Remarque : la prédiction de la thermodynamique.— Nous pouvons comparer ce résultat à celui obtenu dans le cadre de la thermodynamique. Des considérations générales permettent de montrer dans ce cadre que l'entropie du gaz parfait classique est de la forme $S(E, V, N) = Nk_B \ln \left[\frac{V}{N} \left(\frac{E}{N} \right)^\alpha \right] + S_0$. La thermodynamique ne donne cependant aucune indication pour faire une prédiction sur la valeur de l'exposant α , ni sur la constante S_0 , qui doivent être déduits de l'expérience.

IV Relâchement de contraintes – Distribution d'une variable interne

Nous avons introduit précédemment l'entropie microcanonique S^* dont nous avons dit qu'elle jouera un rôle fondamental. La question que nous allons maintenant aborder est de savoir comment extraire l'information physique « intéressante » (pertinente expérimentalement ?) de cette fonction fondamentale. En particulier, si la formule de Sackur-Tetrode (3.63) encode de l'information sur le gaz parfait monoatomique, comme extraire celle-ci ? Par exemple comment obtenir la pression du gaz, i.e. son équation d'état, etc.

IV.A Variables externes et variables internes

Nous n'avons pour l'instant introduit que des variables fixées caractérisant le système isolé : son énergie, le nombre de particules, etc. On parle de « **variables externes** ».

On peut aussi considérer des *variables macroscopiques fluctuantes*, comme le nombre de particules N_v du volume fictif de la figure 3.1. Dans ce cas on parle de « **variables internes** ».

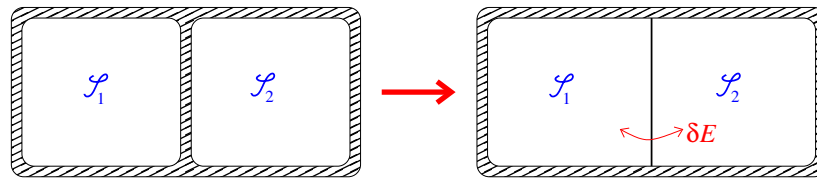


FIGURE 3.12 : **Contact thermique.** Deux systèmes initialement isolés sont mis en contact thermique ; ils peuvent alors échanger de l'énergie.

IV.B Contact thermique entre deux systèmes – Température microcanonique

Nous étudions le contact thermique entre deux systèmes $\mathcal{S}_1$ et $\mathcal{S}_2$ initialement isolés (figure 3.12). Chaque système est caractérisé par son entropie microcanonique $S_1^*(E_1) = k_B \ln \Omega_1(E_1)$ et $S_2^*(E_2) = k_B \ln \Omega_2(E_2)$, où $\Omega_{1,2}$ sont les nombres d'états accessibles des systèmes lorsqu'ils sont isolés. À un certain instant on met les deux systèmes en **contact thermique** (la paroi isolante devient une paroi « diatherme » laissant passer la chaleur), puis **on laisse évoluer le système global vers l'équilibre macroscopique**. Lorsqu'ils sont en contact, les systèmes échangent de l'énergie, E_1 et E_2 deviennent des variables internes, bien que l'énergie totale reste fixée (avant et après le contact), puisque l'ensemble est isolé.

Hypothèse importante : interactions à courte portée

Nous admettons que les interactions entre les deux systèmes sont faibles (mais pas inexistantes sinon ils n'échangeraient pas d'énergie). L'énergie totale est alors la somme de l'énergie de chacun

des sous systèmes

$$E = E_1 + E_2 + E_{1 \leftrightarrow 2}^{\text{int}} \simeq E_1 + E_2 . \quad (3.64)$$

La justification à négliger l'énergie d'interaction $E_{1 \leftrightarrow 2}^{\text{int}}$ repose sur l'hypothèse d'interactions à courte portée. Dans ce cas, on s'attend à ce que l'interaction soit un effet de surface, $E_{1 \leftrightarrow 2}^{\text{int}} \propto N^{2/3}$, négligeable devant les énergies des sous systèmes qui intègrent une énergie en volume, E_1 & $E_2 \propto N$.

L'hypothèse de faibles interactions a une conséquence importante sur le comptage des états accessibles : puisque l'énergie E_1 de $\mathcal{S}_1$ est indépendante de l'état de $\mathcal{S}_2$, ceci reste vrai pour le nombre de microétats accessibles $\Omega_1(E_1)$. Si E_1 et E_2 sont fixées (i.e. *avant* le contact thermique), le nombre de microétats de $\mathcal{S}_1 \otimes \mathcal{S}_2$ accessibles est $\Omega_{1 \otimes 2}|_{\text{avant}} = \Omega_1(E_1) \Omega_2(E_2)$. Si E_1 et E_2 sont des variables internes (*après* le contact thermique), nous devons additionner tous les nombres $\Omega_1(E_1) \Omega_2(E_2)$ compatibles avec la contrainte $E_1 + E_2 = E$:

$$\Omega_{1 \otimes 2}(E) = \sum_{E_1} \Omega_1(E_1) \Omega_2(E - E_1) . \quad (3.65)$$

Clarifions le sens de la somme discrète sur l'énergie en revenant à la formulation en termes des densités d'états. Nous notons $\rho_{1,2}(E_{1,2})$ la densité d'états de $\mathcal{S}_{1,2}$, liée au nombre de microétats accessibles par $\Omega_{1,2}(E_{1,2}) = \rho_{1,2}(E_{1,2}) \delta E$. Nous écrivons naturellement :

$$\Omega_{1 \otimes 2}(E) = \delta E \int dE_1 \rho_1(E_1) \int dE_2 \rho_2(E_2) \overbrace{\delta(E - E_1 - E_2)}^{\rho_{1 \otimes 2}(E) = \sum_{\ell_1, \ell_2} \delta(E - E_{\ell_1} - E_{\ell_2})} \quad (3.66)$$

$$= \delta E \int dE_1 \rho_1(E_1) \rho_2(E - E_1) = \underbrace{\int \frac{dE_1}{\delta E}}_{\ll \sum_{E_1} \gg} \Omega_1(E_1) \Omega_2(E - E_1) . \quad (3.67)$$

Entropie réduite et distribution de E_1

Nous étudions la distribution de l'énergie E_1 (la variable interne). L'hypothèse précédente nous permet d'écrire le nombre de microétats accessibles pour $\mathcal{S}_1 \otimes \mathcal{S}_2$, lorsque E_1 est fixée, comme : $\omega(E_1; E) \stackrel{\text{def}}{=} \Omega_1(E_1) \Omega_2(E - E_1)$. D'après le postulat fondamental, la probabilité $p^*(E_1; E)$ est égale à la fraction des états accessibles ayant l'énergie E_1 :

$$p^*(E_1; E) = \frac{\omega(E_1; E)}{\Omega_{1 \otimes 2}(E)} = \frac{\Omega_1(E_1) \Omega_2(E - E_1)}{\Omega_{1 \otimes 2}(E)} = e^{\frac{1}{k_B} [\tilde{S}^*(E_1; E) - S_{1 \otimes 2}^*(E)]} , \quad (3.68)$$

où nous avons introduit l'entropie réduite :

$$\tilde{S}^*(E_1; E) \stackrel{\text{def}}{=} k_B \ln \omega(E_1; E) = S_1^*(E_1) + S_2^*(E - E_1) , \quad (3.69)$$

qui correspond à l'entropie de $\mathcal{S}_1 \otimes \mathcal{S}_2$, lorsque E_1 est fixée. La condition de normalisation de (3.68) est évidente $\sum_{E_1} p^*(E_1; E) = 1$. Remarquons que *l'additivité des entropies repose sur l'hypothèse d'interaction à courte portée* (i.e. nous avons encore utilisé la propriété (2.50)). Comme nous l'avons vu, les nombres d'états accessibles $\Omega_{1,2}$ sont énormes ; d'autre part ils sont en général des fonctions croissantes de l'énergie.¹³ À cause de la propriété (3.60) ces variations sont très violentes. $\Omega_1(E_1)$ croît (extrêmement vite) avec E_1 alors que $\Omega_2(E - E_1)$ décroît vite (figure 3.13). Le produit va donc atteindre un maximum pour une valeur de l'énergie E_1^{max} , qui correspond à la valeur qui maximise $p^*(E_1; E)$. Cette valeur maximise aussi l'entropie réduite :

$$\frac{\partial \tilde{S}^*}{\partial E_1}(E_1^{\text{max}}; E) = 0 \quad \Rightarrow \quad \frac{\partial S_1^*}{\partial E_1}(E_1^{\text{max}}) = \frac{\partial S_2^*}{\partial E_2}(E - E_1^{\text{max}}) . \quad (3.70)$$

¹³On verra un contre exemple en TD, lorsque l'espace de Hilbert a une dimension finie.

Cette condition permet de trouver la **valeur de $E_1^{\max}$ la plus probable**, caractérisant un **équilibre macroscopique** entre sous systèmes.

Attention toutefois, lorsque nous parlons ici d'équilibre, nous avons en tête une condition fixant la valeur la plus probable de la variable interne (ici E_1), cependant nous devons garder à l'esprit que celle-ci présente des fluctuations que nous allons caractériser plus bas.

