

---

# Cours de physique statistique

Frédéric Faure

Université Grenoble Alpes, France

[frederic.faure@univ-grenoble-alpes.fr](mailto:frederic.faure@univ-grenoble-alpes.fr)

pour Master 1 de physique

(version : 06/01/2024)

---



# Table des matières

<b>I</b>	<b>Probabilités</b>	<b>9</b>
<b>1</b>	<b>Dynamique probabiliste</b>	<b>11</b>
1.1	Mesure de probabilité sur un espace fini $X = \{1, 2, \dots, n\}$	11
1.2	Mesure de probabilité sur $X = \mathbb{R}^d$	15
1.3	Transformations de loi de probabilité	19
1.4	Loi des grands nombres et théorème central limite	22
1.5	Entropie	28
1.6	Dynamique de Markov probabiliste et état d'équilibre	36
<b>II</b>	<b>Systèmes sans interactions</b>	<b>49</b>
<b>2</b>	<b>Mécanique ondulatoire et limite corpusculaire (Hamilton)</b>	<b>51</b>
2.1	Introduction	51
2.2	Mécanique ondulatoire et limite semi-classique de Hamilton	53
2.3	Formule de Weyl semi-classique	56
2.4	Modèle du gaz parfait : $N$ particules libres identiques dans une boîte 3D	60
<b>3</b>	<b>Modèle de Boltzmann Gibbs de l'équilibre thermodynamique</b>	<b>63</b>
3.1	Système thermodynamique et hypothèse ergodique	63
3.2	Exemple du gaz parfait : entropie et équation d'états	69
3.3	Échanges entre deux systèmes thermodynamiques	72
3.4	Exemples de systèmes physique en contact avec un réservoir	78
3.5	Le principe variationnel	84
<b>4</b>	<b>Statistiques quantiques</b>	<b>89</b>
4.1	Introduction	89
4.2	Définition des Fermions et Bosons	95
4.3	Distribution de Bose-Einstein et de Fermi-Dirac	98
4.4	Interprétations et applications de la distribution de Bose-Einstein	101
4.5	Interprétations et applications de la distribution de Fermi-Dirac	108

<b>III</b>	<b>Systèmes en interactions</b>	<b>113</b>
<b>5</b>	<b>Phases et transitions de phases</b>	<b>115</b>
5.1	Introduction . . . . .	115
5.2	Modèle d'Ising du ferromagnétisme . . . . .	117
5.3	Modèle BCS : transition de phase en supraconductivité . . . . .	124
<b>A</b>	<b>Formulaire</b>	<b>143</b>
A.1	Série géométrique . . . . .	143

*Remarque 0.0.1.* Dans la version électronique de ce document (format pdf), les couleurs sur le texte sont souvent des liens vers des pages de [wikipedia](#) pour avoir plus d'informations ou vers d'autres documents ou vidéos.

Les paragraphes avec la mention (\*) sont complémentaires et non présentés en séance de cours.



# Introduction

La physique statistique concerne la description du **comportement collectif et émergent** de plusieurs particules en interactions (ou plusieurs degrés de libertés, comme des ondes).

Le phénomène universel sous-jacent est **la loi des grands nombres** qui explique que la somme de plusieurs comportements individuels aléatoires, imprévisibles, chaotiques, se comporte de façon déterministe et prévisible.

**Exemple 0.0.2.** Considérons un flocon de neige tombant sur le sol depuis un nuage. Sa position au sol est imprévisible. Cependant, si plusieurs flocons tombent, après quelques heures, le sol est recouvert d'une épaisseur uniforme de neige. Cette épaisseur, ainsi que sa vitesse de croissance peut être prévisible. Cela vient du fait que la chute de chaque flocon est statistiquement identique aux autres et indépendante des autres.

## Historique

Références : [Diu \*et al.\* \(1989, p.201\)](#).

Les lois de la thermodynamique ont été découvertes au XIXème siècle. Pour un système donné (ex. un gaz, un fluide dans une enceinte) ces lois sont basées seulement sur une fonction  $S(E, V, N)$  appelée entropie qui dépend de paramètres conservés globalement : l'énergie totale  $E$  de ce système, le volume total  $V$ , le nombre de particules total  $N$ . Les paramètres physique importants de ce système seront donnés par les variations de cette fonction (ses dérivées), qui respectivement sont appelées pour des raison historiques, la température  $T = \left(\frac{\partial S}{\partial E}\right)^{-1}$ , la pression  $p = T \left(\frac{\partial S}{\partial V}\right)$ , le potentiel chimique  $\mu = -T \left(\frac{\partial S}{\partial N}\right)$ .

A la fin du XIXème siècle, la physique statistique a expliqué comment la thermodynamique est une théorie émergente à l'échelle macroscopique, à partir d'une description microscopique : les multiples degrés de libertés des particules qui composent le gaz et interagissent de façon complexes et chaotique. Avec la physique statistique, on a obtenu une explication de l'expression de la fonction entropie  $S(E, V, N)$ . C'est le sujet de ce cours.

## Références et liens conseillés :

- Livres de physique sur la physique statistique : [Diu \*et al.\* \(1989\)](#); [Balian \(2007\)](#), [Couture & Zitoun \(1992\)](#)
- Livres plus avancés : [Pottier \(2000\)](#).

- Livres de mathématiques sur la physique statistique et les probabilités : [Friedli & Velenik \(2017\)](#), [Cover \(1999\)](#).

Première partie

Probabilités



# Chapitre 1

## Dynamique probabiliste

**Introduction :** La physique statistique est basée sur une étude probabiliste des phénomènes collectifs. Dans ce chapitre on étudie quelques notions fondamentales en théorie des probabilités, qui sont indispensables et sont à la base de la physique statistique, mais qui ont des applications plus générales.

### 1.1 Mesure de probabilité sur un espace fini $X = \{1, 2, \dots, n\}$

En physique, la notion de “**probabilité**” et de “variables aléatoire” est une confrontation entre un grand nombre de données expérimentales de nature “aléatoire” et un modèle théorique appelé “loi de probabilité” qui permet d’extraire un comportement prédictif. On présente d’abord les données expérimentales, appelées “suite empirique”.

Dans ce premier paragraphe, pour simplifier, on va considérer un espace fini à  $n \geq 1$  éléments

$$X = \{1, 2, \dots, n\}$$

on dit qu’il est de **cardinal**  $n = \#X$ . On appellera  $X$  l’espace des données ou espace des événements.

**Définition 1.1.1.** Une **suite empirique** (ou données expérimentales) est une suite

$$(x_t)_{t \in \mathbb{N}} = (x_0, x_1, x_2, \dots)$$

où pour  $t \in \mathbb{N}$ ,

$$x_t \in X = \{1, 2, \dots, n\}.$$

*Notation 1.1.2.* La restriction de la suite aux  $t$  premiers éléments sera notée

$$x_{[0,t-1]} := (x_0, x_1, x_2, \dots, x_{t-1}).$$

**Exemple 1.1.3.** “au jeu de dé”,  $X = \{1, 2, \dots, 6\}$  et un tirage pourrait donner par exemple la suite

$$(x_t)_t = (1, 5, 4, 3, 2, 1, 3, \dots)$$

Noter que en pratique cette suite n'est pas infinie, et que chaque valeur de cette suite résulte de phénomènes physique complexes qui sont le mouvement du dé, ses rebonds, l'amortissement, etc.

**Exemple 1.1.4. “suite de Benford”.**  $X = \{1, 2, \dots, 9\}$ . Par définition, pour  $t \in \mathbb{N}$ ,

$$x_t := \text{premier chiffre de } 2^t \text{ en base 10.} \quad (1.1.1)$$

On calcule au préalable la suite

$$(2^t)_{t \in \mathbb{N}} = (1, 2, 4, 8, 16, 32, 64, 128, 256, \dots)$$

En prenant les premier chiffre, on déduit la suite empirique de Benford

$$(x_t)_{t \in \mathbb{N}} = (1, 2, 4, 8, 1, 3, 6, 1, 2, \dots)$$

Remarquer que cette suite résulte d'une formule déterministe simple. On expliquera le comportement de cette suite ci-dessous en section 1.1.14.

