

PHYSIQUE STATISTIQUE :
des processus élémentaires aux phénomènes collectifs

Christophe Texier

Laboratoire de Physique Théorique et Modèles Statistiques

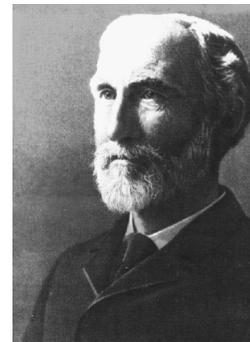
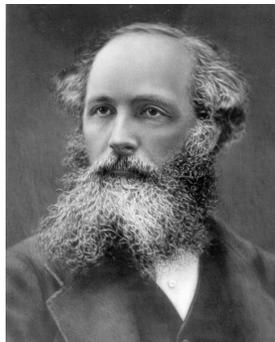


Table des matières

1	Introduction	1
1.1	Le périmètre de la physique statistique	1
1.1.1	À quoi sert la physique statistique ?	1
1.1.2	Processus élémentaires et phénomènes collectifs	2
1.1.3	Fin du réductionnisme et émergence	4
1.2	L’approche probabiliste	8
1.2.1	Nécessité de l’approche probabiliste	8
1.2.2	Quelques éléments de théorie des probabilités	10
1.2.3	Microétats, macroétats, information et entropie statistique	18
1.3	Conseils bibliographiques	23
	Annexe 1.A : Chronologie – D’où vient la physique statistique ?	24
2	Postulat fondamental – Ensemble microcanonique	27
2.1	Description classique : espace des phases	28
2.1.1	Définition de l’espace des phases	28
2.1.2	Pourquoi l’espace des phases ? Théorème de Liouville	28
2.1.3	Ergodicité	30
2.1.4	Formulation classique du postulat fondamental de la physique statistique	31
2.2	Description quantique : espace de Hilbert	34
2.2.1	Postulats de la mécanique quantique	34
2.2.2	Microétats et macroétats, moyennes	35
2.2.3	Densité d’états, limite semi-classique	36
2.2.4	Densité d’états pour des particules indiscernables	39
2.3	Postulat fondamental	41
2.3.1	La distribution microcanonique	41
2.3.2	Entropie microcanonique	42
2.4	Relachement de contraintes	43
2.4.1	Variables externes et variables internes	43
2.4.2	Contact thermique entre deux systèmes – Température microcanonique	43
2.4.3	Généralisation : distribution d’une variable interne	47
2.4.4	Application 1 : échange de volume, pression microcanonique	47
2.4.5	Application 2 : échange de particules, potentiel chimique	48
2.4.6	Irréversibilité	48
2.5	Illustration : gaz parfait monoatomique	50
2.5.1	Thermodynamique du gaz parfait monoatomique	50
2.5.2	Validité du traitement semi-classique de l’indiscernabilité	51
2.5.3	Ordres de grandeur dans un gaz d’Argon	52
2.5.4	Entropie de mélange et paradoxe de Gibbs	53
	Annexe 2.A : La méthode des multiplicateurs de Lagrange	56

3	Système en contact avec un thermostat – Ensemble canonique	57
3.1	Ensemble canonique	57
3.1.1	Distribution canonique	57
3.1.2	Propriétés de l'ensemble canonique	59
3.2	Relachement d'une contrainte	62
3.3	Ensembles microcanonique et canonique	64
3.3.1	Distribution de l'énergie du système	64
3.3.2	Limite thermodynamique	65
3.3.3	La compétition énergie-entropie	65
3.4	Retrouver les « principes » de la thermodynamique	66
3.5	Formulation classique	67
3.5.1	Distribution canonique classique et fonction de partition	67
3.5.2	Distribution de Maxwell des vitesses	68
3.5.3	Théorème d'équipartition de l'énergie	68
3.6	Indiscernabilité et approximation de Maxwell-Boltzmann	69
	Annexe 3.A : Transformations de Legendre et potentiels thermodynamiques	70
4	Application 1 : gaz classique	71
4.1	Gaz parfait monoatomique	71
4.2	Gaz parfaits moléculaires	72
4.3	Gaz réels	72
5	Application 2 : thermodynamique d'un ensemble d'oscillateurs harmoniques	73
5.1	Vibrations d'un corps solide	73
5.2	Thermodynamique d'une ligne de transmission électrique	73
5.3	Thermodynamique du rayonnement	73
6	Application 3 : transitions de phase	75
6.1	La transition de phase liquide-Gaz	75
6.2	Ferromagnétisme	75
7	Système en contact avec un thermostat réservoir de particules – Ensemble grand canonique	77
7.1	Distribution grand canonique	77
7.1.1	Particules sans interaction : décomposition sur les états individuels	78
8	Statistique de Fermi-Dirac	81
8.1	Le gaz de fermions libres	81
9	Statistique de Bose-Einstein	83
9.1	Phonons et photons	83
9.2	Condensation de Bose-Einstein	83
10	Aspects cinétiques	85
10.1	Théorie cinétique	85
10.2	Équation maîtresse	85
A	Formulaire	87
	Bibliographie	88
	Index	91

Chapitre 1

Introduction

1.1 Le périmètre de la physique statistique

1.1.1 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 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 satisfabilité de contraintes,...
11. étude des marchés financiers
12. etc.

1.1.2 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*.

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 (la physique statistique) repose sur une formulation de nature probabiliste était inconcevable (d'un point de vue philosophique) pour de nombreux scientifiques [3].

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 :

<u>Microscopique</u>	⇒	Physique statistique	⇒	<u>Macroscopique</u>
<ul style="list-style-type: none"> • discret • petit nombres de lois simples (postulats de la MQ, équations de Maxwell,...) • description probabiliste (mécanique quantique) • équations réversibles 				<ul style="list-style-type: none"> • continu • diversité, complexité, grand nombre de théories effectives • déterministe • irréversibilité

TABLE 1.1 – L’opposition micro-macro.

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 sur 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.

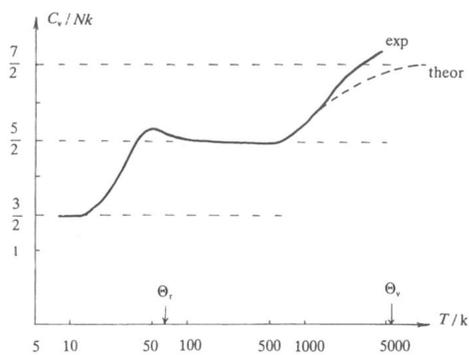


Fig. 8.3. Specific heat of deuterated hydrogen, HD

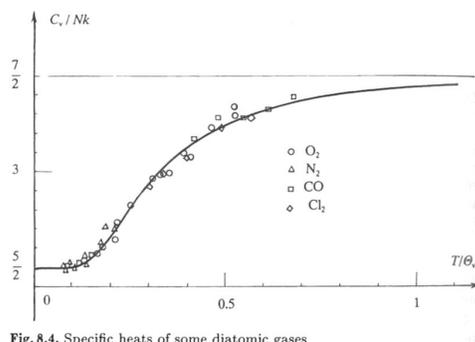


Fig. 8.4. Specific heats of some diatomic gases

FIGURE 1.1 – **Chaleur spécifique de divers gaz diatomiques.** *Figure tirée du livre de R. Balian [3].*

Exemple 2 : Coefficients calorimétriques. – Il est bien connu que les coefficients calorimétriques jouent un rôle très important en thermodynamique (capacités calorifiques $C_V = (\frac{\partial E}{\partial T})_V$ et $C_p = (\frac{\partial E}{\partial T})_p$, compressibilité $\kappa_T = -\frac{1}{V}(\frac{\partial V}{\partial p})_T$, coefficient de dilatation $\alpha = \frac{1}{V}(\frac{\partial V}{\partial T})_p$, etc). La thermodynamique 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 deux chapitres).

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 rappelons 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)¹

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

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 terme 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.

- 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 (théorème H de Boltzmann).

1.1.3 Fin du réductionnisme et émergence

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).

¹La notation dQ désigne une quantité infinitésimale qui n'est pas la différentielle d'une fonction.

2. Ces ingrédients microscopiques seront utilisés, avec des considérations probabilistes, pour construire la théorie physique décrivant la dynamique à l'échelle macroscopique.