Définition de la température microcanonique.— La correspondance avec la thermodynamique nous indique qu'après contact thermique, la condition d'équilibre est l'égalité des températures : la température est définie par l'égalité fondamentale (1.1) : $T = \left(\frac{\partial E}{\partial S}\right)_{V,N}$. La correspondance entre la physique statistique et la thermodynamique (dans une limite qui reste à préciser), nous conduit à introduire la définition générale pour la température microcanonique :

$$\boxed{\frac{1}{T^*(E)} \stackrel{\text{def}}{=} \frac{\partial S^*(E)}{\partial E}} \quad (3.71)$$

Il s'agit de la *température absolue* (introduite par en 1848 par William Thomson, futur Lord Kelvin), dont le zéro correspond à -273.15°C . Notons une petite différence avec la notation de la thermodynamique $1/T = \left(\frac{\partial S}{\partial E}\right)_{V,N}$: l'entropie microcanonique caractérise le système *isolé*, i.e. est naturellement une fonction de E, V, N , donc il sera inutile de préciser que la dérivation partielle de S^* se fait à V et N fixés.

Exemple : température du gaz parfait monoatomique.— Partant de l'entropie (3.63), nous obtenons la température microcanonique du gaz parfait :

$$T^*(E, V, N) = \frac{1}{\partial S^*/\partial E} = \frac{2E}{3Nk_B}. \quad (3.72)$$

Elle est donc *indépendante du volume*; par conséquent, étant intensive, T^* ne peut dépendre que du rapport E/N des deux quantités extensives.

Condition d'équilibre thermique.— Revenons au problème du contact thermique : la condition « d'équilibre » (3.70) peut être re-écrite à l'aide des températures microcanoniques des deux sous systèmes :

$$\boxed{T_1^*(E_1) = T_2^*(E - E_1)} \quad (\text{condition d'équilibre thermique}) \quad (3.73)$$

La solution de cette équation est la valeur la plus probable $E_1 = E_1^{\max}$.

Rq n°1 : fluctuations et stabilité de l'équilibre.— La *stabilité* de l'équilibre, i.e. l'existence d'un maximum de l'entropie exige que celle-ci soit une fonction *concave* et par conséquent la *température est une fonction croissante de l'énergie* :

$$\boxed{\frac{\partial^2 S^*}{\partial E^2} < 0} \quad \Rightarrow \quad \boxed{\frac{\partial T^*}{\partial E} > 0} \quad (3.74)$$

Le lecteur aura peut-être reconnu la chaleur spécifique à volume constant, définie dans le cadre de la thermodynamique comme $C_V \stackrel{\text{def}}{=} \left(\frac{\partial E}{\partial T}\right)_{V,N}$. Pour assurer la correspondance avec la thermodynamique, nous proposons la définition suivante de la chaleur spécifique microcanonique

$$C_V^* \stackrel{\text{def}}{=} \frac{1}{\partial T^*/\partial E} = -\frac{1}{(T^*)^2 \partial^2 S^*/\partial E^2}. \quad (3.75)$$

La stabilité de l'équilibre implique donc la positivité de la chaleur spécifique.¹⁴

On peut garder en tête que E_1 est sujette à des fluctuations caractérisées par la variance

$$\text{Var}(E_1) = -k_B \left[\frac{\partial^2 \tilde{S}^*}{\partial E_1^2}(E_1^{\max}; E) \right]^{-1} = -k_B \left[\frac{\partial^2 S_1^*}{\partial E_1^2}(E_1^{\max}) + \frac{\partial^2 S_2^*}{\partial E_2^2}(E - E_1^{\max}) \right]^{-1}. \quad (3.76)$$

D'après (3.75), les fluctuations sont reliées aux chaleurs spécifiques des deux sous-systèmes

$$(\sigma_{E_1})^2 = \text{Var}(E_1) = k_B T^{*2} \left(\frac{1}{C_{V1}^*} + \frac{1}{C_{V2}^*} \right)^{-1} \quad (3.77)$$

où C_{V1}^* et C_{V2}^* sont les chaleurs spécifiques des deux sous systèmes. Cette expression montre que les fluctuations de l'énergie sont essentiellement contrôlées par le système ayant la plus petite chaleur spécifique : si $C_{V1}^* \ll C_{V2}^*$ alors $\sigma_{E_1} \simeq k_B T^* \sqrt{C_{V1}^*/k_B}$.

Considérons la limite où le nombre de particules est macroscopique : l'entropie et l'énergie sont des quantités extensives, $S_{1,2}^* \propto N$ et $E_{1,2} \propto N$; l'expression précédente montre donc que $1/\text{Var}(E_1) \propto N/N^2$ i.e. $\sigma_{E_1} \propto \sqrt{N}$ (plus simplement : les chaleurs spécifiques sont des quantités extensives). Les fluctuations *relatives* tendent donc vers zéro

$$\frac{\sigma_{E_1}}{E_1^{\max}} \propto 1/\sqrt{N} \xrightarrow{N \rightarrow \infty} 0. \quad (3.78)$$

La limite $N \rightarrow \infty$, pour laquelle on peut **négliger l'existence des fluctuations**, est appelée la **limite thermodynamique**.

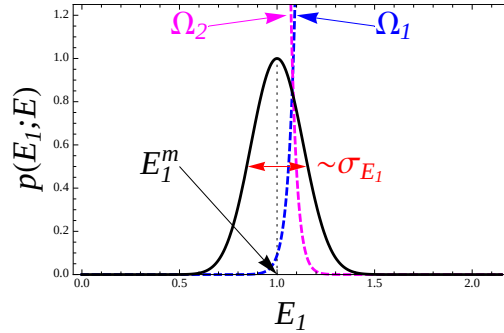


FIGURE 3.13 : **Contact thermique.** *Nombres de microétats accessibles des sous systèmes en fonction de E_1 et distribution de l'énergie.*

Rq n°2 : entropie après contact thermique.— Nous calculons l'entropie du système une fois que le contact thermique a été établi : notre point de départ est l'expression (3.67), dans laquelle nous faisons apparaître l'entropie réduite :

$$\Omega_{1\otimes 2}(E) = \frac{1}{\delta E} \int dE_1 e^{\frac{1}{k_B} [S_1^*(E_1) + S_2^*(E - E_1)]} = \frac{1}{\delta E} \int dE_1 e^{\frac{1}{k_B} \tilde{S}^*(E_1; E)}. \quad (3.79)$$

La somme est dominée par le voisinage de $E_1^{\max}$, la valeur qui maximise l'argument de l'exponentielle. En faisant l'approximation gaussienne (méthode du col, cf. annexe A),

$$\Omega_{1\otimes 2}(E) \simeq \frac{1}{\delta E} \sqrt{\frac{2\pi k_B}{-\frac{\partial^2 \tilde{S}^*}{\partial E_1^2}(E_1^{\max}; E)}} e^{\frac{1}{k_B} \tilde{S}^*(E_1^{\max}; E)} = \sqrt{2\pi} \frac{\sigma_{E_1}}{\delta E} e^{\frac{1}{k_B} \tilde{S}^*(E_1^{\max}; E)} \quad (3.80)$$

¹⁴Notons que ces propriétés reposent sur l'hypothèse d'interactions à courtes portées, centrale dans l'analyse du contact thermique. Lorsque cette hypothèse n'est pas satisfaite, il est possible de trouver des situations où $C_V^* < 0$.

où nous avons utilisé (3.76) et (3.77). On déduit l'expression de l'entropie totale du système $S_{1\otimes 2}^*(E) = k_B \ln \Omega_{1\otimes 2}(E)$:

$$S_{1\otimes 2}^*(E) = \underbrace{S_1^*(E_1^{\max}) + S_2^*(E - E_1^{\max})}_{\substack{\text{entropie à } E_1 \text{ fixée :} \\ \mathcal{O}(N)}} + \underbrace{k_B \ln \left(\frac{\sqrt{2\pi} \sigma_{E_1}}{\delta E} \right)}_{\substack{\text{entropie due aux fluctuations de } E_1 : \\ \mathcal{O}(\ln N)}}. \quad (3.81)$$

Le terme dû aux fluctuations de E_1 traduit un manque d'information sur l'état macroscopique. C'est seulement à la limite thermodynamique, $N \rightarrow \infty$, que l'on peut négliger cette dernière contribution à l'entropie, qui est alors donnée par la somme des entropies des deux sous-systèmes :

$$S_{1\otimes 2}^*(E) \simeq S_1^*(E_1^{\max}) + S_2^*(E - E_1^{\max}). \quad (3.82)$$

Exercice 3.2 : En utilisant la formule de Sackur-Tetrode obtenue plus haut, calculer explicitement $-\frac{\delta E^2}{k_B} \frac{\partial^2 S^*}{\partial E^2}$. Vérifier que le dernier terme de (3.81) est bien $\mathcal{O}(\ln N)$.