**Définition 1.1.5.** Une **observable** (ou fonction test) est une fonction

$$f : X \rightarrow \mathbb{R} \text{ (ou } \mathbb{C} \text{)}$$

et la **moyenne empirique** de  $f$  pour une suite empirique  $(x_t)_t$  est la série

$$\langle f \rangle_{x_{[0, t-1]}} := \frac{1}{t} \sum_{t'=0}^{t-1} f(x_{t'}), \quad t \in \mathbb{N}. \quad (1.1.2)$$

**Exemple 1.1.6.** Par exemple pour une valeur  $k \in X$  donnée, la fonction

$$f_k(x) = \delta_{x=k} = \begin{cases} 1 & \text{si } x = k \\ 0 & \text{sinon} \end{cases} \quad (1.1.3)$$

donnera une moyenne empirique qui est la proportion de valeur  $k$  parmi les  $t$  premiers éléments :

$$\langle f \rangle_{x_{[0, t-1]}} = \frac{1}{t} \# \{x_{t'} = k, \quad t' \in \{0, \dots, t-1\}\} \in [0, 1].$$

Avec ces “fonctions de base”  $(f_k)_{k \in X}$  on peut construire pour chaque instant  $t$ , un histogramme

$$H_t : \begin{cases} X & \rightarrow \mathbb{R}^+ \\ k & \rightarrow \langle f \rangle_{x_{[0, t-1]}} \end{cases}.$$

Jusqu'à présent on a parlé de suite empirique et moyenne empirique. On va maintenant parler de mesures de probabilité et moyenne statistique ou espérance. Ensuite on verra comment ces deux notions sont reliées.

**Définition 1.1.7.** Une **mesure de probabilité** (ou **variable aléatoire** ou **loi de probabilité**) est une fonction positive

$$p : \begin{cases} X & \rightarrow \mathbb{R}^+ \\ x & \rightarrow p(x) \end{cases}$$

telle que

$$\sum_{x \in X} p(x) = 1. \quad (1.1.4)$$

Pour une fonction (observable)  $f : X \rightarrow \mathbb{R}$ , on définit l'**espérance** (ou **moyenne statistique**) par rapport à la mesure  $p$

$$\mathbb{E}_p(f) := \sum_{x \in X} p(x) f(x), \quad (1.1.5)$$

parfois (et souvent en physique) notée  $\langle f \rangle_p = \mathbb{E}_p(f)$ .

*Remarque 1.1.8.* On notera  $\mathcal{P}(X)$  l'**ensemble des mesures de probabilité** sur  $X$ . Il sera utile de considérer la fonction  $p \in \mathcal{P}(X)$  comme un vecteur de composantes  $(p(x))_{x \in X}$  et de même pour la fonction  $f$ , de composantes  $(f(x))_{x \in X}$ . En introduisant le vecteur  $\mathbf{1} = (1, 1, \dots, 1)$ , et le **produit scalaire**  $l^2$  entre deux vecteurs  $u, v$  défini par

$$\langle u|v \rangle := \sum_{x \in X} \bar{u}_x v_x,$$

on observe que la relation de normalisation (1.1.4) s'écrit

$$\langle \mathbf{1}|p \rangle = 1$$

et la définition (1.1.5) s'écrit

$$\mathbb{E}_p(f) = \langle p|f \rangle.$$

La définition suivant confronte les données expérimentales et la loi de probabilité :

**Définition 1.1.9.** On dit qu'une loi de probabilité  $p$  **modélise** une suite empirique  $(x_t)_{t \in \mathbb{N}}$  si pour toute fonction (observable)  $f : X \rightarrow \mathbb{R}$ ,

$$\lim_{t \rightarrow \infty} \langle f \rangle_{x_{[0, t-1]}} = \mathbb{E}_p(f). \quad (1.1.6)$$

On dit aussi que la suite  $(x_t)_t$  est une **suite typique** pour la mesure  $p$ .

*Remarque 1.1.10.* Comme la formule (1.1.6) est linéaire en  $f$ , il suffit de la vérifier pour des fonctions de base comme les fonctions  $f_k$  en (1.1.3).

**Exemple 1.1.11.** (que vous pouvez réaliser vous même). Pour le jeu de dé,  $X = \{1, 2, \dots, 6\}$ , pour un “dé parfait” c’est à dire bien équilibré, on peut supposer que  $p(x) = \frac{1}{6}$  ne dépend pas de  $x$  et modélise toute suite empirique (suite expérimentale que l’on peut faire). On vérifie que  $\sum_{x \in X} p(x) = 1$ . Si on prend la fonction

$$f(x) = \delta_{x=1} + \delta_{x=2}$$

alors la moyenne empirique  $\langle f \rangle_{x_{[0,t-1]}}$  est la proportion des chiffres 1 et 2 parmi les  $t$  premiers lancés et  $\mathbb{E}(f) = \frac{2}{6} = \frac{1}{3}$ . On prédit donc que  $\lim_{t \rightarrow \infty} \langle f \rangle_{x_{[0,t-1]}} = \frac{1}{3}$ .

**Exemple 1.1.12.** La loi de probabilité  $p(x) = \delta_{x=1}$  modélise des suites empiriques où “presque tous les évènements” sont égaux à 1. On peut appeler une telle loi, une loi “presque-déterministe”. (Par définition, une suite suivant une “loi déterministe” aurait dans ce cas, tous ses évènements égaux à 1.)

**Exemple 1.1.13.** La suite empirique  $(x_t)_t = \{0, 1, 0, 1, 0, 1, \dots\}$  de structure périodique est modélisée par la loi de probabilité uniforme  $p(0) = p(1) = \frac{1}{2}$ . Noter que cette suite ne semble pas aléatoire.

**Exemple 1.1.14.** La suite de Benford (1.1.1) est modélisée par la mesure de probabilité sur  $X = \{1, 2, \dots, 9\}$ ,

$$p(x) = \frac{1}{\ln 10} \ln \left( 1 + \frac{1}{x} \right).$$

Ainsi on obtient que  $p(1) = 30\%$ ,  $p(2) = 17\%$ , ...,  $p(9) = 4\%$ . Pour la preuve, voir [Faure \(2014b\)](#). L’explication vient du fait que la mesure est uniforme sur une échelle logarithmique :



Voir [vidéo de Arte sur la loi de Benford](#).

**Exemple 1.1.15.** (\*) Voici une suite empirique qui n’est modélisée par aucune de mesure de probabilité.  $X = \{0, 1\}$ ,

$$(x_t)_{t \in \mathbb{N}} = \left\{ \underbrace{1}_{n_1=1}, \underbrace{0, 0}_{n_2=2}, \underbrace{1, \dots, 1}_{n_3=6}, \underbrace{0, \dots, 0}_{n_4=18}, \underbrace{1, \dots, 1}_{n_5}, \dots \right\}$$

avec

$$n_j = 2 \sum_{k=1}^{j-1} n_k.$$

*Démonstration.* Soit  $N_j = \sum_{k=1}^j n_k$ . On va montrer que

$$\text{Si } j \text{ pair alors } \langle x \rangle_{N_j} \leq \frac{1}{3},$$

$$\text{Si } j \text{ impair alors } \langle x \rangle_{N_j} \geq \frac{2}{3},$$

ce qui implique que  $\langle x \rangle_t = \langle x \rangle_{x_{[0,t-1]}}$  n'a pas de limite pour  $t \rightarrow \infty$  et contredit (1.1.6). En effet, si  $j$  est pair alors

$$\begin{aligned} \langle x \rangle_{N_j} &= \frac{1}{N_j} (n_1 + n_3 + \dots + n_{j-1}) \leq \frac{1}{\left(\sum_{k=1}^{j-1} n_k + n_j\right)} \left(\sum_{k=1}^{j-1} n_k\right) \\ &= \frac{1}{\left(\frac{1}{2}n_j + n_j\right)} \left(\frac{1}{2}n_j\right) = \frac{1}{3}. \end{aligned}$$

et si  $j$  est impair alors

$$\langle x \rangle_{N_j} = \frac{1}{N_j} (n_1 + n_3 + \dots + n_j) \geq \frac{n_j}{N_j} = \frac{n_j}{\left(\frac{1}{2}n_j + n_j\right)} = \frac{2}{3}.$$

□

## 1.2 Mesure de probabilité sur $X = \mathbb{R}^d$

En physique les phénomènes sont très souvent modélisés par des points dans des espaces  $X$  admettant des coordonnées  $x = (x_1, x_2, \dots, x_d) \in \mathbb{R}^d$ . On dit que

$$X = \mathbb{R}^d$$

est un espace réel de dimension  $d$  (finie) (plus généralement on peut considérer une **variété différentielle** réelle de dimension  $d$ ). Par exemple la position d'une particule est  $x = (x_1, x_2, x_3) \in \mathbb{R}^3$ , l'état de la particule en mécanique du point est  $(x_1, x_2, x_3, p_1, p_2, p_3) \in \mathbb{R}^6$  où  $p = (p_1, p_2, p_3)$  sont les coordonnées des l'impulsion, etc.