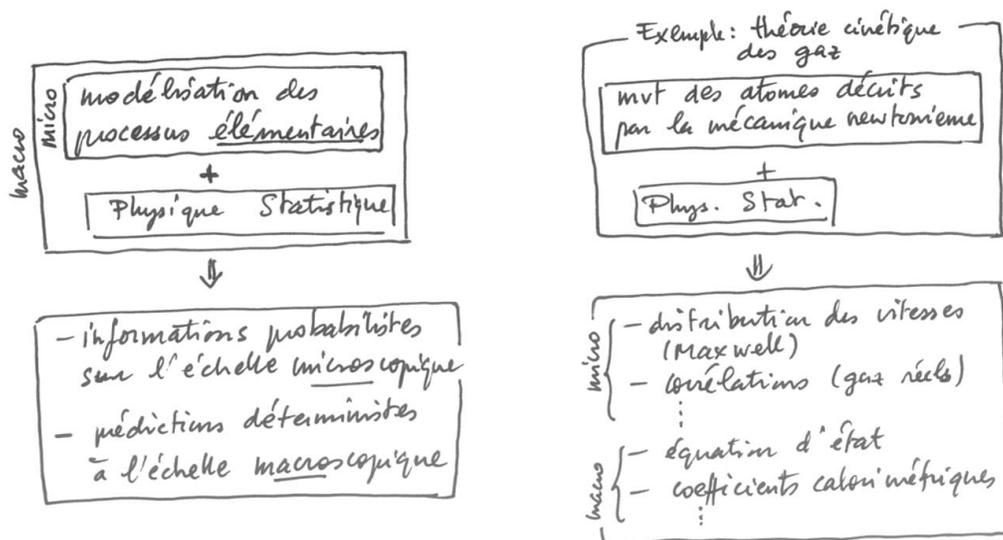


FIGURE 1.2 – La physique statistique est une sur-couche au dessus d'une description microscopique.

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 schématiser : la compréhension de la dynamique des quarks me permet de remonter à celle du noyau, de celle-ci je déduis le comportement des atomes, puis celui des gaz et des solides, etc. Il semble que je puisse ainsi aller jusqu'aux étoiles, puis considérer l'univers dans son ensemble ! Conclusion (erronée) : le comportement de l'univers est contenu dans la dynamique des quarks.

La **physique statistique pulvérise le mythe réductionniste** : 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 d'une partie de l'information de l'échelle microscopique, et une certaine **universalité** de nombreuses propriétés à l'échelle macroscopique [1, 20].

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**, 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) !

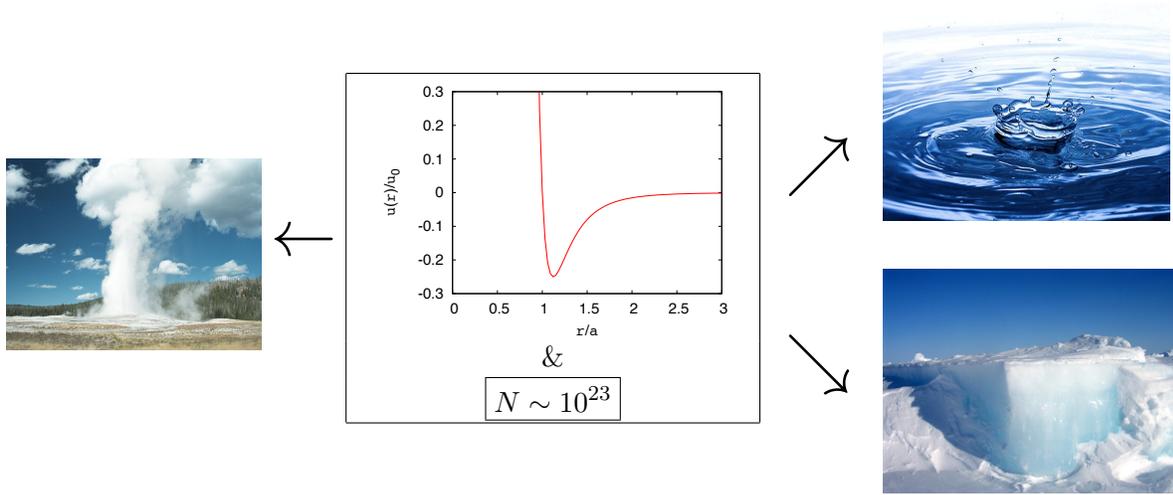


FIGURE 1.3 – **Émergence.** – Les mêmes détails microscopiques (ici un modèle simple d’interaction à deux corps) peuvent conduire à des phénomènes collectifs très différents (lorsque $N \rightarrow \infty$) : des phases gazeuse, liquide ou solide (notons que le potentiel d’interaction entre molécules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

À l’échelle macroscopique, il y a **émergence** de phénomènes collectifs imprévisibles avec la seule connaissance des lois microscopiques.

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].

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.

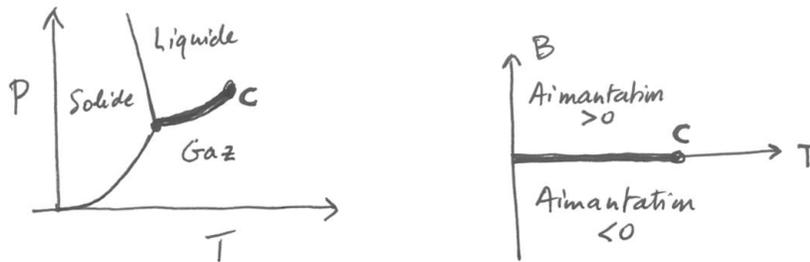


FIGURE 1.4 – Deux diagrammes de phases : liquide-gaz-solide et para-ferromagnétique. Le point critique est repéré par un C.

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 », avec 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, appelé « point critique » et noté C , la distinction entre liquide et gaz disparaît et la chaleur latente de vaporisation également. 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 (ici 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	liquide-gaz		para-ferro	
Chaleur spécifique	$C(T) \sim t ^{-\alpha}$	$\alpha = 0.111(1)$	$C(T) \sim t ^{-\alpha}$	$\alpha = 0.11(3)$
Paramètre d'ordre	densité	$\rho(T) - \rho_c \sim (-t)^\beta$	aimantation	$m(T) \sim (-t)^\beta$
		$\beta = 0.324(2)$		$\beta = 0.325(2)$
Susceptibilité	Compressibilité	$\kappa_T(T) \sim t ^{-\gamma}$	Susceptibilité	$\chi(T) \sim t ^{-\gamma}$
		$\gamma = 1.246(10)$		$\gamma = 1.25(1)$

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).

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

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.

1.2 L'approche probabiliste

1.2.1 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

$$\boxed{\mathcal{N}_A = 6.022\,141\,29(27)\,10^{23}\,\text{mol}^{-1}} \quad (1.2)$$

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.

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

$$N = \frac{10^{-3}}{24} \mathcal{N}_A \simeq 2.5\,10^{20}\text{ atomes} \quad (1.3)$$

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),

$$\text{Mém} = N \times \underbrace{\log_{10}(pLn^{1/3}) \times 6}_{\text{\# d'octets par atome}} \times \underbrace{\frac{\ln 10}{\ln 16}}_{\text{conversion en octets}} \quad (1.4)$$

soit

$$\text{Mém} = \frac{N \ln(p^3 N)}{2 \ln 2} \quad (1.5)$$

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 cm^3 !

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.

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.

1.2.2 Quelques éléments de théorie des probabilités

Évènements et probabilités

Lorsque l'occurrence d'un phénomène présente un caractère stochastique, soit pour des raisons intrinsèque 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 on joue à pile ou face (pour la seconde raison) ou si 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

$$P_m \stackrel{\text{def}}{=} \lim_{N \rightarrow \infty} \frac{N_m}{N} \quad \left(= \lim_{N \rightarrow \infty} \frac{N'_m}{N} \right) \quad (1.6)$$

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

$$\boxed{\sum_{m=1}^M P_m = 1} \quad (\text{normalisation}). \quad (1.7)$$

Variable aléatoire discrète

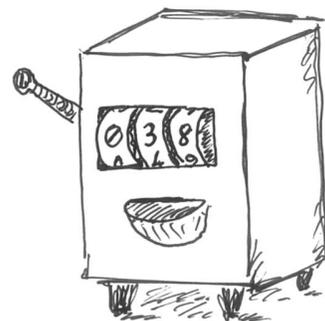
On considère une variable X prenant la valeur $X^{(m)}$ si l'évènement e_m est réalisé (X est entier, ou réel, ou complexe,...) : X est donc une *variable aléatoire*. Sa distribution est reliée aux probabilités $\{P_m\}$ comme

$$P(x) \stackrel{\text{def}}{=} \text{Proba}\{X = x\} = \sum_{m \text{ t.q. } X^{(m)}=x} P_m = \sum_{m=1}^M P_m \delta_{x, X^{(m)}}, \quad (1.8)$$

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 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 1.5)

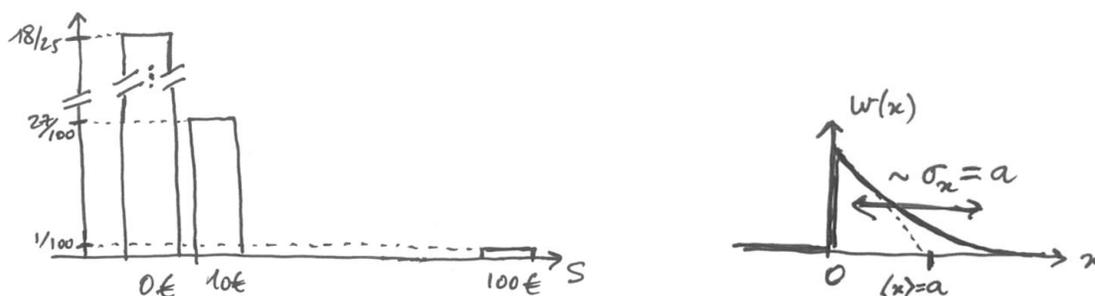


FIGURE 1.5 – **Distributions.** *distribution d'une variable aléatoire discrète (à gauche). À droite : distribution d'une variable continue (loi exponentielle).*

Variable aléatoire continue – Distribution/densité de probabilité

Nous avons jusqu'à présent discuté le cas d'une *variable aléatoire discrète*, prenant des valeurs dans un ensemble dénombrable de nombres, $\{0 \text{ €}, 10 \text{ €}, 100 \text{ €}\}$ 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. 1.6.