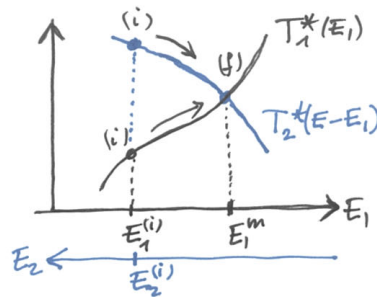


FIGURE 3.14 : **Contact thermique.** Le système ayant initialement la température la plus haute (ici $\mathcal{S}_2$) cède de l'énergie (de la chaleur) à celui ayant initialement la température la plus basse.

Rq n°3 : sens des échanges d'énergie.— Les fonctions $T_1^*(E_1)$ et $T_2^*(E - E_1)$ sont donc respectivement croissante et décroissante en fonction de E_1 . Si le point d'intersection correspond à l'équilibre, nous constatons donc que le système ayant initialement la température la plus haute cède de l'énergie à celui ayant initialement la température la plus basse (figure 3.14).¹⁵

Conclusion : En conclusion, nous pouvons formuler le « théorème » suivant :

Théorème : Lorsqu'une contrainte est relâchée dans un système isolé, celui-ci évolue spontanément vers un état d'équilibre macroscopique qui **maximise son entropie** réduite.

IV.C Échange de volume – Pression microcanonique

Considérons le problème de la **détente de Joule** (figure 3.15). Deux gaz occupent deux parties d'un récipient séparées par une paroi isolante (« adiabatique »), bloquée par un taquet. Le

¹⁵ Nous rencontrerons des situations où la température absolue peut devenir négative (si la dimension de l'espace des états est finie, l'entropie peut décroître en fonction de l'énergie pour les énergies les plus élevées : par exemple pour un ensemble de spins). Dans ce cas nous verrons que les températures absolues négatives sont plus « chaudes » que les températures absolues positives.

volume total est $V = V_1^{(i)} + V_2^{(i)}$. À un instant on retire le taquet pour permettre à la paroi de se déplacer librement. On permet également aux systèmes d'échanger de l'énergie, i.e. la paroi devient « diatherme ». Les énergies E_1 et E_2 ainsi que les volumes V_1 et V_2 deviennent alors des variables internes, avec les contraintes $E = E_1 + E_2$ et $V = V_1 + V_2$. L'analyse est analogue à celle de l'échange thermique, à la différence que l'on doit maintenant introduire une entropie réduite fonction de deux variables internes $\tilde{S}^*(E_1, V_1; E, V) = S_1^*(E_1, V_1) + S_2^*(E - E_1, V - V_1)$. La condition d'équilibre est

$$\frac{\partial S_1^*}{\partial E_1}(E_1^{\max}, V_1^{\max}) = \frac{\partial S_2^*}{\partial E_2}(E - E_1^{\max}, V - V_1^{\max}) \quad \text{et} \quad \frac{\partial S_1^*}{\partial V_1}(E_1^{\max}, V_1^{\max}) = \frac{\partial S_2^*}{\partial V_2}(E - E_1^{\max}, V - V_1^{\max}). \quad (3.83)$$

Nous savons par ailleurs que l'équilibre mécanique est un équilibre entre forces, i.e. entre pressions. Nous définissons donc la pression microcanonique en accord avec la définition de la thermodynamique, éq. (1.1) :

$$\boxed{\frac{p^*}{T^*} \stackrel{\text{def}}{=} \frac{\partial S^*}{\partial V}}. \quad (3.84)$$

Remarquons que, bien que nous ne l'ayons pas écrit, p^* comme T^* sont donc des fonctions des arguments de l'entropie microcanonique $S^*(E, V, N, \dots)$. La condition d'équilibre de la paroi prend finalement la forme

$$\boxed{T_1^* = T_2^* \quad \text{et} \quad p_1^* = p_2^* \quad (\text{condition d'équilibre})} \quad (3.85)$$

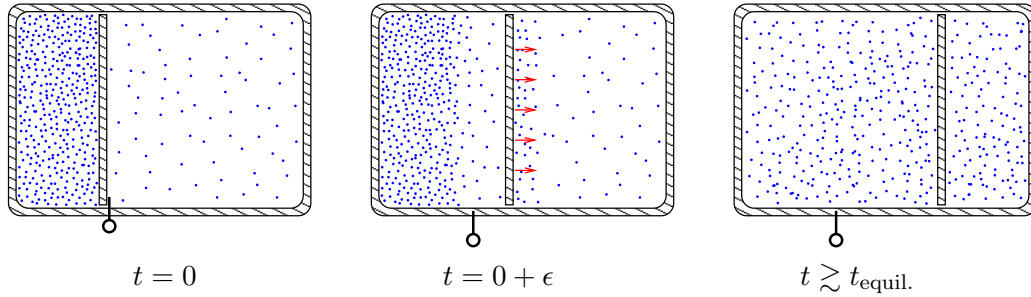


FIGURE 3.15 : **Détente de Joule.** À un instant initial, $t = 0$, on libère la paroi, puis on attend jusqu'à ce que l'équilibre (macroscopique) s'établisse.

On pourra vérifier que le sens de l'échange de volume est le suivant : le système ayant la pression la plus basse cède du volume à celui ayant la pression la plus haute (figure 3.15).

Notons que l'existence d'un équilibre est assurée par la propriété de convexité $\frac{\partial^2 S^*}{\partial V^2} < 0$, que nous pouvons encore interpréter $\frac{\partial p^*}{\partial V} < 0$.

Exemple : pression du gaz parfait – équation d'état. – Nous pouvons injecter la formule de Sackur-Tetrode dans la définition de la pression. Remarquons toutefois que l'expression complète de l'entropie n'est pas nécessaire. Il nous suffit de remarquer que, pour des particules non corrélées (gaz parfait classique) le nombre de microétats dépend du volume comme

$$\Omega \propto V^N \quad (3.86)$$

et par conséquent $S^* = Nk_B \ln V + f(E, N)$. Il s'ensuit immédiatement que $p^* = \frac{Nk_B T^*}{V}$, autrement dit nous avons retrouvé l'équation d'état du gaz parfait classique : la loi de Boyle-Mariotte ou loi *universelle* des gaz parfaits (classiques)

$$\boxed{p^* V = Nk_B T^*} \quad (3.87)$$

Le point remarquable est l'**universalité** de ce résultat, indépendant des détails microscopiques (nature du gaz) et même des détails du modèle. Nous verrons qu'elle s'applique également à des gaz parfaits polyatomiques. Ses limitations viendront de la présence des corrélations, soit sous l'effet de la mécanique quantique (gaz parfait quantique) soit sous l'effet des interactions (gaz réels). En utilisant l'expression de la température, nous pouvons finalement exprimer la pression en termes des paramètres naturels de l'ensemble microcanonique : nous obtenons une expression

$$p^*(E, V, N) = \frac{2}{3} \frac{E}{V} \quad (3.88)$$

indépendante du nombre de particules. Cette écriture de la pression est toutefois une forme *non universelle* (le $2/3$ dépend de la nature du gaz, ici monoatomique).

Remarque : Nous verrons que la relation (3.88) est plus générale qu'il n'y paraît : dans le problème (annexe 3.C page 77) nous montrons qu'elle est valable aussi bien dans le régime classique que dans le régime quantique, pour des bosons ou des fermions.

Remarque 2 – détente de Joule avec une paroi adiabatique : La détente de Joule avec une paroi adiabatique est plus délicate à discuter (cf. [6] ; exercice II.16 de [11]). Même si les deux systèmes ne peuvent pas échanger de *chaleur* à travers la paroi, ils peuvent échanger du *travail*, cependant les variations d'énergies et de volumes ne sont pas indépendantes. Montrons que l'équilibre est encore caractérisé par la condition d'égalité des pressions. Pour cela nous supposons que le système est dans son état d'équilibre et cherchons une propriété permettant de définir cet équilibre. Si l'on écarte le piston de section A d'une distance infinitésimale dx , la variation de l'énergie du gaz 1 est donnée par le travail des forces exercées sur le piston (par le gaz 2) : ¹⁶ $dE_1 = dW_1 = -F_{2 \rightarrow 1} dx = -p_2^* dV_1$ où $dV_1 = A dx$. De même $dE_2 = -p_1^* dV_2$. Les variations d'énergies et de volumes ne sont pas indépendantes. La variation d'entropie (réduite) est donc complètement contrôlée par $dV_1 = -dV_2$:

$$d\tilde{S}^* = dS_1^* + dS_2^* = \frac{1}{T_1^*} dE_1 + \frac{p_1^*}{T_1^*} dV_1 + \frac{1}{T_2^*} dE_2 + \frac{p_2^*}{T_2^*} dV_2 = \left(\frac{1}{T_1^*} + \frac{1}{T_2^*} \right) (p_1^* - p_2^*) dV_1. \quad (3.89)$$

L'équilibre correspond au maximum de l'entropie (réduite), i.e. $p_1^* = p_2^*$. QED.