On note

$$dx = dx_1 dx_2 \dots dx_d$$

l'élément de volume associé à ces coordonnées sur cet espace  $X$ . La mesure  $dx$  est appelée la **mesure de Lebesgue** et servira de "mesure de référence". Dans la suite on utilisera comme exemple la droite réelle  $X = \mathbb{R}$  avec la mesure de Lebesgue  $dx$ .

L'espace  $X$  contient donc un nombre infini de points  $x$  et **ce qui change avec la section précédente est que les sommes  $\sum_{x \in X} \dots$  sont remplacées par des intégrales  $\int_{x \in X} \dots dx$ .** On va préciser cela, en reprenant les définitions du paragraphe précédent.

On rappelle que si  $f, g : X \rightarrow \mathbb{C}$  sont deux fonctions dérivables à support compact (i.e. nulles hors d'un intervalle, et on note  $f, g \in C_c^\infty(X)$ ) alors on définit leur **produit scalaire**  $L^2$  par

$$\langle f|g \rangle_{L^2(X, dx)} := \int_X \overline{f(x)} g(x) dx. \quad (1.2.1)$$

**Définition 1.2.1.** Une **mesure de probabilité** (ou **variable aléatoire** ou **loi de probabilité**) est une mesure  $p(x) dx$  avec la fonction positive (ou distribution positive) appelée **densité de probabilité**

$$p : \begin{cases} X & \rightarrow \mathbb{R}^+ \\ x & \rightarrow p(x) \end{cases}$$

telle que

$$\int_X p(x) dx = 1. \quad (1.2.2)$$

(ce qui s'écrit aussi  $\langle \mathbf{1}|p \rangle = 1$  avec la fonction  $\mathbf{1}(x) = 1, \forall x$ ). Pour une fonction (appelée observable ou fonction test)  $f : X \rightarrow \mathbb{R}, f \in C_c^\infty(X)$ , on définit l'**espérance de  $f$**  (ou **moyenne statistique**)

$$\mathbb{E}_p(f) := \int_X p(x) f(x) dx = \langle p|f \rangle, \quad (1.2.3)$$

aussi notée en physique  $\langle f \rangle_p = \mathbb{E}_p(f)$ .

*Remarque 1.2.2.* On notera  $\mathcal{P}(X)$  l'**ensemble des mesures de probabilité** sur  $X$ . Attention à la notation et au vocabulaire :  $p(x) dx$  est appelée la mesure de probabilité et le pré-facteur  $p(x)$  est appelé densité de probabilité. On verra plus tard (au chapitre loi induite) l'utilité de garder le symbole  $dx$  dans la mesure lorsque l'on fait des changements de variable.

*Remarque 1.2.3.* Attention une mesure de probabilité n'est pas toujours une fonction continue, c'est en général ce qui est appelé une **distribution**  $p$  sur  $X$ , vérifiant  $\forall f \in C_c^\infty(X, \mathbb{R}^+)$ , on a  $\langle f|p \rangle \geq 0$  et  $\langle \mathbf{1}|p \rangle = 1$ . Voir ci-dessous l'exemple de la distribution de Dirac qui ne correspond pas à une fonction.

La définition suivante confronte les données expérimentales et la loi de probabilité. Une suite empirique  $(x_t)_{t \in \mathbb{N}}$  est une suite de valeurs  $x_t \in X$ .

**Définition 1.2.4.** On dit qu'une loi de probabilité  $p$  **modélise** une suite empirique  $(x_t)_{t \in \mathbb{N}}$  si pour toute fonction (observable)  $f : X \rightarrow \mathbb{R}$ ,

$$\lim_{t \rightarrow \infty} \langle f \rangle_{x_{[0, t-1]}} = \mathbb{E}_p(f). \quad (1.2.4)$$

On dit aussi que la suite  $(x_t)_t$  est une **suite typique** pour la mesure  $p$ . (Définition précise en 1.5.6).

**Exemple 1.2.5.** Sur  $X = \mathbb{R}$ , la densité de probabilité  $p(x) = \frac{1}{\pi^{1/2}} e^{-\frac{1}{2}x^2}$  donne la mesure de probabilité  $p(x) dx$  appelé loi normale  $\mathcal{N}(0, 1)$ . Plus généralement, pour une variable réelle aléatoire  $x \in \mathbb{R}$ , la **loi normale**  $\mathcal{N}(\mu, \sigma^2)$  de paramètres  $\mu \in \mathbb{R}$  (moyenne),  $\sigma^2 > 0$  (variance), est la **mesure de probabilité**

$$p(x) dx = \frac{1}{Z} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx$$

avec la "constante de normalisation"  $Z > 0$  tel que  $\int_{\mathbb{R}} p(x) dx = 1$ .

**Exemple 1.2.6.** Sur  $X = \mathbb{R}$ , et pour  $a < b$ , la densité de probabilité  $p(x) = \frac{1}{(b-a)}$  si  $x \in [a, b]$  et  $p(x) = 0$  sinon, donne la mesure de probabilité  $p(x) dx$  appelé "**loi uniforme** sur l'intervalle  $[a, b]$ ".

**Exemple 1.2.7.** Sur  $X = \mathbb{R}$ , si on considère la fonction (observable)  $f(x) = 1$  si  $x \in [a, b]$  et  $f(x) = 0$  sinon, appelée "fonction caractéristique ou **fonction indicatrice** de l'intervalle  $[a, b]$ ", alors

$$\mathbb{E}_p(f) = \int_a^b p(x) dx = \text{"Proba}(x \in [a, b])"$$

**Exemple 1.2.8.** Sur  $X = \mathbb{R}$ , pour un point  $x_0 \in \mathbb{R}$ , la **mesure de Dirac (ou distribution de Dirac)**  $\delta_{x_0}$  est définie par, pour toute fonction test  $f \in C_c^\infty(\mathbb{R})$ ,

$$\langle \bar{f} | \delta_{x_0} \rangle = f(x_0) = \int f(x) \delta_{x_0}(x) dx,$$

Remarquer que  $\langle \mathbf{1} | \delta_{x_0} \rangle = 1$ , donc  $\delta_{x_0}$  est une mesure de probabilité.

Voici quelques exemples importants utilisés en probabilité, de fonctions  $f$  et des moyennes associées  $\mathbb{E}_p(f)$  (défini en (1.2.3)).

**Définition 1.2.9.** Pour la fonction  $f(x) = x$ , la moyenne  $\mathbb{E}_p(f)$  de  $f$  s'appelle la **moyenne** de la variable aléatoire  $x$  :

$$\mu = \mathbb{E}_p(x) = \int xp(x) dx. \quad (1.2.5)$$

Pour la fonction  $f(x) = (x - \mathbb{E}_p(x))^2$ , la moyenne  $\mathbb{E}_p(f)$  de  $f$  s'appelle la **variance** de la variable aléatoire  $x$  :

$$\begin{aligned} \text{Var}(x) &:= \mathbb{E}_p((x - \mathbb{E}_p(x))^2) = \int (x - \mathbb{E}_p(x))^2 p(x) dx & (1.2.6) \\ &= \int (x^2 - 2x\mathbb{E}_p(x) + \mathbb{E}_p(x)^2) p(x) dx \\ &= \mathbb{E}_p(x^2) - 2\mathbb{E}_p(x)^2 + \mathbb{E}_p(x)^2 \\ &= \mathbb{E}_p(x^2) - \mathbb{E}_p(x)^2 \end{aligned}$$

et

$$\sigma = (\text{Var}(x))^{1/2} \quad (1.2.7)$$

s'appelle l'**écart type**.

Pour la fonction  $f(x) = (x - \mathbb{E}_p(x))^m$ , la moyenne  $\mathbb{E}_p(f)$  de  $f$  s'appelle le **moment d'ordre  $m$** .