Nous avons déjà introduit la notion de distribution d'une variable aléatoire discrète : (1.8) 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 évènements, 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 :

$$w(x) dx \stackrel{\text{def}}{=} \text{Proba}\{X \in [x, x + dx]\} . \quad (1.9)$$

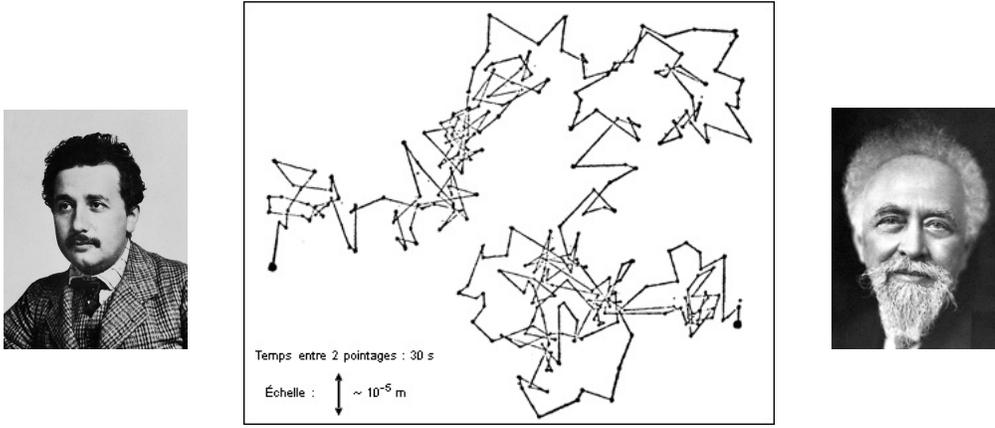


FIGURE 1.6 – Mouvement brownien d'un grain de pollen.

La condition de normalisation est alors

$$\int w(x) dx = 1 \quad (1.10)$$

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 ?

$$p(y) dy = \text{Proba}\{X \in [f(y), f(y) + \underbrace{f'(y)dy}_{dx}]\} = w(f(y)) f'(y) dy \quad \Rightarrow p(y) = w(f(y)) f'(y) \quad (1.11)$$

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 $p(E) = \sqrt{E/m} f(\sqrt{2mE})$.

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.

$$w(x_1, \dots, x_n) dx_1 \cdots dx_n = \text{Proba}\{X_1 \in [x_1, x_1 + dx_1] \& \cdots \& X_n \in [x_n, x_n + dx_n]\}. \quad (1.12)$$

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 :

$$w_1(x_1) dx_1 = dx_1 \int dx_2 \cdots \int dx_n w(x_1, \dots, x_n) \quad (1.13)$$

$$= \text{Proba}\{X_1 \in [x_1, x_1 + dx_1] \text{ sachant que les } n-1 \text{ autres variables peuvent valoir n'importe quoi}\}. \quad (1.14)$$

Moyenne, variance, moments, cumulants, corrélations,...

Supposons que l'on réalise N expériences dont l'espace des évènements 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 évènements 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 :

$$\bar{X}^{(N)} = \frac{1}{N} \sum_{i=1}^N X^{(m_i)} = \frac{1}{N} \sum_{m=1}^M N_m X^{(m)} \quad (1.15)$$

La première somme porte sur les N évènements qui se sont effectivement réalisés (la suite pile, face, face, pile, etc), alors que la seconde porte sur l'ensemble des M évènements (les deux possibilités pile ou face), ce pourquoi nous avons réintroduit le nombre d'occurrences N_m de l'évènement e_m . Remarquons que si nous réalisons un autre ensemble de N expériences, nous obtenons une suite d'évènements 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'évènements coïncident, ce qui sera prouvé plus bas). Nous déduisons que la moyenne probabiliste peut s'exprimer en terme de la distribution des évènements

$$\langle X \rangle = \sum_{m=1}^M P_m X^{(m)} \quad (1.16)$$

ou plutôt en terme de la distribution de la variable aléatoire :

$$\langle X \rangle = \sum_{v=1}^V P(x_v) x_v \quad (1.17)$$

où la somme porte sur les valeurs distinctes prises par la variable aléatoire.

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

$$\langle X \rangle = \int dx w(x) x. \quad (1.18)$$

Remarquons finalement que les formules (1.17) et (1.18) peuvent être généralisées pour calculer la moyenne de n'importe quelle fonction de la variable aléatoire,

$$\langle f(X) \rangle = \sum_{v=1}^V P(x_v) f(x_v) \quad (1.19)$$

pour une variable aléatoire discrète et

$$\langle f(X) \rangle = \int dx w(x) f(x) \quad (1.20)$$

pour une variable aléatoire continue.

Variance.— $\langle X \rangle$ 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 ²

$$\boxed{\text{Var}(X) \stackrel{\text{def}}{=} \langle (X - \langle X \rangle)^2 \rangle = \langle X^2 \rangle - \langle X \rangle^2} \quad (1.21)$$

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)

$$\boxed{\sigma_X \stackrel{\text{def}}{=} \sqrt{\text{Var}(X)}} \quad (1.22)$$

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 1.5). 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 indépendants (non corrélés) la probabilité d'occurrence des deux événements est donnée par le produit des probabilités

$$P(A \& B) = P(A) P(B) \quad (\text{définition de l'indépendance statistique}); \quad (1.23)$$

ex : quelle est la probabilité pour que moi et mon ami à Tokyo rations un train, sachant que nous n'interagissons pas. Cette propriété ne sera pas respectée si les deux événements sont corrélés, $P(A \& B) \neq P(A) P(B)$ (ex : quelle est la probabilité pour que moi et mon ami à Tokyo rations un train, sachant que nous venons de nous parler au téléphone, ce qui peut nous avoir retardé tous les deux).

Considérons deux variables aléatoires X et Y distribuées par la loi jointe $p(x, y)$.

Si les variables sont *statistiquement indépendantes (non corrélées)*, la distribution est donc factorisée $p(x, y) = \pi(x) \varpi(y)$, où $\pi(x)$, resp. $\varpi(y)$, est la loi marginale de X , resp. de Y . Dans ce cas $\langle XY \rangle = \int dx dy p(x, y) xy = \int dx \pi(x) x \int dy \varpi(y) y = \langle X \rangle \langle Y \rangle$.

En général, afin de quantifier les corrélations entre les deux variables on introduit la corrélation $C \stackrel{\text{def}}{=} \langle XY \rangle - \langle X \rangle \langle Y \rangle = \langle (X - \langle X \rangle)(Y - \langle Y \rangle) \rangle$. On a $C = 0$ lorsque les variables sont indépendantes, $C > 0$ ou $C < 0$ lorsqu'elles sont corrélées ou anti-corrélées.

On peut généraliser la notion à un ensemble de N variables aléatoires X_1, \dots, X_N . Nous introduisons la matrice $N \times N$ des covariances $C_{ij} \stackrel{\text{def}}{=} \langle (X_i - \langle X_i \rangle)(X_j - \langle X_j \rangle) \rangle$ qui encode les *fluctuations* C_{ii} (variances) et les *corrélations* entre les variables.

Moments et cumulants.— La notion de « *moment* » généralise la notion de moyenne et de valeur quadratique moyenne. Le n ème 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 : $\langle X^2 \rangle = \langle X \rangle^2 + \text{Var}(X) \simeq \langle X \rangle^2$.

²Nous 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.

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.
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).
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 *déviatoin 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).

Évidemment cette logique ne semble pas très évidente à suivre, en particulier il n'est pas très clair de justifier les définitions de c_3 et c_4 données ci-dessus. Nous expliquons dans le paragraphe suivant comme définir systématiquement les cumulants.

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 $P(x)$, 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* :

$$g(p) \stackrel{\text{def}}{=} \langle e^{-pX} \rangle = \int dx w(x) e^{-px} \quad (1.24)$$

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

$$g(p) = \sum_{n=0}^{\infty} \frac{(-p)^n}{n!} \mu_n \quad \Rightarrow \quad \mu_n = (-1)^n g^{(n)}(0). \quad (1.25)$$

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

$$w(p) \stackrel{\text{def}}{=} \sum_{n=1}^{\infty} \frac{(-p)^n}{n!} c_n \quad \Rightarrow \quad c_n = (-1)^n w^{(n)}(0). \quad (1.26)$$

On prendra maintenant la relation entre les deux fonctions génératrices

$$w(p) = \ln g(p) \quad (1.27)$$

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.