Cependant, l'égalité des pressions ne suffit pas à déterminer complètement l'état d'équilibre. En effet, puisque de l'énergie a pu s'échanger, on doit déterminer quatre inconnues, E_1 , E_2 , V_1 et V_2 , alors que nous ne disposons que de trois équations : $E_1 + E_2 = E$, $V_1 + V_2 = V$ et $p_1^*(E_1, V_1) = p_2^*(E_2, V_2)$. L'état d'équilibre dépend des détails de la dynamique au cours de la transformation.

IV.D Échange de particules – Potentiel chimique microcanonique

Enfin, sans le détailler puisque la logique est encore similaire, nous pouvons considérer une situation où les deux systèmes échangent des particules. On relâche cette fois une contrainte sur le nombre de particules de chaque sous-système, le nombre total $N = N_1 + N_2$ restant conservé. En définissant le potentiel chimique microcanonique

$$\boxed{\frac{\mu^*}{T^*} \stackrel{\text{def}}{=} -\frac{\partial S^*}{\partial N}} \quad (3.90)$$

¹⁶ Rappelons que le **travail** reçu par un système est le travail des forces *extérieures*. Pour un gaz séparé de l'extérieur par un piston, le travail infinitésimal est $dW = -p^{\text{ext}} dV$ où p^{ext} est la pression extérieure. Si la transformation est réversible, i.e. le système passe par une succession d'états d'équilibre, il y a égalité entre pressions intérieure et extérieure $p = p^{\text{ext}}$ et le travail infinitésimal prend la forme $dW^{\text{rev}} = -p dV$.

puis en maximisant l'entropie réduite, on arrive à la condition d'équilibre

$$\frac{\mu_1^*}{T_1^*} = \frac{\mu_2^*}{T_2^*}. \quad (3.91)$$

Remarquons que si l'on échange des particules, on échange également de l'énergie, et donc l'équilibre est également caractérisé par l'égalité des températures [6]

$$\boxed{T_1^* = T_2^* \quad \text{et} \quad \mu_1^* = \mu_2^* \quad (\text{condition d'équilibre})} \quad (3.92)$$

Ces relations nous permettent de comprendre le sens physique du potentiel chimique : le « paramètre conjugué de N » joue le même rôle qu'une « pression », mais pour les populations de particules. On vérifie cette fois que les particules se déplacent du système ayant initialement le potentiel chimique le plus élevé, vers celui l'ayant le plus bas. Ceci se comprend bien en revenant à l'égalité fondamentale (1.1) : lorsque $S = \text{cste}$ et $V = \text{cste}$, on a $dE = \mu dN$, i.e. le potentiel chimique est alors l'énergie gagnée par particule : $N \rightarrow N + 1$ implique alors $E \rightarrow E + \mu$. Le flot de particules vers la région de potentiel chimique le plus bas permet donc de minimiser l'énergie totale.

Exemple : potentiel chimique du gaz parfait.— Nous calculons le potentiel chimique à partir de la formule de Sackur-Tetrode :

$$\mu^*(E/N, V/N) = k_B T^* \ln \left(\frac{N}{V} \left[\frac{3\pi \hbar^2 N}{mE} \right]^{3/2} \right). \quad (3.93)$$

Grandeur	Force conjuguée	Équilibre thermo.	Sens des échanges ^(*)
Énergie E	$\frac{1}{T^*} = \frac{\partial S^*}{\partial E}$	$T_1^* = T_2^*$	si $T_1^{*(i)} > T_2^{*(i)}$ $\mathcal{S}_1 \xrightarrow{\text{énergie}} \mathcal{S}_2$
Volume V	$\frac{p^*}{T^*} = \frac{\partial S^*}{\partial V}$	$T_1^* = T_2^* \ \& \ p_1^* = p_2^*$	si $p_1^{*(i)} > p_2^{*(i)}$ $\mathcal{S}_1 \xleftarrow{\text{volume}} \mathcal{S}_2$
# de particules N	$\frac{\mu^*}{T^*} = -\frac{\partial S^*}{\partial N}$	$T_1^* = T_2^* \ \& \ \mu_1^* = \mu_2^*$	si $\mu_1^{*(i)} > \mu_2^{*(i)}$ $\mathcal{S}_1 \xrightarrow{\text{particules}} \mathcal{S}_2$

TABLE 3.1 : ^(*) : les sens d'échange de volume et de particules sont donnés en supposant égalité des températures initiales $T_1^{*(i)} = T_2^{*(i)}$.

IV.E Cas général : distribution d'une variable interne quelconque

Nous généralisons et résumons la discussion précédente. Nous considérons une *variable interne* X , i.e. une grandeur macroscopique extensive, susceptible de *fluctuer* (par exemple $X \rightarrow E_1$, ou $X \rightarrow N_v$ de la figure 3.1, etc). Nous introduisons

$$\omega(x; E) \stackrel{\text{def}}{=} \{ \# \text{ de microétats accessibles tels que } X = x \} \quad (3.94)$$

La distribution de la variable aléatoire est alors donnée, d'après le postulat fondamental, par :

$$p^*(x; E) = \frac{\omega(x; E)}{\Omega(E)} = \exp \frac{1}{k_B} \left[\tilde{S}^*(x; E) - S^*(E) \right] \quad \text{où } \tilde{S}^*(x; E) \stackrel{\text{def}}{=} k_B \ln \omega(x; E) \quad (3.95)$$

où est l'entropie réduite. Nous cherchons maintenant la valeur $x^{\max}$ la plus probable : $\frac{\partial}{\partial x} p^*(x; E) = 0$ implique

$$\boxed{\frac{\partial \tilde{S}^*}{\partial x}(x^{\max}; E) = 0} \quad (3.96)$$

On peut répéter la conclusion tirée de l'étude du contact thermique : les propriétés d'extensivité de l'entropie nous permettent de montrer que les fluctuations relatives de la variable interne sont négligeables, $\sigma_X \propto \sqrt{N}$, comparativement à sa valeur moyenne, $x^{\max} \propto N$. À la limite thermodynamique $N \rightarrow \infty$, les variables extensives prennent une valeur certaine (on dit qu'elles sont « auto-moyennantes »). Ce point explique **l'opposition entre l'aléatoire à l'échelle microscopique et le déterminisme à l'échelle macroscopique** (tableau 1.1).

Finalement l'entropie du système est $S^*(E) \simeq \hat{S}^*(x^{\max}; E)$, à une correction sous extensive $\mathcal{O}(\ln N)$ près.

Ces propriétés sont résumées dans le tableau.

V Irréversibilité

Les différents exemples de relâchement de contraintes que nous avons analysés nous ont montré que l'évolution spontanée du système isolé s'accompagne d'une *augmentation* de l'entropie. C'est l'essence du second principe de la thermodynamique. Pour comprendre plus précisément l'origine de l'irréversibilité, nous considérons un autre cas précis : un récipient est séparé en deux volumes égaux par une paroi amovible. Un volume contient initialement $N_1^{(i)} = N$ molécules alors que l'autre est vide $N_2^{(i)} = 0$. À un instant on retire la paroi, les molécules se répartissent rapidement dans les deux volumes (figure 3.16). Puisqu'elles se déplacent dans tous les sens aléatoirement, le nombre de molécules N_1 est une variable interne (fluctuante), qui vaut en moyenne $N/2$. Une molécule ayant une probabilité $1/2$ de se trouver dans le volume de gauche, la distribution du nombre de molécules dans le volume de gauche est donnée par

$$p^*(N_1) = \frac{N!}{2^N N_1! (N - N_1)!} . \quad (3.97)$$

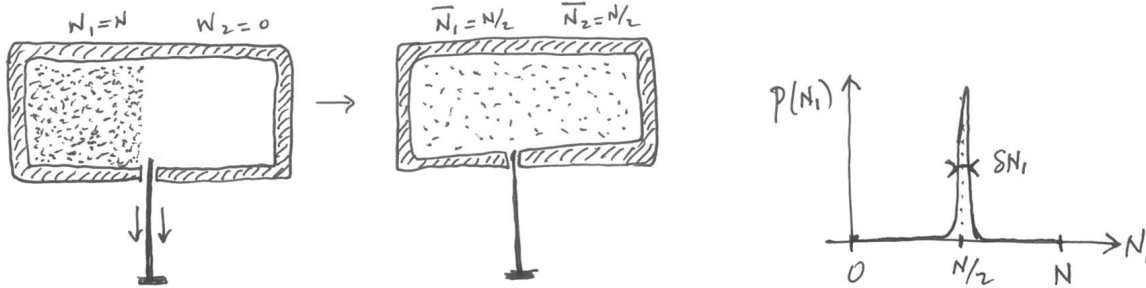


FIGURE 3.16 : **Irréversibilité.**

La configuration la plus probable (la valeur de N_1 qui maximise la distribution) est $N_1^{\max} = N/2$. Cependant, rien n'interdit *a priori* que toutes les molécules ne se remettent toutes dans le volume de gauche à un instant donné, et reviennent dans leur état initial. La probabilité d'un tel évènement est $p^*(N) = 2^{-N}$. Comparons cette probabilité avec la probabilité de la situation la plus probable : en utilisant la formule de Stirling (A.12), nous obtenons :

$$\frac{\min[p^*(N_1)]}{\max[p^*(N_1)]} = \frac{p^*(N)}{p^*(N/2)} = \frac{[(N/2)!]^2}{N!} \simeq \sqrt{\frac{\pi N}{2}} 2^{-N} \sim 2^{-N} = 10^{-N \frac{\ln 2}{\ln 10}} . \quad (3.98)$$

L'application numérique pour $v \sim 1\text{cm}^3$ de gaz parfait dans les conditions normales de température et de pression, $N \sim 3 \cdot 10^{20}$, donc

$$\frac{p^*(N)}{p^*(N/2)} \sim 10^{-10^{20}} = \underbrace{0.000 \dots 0001}_{10^{20} \text{ chiffres}} \quad (3.99)$$

Encore une fois rien n'interdit que la situation où $N_1 = N$ ne se produise, cependant, sa probabilité est si faible, relativement à la probabilité de l'état le plus probable, qu'on peut considérer qu'elle ne se réalise jamais en pratique.