**Définition 1.2.10.** A partir de données empiriques (expérimentales)  $(x_t)_t = (x_0, x_1, \dots)$ , on peut construire pour chaque valeur de  $t \in \mathbb{N}$  un **histogramme**  $H_t : k \in [1, K] \rightarrow \mathbb{R}$  définit par la moyenne empirique

$$H_t(k) = \langle f_k \rangle_{x_{[0, t-1]}}$$

où pour chaque  $k \in [1, K]$ ,  $f_k : X \rightarrow \mathbb{R}$  est une fonction (observable) telle que cette famille de fonction  $(f_k)_k$  vérifie la relation appelée "partition de l'unité" :

$$\sum_{k=1}^K f_k(x) = 1, \forall x \in X.$$

Un exemple habituel est pour  $X = \mathbb{R}$ , de choisir un "pas"  $\delta > 0$  et définir pour  $k \in \mathbb{Z}$

$$f_k(x) = \begin{cases} 1 & \text{si } x \in [k\delta, (k+1)\delta] \\ 0 & \text{sinon} \end{cases}$$

qui est la fonction caractéristique de l'intervalle  $[k\delta, (k+1)\delta]$ .

## 1.3 Transformations de loi de probabilité

On présente quelques opérations élémentaires utiles sur les lois de probabilité.

### 1.3.1 Probabilités d'évènements indépendants

La notion d'évènements indépendants est très importante et très simple : elle correspond à la géométrie d'espace produit. Voici la définition.

**Définition 1.3.1.** Si pour  $j = 1, 2$  on a deux mesures de probabilités

$$p_j : \begin{cases} X_j \rightarrow \mathbb{R}^+ \\ x \rightarrow p_j(x) \end{cases}$$

Alors sur l'espace produit  $X = X_1 \times X_2$  la mesure

$$p : \begin{cases} X_1 \times X_2 & \rightarrow \mathbb{R}^+ \\ x = (x_1, x_2) & \rightarrow p(x) = p_1(x) p_2(x) \end{cases} \quad (1.3.1)$$

est appelée **mesure produit** et modélise l'**indépendance** des variables  $x_1$  et  $x_2$ .

Au contraire, si  $p(x) \neq p_1(x_1) p_2(x_2)$  ne se factorise pas, on dit que les variables  $x_1$  et  $x_2$  sont **corrélées**. Une caractérisation (partielle) de la corrélation se fait par le **degré de corrélation** :

$$r = \frac{\text{Cov}_{1,2}}{V_1 V_2}$$

avec la **variance**, pour  $j = 1, 2$ ,

$$V_j := \mathbb{E}((x_j - \mathbb{E}(x_j))^2)$$

et la **covariance**

$$\text{Cov}_{1,2} := \mathbb{E}((x_1 - \mathbb{E}(x_1))(x_2 - \mathbb{E}(x_2)))$$

### Exercice 1.3.2.

1. Montrer que

$$V_j = \mathbb{E}((x_j - \mathbb{E}(x_j))^2) = \mathbb{E}(x_j^2) - \mathbb{E}(x_j)^2$$

$$\text{Cov}_{1,2} := \mathbb{E}((x_1 - \mathbb{E}(x_1))(x_2 - \mathbb{E}(x_2))) = \mathbb{E}(x_1 x_2) - \mathbb{E}(x_1) \mathbb{E}(x_2)$$

2. Montrer que si les variables  $x_1$  et  $x_2$  sont indépendantes alors  $\text{Cov}_{1,2} = 0$ , donc le degré de corrélation est  $r = 0$ .
3. Donner un exemple de variables non indépendantes mais avec  $r = 0$ .

### 1.3.2 Loi induite

Avant de donner une définition générale, commençons par considérer des exemples simples, parlant et utiles.

**Exemple 1.3.3.** Si pour le jeu de dés,  $x_1, x_2 \in X = \{1, 2, \dots, 6\}$  sont deux variables indépendantes de même loi  $p(x) = \frac{1}{6}$ , alors

$$p(x_1, x_2) = p(x_1)p(x_2) = \frac{1}{36}.$$

Si on considère la somme

$$S = x_1 + x_2 \in \{2, 3, \dots, 11, 12\},$$

la loi de  $S$  notée  $p_S$ , appelée loi induite est donnée par la formule générale :

$$p_S(S) = \sum_{\substack{x_1, x_2 \\ S=x_1+x_2 \text{ fixe}}} p(x_1, x_2) \quad (1.3.2)$$

Cela donne simplement par exemple d'après le schéma ci-dessous,

$$p_S(4) = 3 \times \frac{1}{36} = \frac{1}{12}, \quad p_S(7) = 6 \times \frac{1}{36} = \frac{1}{6},$$

6	7	8	9	10	11	12
5	6	7	8	9	10	11
4	5	6	7	8	9	10
3	4	5	6	7	8	9
2	3	4	5	6	7	8
1	2	3	4	5	6	7
	1	2	3	4	5	6

**Exemple 1.3.4.** Considérons deux variables aléatoires  $x_1, x_2 \in X = \mathbb{R}$  indépendantes, chacune de loi normale

$$p(x) = \frac{1}{Z} e^{-\frac{x^2}{2\sigma^2}}$$

avec un écart type  $\sigma > 0$  et  $Z > 0$  qui est une constante de normalisation. On a alors

$$p(x_1, x_2) = p(x_1)p(x_2) = \frac{1}{Z^2} e^{-\frac{1}{2\sigma^2}(x_1^2+x_2^2)} = \frac{1}{Z^2} e^{-\frac{1}{2\sigma^2}r^2} \quad (1.3.3)$$

avec les coordonnées polaires  $r, \theta$  définies par

$$x_1 = r \cos \theta, \quad x_2 = r \sin \theta,$$

on a la mesure de Lebesgue

$$dx_1 dx_2 = r d\theta dr \quad (1.3.4)$$

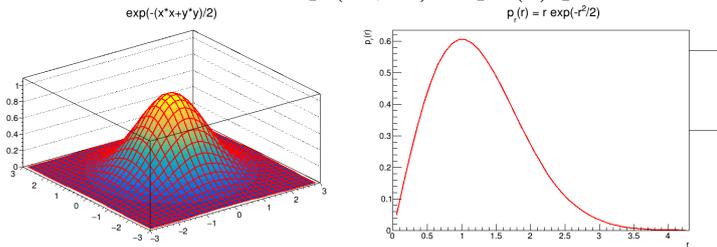
Supposons que l'on cherche la loi induite  $p_r(r)$  sur la variable  $r$ . Pour cela, on écrit par analogie avec (1.3.2) (attention, on écrit la mesure  $dr$  mais ici  $r$  est fixé)

$$\begin{aligned} p_r(r) dr &= \int_{\substack{x_1, x_2 \\ r \text{ fixe}}} p(x_1, x_2) dx_1 dx_2 \\ &= \int_{\substack{\theta \\ r \text{ fixe}}} \frac{1}{Z^2} e^{-\frac{1}{2\sigma^2}r^2} d\theta dr \\ &= \frac{2\pi}{Z^2} e^{-\frac{1}{2\sigma^2}r^2} r dr \end{aligned} \quad (1.3.5)$$

Noter que la dernière intégrale porte seulement sur la variable  $\theta$ . On obtient le résultat attendu que  $p_r(0) = 0$ , i.e. le cercle de rayon  $r = 0$  est "non probable", bien que le point origine  $(x_1, x_2) = (0, 0)$  soit le maximum de la fonction  $p(x_1, x_2)$ , i.e. le point "le plus probable". Le maximum de  $p_r(r)$  se situe à une valeur  $r_{\max} > 0$  telle que

$$\frac{dp_r}{dr} = 0 \Leftrightarrow -\frac{1}{\sigma^2} r_{\max}^2 + 1 = 0 \Leftrightarrow r_{\max} = \sigma.$$

Voici l'allure des fonctions  $p(x_1, x_2)$  et  $p_r(r)$  pour  $\sigma = 1$  :



### 1.3.2.1 (\*) Définition générale

On présente le formalisme qui permet de considérer une loi induite en général. Cela contiendra les exemples précédents. Supposons donné une loi  $p_X(x) dx$  sur  $x \in X$  et une application  $C^\infty$

$$\pi : X \rightarrow Y.$$

On note  $\pi^\circ$  l'opérateur de composition (ou **pull-back**) des fonctions par  $\pi$  :

$$\pi^\circ \begin{cases} C^\infty(Y) & \rightarrow C^\infty(X) \\ u & \rightarrow u \circ \pi \end{cases}$$

On suppose qu'il y a des mesures  $dx, dy$  respectivement sur les espaces  $X, Y$ , cela permet de définir les produits scalaires  $L^2(X, dx)$  et  $L^2(Y, dy)$  par (1.2.1). On définit alors l'opérateur  **$L^2$ -adjoint**

$$(\pi^\circ)^\dagger : C^\infty(X) \rightarrow C^\infty(Y)$$

par

$$\langle u | (\pi^\circ)^\dagger v \rangle_{L^2(X, dx)} = \langle \pi^\circ u | v \rangle_{L^2(Y, dy)},$$

et finalement la **loi induite**  $p_Y(y)$  sur  $Y$  (aussi appelée **mesure push-forward**) est définie par

$$p_Y := (\pi^\circ)^\dagger p_X. \quad (1.3.6)$$

On vérifie que c'est une loi de probabilité car

$$\langle \mathbf{1} | p_Y \rangle = \langle \mathbf{1} | (\pi^\circ)^\dagger p_X \rangle = \langle \pi^\circ \mathbf{1} | p_X \rangle = \langle \mathbf{1} | p_X \rangle = 1.$$

**Exercice 1.3.5.** Dans l'exemple précédent, montrer explicitement comment la définition générale (1.3.6) donne (1.3.5).