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 $g_\nu(p) = \exp[\nu e^{-p}]$ et $w_\nu(p) = \nu e^{-p}$, 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 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 $w_{a,\sigma}(p) = -ap + \frac{1}{2}\sigma^2 p^2$, 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.

Théorème de la limite centrale

Afin de définir la moyenne probabiliste $\langle X \rangle$ nous avons introduit la moyenne statistique $\bar{X}^{(N)}$ sur un ensemble d'évènements réalisés et avons noté que cette dernière fluctue lorsqu'elle est calculée sur des ensembles finis d'évènements distincts. Nous allons prouver cette assertion dans ce paragraphe.

Nous posons le problème comme suit. Considérons un ensemble de variables aléatoires

- *statistiquement indépendantes*, non corrélées (hypothèse 1).
- 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, sinon que ses deux premiers moments existent $\langle x \rangle < \infty$ et $\langle x^2 \rangle < \infty$.

Un exemple d'un ensemble de telles variables : les vitesses des atomes d'un gaz parfait classique (« parfait » + « classique » = particules non corrélées). Nous notons x_1, \dots, x_N ces N variables et considérons leur somme

$$S = \sum_{n=1}^N x_n \quad (1.28)$$

(notons qu'à un $1/N$ près la somme coïncide avec la moyenne statistique $\bar{x}^{(N)}$ des N variables introduite plus haut).

Nous posons la question : quelle est la distribution $P_N(S)$ de la variable S ? La manière la plus simple d'aborder le problème 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 $G_N(p) = \prod_{n=1}^N \langle e^{-px_n} \rangle$, et puisque ces variables sont identiquement distribuées (hyp.2) :

$$G_N(p) = [g(p)]^N. \quad (1.29)$$

Cette relation nous suggère d'écrire la fonction génératrice sous une forme exponentielle $g(p) = \exp w(p)$, i.e. fonction des *cumulants*, ou encore de relier directement les fonctions génératrices des cumulants de x et de S . Nous introduisons $W_N(p) \stackrel{\text{def}}{=} \sum_{k=1}^{\infty} \frac{(-p)^k}{k!} C_k = \ln G_N(p)$, alors

$$W_N(p) = N w(p) = N \sum_{k=1}^{\infty} \frac{(-p)^k}{k!} c_k \quad \text{i.e.} \quad \boxed{C_k = N c_k} \quad (1.30)$$

Pour les premiers :

$$\langle S \rangle = N \langle x \rangle \quad (\text{évident}), \quad (1.31)$$

$$\text{Var}(S) = N \text{Var}(x) \quad (\text{déjà moins trivial}), \quad (1.32)$$

etc.

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

$$\boxed{\sigma_S = \sqrt{N} \sigma_x} \quad (1.33)$$

Les fluctuations de la somme de N variables indépendantes croissent 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 $k > 2$:

$$G_N(p) \simeq e^{-Nc_1 p + \frac{1}{2} Nc_2 p^2} \quad \text{pour } N \rightarrow \infty \quad \text{i.e.} \quad \boxed{P_N(S) \simeq \frac{1}{\sqrt{2\pi N \text{Var}(x)}} \exp -\frac{(S - N \langle x \rangle)^2}{2N \text{Var}(x)}} \quad (1.34)$$

La distribution de la somme est *gaussienne* (quelle que soit la loi $p(x)$, tant que sa moyenne et sa variance sont finies).

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) $\sigma_S = \sqrt{N} \sigma_x$.

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

$$y \stackrel{\text{def}}{=} \frac{S - N\langle x \rangle}{\sqrt{N\text{Var}(x)}}. \quad (1.35)$$

Nous calculons la fonction génératrice de y :

$$\gamma_N(p) \stackrel{\text{def}}{=} \langle e^{-py} \rangle = \left\langle \exp -p \frac{\sum_n (x_n - c_1)}{\sqrt{Nc_2}} \right\rangle \quad (1.36)$$

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

$$\gamma_N(p) = \left[\tilde{g} \left(\frac{p}{\sqrt{Nc_2}} \right) \right]^N = \exp N \left[\frac{1}{2} c_2 \left(\frac{p}{\sqrt{Nc_2}} \right)^2 - \frac{1}{6} c_3 \left(\frac{p}{\sqrt{Nc_2}} \right)^3 + \dots \right] = \exp \left[\frac{1}{2} p^2 + \mathcal{O} \left(\frac{p^3}{\sqrt{N}} \right) \right] \quad (1.37)$$

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.

Exemple Revenons à la motivation du début du paragraphe concernant la distribution de la moyenne statistique $\bar{x}^{(N)} = \frac{1}{N} S = \frac{1}{N} \sum_{n=1}^N x_n$ dans la limite $N \rightarrow \infty$. Le théorème montre que la distribution $\pi_N(\bar{x})$ de $\bar{x}^{(N)}$ est une distribution gaussienne centrée sur $\langle \bar{x}^{(N)} \rangle = \langle x \rangle$ et de largeur $\text{Var}(\bar{x}^{(N)}) \propto 1/\sqrt{N}$. Autrement dit $\pi_N(\bar{x})$ est de plus en plus étroite au fur et à mesure que N croît. À la limite elle tend vers une distribution de Dirac :

$$\pi_N(\bar{x}) \simeq \sqrt{\frac{N}{2\pi\text{Var}(x)}} \exp -\frac{N(\bar{x} - \langle x \rangle)^2}{2\text{Var}(x)} \xrightarrow{N \rightarrow \infty} \delta(\bar{x} - \langle x \rangle). \quad (1.38)$$

On dit que $\bar{x}^{(N)}$ est une variable aléatoire auto-moyennante dans la limite $N \rightarrow \infty$.

1.2.3 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*.

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 (j'ai repris la notation de la section 1.2.2, i.e. un microétat est ici vu comme une éventualité de l'état du système parmi un ensemble d'états possibles).

- **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).

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.

L'information manquante : entropie statistique

Nous voyons que la notion de micro ou macroétat fait référence à l'*information* dont nous disposons sur le système : soit toute l'information nécessaire permettant de spécifier l'état microscopique (pile ou face), soit l'information partielle spécifiant avec quelles probabilités sont occupés les microétats (l'ensemble des probabilités $\{P_{\text{pile}}, P_{\text{face}}\}$).

Un outil essentiel est donc une mesure du *manque d'information* dû à l'occupation aléatoire des microétats. Le manque d'information est mesuré par « l'entropie statistique » définie comme :

$$S(\{P_m\}) \stackrel{\text{def}}{=} -k_B \sum_m P_m \ln P_m \quad (1.39)$$

où la constante en facteur est arbitraire et mesure le « quantum d'entropie » (i.e. le quantum d'information manquante) ; nous la choisirons égale à la constante de Boltzmann afin d'assurer la correspondance avec l'entropie de la thermodynamique : ³

$$k_B = 1.380\,6505(24) \times 10^{-23} \text{ J.K}^{-1} \quad (1.40)$$

(k_B est reliée à la constante des gaz parfaits $k_B = R/\mathcal{N}_A$).

Propriétés de l'entropie

D'où vient la définition de l'entropie ? Historiquement la définition (1.39) a été proposée en recherchant une fonction obéissant à un certain nombre de propriétés que nous énonçons maintenant. Autrement dit on peut montrer que l'entropie de von Neumann (1.39) est un bon choix de fonction de l'ensemble des probabilités $\{P_m\}$ qui obéit à toutes ces propriétés « raisonnables ». ⁴

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

$$S_{\min} = S(\text{microétat}) = 0 \quad (1.41)$$

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

$$S_{\max} = S(1/M, \dots, 1/M) = k_B \ln M \quad (1.42)$$

(pour une démonstration, cf. annexe 2.5.4).

³En 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.

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 :

$$S(P_1, \dots, P_M, 0, \dots, 0) = S(P_1, \dots, P_M) \quad (1.43)$$

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.

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

$$S(P_1, \dots, P_M) = \overbrace{S(q_A, q_B)}^{\text{choix entre } A \text{ et } B} + q_A \underbrace{S(P_1/q_A, \dots, P_m/q_A)}_{\text{entropie du sous ensemble } A} + q_B \underbrace{S(P_{m+1}/q_B, \dots, P_M/q_B)}_{\text{entropie du sous ensemble } B} \quad (1.44)$$

L'équation (1.44), 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

$$p_n^A = \sum_{m \in B} p_{n,m} \quad \text{resp.} \quad p_m^B = \sum_{n \in A} p_{n,m} \quad (1.45)$$

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 nous donnant l'entropie

$$S(\{p_{n,m}\}) = S(\{p_n^A\}) + S(\{p_m^B\}) - \underbrace{k_B \sum_{n,m} p_{n,m} \ln \left(\frac{p_{n,m}}{p_n^A p_m^B} \right)}_{\text{corrélacion } \mathcal{C}(A,B) \geq 0}, \quad (1.46)$$

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

$$\boxed{S(\{p_{n,m}\}) \leq S(\{p_n^A\}) + S(\{p_m^B\})} \quad (1.48)$$

⁵Nous avons noté 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).