L'irréversibilité correspond à l'évolution d'un système d'un état très peu probable vers un état beaucoup plus probable.

La flèche du temps : L'augmentation de l'entropie au cours de cette évolution est notre manière d'orienter la flèche du temps, i.e. notre définition du futur.

Rq : Retour sur le « principe d'entropie maximale »

Nous avons évoqué à plusieurs reprises qu'un principe pouvant être pris comme *le cœur de la physique statistique des systèmes à l'équilibre* est le « principe d'entropie maximale », permettant de construire le macroétat, i.e. de spécifier les probabilités d'occupation des microétats. Remarquons que les considérations sur l'irréversibilité ont bien illustré pourquoi le système choisi l'état correspondant à une entropie (réduite) maximale.

VI Quelques remarques sur le rôle de l'indiscernabilité

VI.A Validité du traitement semi-classique de l'indiscernabilité – cas du gaz parfait monoatomique

Nous discutons la validité du traitement semi-classique de l'indiscernabilité des particules dans l'étude du gaz parfait, i.e. de la formule de Sackur-Tetrode (3.63).

Nous avons obtenu la température du gaz en fonction de l'énergie. Cette relation peut être inversée pour déduire l'énergie en fonction de la température microcanonique

$$E = \frac{3}{2} N k_B T^*, \quad (3.100)$$

expression que nous retrouverons ultérieurement par un autre moyen. L'entropie (3.63) peut ainsi être exprimée en fonction de la température :

$$S^* = N k_B \ln \left(e^{5/2} \frac{V}{N} \left[\frac{m k_B T^*}{2\pi \hbar^2} \right]^{3/2} \right). \quad (3.101)$$

Cette expression fait mieux apparaître les limites de la formule de Sackur-Tétrode : elle diverge aux basses températures $S^* \rightarrow -\infty$ lorsque $T^* \rightarrow 0$, ce qui ne respecte pas le théorème de Nernst, parfois appelé 3ème principe de la thermodynamique,

$$\lim_{T \rightarrow 0} S^*(T) = 0. \quad (3.102)$$

Nous verrons plus tard que (3.63) et (3.101) ne sont valables que dans une limite *diluée* ou de relativement haute température, lorsque

$$\frac{V}{N} \left[\frac{m E}{3\pi \hbar^2 N} \right]^{3/2} = \frac{V}{N} \left[\frac{m k_B T^*}{2\pi \hbar^2} \right]^{3/2} \gg 1, \quad (3.103)$$

ce qui assure en gros $S^* \gtrsim 0$. Si cette condition n'est pas respectée, les effets du postulat de symétrisation, que nous avons traités de façon approchée sans distinguer les cas bosonique et fermionique, se font sentir fortement.

Pour mieux comprendre le sens de cette condition, nous écrivons (3.63) sous une autre forme en introduisant

$$\Delta x \stackrel{\text{def}}{=} \left(\frac{V}{N} \right)^{1/3} \quad \text{et} \quad \Delta p \stackrel{\text{def}}{=} \sqrt{2m \frac{E}{N}}. \quad (3.104)$$

$\Delta x^3 = V/N$ représente le volume disponible par particule dans l'espace physique, et Δx la distance typique entre deux particules voisines. Le fait que le **comptage des états des particules libres fasse intervenir le volume V/N , plutôt que le volume V qu'elles peuvent explorer, est une manifestation de l'indiscernabilité**, i.e. de l'invariance du microétat par permutation entre particules. D'autre part Δp représente l'impulsion typique d'une particule; puisque l'impulsion d'une particule fluctue, Δp^3 s'interprète comme le volume exploré par une particule dans l'espace des impulsions. On vérifie facilement que la formule de Sackur-Tetrode (3.63) devient

$$S^* = 3Nk_B \ln \left(a \frac{\Delta x \Delta p}{h} \right) \quad \text{avec} \quad a = \sqrt{6\pi} e^{5/6} = 9.9865... \simeq 10. \quad (3.105)$$

$\Omega = (a\Delta x \Delta p)^{3N} / h^{3N}$ est le nombre d'états quantiques accessibles (compte tenu de l'indiscernabilité). Pour que la règle de comptage semi-classique soit valable, il faut que le nombre d'états soit très grand $\Omega \gg 1$, ce qui exige que

$$a\Delta x \Delta p \gg h \quad (3.106)$$

C'est est une autre écriture de la condition (3.103). Lorsque cette condition n'est pas respectée, les effets du postulat de symétrisation deviennent importants. Examinons les conséquences pratiques.

Cas des bosons.— Dans le cas des bosons (chapitre 10), la condition (3.103), ou (3.106), est violée lorsqu'une fraction macroscopique des particules occupent l'état fondamental (condensation de Bose), ce qui se produit à très basse température (typiquement $T \lesssim \mu\text{K}$ dans les gaz dilués; en matière condensée, la superfluidité ou la supraconductivité s'apparentent à des phénomènes de condensation qui apparaissent pour $T \lesssim \text{K}$). Les effets quantiques liés à l'indiscernabilité des bosons se manifestent donc à très basse température.

Cas des fermions.— Dans le cas des fermions (chapitre 9), le principe de Pauli impose aux particules de s'empiler dans des états individuels distincts, ce qui confère une énergie énorme au système. Par exemple, si on considère les électrons d'un métal, la condition (3.103), ou (3.106), est respectée pour $T \gtrsim 10000\text{K}$. Autrement dit les électrons d'un métal ne sont jamais dans une limite diluée. C'est seulement lorsque le gaz de fermions est très dilué (gaz atomiques) qu'il sera justifié de considérer que le régime classique est atteint.

VI.B Entropie de mélange et paradoxe de Gibbs

Position du problème

À la fin du XIX^e siècle (période pré-quantique), la question de l'indiscernabilité des particules identiques n'était pas clarifiée, ce qui posait des problèmes relatifs aux propriétés d'extensivité de l'entropie du gaz parfait : la fonction $S_{\text{disc}}^* \simeq k_B \ln [\Phi_{\text{disc}}] = Nk_B \left[\frac{3}{2} + \ln \left(\frac{V}{N} \left[\frac{mE}{3\pi\hbar^2} \right]^{3/2} \right) \right]$ ne présente pas les bonnes propriétés d'extensivité. Gibbs avait proposé d'introduire un facteur $1/N!$ dans le comptage des microétats, de manière *ad hoc* pensait-on à l'époque, ce qui conduit à une expression $S^* \simeq k_B \ln [\Phi_{\text{indisc}}] = k_B \ln \left[\frac{1}{N!} \Phi_{\text{disc}} \right] = Nk_B \left[\frac{5}{2} + \ln \left(\frac{V}{N} \left[\frac{mE}{3\pi\hbar^2 N} \right]^{3/2} \right) \right]$ qui possède les bonnes propriétés d'extensivité.

Ces difficultés sont bien illustrées à travers le « paradoxe de Gibbs » que nous discutons maintenant. Considérons l'expérience suivante : deux gaz parfaits de natures différentes, par exemple un gaz d'hélium et un gaz d'argon, sont disposés dans deux volumes V_1 et V_2 ne communiquant pas. Les deux gaz ont initialement *mêmes températures et pressions* : $T_1^* = T_2^* \equiv T^*$ et $p_1^* = p_2^* \equiv p^*$. À un instant on retire la paroi pour permettre aux deux gaz de se mélanger (figure 3.17). La pression et la température du mélange est celle des gaz initialement séparés puisque $p^*V_1 = N_1k_B T^*$ et $p^*V_2 = N_2k_B T^*$ implique $p^*(V_1 + V_2) = (N_1 + N_2)k_B T^*$. Il est utile pour la suite d'écrire

$$\frac{V_1}{N_1} = \frac{V_2}{N_2} = \frac{V_1 + V_2}{N_1 + N_2} . \quad (3.107)$$

L'entropie a néanmoins augmentée, $\Delta S_{\text{mélange}} \stackrel{\text{def}}{=} S^{(\text{final})} - S^{(\text{initial})} > 0$, du fait de l'incertitude liée à la position des atomes des deux espèces dans l'état final, par rapport à l'état initial : il y a $\frac{(N_1 + N_2)!}{N_1!N_2!}$ manières de répartir aléatoirement les atomes des deux espèces dans les deux volumes, ce qui correspond donc à une augmentation de l'entropie du système

$$\Delta S_{\text{mélange}} = k_B \ln \left[\frac{(N_1 + N_2)!}{N_1!N_2!} \right] \simeq k_B [(N_1 + N_2) \ln(N_1 + N_2) - N_1 \ln N_1 - N_2 \ln N_2] . \quad (3.108)$$

Si l'on considère maintenant que les deux gaz sont de même nature, la présence de la paroi ne change rien aux propriétés du système et l'entropie ne devrait pas varier. Or il n'est à première vue pas évident de comprendre pourquoi le raisonnement qui a conduit à (3.108) ne s'applique pas (paradoxe de Gibbs).