## 1.4 Loi des grands nombres et théorème central limite

La loi des grands nombres est fondamentale en physique et montre l'émergence d'un résultat "presque sûr" (presque déterministe) à partir de la contribution de phénomènes aléatoires. Le théorème central limite est plus fin et étudie les fluctuations autour de ce résultat.

**Question 1.4.1. "Somme de variables aléatoires".** Si  $(x_t)_{t \in \mathbb{N}}$  sont des variables aléatoires indépendantes de même mesure de probabilité  $p(x) dx$  (i.e. "identiquement distribuées"), on souhaite décrire pour  $t \gg 1$ , la loi  $P_{S_t}(s) ds$  de la somme

$$S_t = \sum_{t'=0}^{t-1} x_{t'} = x_0 + x_1 + \dots + x_{t-1}.$$

**Exemple 1.4.2.** Voici quelques exemples concrets d'applications :

- Ce problème concerne par exemple la "**marche aléatoire**",  $x_t$  est un déplacement aléatoire à chaque seconde d'un point et  $S_t$  est la somme des déplacements de ce point, donc le déplacement total à la date  $t$ .
- En bourse, le cours d'une valeur subit des variations aléatoires  $x_t$  et la cours total  $S_t$  est la somme de ces variations.

- Lorsqu'il neige, l'épaisseur de neige au sol  $S_t$  est la somme des contributions de chaque flocon  $x_t$  qui se pose de façon aléatoire.
- L'énergie d'un ensemble de particules est la somme des énergies de chaque particule.

Voici maintenant un exemple illustratif qui montre les phénomènes que l'on voudra énoncer et expliquer.

**Exemple 1.4.3.** Considérons la variable aléatoire  $x \in \mathbb{R}$  avec la loi discrète

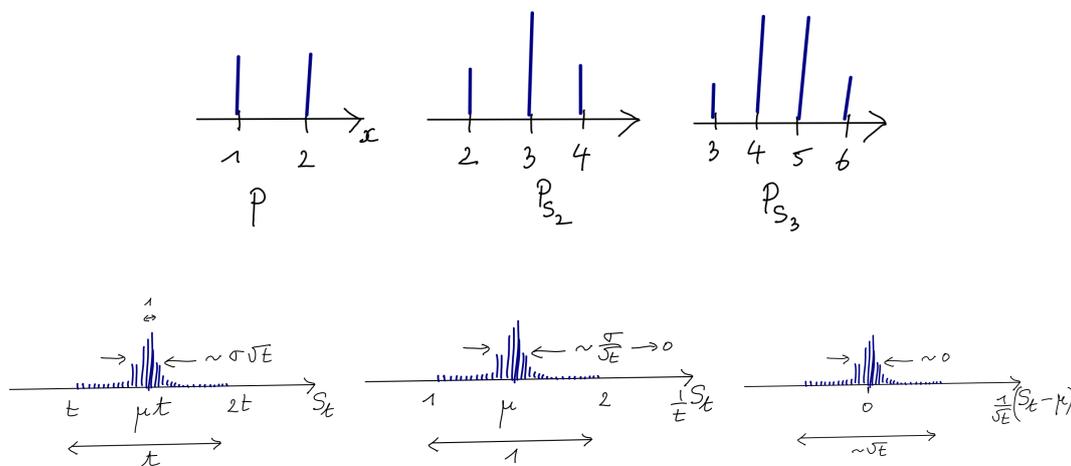
$$p(x) = \frac{1}{2} (\delta_1(x) + \delta_2(x)).$$

Cette loi signifie que sa valeur est 1 avec probabilité 1/2 ou 2 avec probabilité 1/2. Alors on déduit que pour la somme  $S_2 = x_0 + x_1$  formée par deux variables identiques, indépendantes de même loi  $p(x)$ , on a

$$P_{S_2} = \frac{1}{4} (\delta_2 + 2\delta_3 + \delta_4)$$

et de même pour la somme  $S_3 = x_0 + x_1 + x_2$ ,

$$P_{S_3} = \frac{1}{8} (\delta_3 + 3\delta_4 + 3\delta_5 + \delta_6).$$



Observations : on pourrait continuer avec les formules de combinatoire et observer que la distribution  $P_{S_t}$  pour  $t \gg 1$ , est de support dans l'intervalle  $[t, 2t]$ , avec une concentration de pics en  $s = \mu t$  avec la moyenne

$$\mu = \mathbb{E}_p(x) = \int xp(x) dx = 1\frac{1}{2} + 2\frac{1}{2} = \frac{3}{2},$$

et plus précisément, que ce pic a une forme Gaussienne, centré en  $\mu t$  et de largeur  $\sigma\sqrt{t}$  avec l'écart type (1.2.7)

$$\sigma = (\mathbb{E}_p((x - \mu)^2))^{1/2}.$$

Si regarde de plus près, cette gaussienne est en fait une contribution de pics de Dirac, distants de 1. Hors de ce groupe central les évènements sont extrêmement rares et sont décrit par la **théorie des grandes déviations**<sup>1</sup>. Voir **ces vidéos**.

On peut faire la même observation à partir d'une loi  $p(x)$  continue, et bien que dans ce cas on ne puisse pas faire d'étude combinatoire, on observe les mêmes phénomènes que précédemment.

Voici un énoncé précis et une preuve des ces phénomènes observés.

**Théorème 1.4.4.** « *loi des grands nombres* ». Si  $x_1, x_2, \dots$  sont des variables aléatoires indépendantes de même mesure de probabilité  $p(x) dx$ , d'espérance (i.e. moyenne) finie  $\mu = \mathbb{E}(x) < \infty$  et de variance finie  $\sigma = \sqrt{\mathbb{E}(x - \mu)^2} < \infty$  alors la somme  $S_t = x_0 + x_1 + \dots + x_{t-1}$  renormalisée par  $1/t$

$$R_t = \frac{1}{t} S_t$$

a une mesure de probabilité  $p_{R_t}(s)$  qui **converge** pour  $t \rightarrow \infty$  vers la mesure de Dirac en  $\mu$  (i.e. loi déterministe)

$$p_{R_t}(s) ds \xrightarrow[t \rightarrow \infty]{*} \delta_\mu \quad (1.4.1)$$

indépendamment de la mesure initiale  $p$ , au sens où pour toute observable  $f \in C_c^\infty(X)$ ,

$$\langle f | P_{R_t} \rangle \xrightarrow[t \rightarrow \infty]{} \langle f | \delta_\mu \rangle = \bar{f}(\mu) \quad (1.4.2)$$

*Remarque 1.4.5.* Pour avoir le résultat (1.4.2) il a été important de considérer non pas la variable  $S_t$  mais la variable  $\frac{1}{t}S_t$  qui est divisée par  $t$ , comme observé sur la figure .

*Remarque 1.4.6.* La convergence (1.4.2) s'appelle “**convergence au sens des distributions**”, ou “**convergence faible**”, ou “**convergence en loi**” et est notée avec une étoile \* comme en (1.4.1).