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

$$\mathcal{C}(A, B) \stackrel{\text{def}}{=} k_B \sum_{n,m} p_{n,m} \ln \left(\frac{p_{n,m}}{p_n^A p_m^B} \right) \geq k_B \sum_{n,m} p_{n,m} \left(1 - \frac{p_n^A p_m^B}{p_{n,m}} \right) = k_B \sum_{n,m} (p_{n,m} - p_n^A p_m^B) = 0. \quad (1.47)$$

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, alors 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 .

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

$$p_{\sigma,\sigma'} : \frac{1}{4} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \quad (1.49)$$

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

$$\tilde{p}_{\sigma,\sigma'} : \frac{1}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad (1.50)$$

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.

Remarquons que lorsque les $A \otimes B$ 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 ».

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 [10],...) qui permet de décrire les deux types d'états sur un pied d'égalité (cf. note de bas de page n°8 page 36).

☺ Les idées importantes

- La distinction entre processus élémentaires et phénomènes collectifs.
- 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 [13].
- 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.

1.3 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 [27].
- Un livre au classicisme très français plutôt synthétique et plaisant : R. Castaing [7].
- Pédagogique et plutôt synthétique : C. & H. Ngô [24].
- L'excellent photocopié de cours d'A. Georges & M. Mézard [11] (difficile à se procurer ?)
- Très (trop ?) complet et pédagogique : B. Diu, C. Guthmann, D. Lederer & B. Roulet [10]. 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 [25].
- Les ouvrages récents de M. Kardar méritent très certainement d'être mentionnés [15, 14] (le premier tome correspond davantage à notre programme).
- Deux ouvrages intéressants sont ceux de K. Huang [12] et plus original, celui de S.-K. Ma [23].
- Pour rester du côté asiatique on peut mentionner le classique livre de R. Kubo *et alter* [30] et, sur des sujets plus avancés (aspects hors équilibre,...) [18].

Ouvrages de référence :

- Re-citons : B. Diu, C. Guthmann, D. Lederer & B. Roulet [10].
- Un livre classique très complet est celui de F. Reif [26].
- Un autre classique (plus difficile) : le « Landau & Lifchitz » [19]. Discussion détaillée des symétries.
- Très profond et complet, mais également plus difficile : R. Balian [3, 4].

Exercices :

- L. Couture, C. Chahine & R. Zitoun [9]
- Récemment : H. Krivine & J. Treiner [17].

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. [21].
- Encore plus récent, assez condensé et avec une légère coloration « matière condensée/physique mésoscopique », [28] (fait spécialement pour le magistère d'Orsay ☺).

C. Physique du solide

- Un livre assez descriptif et agréable : C. Kittel [16].
- Un classique assez détaillé : Ashcroft & N. Mermin [2].



Annexe 1.A : Chronologie – D’où vient la physique statistique ?

Je reprends l’introduction des livres de Balian et de Pathria [3, 25]. La physique statistique s’est développée dans la seconde moitié du XIX^e siècle. Les précurseurs sont Daniel Bernouilli (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), John 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 Bernouilli 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 (équivalence travail-chaleur définitivement 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 ?
- 1875 : Paradoxe de Gibbs.

- 1900 : Planck propose la loi éponyme décrivant la thermodynamique du rayonnement. 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.
- 1912 : Debye améliore la théorie d'Einstein et explique les comportements de basse température de la chaleur spécifique $C_V^{\text{vib}} \propto T^3$.
- 1912 : Smoluchowski lève le paradoxe du démon de Maxwell pour un démon automatique.
- 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).
- 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).
- 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).
- 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).
- 1948 : Théorie de l'information (Shannon & Weaver).
- 1950 : En interprétant le "paradoxe du démon de Maxwell" avec les idées de la théorie de 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

Postulat fondamental – Ensemble microcanonique

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.

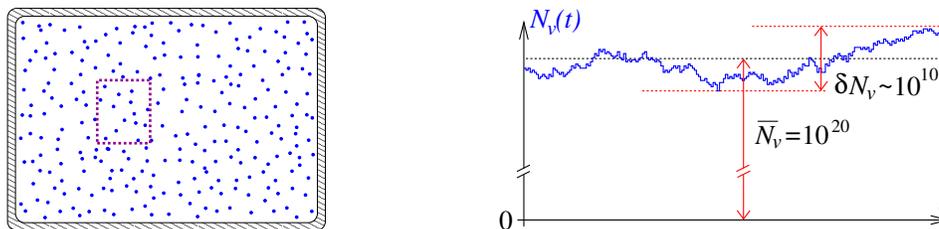


FIGURE 2.1 – **Équilibre macroscopique et fluctuations.** – On mesure le nombre de particules $N_v(t)$ contenues dans un petit volume d'un gaz. Pour un volume $v \sim 1 \text{ cm}^3$ de gaz dans les conditions de température et de pression normales, nous avons vu que $\bar{N}_v \sim 10^{20}$.

Nous isolons par la pensée un volume v dans un gaz (figure 2.1), 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*

$$\overline{N_v(t)}^{(T,t_0)} = \frac{1}{T} \int_{t_0}^{t_0+T} dt N_v(t) \simeq n v \quad \text{pour } T \text{ "grand"}, \quad (2.1)$$

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 la condition *d’équilibre macroscopique* requiert (i) la *stationnarité* des quantités macroscopiques, et (ii) *l’absence de flux*.¹

2.1 Description classique : espace des phases

2.1.1 Définition de l’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*

$$\dot{q}_i(t) = \frac{\partial H}{\partial p_i}(\{q_i(t), p_i(t)\}) \quad (2.2)$$

$$\dot{p}_i(t) = -\frac{\partial H}{\partial q_i}(\{q_i(t), p_i(t)\}) \quad (2.3)$$

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 $(\{q_i, p_i\})$.

2.1.2 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 :

$$q_i(t + \delta t) \simeq q_i(t) + \dot{q}_i(t) \delta t \quad (2.4)$$

$$p_i(t + \delta t) \simeq p_i(t) + \dot{p}_i(t) \delta t \quad (2.5)$$

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

²Il 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.

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

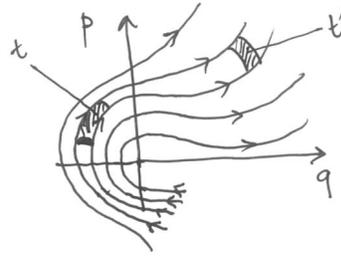


FIGURE 2.2 – **Théorème de Liouville.** On a représenté des trajectoires physiques dans l'espace des phases. L'évolution temporelle conserve le volume dans l'espace des phases, i.e. les deux volumes hachurés, représentant un ensemble de microétats à deux instants de leur évolution, sont égaux.

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

$$\begin{pmatrix} \vdots \\ q_i \\ p_i \\ \vdots \end{pmatrix} \equiv \begin{pmatrix} \vdots \\ q_i(t) \\ p_i(t) \\ \vdots \end{pmatrix} \rightarrow \begin{pmatrix} \vdots \\ \tilde{q}_i \\ \tilde{p}_i \\ \vdots \end{pmatrix} \equiv \begin{pmatrix} \vdots \\ q_i(t + \delta t) \\ p_i(t + \delta t) \\ \vdots \end{pmatrix} = \begin{pmatrix} \vdots \\ q_i + \frac{\partial H}{\partial p_i}(\{q_i, p_i\}) \delta t \\ p_i - \frac{\partial H}{\partial q_i}(\{q_i, p_i\}) \delta t \\ \vdots \end{pmatrix} \quad (2.6)$$

Le Jacobien de cette transformation est donné par

$$J = \det(1 + K \delta t) \quad \text{où} \quad K \stackrel{\text{def}}{=} \begin{pmatrix} \vdots & \vdots \\ \cdots & \frac{\partial^2 H}{\partial q_i \partial p_i} & \frac{\partial^2 H}{\partial p_i^2} & \cdots \\ \cdots & -\frac{\partial^2 H}{\partial q_i^2} & -\frac{\partial^2 H}{\partial q_i \partial p_i} & \cdots \\ \vdots & \vdots \end{pmatrix} \quad (2.7)$$

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

$$J \simeq 1 + \sum_i \left(\frac{\partial^2 H}{\partial q_i \partial p_i} - \frac{\partial^2 H}{\partial q_i \partial p_i} \right) \delta t = 1. \quad (2.8)$$

Autrement dit (théorème de Liouville ⁵)

$$\boxed{\prod_i d\tilde{q}_i d\tilde{p}_i = \prod_i dq_i dp_i} \quad (2.9)$$

⁴On 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 $\vec{V} = (\cdots, \dot{q}_i, \dot{p}_i, \cdots)$ 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.

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 2.2).

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 $\rho(\dots; t)$.

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 $d\theta dp_\theta d\varphi dp_\varphi$.