Plutôt que d'utiliser des arguments combinatoires (physique statistique), on peut comprendre l'augmentation de l'entropie comme une conséquence du second principe de la thermodynamique. Lorsque les gaz sont de natures différentes, la transformation est *irréversible* et donc $\Delta S_{\text{mélange}} > 0$: si on remet la paroi, on ne revient pas dans l'état initial. En revanche si les gaz sont de même nature, la transformation est *réversible* : re-introduire la paroi nous place à nouveau dans la même situation qu'au départ de l'expérience. On attend alors $\Delta S_{\text{mélange}} = 0$.

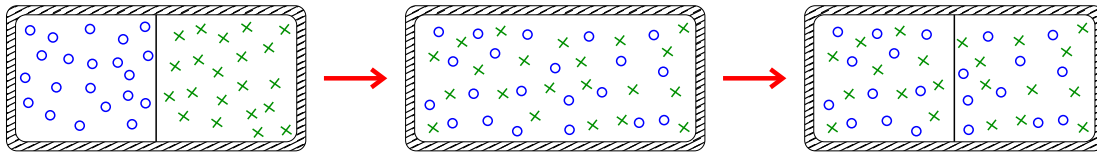


FIGURE 3.17 : **Entropie de mélange.** Deux gaz de natures différentes occupent deux volumes. À un certain instant on autorise les gaz à se mélanger (au centre). Si on réintroduit la paroi (figure de droite), on obtient une situation différente de la configuration initiale.

Résolution du paradoxe : indiscernabilité

La résolution du « paradoxe » se trouve dans l'indiscernabilité quantique des atomes du gaz. Si les atomes étaient discernables, alors il y aurait une perte d'information lorsque les volumes sont mis en communication. Parce que les atomes sont identiques, donc indiscernables, on ne peut pas plus les distinguer en présence ou en l'absence de la paroi, ce qui explique que l'entropie de mélange est nulle lorsque les deux gaz sont de même nature.

Nous vérifions cette assertion à partir de la formule de Sackur-Tetrode obtenue plus haut, qui intègre correctement l'effet de l'indiscernabilité. Puisque la température des gaz est inchangée avant et après la transformation, il est plus commode de considérer (3.101). Allons plus loin dans

la simplification : puisque l'entropie de mélange est purement liée à la configuration spatiale des gaz, il suffit de re-écrire la formule de Sackur-Tetrode comme :

$$S^* = Nk_B \ln(cV/N) . \quad (3.109)$$

où c est une constante (fonction de T^* et de la masse m des atomes).

Gaz de différentes natures Lorsque les gaz sont de natures différentes, la variation d'entropie est donc donnée en utilisant (3.109) (écrite pour chaque gaz) :

$$\Delta S_{\text{mélange}} = \overbrace{N_1 k_B \ln \left[c_1 \frac{V_1 + V_2}{N_1} \right] + N_2 k_B \ln \left[c_2 \frac{V_1 + V_2}{N_2} \right]}^{\text{final : 2 gaz dans } V_1 + V_2} - \overbrace{N_1 k_B \ln \left[c_1 \frac{V_1}{N_1} \right] - N_2 k_B \ln \left[c_2 \frac{V_1}{N_2} \right]}^{\text{initial}} \quad (3.110)$$

$$= N_1 k_B \ln \left[\frac{N_1 + N_2}{N_1} \right] + N_2 k_B \ln \left[\frac{N_1 + N_2}{N_2} \right] \quad (3.111)$$

où nous avons utilisé (3.107). Cette dernière expression correspond précisément à (3.108).

Gaz de même nature Lorsque les gaz sont de même nature, le calcul de l'entropie $S^{(\text{initial})}$ est le même que précédemment, en faisant $c_1 = c_2 = c$. En revanche $S^{(\text{final})}$ n'est pas la somme des entropies de deux gaz occupant le volume $V_1 + V_2$ comme précédemment, mais l'entropie d'un gaz de $N_1 + N_2$ particules dans ce volume :

$$\Delta S_{\text{mélange}} = \overbrace{(N_1 + N_2) k_B \ln \left[c \frac{V_1 + V_2}{N_1 + N_2} \right]}^{\text{final : 1 gaz dans } V_1 + V_2} - \overbrace{N_1 k_B \ln \left[c \frac{V_1}{N_1} \right] - N_2 k_B \ln \left[c \frac{V_1}{N_2} \right]}^{\text{initial}} = 0 \quad (3.112)$$

en vertu de (3.107), ce qui montre qu'il n'y a pas de paradoxe.

Conclusion

Le paradoxe de Gibbs (1875), i.e. la nécessité du facteur $1/N!$ dans $\Omega_{\text{indisc}}(E)$ dans la limite diluée, permettant de restaurer les propriétés d'extensivité de l'entropie, est la manifestation à l'échelle macroscopique de l'énorme réduction de la dimension de l'espace de Hilbert d'un nombre macroscopique de particules indiscernables par rapport à des particules discernables, i.e. conséquence du postulat de symétrisation de la mécanique quantique.

☺ Les idées importantes

- Description classique et quantique des microétats et macroétats.
 - Notion d'équilibre macroscopique.
 - Évidemment le **postulat fondamental**, énoncé et distribution microcanonique (3.55).
 - Entropie microcanonique (3.56).
 - Notion de densité d'états. Important : calcul de densité d'états à l'approximation semi-classique, équation (3.42).
 - Indiscernabilité, règle (3.52).
 - Calcul de l'entropie à l'aide de la densité d'états à l'approximation semi-classique ; on pourra se rappeler que si le nombre de particules est macroscopique, il est plus facile d'écrire : $S^* = k_B \ln [\rho(E)\delta E] \simeq k_B \ln [\Phi(E)]$
 - Pour retrouver les définitions (température, pression,...) on se rappellera de la correspondance à la thermodynamique dictée par la relation fondamentale $dE = TdS - pdV + \mu dN \Rightarrow dS = \frac{1}{T}dE + \frac{p}{T}dV - \frac{\mu}{T}dN$.
 - Les problèmes de relâchement de contraintes, le « théorème » sur l'augmentation de l'entropie.
 - La notion de limite thermodynamique.
 - Un exercice d'application important : le gaz parfait monoatomique.
- Un autre exercice assez simple et instructif : le cristal de spin 1/2.

🔍 Pour en savoir plus

- Sur la présentation classique (début du chapitre) : le début de l'ouvrage [7].
- Sur la limite semi-classique, consulter le chapitre 2 du livre [3].
- Sur le postulat de symétrisation, chapitre 11 de [34].
- Sur l'introduction de la notion d'entropie dans le cadre de la *thermodynamique*, cf. [25].
- Sur l'origine de la « définition » de la pression microcanonique $p^* = T^* \frac{\partial S^*}{\partial V}$, on pourra consulter le complément II.D de [11].
- Je recommande l'excellente discussion sur l'entropie du chapitre 5 de Sethna [32].

Annexe A

Formulaire

Fonction Gamma d'Euler

Pour $\operatorname{Re} z > 0$ on a la représentation de la fonction Gamma :

$$\Gamma(z) \stackrel{\text{def}}{=} \int_0^\infty dt t^{z-1} e^{-t}. \quad (\text{A.1})$$

La relation fonctionnelle

$$\Gamma(z+1) = z \Gamma(z) \quad (\text{A.2})$$

permet d'étendre la définition de la fonction dans la seconde moitié du demi-plan complexe de z . Deux valeurs particulières :

$$\Gamma(1) = 1 \quad \text{et} \quad \Gamma(1/2) = \sqrt{\pi} \quad (\text{A.3})$$

d'où $\Gamma(n+1) = n!$ et $\Gamma(n + \frac{1}{2}) = \frac{\sqrt{\pi}}{2^n} (2n-1)!!$.