*Démonstration.* On considère pour simplifier l'espace des évènements  $X = \mathbb{R}$ . Pour démontrer (1.4.2), on va voir qu'il est très naturel et utile d'utiliser la transformée de Fourier d'une fonction  $f(x)$  définie par

$$(\mathcal{F}f)(k) := \frac{1}{\sqrt{2\pi}} \int e^{-ikx} f(x) dx = \frac{1}{\sqrt{2\pi}} \langle e^{ik \cdot} | f \rangle_{L^2}. \quad (1.4.3)$$

1. Références conseillées : introduction **An Introduction to Large Deviations for. Teletraffic Engineers. John T. Lewis**<sup>1</sup>, un livre **Den Hollander** (2008, chap1, chap2, section1.), plus développé : **The large deviation approach to statistical mechanics** by Hugo Touchette.

où  $k \in \mathbb{R}$  est appelé variable de Fourier ou vecteur d'onde en physique, c'est une fréquence spatiale (ou vecteur cotangent en géométrie différentielle).

On rappelle le **théorème de Parseval** ou de **Plancherel** que pour toute fonctions  $f, g$  on a

$$\langle f|g \rangle_{L^2} = \langle \mathcal{F}f | \mathcal{F}g \rangle_{L^2}. \quad (1.4.4)$$

Alors

$$\langle f|p_{R_t} \rangle \stackrel{(1.4.4)}{=} \langle \mathcal{F}f | \mathcal{F}p_{R_t} \rangle$$

Or  $s = \frac{1}{t}(x_0 + x_1 + \dots + x_{t-1})$  et comme en (1.3.5) on écrit

$$p_{R_t}(s) ds = \int_{s \text{ fixe}} p(x_0, \dots, x_{t-1}) dx_0 \dots dx_{t-1}.$$

D'après l'indépendance

$$p(x_0, \dots, x_{t-1}) \stackrel{(1.3.1)}{=} p(x_0) \dots p(x_{t-1}).$$

Alors pour  $k \in \mathbb{R}$  fixé,

$$\begin{aligned} (\mathcal{F}p_{R_t})(k) &\stackrel{(1.4.3)}{=} \frac{1}{\sqrt{2\pi}} \int e^{-iks} p_{R_t}(s) ds \\ &= \frac{1}{\sqrt{2\pi}} \int e^{-ik\frac{1}{t}(x_0+\dots+x_{t-1})} p(x_0) \dots p(x_{t-1}) dx_0 \dots dx_{t-1} \\ &= \frac{1}{\sqrt{2\pi}} \left( \int e^{-i\frac{k}{t}x} p(x) dx \right)^t \\ &\stackrel{(1.4.3)}{=} \frac{1}{\sqrt{2\pi}} \left( \sqrt{2\pi} (\mathcal{F}p) \left( \frac{k}{t} \right) \right)^t \end{aligned} \quad (1.4.5)$$

Or pour  $t \rightarrow \infty$ , on a  $\frac{k}{t} \rightarrow 0$  donc on considère le **développement de Taylor** qui s'écrit en général  $f(x) = f(0) + xf'(0) + O(x^2)$ , ici

$$(\mathcal{F}p) \left( \frac{k}{t} \right) = (\mathcal{F}p)(0) + \left( \frac{k}{t} \right) (\mathcal{F}p)'(0) + O\left( \frac{k^2}{t^2} \right).$$

On calcule

$$\begin{aligned} (\mathcal{F}p)(0) &\stackrel{(1.4.3)}{=} \frac{1}{\sqrt{2\pi}} \int e^{-i0x} p(x) dx \stackrel{(1.2.2)}{=} \frac{1}{\sqrt{2\pi}} \\ (\mathcal{F}p)'(0) &\stackrel{(1.4.3)}{=} \frac{1}{\sqrt{2\pi}} \int (-ix) e^{-i0x} p(x) dx \stackrel{(1.2.2)}{=} \frac{-i}{\sqrt{2\pi}} \int xp(x) dx \stackrel{(1.2.5)}{=} \frac{-i\mu}{\sqrt{2\pi}}. \end{aligned}$$

Donc

$$(\mathcal{F}p_{R_t})(k) \stackrel{(1.4.5)}{=} \frac{1}{\sqrt{2\pi}} \left( 1 - i\mu \left( \frac{k}{t} \right) + O\left( \frac{k^2}{t^2} \right) \right)^t \quad (1.4.6)$$

Ensuite on utilise que  $x^t = (e^{\ln x})^t = e^{t \ln x}$  et que  $\ln(1+x) = x + O(x^2)$ ,

$$\begin{aligned} (\mathcal{F}p_{R_t})(k) &\stackrel{(1.4.5)}{=} \frac{1}{\sqrt{2\pi}} \exp\left(t \ln\left(1 - i\mu\left(\frac{k}{t}\right) + O\left(\frac{k^2}{t^2}\right)\right)\right) \\ &= \frac{1}{\sqrt{2\pi}} \exp\left(t\left(-i\mu\left(\frac{k}{t}\right) + O\left(\frac{k^2}{t^2}\right)\right)\right) \\ &= \frac{1}{\sqrt{2\pi}} \exp\left(-i\mu k + O\left(\frac{k^2}{t}\right)\right) \end{aligned}$$

Finalement on observe que la transformée de Fourier de la mesure de Dirac est

$$(\mathcal{F}\delta_\mu)(k) \stackrel{(1.4.3)}{=} \frac{1}{\sqrt{2\pi}} \langle e^{ikx} | \delta_\mu(x) \rangle = \frac{1}{\sqrt{2\pi}} e^{-ik\mu}$$

donc à  $k \in \mathbb{R}$  fixé,

$$(\mathcal{F}p_{R_t})(k) \xrightarrow{t \rightarrow \infty} \frac{1}{\sqrt{2\pi}} \exp(-i\mu k) = (\mathcal{F}\delta_\mu)(k).$$

De cela on déduit (cela est vrai car  $(p_{R_t})_t$  est une suite de mesures, c'est le **théorème de Levy**) que  $\forall \tilde{f} \in \mathcal{S}(\mathbb{R}_k)$ ,  $\langle \tilde{f} | \mathcal{F}p_{R_t} \rangle \xrightarrow{t \rightarrow \infty} \langle \tilde{f} | \mathcal{F}\delta_\mu \rangle$  et donc  $\forall f \in \mathcal{S}(\mathbb{R})$ ,

$$\langle f | p_{R_t} \rangle \stackrel{(1.4.4)}{=} \langle \mathcal{F}f | \mathcal{F}p_{R_t} \rangle \xrightarrow{t \rightarrow \infty} \langle \mathcal{F}f | \mathcal{F}\delta_\mu \rangle \stackrel{(1.4.4)}{=} \langle f | \delta_\mu \rangle.$$

Ce qui montre (1.4.2). □

**Théorème 1.4.7.** « *théorème central limite* ». Si  $x_1, x_2, \dots$  sont des variables aléatoires indépendantes de même mesure de probabilité  $p(x) dx$ , d'espérance (i.e. moyenne) finie  $\mu = \mathbb{E}(x) < \infty$  et de variance finie  $\sigma = \sqrt{\mathbb{E}(x - \mu)^2} < \infty$  alors la somme  $S_t = x_0 + x_1 + \dots + x_{t-1}$  décalée de  $\mu t$  et renormalisée par  $\frac{1}{\sqrt{t}}$ ,

$$R_t = \frac{1}{\sqrt{t}} (S_t - t\mu)$$

a une mesure de probabilité  $p_{R_t}(s)$  qui **converge** pour  $t \rightarrow \infty$  vers la loi normale  $\mathcal{N}(0, \sigma^2)$  :

$$p_{R_t}(s) ds \xrightarrow{t \rightarrow \infty} \frac{1}{Z} e^{-\frac{1}{2}\left(\frac{s}{\sigma}\right)^2} ds \tag{1.4.7}$$

indépendamment de la mesure initiale  $p$ .