2.1.3 Ergodicité

Moyenne probabiliste (moyenne d'ensemble)

Nous avons vu plus haut qu'un macroétat est décrit par la donnée d'une distribution $\rho(\{q_i, p_i\})$ dans l'espace des phases. Considérons une grandeur physique, fonction des coordonnées dans l'espace des phases $A(\{q_i, p_i\})$. La moyenne (probabiliste) de cette grandeur physique, lorsque le système est dans le macroétat, est donnée par

$$\langle A \rangle \stackrel{\text{def}}{=} \int \prod_i dq_i dp_i \rho(\{q_i, p_i\}) A(\{q_i, p_i\}). \quad (2.10)$$

Moyenne temporelle

Si maintenant la grandeur physique est mesurée au cours d'une expérience, on obtient *a priori* une certaine fonction du temps (qui dépend des conditions initiales), que nous notons simplement $A(t)$. Cette fonction pourrait être l'énergie, ou le nombre de particules $N_v(t)$ dans un sous volume du gaz (figure 2.1). 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 :

$$\bar{A}^{(T)} = \frac{1}{T} \int_0^T dt A(\{q_i(t), p_i(t)\}), \quad (2.11)$$

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

Hypothèse ergodique

Il est raisonnable de supposer que la moyenne (2.11) tend vers une limite lorsque $T \rightarrow \infty$ (cf. figure 2.1). 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. chapitre précédent et justification ci-dessous). Cette hypothèse nous permet de **remplacer la moyenne temporelle** par une **moyenne d'ensemble** (moyenne probabiliste) :

$$\boxed{\lim_{T \rightarrow \infty} \overline{A}^{(T)} \stackrel{\text{hyp. ergod.}}{=} \langle A \rangle} \quad (2.12)$$

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.

2.1.4 Formulation classique du postulat fondamental de la physique statistique

La question est maintenant de déterminer la mesure 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 29, 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 :

$$\text{densité stationnaire} \Rightarrow \rho(\{q_i, p_i\}) = \text{fct}(H(\{q_i, p_i\})) \quad (2.13)$$

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

$$E = H(\{q_i, p_i\}) \quad (2.14)$$

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

$$\rho^*(\{q_i, p_i\}) = \frac{1}{\Sigma(E)} \delta[E - H(\{q_i, p_i\})] \quad \text{où} \quad \Sigma(E) = \int \prod_i dq_i dp_i \delta[E - H(\{q_i, p_i\})] \quad (2.15)$$

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 :

$$\rho^*(\{q_i, p_i\}) = \begin{cases} \frac{1}{\mathcal{V}(E, \delta E)} & \text{si } E \leq H(\{q_i, p_i\}) \leq E + \delta E \\ 0 & \text{sinon} \end{cases} \quad \text{où } \mathcal{V}(E, \delta E) = \Sigma(E) \delta E \quad (2.16)$$

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 \rightarrow classique. L'équation (2.16) 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 :

$$\Sigma(E) = \int dqdp \underbrace{\delta\left(E - \frac{1}{2}(p^2 + q^2)\right)}_{\frac{1}{|p|}[\delta(p - \sqrt{2E - q^2}) + \delta(p + \sqrt{2E - q^2})]} = \int_{-\sqrt{2E}}^{+\sqrt{2E}} \frac{2dq}{\sqrt{2E - q^2}} = 2 \int_0^1 \frac{dt}{\sqrt{t}\sqrt{1-t}} \quad (2.17)$$

$$= 2 B(1/2, 1/2) = 2\pi \quad (2.18)$$

où nous avons utilisé (A.13). La dernière intégrale est reliée à la fonction Beta d'Euler (A.4). Finalement :

$$\rho^*(q, p) = \frac{1}{\pi} \delta(2E - p^2 - q^2) . \quad (2.19)$$

Nous déduisons la probabilité de présence au point q :

$$w(q) \stackrel{\text{def}}{=} \int dp \rho^*(q, p) = \frac{1}{\pi \sqrt{2E - q^2}} \quad \text{pour } |q| \leq \sqrt{2E} \quad (2.20)$$

(le calcul utilise à nouveau (A.13)). La distribution dans l'espace physique nous permet de calculer des moyennes de fonctions de la position. Par exemple :

$$\frac{1}{2} \langle q^2 \rangle = \int dq w(q) q^2 = \frac{1}{2} E \quad (2.21)$$

(l'intégrale est à nouveau donnée par (A.4)). Finalement nous reconnaissons un résultat bien connu : nous avons retrouvé la moyenne *temporelle* de l'énergie potentielle de l'oscillateur harmonique, sur une période d'oscillation $\overline{E_{\text{pot}}(t)} = \frac{1}{2} \overline{q(t)^2} = \frac{1}{2} E$. 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).

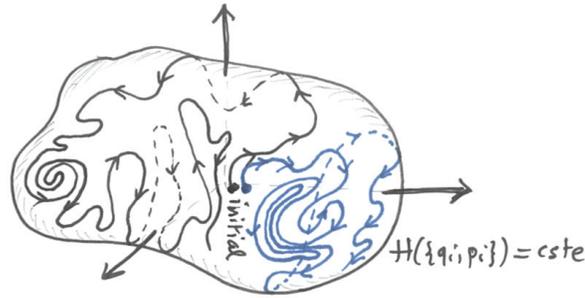


FIGURE 2.3 – **Ergodicité.** Deux trajectoires initialement très proches évoluent de manière très différentes sur la couche d'iso-énergie (attention, en principe l'espace des phases a en principe une dimension paire). Il a été prouvé rigoureusement que la distance entre les deux trajectoires croît comme $\exp \gamma t$ (où γ est un exposant de Lyapunov).

Tentative de justification de l'hypothèse ergodique.— Essayons de comprendre comment nous pouvons justifier le passage de la moyenne temporelle (2.11) 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 (2.11) est très grand devant le temps τ_{corr}^A et découpons l'intégrale en \mathcal{N}_T morceaux associés à des intervalles de temps $\Delta T = T/\mathcal{N}_T \gtrsim \tau_{\text{corr}}^A$:

$$\bar{A}^{(T)} = \frac{1}{\mathcal{N}_T} \left[\int_0^{\Delta T} \frac{dt}{\Delta T} A(t) + \int_{\Delta T}^{2\Delta T} \frac{dt}{\Delta T} A(t) + \cdots + \int_{(\mathcal{N}_T-1)\Delta T}^{\mathcal{N}_T\Delta T} \frac{dt}{\Delta T} A(t) \right]. \quad (2.22)$$

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

$$\bar{A}^{(T)} = \langle A \rangle + \underbrace{\mathcal{O}(1/\sqrt{T})}_{\text{terme aléatoire}} \quad (2.23)$$

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.

2.2 Description quantique : espace de Hilbert

2.2.1 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, 22, 28] 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 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*). La mesure est donc une évolution *stochastique et irréversible* de l'état quantique.
- *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 terme 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, en terme de fonction d'onde à N corps. 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.⁶

2.2.2 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.⁷ Ceci est bien illustré en écrivant les règles de projection pour un état pur et un mélange statistique :

$$\text{Proba}(\phi \leftarrow \text{état pur } \psi_m) = |\langle \phi | \psi_m \rangle|^2 \quad (2.24)$$

$$\text{Proba}(\phi \leftarrow \text{mélange statistique } \{P_m\}) = \sum_m P_m |\langle \phi | \psi_m \rangle|^2 \quad (2.25)$$

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*) :

$$\langle A \rangle_{\substack{\text{mélange} \\ \text{statistique}}} = \underbrace{\sum_m P_m}_{\text{moyenne statistique}} \overbrace{\langle \psi_m | A | \psi_m \rangle}^{\text{moyenne quantique}} \quad (2.26)$$

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

$$S_z = \frac{\hbar}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (2.27)$$

Nous étudions la moyenne de la composante du spin S_x :

$$S_x = \frac{\hbar}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \quad (2.28)$$

⁶En 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).

⁷**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).

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 $\text{Proba}(\uparrow \leftarrow \psi)$ $\text{Proba}(\downarrow \leftarrow \psi)$ 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. ⁸ Cf. [3] ou [10].

2.2.3 Densité d’états, limite semi-classique

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 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.

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 :

$$\rho(E) \stackrel{\text{def}}{=} \sum_n \delta(E - E_n) \quad (2.29)$$

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). Cette quantité mesure le nombre d’états dans un intervalle :

$$\boxed{\rho(E) dE = \# \text{ d'états dans } [E, E + dE]} \quad (2.30)$$

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*

$$\Phi(E) \stackrel{\text{def}}{=} \int_{-\infty}^E dE' \rho(E') = \# \text{ d'états d'énergies inférieures à } E, \quad (2.31)$$

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

⁸ **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 (1.39) s’écrit $S = -k_B \text{tr}\{\rho \ln \rho\}$.

☞ **Exercice 2.1 :** 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 $\text{tr}\{\rho_p S_x\}$.

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.