Fonction Beta d'Euler

$$B(\mu, \nu) = \int_0^1 dt t^{\mu-1} (1-t)^{\nu-1} = 2 \int_0^{\pi/2} d\theta \sin^{2\mu-1} \theta \cos^{2\nu-1} \theta = \frac{\Gamma(\mu)\Gamma(\nu)}{\Gamma(\mu+\nu)} \quad (\text{A.4})$$

Fonction Zeta de Riemann

Une représentation intégrale

$$\zeta(\alpha) = \frac{1}{\Gamma(\alpha)} \int_0^\infty dt \frac{t^{\alpha-1}}{e^t - 1} \quad \text{valable pour } \operatorname{Re} \alpha > 1 \quad (\text{A.5})$$

(on peut étendre le domaine de définition de $\zeta(\alpha)$ dans le plan complexe de la variable α en utilisant des relations fonctionnelles). Par développement de la fraction en série géométrique, on montre que

$$\zeta(\alpha) = \sum_{n=1}^\infty n^{-\alpha} \quad \text{pour } \operatorname{Re} \alpha > 1. \quad (\text{A.6})$$

Quelques valeurs particulières utiles : $\zeta(1) = \infty$, $\zeta(3/2) \simeq 2.612$, $\zeta(2) = \pi^2/6$, $\zeta(5/2) \simeq 1.341$, $\zeta(3) \simeq 1.202$, $\zeta(4) = \pi^4/90$, etc.

D'autres représentations intégrales utiles sont

$$(1 - 2^{1-\alpha})\Gamma(\alpha)\zeta(\alpha) = \int_0^\infty dt \frac{t^{\alpha-1}}{e^t + 1} \quad \text{valable pour } \operatorname{Re} \alpha > 1 \quad (\text{A.7})$$

Intégrales gaussiennes

$$\int_{-\infty}^{+\infty} dx e^{-x^2} = \sqrt{\pi} \quad (\text{A.8})$$

Plus généralement

$$\int_0^{+\infty} dx x^n e^{-\frac{1}{2}ax^2} = \frac{1}{2} \left(\frac{2}{a}\right)^{\frac{n+1}{2}} \Gamma\left(\frac{n+1}{2}\right) \quad (\text{A.9})$$

Méthode du col

Soit l'intégrale

$$I(\lambda) = \int_a^b dx e^{-\lambda\varphi(x)} \quad (\text{A.10})$$

où la fonction $\varphi(x)$ possède un unique minimum (absolu) en $x_* \in]a, b[$, « loin » des bords ; on suppose de plus que $\varphi''(x_*) \neq 0$. Dans la limite $\lambda \rightarrow \infty$ l'intégrale est dominée par les contributions du voisinage du minimum :

$$I(\lambda) \underset{\lambda \rightarrow \infty}{\simeq} \sqrt{\frac{2\pi}{\lambda\varphi''(x_*)}} e^{-\lambda\varphi(x_*)} \quad (\text{A.11})$$

(les restrictions sont faciles à lever ; on pourra facilement obtenir des généralisations de la formule précédente lorsque les hypothèses ne sont pas satisfaites).

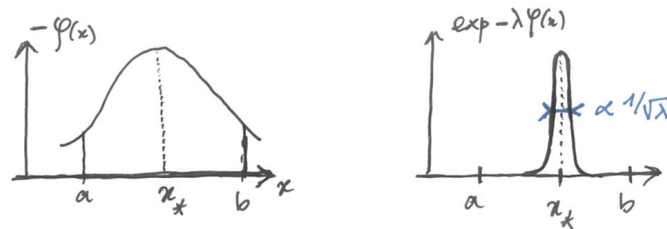


FIGURE A.1 : Méthode du col.

Formule de Stirling On peut appliquer la formule précédente à la fonction Gamma en l'écrivant $\Gamma(z+1) = z^{z+1} \int_0^\infty dt e^{-z\varphi(t)}$ avec $\varphi(t) = t - \ln t$. On obtient la formule de Stirling

$$\Gamma(z+1) \underset{z \rightarrow \infty}{\simeq} \sqrt{2\pi} z^{z+\frac{1}{2}} e^{-z} \quad (\text{A.12})$$

qu'on utilisera souvent sous la forme

$$\ln \Gamma(z+1) \underset{z \rightarrow \infty}{=} z \ln z - z + \frac{1}{2} \ln(2\pi z) + \mathcal{O}(1/z) \quad (\text{A.13})$$

Remarquons qu'il s'agit des premiers termes d'une série asymptotique (i.e. si on poursuivait le développement en complétant par des termes c_n/z^n pour $n > 1$, la série $\sum_{n=1}^\infty c_n/z^n$ ne converge pas).

Distribution de Dirac

Rappelons la formule définissant l'action de la distribution de Dirac. Soit $f(x)$ une fonction continue à l'origine :

$$\int dx f(x) \delta(x) = f(0) , \quad (\text{A.14})$$

où l'intégrale inclut l'origine. Autrement dit $\delta(x)$ est une fonction de « poids $\int_{\mathbb{R}} dx \delta(x) = 1$ » concentré à l'origine.

La transformée de Fourier de la distribution de Dirac est une constante (δ est la superposition d'ondes planes à toutes les fréquences) i.e.

$$\delta(x) = \int_{-\infty}^{+\infty} \frac{dk}{2\pi} e^{ikx} . \quad (\text{A.15})$$

Une petite propriété utile : soit $\lambda \in \mathbb{R}^*$, on a $\delta(\lambda x) = \frac{1}{|\lambda|} \delta(x)$. La propriété montre que la distribution a la dimension physique inverse de son argument $[\delta(x)] = 1/[x]$.

Soit $f(x)$ une fonction ayant un ensemble de racines *simples* notées x_n . On a

$$\delta(f(x)) = \sum_{\substack{x_n \text{ tq} \\ f(x_n) = 0}} \frac{\delta(x - x_n)}{|f'(x_n)|} = \frac{1}{|f'(x)|} \sum_{\substack{x_n \text{ tq} \\ f(x_n) = 0}} \delta(x - x_n) . \quad (\text{A.16})$$

Exemple 1 : $\delta(x^2 - x_0^2) = \frac{1}{2x_0} [\delta(x - x_0) + \delta(x + x_0)]$.

Exemple 2 : $\delta(\sin x) = \sum_{n \in \mathbb{Z}} \delta(x - n\pi)$ est un peigne de Dirac.

Bibliographie

- [1] P. W. Anderson, More is different, *Science* **177**(n°4047), 393 (4 August 1972).
- [2] N. W. Ashcroft and N. D. Mermin, *Solid State Physics*, Saunders College (1976).
- [3] R. Balian, *From microphysics to macrophysics. Volume I*, Springer, Berlin (1992).
- [4] R. Balian, *From microphysics to macrophysics. Volume II*, Springer, Berlin (1992).
- [5] J.-L. Basdevant and J. Dalibard, *Mécanique quantique*, Les éditions de l'École polytechnique, Palaiseau (2004).
- [6] H. B. Callen, *Thermodynamics and an introduction to thermostatistics*, John Wiley & Sons, New York (1985), 2nd edition.
- [7] R. Castaing, *Thermodynamique statistique*, Masson, Paris (1970).
- [8] C. Cohen-Tannoudji, B. Diu and F. Laloë, *Mécanique Quantique*, Hermann, Paris (1973), tomes 1 et 2.
- [9] C. Cohen-Tannoudji, J. Dupont-Roc and G. Grynberg, *Photons et atomes : Introduction à l'électrodynamique quantique*, Interéditions/Éditions du CNRS (1987).
- [10] L. Couture, C. Chahine and R. Zitoun, *Thermodynamique - cours, exercices et problèmes résolus*, Dunod (1993).
- [11] B. Diu, C. Guthman, D. Lederer and B. Roulet, *Physique Statistique*, Hermann, Paris (1989).
- [12] A. Georges and M. Mézard, *Physique Statistique*, édité par l'École Polytechnique, Palaiseau (2008), ISBN 978-2-7302-1484-1.
- [13] K. Huang, *Statistical mechanics*, John Wiley & Sons, New York (1963).
- [14] E. T. Jaynes, Information Theory and Statistical Mechanics, *Phys. Rev.* **106**, 620–630 (1957).
- [15] M. Kardar, *Statistical physics of fields*, Cambridge University Press (2007).
- [16] M. Kardar, *Statistical physics of particles*, Cambridge University Press (2007).
- [17] C. Kittel, *Physique de l'état solide*, Dunod (1994), 5ème édition.
- [18] W. Krauth, *Statistical Mechanics : Algorithms and Computations*, Oxford University Press (2006).
- [19] H. Krivine and J. Treiner, *La physique statistique en exercices*, Vuibert, Paris (2003).