*Démonstration.* (\*) en TD. On suit la même démarche que la preuve du théorème 1.4.4. Les changements sont que à cause du décalage de  $\mu t$ , on peut supposer que la moyenne est

nulle et au lieu du facteur  $\frac{1}{t}$ , il faut mettre le facteur  $\frac{1}{\sqrt{t}}$ . Au lieu de (1.4.5), on a

$$(\mathcal{F}p_{R_t})(k) = \frac{1}{\sqrt{2\pi}} \left( \sqrt{2\pi} (\mathcal{F}p) \left( \frac{k}{\sqrt{t}} \right) \right)^t,$$

on fait un développement de Taylor à l'ordre deux (au lieu de ordre 1 dans la preuve de la loi des grands nombres)

$$(\mathcal{F}p) \left( \frac{k}{\sqrt{t}} \right) = (\mathcal{F}p)(0) + \left( \frac{k}{\sqrt{t}} \right) (\mathcal{F}p)'(0) + \frac{1}{2} \left( \frac{k}{\sqrt{t}} \right)^2 (\mathcal{F}p)''(0) + O \left( \left( \frac{k}{\sqrt{t}} \right)^3 \right).$$

On a toujours  $(\mathcal{F}p)(0) = \frac{1}{\sqrt{2\pi}}$ . La moyenne est nulle donc  $(\mathcal{F}p)'(0) = 0$ . On calcule

$$(\mathcal{F}p)''(0) \stackrel{(1.4.3)}{=} \frac{1}{\sqrt{2\pi}} \int (-ix)^2 e^{-i0x} p(x) dx \stackrel{(1.2.2)}{=} \frac{-1}{\sqrt{2\pi}} \int x^2 p(x) dx \stackrel{(1.2.5)}{=} \frac{-\sigma^2}{\sqrt{2\pi}}.$$

Ensuite on utilise que  $x^t = e^{t \ln x}$  et que  $\ln(1+x) = x + O(x^2)$ ,

$$(\mathcal{F}p_{R_t})(k) = \frac{1}{\sqrt{2\pi}} \left( 1 - \frac{1}{2} \left( \frac{k}{\sqrt{t}} \right)^2 \sigma^2 + O \left( \left( \frac{k}{\sqrt{t}} \right)^3 \right) \right)^t \quad (1.4.8)$$

$$= \frac{1}{\sqrt{2\pi}} \exp \left( t \ln \left( 1 - \frac{1}{2} \left( \frac{k}{\sqrt{t}} \right)^2 \sigma^2 + O \left( \left( \frac{k}{\sqrt{t}} \right)^3 \right) \right) \right) \quad (1.4.9)$$

$$= \frac{1}{\sqrt{2\pi}} \exp \left( t \left( -\frac{1}{2} \left( \frac{k}{\sqrt{t}} \right)^2 \sigma^2 + O \left( \left( \frac{k}{\sqrt{t}} \right)^3 \right) \right) \right) \quad (1.4.10)$$

$$= \frac{1}{\sqrt{2\pi}} \exp \left( -\frac{1}{2} k^2 \sigma^2 + O \left( \frac{k^3}{\sqrt{t}} \right) \right) \quad (1.4.11)$$

Finalement on observe que la transformée de Fourier de la Gaussienne  $G_\sigma(s) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2} \frac{s^2}{\sigma^2}}$  est

$$(\mathcal{F}G)(k) \stackrel{(1.4.3)}{=} \frac{1}{\sqrt{2\pi}} \langle e^{-iks} | G(s) \rangle = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2} k^2 \sigma^2}$$

donc

$$(\mathcal{F}p_{R_t})(k) \xrightarrow{t \rightarrow \infty} \frac{1}{\sqrt{2\pi}} \exp \left( -\frac{1}{2} k^2 \sigma^2 \right) = (\mathcal{F}G_\sigma)(k).$$

Finalement

$$\langle f | p_{R_t} \rangle \stackrel{(1.4.4)}{=} \langle \mathcal{F}f | \mathcal{F}p_{R_t} \rangle \xrightarrow{t \rightarrow \infty} \langle \mathcal{F}f | \mathcal{F}G_\sigma \rangle \stackrel{(1.4.4)}{=} \langle f | G_\sigma \rangle.$$

Ce qui montre (1.4.2). □

**Exercice 1.4.8.** On considère le modèle suivant très simplifié de chute de neige. Le temps est discrétisé et à chaque instant  $t \in \mathbb{N}$  un flocon tombe à une certaine position discrète

$j \in \{1, 2, \dots, n\}$ , selon une mesure de probabilité uniforme. On note  $S_t(j) \in \mathbb{N}$  le nombre de flocons qui sont tombés à la position  $j$  et sur l'intervalle de temps  $[0, t - 1]$ , ainsi la fonction  $S_t : j \rightarrow S_t(j)$  représente le manteau neigeux à la date  $t$ . En utilisant la loi des grands nombres et le théorème central limite, décrire pour  $j$  fixé, à différentes échelles la loi de probabilité de  $S_t(j)$  pour  $t \gg 1$ .

## 1.5 Entropie

### 1.5.1 Introduction

En termes simples on dit que l'entropie mesure le “désordre” ou “l'incertitude sur l'avenir”. L'entropie est une valeur  $S \geq 0$  sans dimension.

- Une valeur d'entropie élevée  $S > 0$  signifie qu'il y a beaucoup de désordre ou une grande imprévisibilité (par exemple l'évènement “pile” au jeu de pile ou face ou l'évènement “jour ensoleillé” pour les jours à venir à Grenoble)
- au contraire  $S \approx 0$  signifie que les évènements à venir sont prédictibles (par exemple l'évènement “jour ensoleillé” pour les jours à venir à Dakar).

Précisément, on va définir l'**entropie relative** d'une mesure de probabilité par rapport à une autre mesure de référence, comme le taux d'incertitude pour une suite typique<sup>2</sup> d'évènements issue de cette loi de probabilité. L'entropie est à la base de la physique statistique.

Voici la question à laquelle on va répondre ensuite.

**Question 1.5.1.** Soit une mesure de probabilité donnée  $p : X \rightarrow \mathbb{R}^+$ , et  $\vec{x}_t = (x_1, x_2, \dots, x_t)$  une suite de données avec  $x_j \in X$  variables aléatoires indépendantes identiquement distribuées selon la loi  $p$ . La question est d'estimer le nombre  $\mathcal{N}_t$  de telles suites typiques  $\vec{x}_t$  de longueur  $t$ , c'est à dire qui apparaissent de façon probable pour  $t \gg 1$ . Montrer que  $\mathcal{N}_t \approx e^{tS}$  avec un taux  $S \geq 0$  appelé **entropie** que l'on souhaite exprimer à partir de la loi  $p$ .

Explorons tout d'abord cette question sur des exemples.

**Exemple 1.5.2.**  $X = \{1, 2\}$  et on considère la loi de probabilité déterministe  $p(1) = 1$ ,  $p(2) = 0$ . Ainsi il y a une suite typique de longueur  $t$  qui est  $\vec{x}_t = (1, 1, \dots, 1)$ . Donc  $\mathcal{N}_t = 1 = e^{0 \times t}$ , donnant  $S = 0$ .

**Exemple 1.5.3.**  $X = \{1, 2\}$  et on considère la loi de probabilité uniforme  $p(1) = \frac{1}{2}$ ,  $p(2) = \frac{1}{2}$ . Alors à chaque étape les 2 évènements 1, 2 sont équiprobables, ainsi il y a  $\mathcal{N}_t = 2^t = e^{(\ln 2)t}$  suites typiques de longueur  $t$ , par exemple  $\vec{x}_t = (1, 2, 2, 1, \dots, 2)$ . Donc  $S = \ln 2 \approx 0.69$ .

---

2. On rappelle que l'on a défini une suite typique dans la définition 1.1.9.

**Exemple 1.5.4.** Cas plus général.  $X = \{1, 2\}$  et on considère la loi de probabilité quelconque  $p(1) = 1 - \epsilon$ ,  $p(2) = \epsilon$  avec  $0 \leq \epsilon \leq 1$ . Une suite typique de longueur  $t$  ressemble à  $\vec{x}_t = (1, 1, 1, 2, 1, \dots, 1, 2, 1)$  avec une proportion  $p(1) = 1 - \epsilon$  de 1 et  $p(2) = \epsilon$  de 2. On démontre ci-dessous en section 1.5.1.1, (voir aussi en TD avec une méthode combinatoire) que le nombre de telles suites typiques est

$$\mathcal{N}_t \approx e^{St}$$

avec

$$\begin{aligned} S &= \sum_{x \in X} p(x) \ln \frac{1}{p(x)} \\ &= p(1) \ln \frac{1}{p(1)} + p(2) \ln \frac{1}{p(2)} \\ &= (1 - \epsilon) \ln \frac{1}{1 - \epsilon} + \epsilon \ln \frac{1}{\epsilon} \end{aligned} \tag{1.5.1}$$

Observons que (utilisons que  $(\ln f)' = \frac{f'}{f}$  donc  $(\ln \frac{1}{f})' = -\frac{f'}{f}$ )