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

$$E_n = \hbar\omega \left(n + \frac{1}{2} \right) \quad n \in \mathbb{N}. \quad (2.32)$$

Dans ce cas la densité d'états intégrée $\Phi(E)$ est une fonction en marches d'escalier régulières. 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* ».

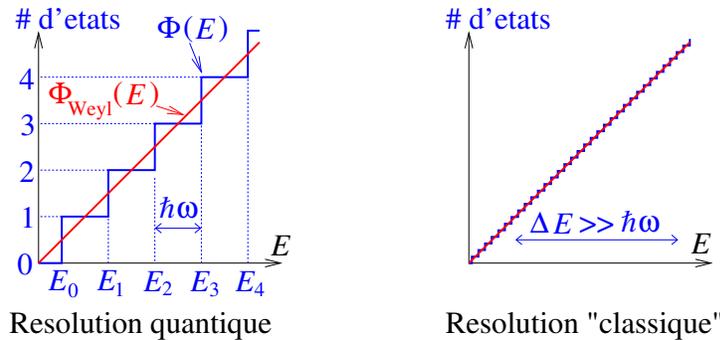


FIGURE 2.4 – **Densité d'états intégrée de l'oscillateur harmonique.** *Comparaison de la densité d'états intégrée $\Phi(E)$ avec son approximation semi-classique $\Phi_{\text{Weyl}}(E)$. « Vue de près » (gauche) et « vue de loin » (droite).*

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

$$\sum_n f(E_n) = \int dE \rho(E) f(E). \quad (2.33)$$

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

$$\sum_n f(E_n) \simeq \int dE \rho_{\text{Weyl}}(E) f(E), \quad (2.34)$$

où l'intégrale réalise la moyenne locale sur une échelle plus grande que les écarts entre niveaux. Il sera en général plus facile de calculer l'intégrale (2.34) que la somme (2.33).

Densité d'états 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

$$\psi_{\vec{k}}(\vec{r}) = \frac{1}{\sqrt{V}} e^{i\vec{k}\cdot\vec{r}} \quad \text{où} \quad \vec{k} = \frac{2\pi}{L}(n_1, \dots, n_d) \quad (2.35)$$

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

$$\Phi(E) = \sum_{\vec{k}} \theta_{\text{H}} \left(E - \frac{\hbar^2 \vec{k}^2}{2m} \right) \quad (2.36)$$

où la somme porte sur tous les vecteurs d'onde quantifiés. $\theta_{\text{H}}(x)$ est la fonction de Heaviside. Si le volume est « grand », nous faisons une approximation continue et remplaçons la somme par une intégrale : puisqu'un état occupe un volume $(\frac{2\pi}{L})^d$ dans l'espace des vecteurs d'onde, nous avons la correspondance :

$$\boxed{\sum_{\vec{k}} \rightarrow \frac{V}{(2\pi)^d} \int d\vec{k}} \quad (2.37)$$

valable lorsque la somme porte sur une fonction de \vec{k} variant lentement à l'échelle de $2\pi/L$. 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 :

$$\Phi_{\text{Weyl}}(E) = \frac{1}{h^d} \int_V d\vec{r} \int d\vec{p} \theta_{\text{H}} \left(E - \frac{\vec{p}^2}{2m} \right) \quad (2.38)$$

Nous interprétons donc ce résultat comme

$$\boxed{\Phi_{\text{Weyl}}(E) = \frac{1}{h^d} (\text{volume occupé par les microétats de l'espace des phases d'énergies } < E)} \quad (2.39)$$

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 *classiques*. 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 :⁹

$$\Phi_{\text{Weyl}}(E) = \frac{V}{\Gamma(\frac{d}{2} + 1)} \left(\frac{mE}{2\pi\hbar^2} \right)^{d/2}. \quad (2.40)$$

⁹Le volume de l'hypersphère de rayon 1 est $\mathcal{V}_d = \pi^{d/2}/\Gamma(\frac{d}{2} + 1)$.

Cas général Nous admettrons que la règle semi-classique (2.39) s'applique pour un hamiltonien général $H(\{q_i, p_i\})$:

$$\Phi_{\text{Weyl}}(E) = \int \prod_{i=1}^d \frac{dq_i dp_i}{h} \theta_{\text{H}}(E - H(\{q_i, p_i\})) \quad (2.41)$$

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

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

$$\Phi_{\text{Weyl}}(E) = \int \frac{dq dp}{h} \theta_{\text{H}}(E - H(q, p)) = \frac{\sqrt{2m}}{h} \sqrt{\frac{2}{m\omega^2}} \underbrace{\int dQ dP \theta_{\text{H}}(E - P^2 - Q^2)}_{\text{surface disque } = \pi E} = \frac{1}{\hbar\omega} E. \quad (2.42)$$

Nous avons retrouvé le résultat discuté au début du paragraphe (figure 2.4).

2.2.4 Densité d'états pour des particules 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 intéresse essentiellement, le *comptage des états quantiques* à travers la densité d'états, son effet est toutefois plus simple à discuter. 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, les $N!$ permutations des indices correspondent au même état quantique. Dans le cas des bosons, on a

$$\Phi_{\text{bosons}}(E) = \sum_{\vec{k}_1, \dots, \vec{k}_N} \frac{n_{\vec{k}_1}! \cdots n_{\vec{k}_N}!}{N!} \theta_{\text{H}}\left(E - \sum_{i=1}^N \epsilon_{\vec{k}_i}\right) \quad (2.43)$$

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) :

$$\Phi_{\text{fermions}}(E) = \frac{1}{N!} \sum_{\substack{\vec{k}_1, \dots, \vec{k}_N \\ \text{tous différents}}} \theta_{\text{H}} \left(E - \sum_{i=1}^N \epsilon_{\vec{k}_i} \right). \quad (2.44)$$

La difficulté du calcul réside dans le fait que les sommations sur les vecteurs d'onde 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 révèlera 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 :

$$\Phi_{\text{indisc}}(E) = \sum_{\substack{\{n_{\vec{k}}\} \text{ tq.} \\ \sum_{\vec{k}} n_{\vec{k}} = N}} \theta_{\text{H}} \left(E - \sum_{\vec{k}} n_{\vec{k}} \epsilon_{\vec{k}} \right). \quad (2.45)$$

Contrairement aux expressions (2.43) et (2.44), 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. 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 : $n_{\vec{k}_i}$ vaut la plupart du temps 0, rarement 1 et est encore plus rarement supérieur à 1 ; autrement dit le nombre d'états accessibles est beaucoup plus grand que le nombre de particules, ce qui se produit si E devient « grande ». On peut alors négliger les contraintes sur les nombres quantiques lors des sommations dans (2.43) et (2.44). Nous obtenons alors un résultat identique pour les bosons et les fermions

$$\Phi_{\text{indisc}}(E) \simeq \frac{1}{N!} \sum_{\vec{k}_1, \dots, \vec{k}_N} \theta_{\text{H}} \left(E - \sum_{i=1}^N \frac{\hbar^2 \vec{k}_i^2}{2m} \right) \quad (2.46)$$

Autrement dit

$$\boxed{\Phi_{\text{indisc}}(E) \simeq \frac{1}{N!} \Phi_{\text{disc}}(E)} \simeq \Phi_{\text{bosons}}(E) \simeq \Phi_{\text{fermions}}(E) \quad (2.47)$$

où dans l'approximation semi-classique, $\Phi_{\text{disc}}(E)$ est donné par (2.41). L'approximation faite ici est équivalente à « l'approximation de Maxwell-Boltzmann » qui sera discutée plus tard (les expressions (2.43, 2.44, 2.45, 2.47) deviendront plus claires aux chapitres 8 et 9).

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 (2.36), et donc

$$\Phi_{\text{indisc}}^{\text{libres}}(E) \simeq \frac{1}{N!} \frac{V^N}{\Gamma(\frac{3N}{2} + 1)} \left(\frac{mE}{2\pi\hbar^2} \right)^{3N/2} \quad (2.48)$$

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 $\Gamma(n+1) = n!$.

Fonction Beta d'Euler

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

Intégrales gaussiennes

$$\int_{-\infty}^{+\infty} dx e^{-x^2} = \sqrt{\pi} \quad (\text{A.5})$$

Plus généralement

$$\int_0^{+\infty} dx x^n e^{-\frac{1}{2}ax^2} = \frac{1}{2} \left(\frac{2}{a}\right)^{\frac{n+1}{2}} \Gamma\left(\frac{n+1}{2}\right) \quad (\text{A.6})$$

Méthode du col

Soit l'intégrale

$$I(\lambda) = \int_a^b dx e^{-\lambda\varphi(x)} \quad (\text{A.7})$$

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 :

$$I(\lambda) \underset{\lambda \rightarrow \infty}{\simeq} \sqrt{\frac{2\pi}{\lambda\varphi''(x_*)}} e^{-\lambda\varphi(x_*)} \quad (\text{A.8})$$

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

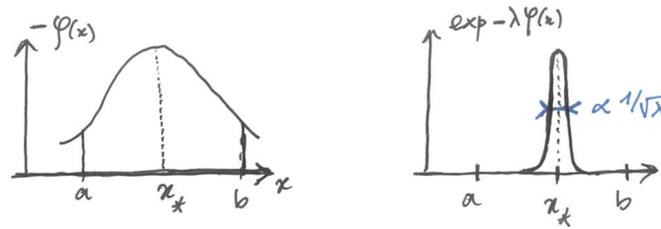


FIGURE A.1 – Méthode du col.