- [20] R. Kubo, M. Toda and N. Hashitsume, *Statistical physics II : non-equilibrium statistical mechanics*, Springer-Verlag (1992).
- [21] L. D. Landau and E. Lifchitz, *Physique statistique*, Mir (1966), tome 5.
- [22] R. B. Laughlin, *Un univers différent*, Fayard (2005).
- [23] M. Le Bellac, *Physique quantique*, EDP Sciences – CNRS Éditions (2003).
- [24] J.-M. Lévy-Leblond and F. Balibar, *Quantique – Rudiments*, CNRS-InterÉditions, Paris (1984).
- [25] S.-K. Ma, *Statistical mechanics*, World Scientific, Singapore (1985).
- [26] K. Mallick, Some recent developments in non-equilibrium statistical physics, *Pramana J. Phys.* **73**(3), 417–451 (2009), available at <http://www.ias.ac.in/pramana/v73/p417/fulltext.pdf>.
- [27] C. Ngô and H. Ngô, *Physique Statistique*, Dunod, Paris (2008), 3ème édition.
- [28] R. K. Pathria and P. D. Beale, *Statistical mechanics*, Academic Press, Elsevier (2011), 3rd edition.
- [29] C. J. Pethick and H. Smith, *Bose-Einstein condensation in dilute gases*, Cambridge University Press (2002).
- [30] F. Reif, *Fundamentals of statistical and thermal physics*, McGraw-Hill, New York (1965).
- [31] F. Reif, *Physique statistique*, Armand Colin, Paris (1972), Berkeley : cours de physique, volume 5.
- [32] J. P. Sethna, *Statistical Mechanics : Entropy, order parameter and complexity*, Oxford University Press (2006).
- [33] C. Texier, *Physique statistique (faiblement) hors équilibre : formalisme de la réponse linéaire. Application à l'étude de la dissipation quantique et du transport életronique* (2013), notes de cours du DEA de physique quantique, École Normale Supérieure, disponible à l'adresse http://www.lptms.u-psud.fr/christophe_texier/.
- [34] C. Texier, *Mécanique quantique*, Dunod, Paris, second edition (2015).
- [35] M. Toda, R. Kubo and N. Saitô, *Statistical physics I : equilibrium statistical mechanics*, Springer-Verlag (1992).

Index

A

Agitation thermique 38
 Aléatoire (variable) 18
 Argon 107
 Avogadro (nombre d') 15

B

Bernouilli, Daniel 11
 Binomiale (distribution) 25
 Bit 16
 Blocage de Pauli 188
 Boltzmann, Ludwig 11
 Born-Oppenheimer (approximation de) ... 108
 Bose-Einstein (distribution de) 119, 173
 Boson 48, 75
 Brisure spontanée de symétrie 154, 159

C

Capacité calorifique (voir Chaleur spécifique) 3
 Casimir (pression) 132
 Chaleur 93
 Chaleur latente 148
 de condensation de Bose-Einstein 200
 Chaleur spécifique 3, 62, **84**
 gaz diatomique 3, 110, 111
 gaz parfait monoatomique 106
 Champ moyen (approximation de) 152
 Chaos 17
 Clausius, Rudolf 11, 93
 COBE 130
 Col (méthode du) 32, 214
 Collision 107
 Compressibilité 3, 7, 87
 Conjugué
 paramètre conjugué 85, 87
 variable canoniquement conjugué 39
 Constante de Boltzmann 29
 Contact thermique 60
 Corps noir 128
 Corrélations 21
 Covariances (matrice des) 22
 Crochets de Poisson 40, 44
 Cumulants 23, 84

Cumulants (définition) 24

D

De Broglie (longueur thermique de) ... 97, 194
 Debye (modèle de) 127
 Debye, Petrus 12
 Dégénérescence 57
 Densité de modes 122
 Détente de Joule 64
 Développement du viriel 145, 171
 Diatherme 60
 Diffusion
 des molécules dans un gaz 107
 Dilatation (coefficient de) 3
 Distribution de Bose-Einstein 119
 Distribution de probabilité 19
 Dulong et Petit (loi de) 124

E

Échange (interaction d') 151
 Einstein, Albert 12
 Élémentaire 2
 Émergence **6**, 27, 123, 192
 Énergie de Fermi 183
 Énergie libre **82**
 Énergie libre (interprétation thermo.) 82
 Enthalpie 3
 Enthalpie libre (fonction de Gibbs) 147
 Entropie 11
 canonique 84
 de Rényi 29, 102
 extensivité 58, 59, 70
 microcanonique 57
 principe d'entropie maximale ... 46, 57, 69
 réduite 61, 63–65, 67, 69
 Sackur-Tetrode (formule de) .. 59, 106, 200
 validité 69
 statistique (de von Neumann) **29**
 Équilibre macroscopique 38
 Espérance 20
 État pur 28
 Étoiles
 à neutrons 180

- Naines blanches 180
 Évènement 17
 Exposants critiques 160
 Extensivité **58**, 70, 83
- F**
 Facteurs d'occupation 55, 172
 Fermi (énergie de) 183
 Fermi-Dirac (distribution de) 174
 Fermion 48, 75
 Fluctuation-dissipation 84, 106
 Fonction de partition 81
 Fonction génératrice 24, 81, 167
 Fond cosmologique 130
 Fugacité 167, 194
- G**
 Gap 192
 Gaussienne (distribution) 25
 Gaz parfait 11, 16, 26
 monoatomique 59, 69
 potentiel chimique 67
 pression 65
 Sackur-Tetrode (formule de) **59**, 69
 validité 69
 Gibbs (fonction de) 147
 Gibbs, J. Willard 11, 71
 Gibbs-Duhem (relation de) 101
 Grandes déviations **33**
 Gumbel (loi de) 23
- H**
 Heisenberg (modèle de) 151
 Herapath, John 11
 hétéronucléaires (molécules) 113
 homonucléaires (molécules) 113
- I**
 Indiscernabilité 48, **53**, 70, 71, **75**, 95
 Intrication 30
 Irréversibilité 68
 Ising (modèle d') 12, **151**
- J**
 Joule, James P. 11
- K**
 Kelvin (William Thomson) 62
 Kronecker (symbole de) 18
- L**
 Landau, Lev D. 12
- Lennard-Jones (potentiel de) 142
 Libre parcours moyen 107
 Limite centrale (théorème de la) **26**, 33, 44, 46
 Limite thermodynamique **63**, **92**, 100
 Longueur thermique 97, 194
- M**
 Macroétat 28
 Magnon 192
 Maxwell (relations de) 101
 Maxwell, John C. 11
 Maxwell-Boltzmann (approximation de) **56**, **96**
 Mayer (relation de) 4, 106
 Mélange statistique 28
 Mermin-Wagner (théorème) 135
 Méthode du col 32, 214
 Microétat 28
 Modes propres 135
 Moments (définition) 23
 Mouvement brownien 19
 Moyenne (définition) 20
- N**
 Nernst (théorème de) 69, 94, 188
- O**
 Octet 16
 Onsager, Lars 12
 Opalescence critique 7
 Opérateur densité 49
 Orho-hydrogène 113
- P**
 Para-hydrogène 113
 Paradoxe de Gibbs 11, 71
 Paramètre conjugué 85
 Paramètre d'ordre 153, 159
 Pauli (blocage de) 188
 Perrin, Jean 19
 Phase stationnaire (approximation de la) ... 32
 Phénomènes collectifs 2
 Phonons 123
 Photons 129
 Planck (loi de) 130
 Planck, Max 11
 Poincaré (temps de) 46
 Point critique (Van der Waals) 146, 148
 Poisson (crochets de) 40, 44
 Pomeranchuk (effet) 180
 Postulat de symétrisation 48, **75**
 Postulat fondamental 57

- Postulat fondamental de la thermodynamique 57
- Potentiel chimique
 canonique.....85, 87
 microcanonique.....66
- Potentiel chimique (nullité du).....192
- Pression
 canonique.....85, 87
 Casimir.....132
 de radiation.....132
 microcanonique.....65
- Principe de Pauli.....55
- Probabilités.....15, 17
- Q**
- Quantique (mécanique).....10, 17
- R**
- Radiation (pression de).....132
- Réductionnisme.....5
- Relation de Gibbs-Duhem.....101
- Relations de Maxwell.....101
- Rényi (entropie de).....29, 102
- Réversible (transformation).....66, **85**, 92
- S**
- Sackur-Tetrode (formule de)..**59**, 59, 106, 200
 validité.....69
- Section efficace.....107
- Séparable.....96
- Sinai (problème de).....41
- Sphères dures.....41
- Spin.....48
- Stefan-Boltzmann (loi de).....131
- Stirling (formule de).....214
- Superfluidité.....98
- T**
- Température de Fermi.....183
- Température
 absolue.....62
 canonique.....81
 microcanonique.....62
- Théorème d'équipartition.....95, 106
- Théorème de Feynman-Hellmann.....85
- Théorème de la limite centrale..**26**, 33, 44, 46
- Théorème de Liouville.....39, 40, 44
- Théorème de Mermin-Wagner.....135
- Théorème de Nernst.....69, 94, 188
- Théorème H11
- Théorie ergodique.....47
- Thermodynamique.....3
 1er principe.....4, 11
 2ème principe.....5, 11
 3ème principe.....69, 94, 188
 limite thermodynamique.....**63**, **92**, 100
 postulat fondamental.....4
 relation fondamentale.....4
- Thomson, William (Lord Kelvin).....62
- Transition de phase
 du 1er ordre.....7
 du 2nd ordre.....7
 liquide-gaz.....146
 para-ferro.....150
- Travail.....66, 93
- Tsallis, Constantino.....102
- U**
- Universalité.....6, 7
- V**
- Van der Waals.....**144**
- Variable conjugué.....85
- Variable externe.....60
- Variable interne.....60
- Variance.....21
- Viriel (développement du).....145
- Viriel (developpement du).....145, 171
- von Neumann, John.....29
- W**
- Weyl (terme de).....50
- Z**
- Zone de Brillouin.....135