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

$$\Gamma(z + 1) \underset{z \rightarrow \infty}{\simeq} \sqrt{2\pi} z^{z+\frac{1}{2}} e^{-z} \tag{A.9}$$

qu'on utilisera souvent sous la forme

$$\ln \Gamma(z + 1) \underset{z \rightarrow \infty}{=} z \ln z - z + \frac{1}{2} \ln(2\pi z) + \mathcal{O}(1/z) \tag{A.10}$$

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 :

$$\int dx f(x) \delta(x) = f(0), \tag{A.11}$$

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.

$$\delta(x) = \int_{-\infty}^{+\infty} \frac{dk}{2\pi} e^{ikx}. \tag{A.12}$$

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

$$\delta(f(x)) = \sum_{\substack{x_n \text{ tq} \\ f(x_n) = 0}} \frac{\delta(x - x_n)}{|f'(x_n)|} = \frac{1}{|f'(x)|} \sum_{\substack{x_n \text{ tq} \\ f(x_n) = 0}} \delta(x - x_n). \tag{A.13}$$

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] L. Couture, C. Chahine and R. Zitoun, *Thermodynamique - cours, exercices et problèmes résolus*, Dunod (1993).
- [10] B. Diu, C. Guthman, D. Lederer and B. Roulet, *Physique Statistique*, Hermann, Paris (1989).
- [11] A. Georges and M. Mézard, *Physique Statistique*, édité par l'École Polytechnique, Palaiseau (2008), ISBN 978-2-7302-1484-1.
- [12] K. Huang, *Statistical mechanics*, John Wiley & Sons, New York (1963).
- [13] E. T. Jaynes, Information Theory and Statistical Mechanics, *Phys. Rev.* **106**, 620-630 (1957).
- [14] M. Kardar, *Statistical physics of fields*, Cambridge University Press (2007).
- [15] M. Kardar, *Statistical physics of particles*, Cambridge University Press (2007).
- [16] C. Kittel, *Physique de l'état solide*, Dunod (1994), 5ème édition.
- [17] H. Krivine and J. Treiner, *La physique statistique en exercices*, Vuibert, Paris (2003).
- [18] R. Kubo, M. Toda and N. Hashitsume, *Statistical physics II : non-equilibrium statistical mechanics*, Springer-Verlag (1992).
- [19] L. D. Landau and E. Lifchitz, *Physique statistique*, Mir (1966), tome 5.
- [20] R. B. Laughlin, *Un univers différent*, Fayard (2005).

- [21] M. Le Bellac, *Physique quantique*, EDP Sciences – CNRS Éditions (2003).
- [22] J.-M. Lévy-Leblond and F. Balibar, *Quantique – Rudiments*, CNRS-InterÉditions, Paris (1984).
- [23] S.-K. Ma, *Statistical mechanics*, World Scientific, Singapore (1985).
- [24] C. Ngô and H. Ngô, *Physique Statistique*, Dunod, Paris (2008), 3ème édition.
- [25] R. K. Pathria and P. D. Beale, *Statistical mechanics*, Academic Press, Elsevier (2011), 3rd edition.
- [26] F. Reif, *Fundamentals of statistical and thermal physics*, McGraw-Hill, New York (1965).
- [27] F. Reif, *Physique statistique*, Armand Colin, Paris (1972), Berkeley : cours de physique, volume 5.
- [28] C. Texier, *Mécanique quantique*, Dunod, Paris (2011).
- [29] C. Texier, *Physique statistique (faiblement) hors équilibre : formalisme de la réponse linéaire. Application à l'étude de la dissipation quantique et du transport électronique* (2011), notes de cours du DEA de physique quantique, École Normale Supérieure, disponible à l'adresse http://www.lptms.u-psud.fr/userpage/christophe_texier/.
- [30] M. Toda, R. Kubo and N. Saitô, *Statistical physics I : equilibrium statistical mechanics*, Springer-Verlag (1992).
- [31] C. N. Yang and T. D. Lee, Statistical Theory of Equations of State and Phase Transitions. I. Theory of Condensation, *Phys. Rev.* **87**, 404–409 (1952).

Index

A	
Agitation thermique	27
Aléatoire (variable)	10
Argon	52
Avogadro (nombre d')	8
B	
Bernouilli, Daniel	24
Binomiale (distribution)	16
Bit	9
Boltzmann, Ludwig	24
Bose-Einstein (distribution de)	78
Boson	34
C	
Capacité calorifique (voir Chaleur spécifique)	3
Chaleur	67
Chaleur spécifique	3, 60
gaz diatomique	3
gaz parfait monoatomique	71
Chaos	9
Clausius, Rudolf	24
Col (méthode du)	87
Collision	52
Compressibilité	7, 63
Conjugué	
paramètre conjugué	61, 63, 66
variable canoniquement conjugué	28
Constante de Boltzmann	19
Contact thermique	43
Corrélations	14
Covariances	14
Cumulants	15, 60
Cumulants (définition)	16
D	
de Broglie (longueur thermique)	68
Debye, Petrus	25
Dégénérescence	41
Détente de Joule	47
Diffusion	
des molécules dans un gaz	53
Distribution de probabilité	11
E	
Einstein, Albert	25
Élémentaire	2
Émergence	5
Énergie libre	59
Entropie	24
canonique	61
de Rényi	19, 59
extensivité	42
microcanonique	42
principe d'entropie maximale ...	32, 42, 49
Sackur-Tetrode (formule de)	50, 72
validité	51
statistique (de von Neumann)	19
Équilibre macroscopique	27
Espérance	13
État pur	18
Évènement	10
F	
Facteurs d'occupation	40, 78
Fermi-Dirac (distribution de)	79
Fermion	34
Fluctuation-dissipation	61, 71
Fonction de partition	58
Fonction génératrice	15, 58, 77
Fugacité	77
G	
Gaussienne (distribution)	16
Gaz parfait	8, 16, 24
monoatomique	50
potentiel chimique	51
pression	51
Sackur-Tetrode (formule de)	50, 51
validité	51
Gibbs, J. Willard	24, 53
H	
Herapath, John	24
I	
Indiscernabilité	34, 39, 52-54

- Intrication 21
 Irréversibilité 48
- J**
- Joule, James P. 24
- L**
- Landau, Lev D. 25
 Libre parcours moyen 52
 Limite thermodynamique 45, 65, 70
 Longueur thermique (de de Broglie) 68
- M**
- Macroétat 18
 Maxwell, John C. 24
 Maxwell-Boltzmann (approximation de) 40, 69
 Mayer (relation de) 3, 71
 Mélange statistique 18
 Méthode du col 87
 Microétat 18
 Moments (définition) 14
 Mouvement brownien 11
 Moyenne (définition) 13
- N**
- Nernst (théorème de) 51
- O**
- Octet 9
 Opalescence critique 7
 Opérateur densité 36
- P**
- Paradoxe de Gibbs 24, 53
 Paramètre conjugué 61, 66
 Perrin, Jean 11
 Phénomènes collectifs 2
 Planck, Max 25
 Poincaré (temps de) 33
 Postulat de symétrisation 34
 Postulat fondamental 41
 Postulat fondamental de la thermodynamique
 42
 Potentiel chimique
 canonique 62, 63
 microcanonique 48
 Pression
 canonique 62, 63
 microcanonique 47
 Principe de Pauli 40
 Probabilités 8, 10
- Q**
- Quantique (mécanique) 9, 23
- R**
- Réductionnisme 5
 Rényi (entropie de) 19, 59
 Réversible (transformation) 66
- S**
- Sackur-Tetrode (formule de) 50
 validité 51
 Section efficace 52
 Spin 34, 35
 Stirling (formule de) 88
- T**
- Température
 absolue 45
 canonique 58
 microcanonique 45
 Théorème d'équipartition 68, 71
 Théorème de Feynman-Hellmann 61, 66
 Théorème de la limite centrale 16
 Théorème de Liouville 28, 29, 31
 Théorème de Nernst 51
 Théorème H 24
 Théorie ergodique 34
 Thermodynamique 3
 1er principe 4, 24
 2ème principe 4, 24
 3ème principe 51
 limite thermodynamique 45, 65, 70
 postulat fondamental 4
 relation fondamentale 4
 Thomson, William (Lord Kelvin) 45
 Transition de phase
 du 1er ordre 6
 du 2nd ordre 7
 Travail 66
 Tsallis, Constantino 59
- U**
- Universalité 5, 6
- V**
- Variable conjugué 61
 Variable externe 43
 Variable interne 43
 Variance 14
 von Neumann, John 19
- W**
- Weyl (terme de) 37