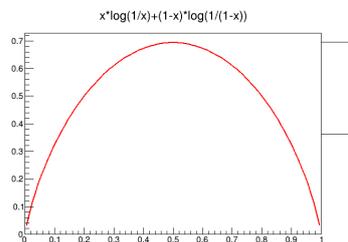
$$\frac{dS}{d\epsilon} = -\ln \frac{1}{1 - \epsilon} + (1 - \epsilon) \frac{1}{(1 - \epsilon)} + \ln \frac{1}{\epsilon} - \epsilon \frac{1}{\epsilon} = -\ln \frac{1}{1 - \epsilon} + \ln \frac{1}{\epsilon}$$

$$\frac{d^2S}{d\epsilon^2} = -\frac{1}{1 - \epsilon} - \frac{1}{\epsilon} < 0$$

ainsi la fonction  $S$  est concave et le maximum est

$$\frac{dS}{d\epsilon} = 0 \Leftrightarrow \frac{1}{\epsilon} = \frac{1}{1 - \epsilon} \Leftrightarrow \epsilon = \frac{1}{2}$$

qui correspond à la distribution uniforme donnant  $S = \ln 2$ . La dérivée diverge aux bords :

$$\frac{dS}{d\epsilon} \xrightarrow{\epsilon \rightarrow 0,1} \pm \infty.$$


*Remarque 1.5.5.* Historiquement, en 1877, Boltzmann a introduit l'entropie et la formule  $S = k \ln W$  est gravée sur sa tombe. Par rapport à ce qui précède, ici  $W$  est le nombre d'état supposés équiprobables pour un ensemble de particules de taille macroscopique ( $W = 2$  dans le jeu de pile/face), et le résultat  $\ln W$  étant cependant énorme, on rajoute en physique

une constant  $k$  très petite pour obtenir une valeur  $S = k \ln W$  de l'ordre de 1. . En 1948 Shannon ingénieur en télécommunications veut estimer combien de copies  $\mathcal{N}_t$  du même message il faut envoyer au minimum, pour compenser les modifications aléatoires des caractères lors du transfert. Il donne alors la définition et les propriétés que l'on va voir ci-dessous [Cover \(1999\)](#).

### 1.5.1.1 Preuve rapide de la formule (1.5.1)

D'après (1.3.1), la probabilité d'apparition d'une suite  $\vec{x}_t = (x_1, x_2, \dots, x_t)$  contenant des évènements indépendants de même loi  $p$  est le produit des probabilité. On va utiliser  $\ln(\cdot)$  pour transformer ce produit en somme et ensuite utiliser la loi des grands nombres (1.4.1).

$$\begin{aligned} p(\vec{x}_t) &\stackrel{(1.3.1)}{=} p(x_1) p(x_2) \dots p(x_t) = \prod_{j=1}^t p(x_j) \\ &= e^{\sum_{j=1}^t \ln p(x_j)} = e^{-tS_t} \end{aligned}$$

avec

$$S_t = \frac{1}{t} \sum_{j=1}^t \ln \frac{1}{p(x_j)}.$$

D'après la loi des grands nombres (1.4.1), la mesure de probabilité  $p_{S_t}$  de  $S_t$  vérifie

$$p_{S_t}(s) ds \xrightarrow[t \rightarrow \infty]{*} \delta_S$$

avec la moyenne

$$S := \mathbb{E} \left( \ln \left( \frac{1}{p(x)} \right) \right) = \sum_{x \in X} p(x) \ln \frac{1}{p(x)},$$

qui est l'entropie. Cela montre que pour  $t \gg 1$ , les suites typiques ont toutes la même probabilité  $p(\vec{x}_t) \approx e^{-tS}$  et toujours d'après la loi des grands nombre, cela est vrai pour presque toutes les suites donc

$$\mathcal{N}_t \times e^{-tS} \approx 1 \Leftrightarrow \mathcal{N}_t \approx e^{tS}.$$

### 1.5.2 (\*) Cas d'un espace $X$ fini

Dans cette section on considère une mesure de probabilité  $p(x)$  sur un espace fini  $X = \{1, 2, \dots, n\}$ . Rappel : on a défini en (1.1.6) que une suite empirique  $(x_t)_{t \in \mathbb{N}} \in X^{\mathbb{N}}$  est typique pour la mesure  $p$  si pour toute observable  $f : X \rightarrow \mathbb{R}$ ,

$$\lim_{t \rightarrow \infty} \langle f \rangle_{x_{[0, t-1]}} = \mathbb{E}_p(f).$$

Ici on va devoir préciser la vitesse de cette convergence :

**Définition 1.5.6.**  $(x_t)_{t \in \mathbb{N}} \in X^{\mathbb{N}}$  est **typique** pour la mesure  $p$  si pour toute observable  $f : X \rightarrow \mathbb{R}$ ,

$$\left| \langle f \rangle_{x_{[0,t-1]}} - \mathbb{E}_p(f) \right| \leq \|f\|_{L^1} t^{-\frac{1}{3}}, \quad \forall t \in \mathbb{N}. \quad (1.5.2)$$

On note  $\mathcal{S}_p$  l'ensemble des suites typiques pour la mesure  $p$ .

*Remarque 1.5.7.* L'exposant  $-\frac{1}{3}$  dans (1.5.2) est choisit dans l'intervalle  $]-\frac{1}{2}, 0[$  et conviendra d'après le théorème central limite.

Pour une suite  $(x_t)_{t \in \mathbb{N}} = (x_0, x_1, \dots)$  on note

$$x_{[0,t-1]} = (R_t x) = (x_0, x_1, \dots, x_{t-1}) \quad (1.5.3)$$

sa restriction aux  $t$  premiers termes, i.e.  $R_t$  est l'opérateur restriction. Ainsi on note  $R_t(\mathcal{S}_p)$  les suites  $(x_0, x_1, \dots, x_{t-1})$  à  $t$  éléments qui sont les restrictions des suites typiques de  $p$ .

Par ailleurs, la probabilité de la suite  $x_{[0,t-1]} = (x_0, x_1, \dots, x_{t-1})$  est

$$p(x_{[0,t-1]}) \stackrel{(1.3.1)}{=} p(x_0) p(x_1) \dots p(x_{t-1}). \quad (1.5.4)$$

On notera

$$p(R_t(\mathcal{S}_p)) = \sum_{x_{[0,t-1]} \in R_t(\mathcal{S}_p)} p(x_{[0,t-1]}) \quad (1.5.5)$$

la probabilité de l'ensemble  $R_t(\mathcal{S}_p)$ . Un premier résultat important est le suivant

**Théorème 1.5.8.** *Cover (1999, thm 3.1.2 p.51) Soit  $p \in \mathcal{P}(X)$  une loi de probabilité et  $\mathcal{S}_p$  les suites typiques associées.*

$$\lim_{t \rightarrow \infty} p(R_t(\mathcal{S}_p)) = 1, \quad (1.5.6)$$

*c'est à dire que une longue suite empirique est une suite typique presque sûrement.*

Il devient alors important de dénombrer le nombre de suites typiques  $x_{[0,t-1]} \in R_t(\mathcal{S}_p)$ , autrement dit de calculer le cardinal  $\#R_t(\mathcal{S}_p)$  et calculer la probabilité individuelle  $(p(x_{[0,t-1]}))_t$  d'une suite typique. C'est donné par les résultats suivants et parfois appelé "AEP : **asymptotic equipartition property**" Cover (1999, chap.3).

**Théorème 1.5.9.** *Cover (1999, thm 15.7.1 p.475) "Shanon-MacMullen-Bricman"*  
 Soit  $p \in \mathcal{P}(X)$  une loi de probabilité sur un espace fini  $X$ . Alors pour toute suite typique  $(x_t)_t \in \mathcal{S}_p$ ,

$$\lim_{t \rightarrow \infty} \frac{1}{t} \ln \left( \frac{1}{p(x_{[0,t-1]})} \right) = S(p) \quad (1.5.7)$$

avec

$$S(p) = \sum_{x=1}^n p(x) \ln \left( \frac{1}{p(x)} \right), \quad (1.5.8)$$

appelée **entropie de la mesure**  $p$ . Eq.(1.5.7) signifie que pour  $t \gg 1$  les suites typiques ont des probabilités "presque" identiques :  $p(x_{[0,t-1]}) = e^{-t(S(p)+o(1))}$ .

Conséquences :

$$S(p) = \lim_{t \rightarrow \infty} \frac{1}{t} \ln (\#R_t(\mathcal{S}_p)) \quad (1.5.9)$$

qui signifie que le nombre de suites typiques croît comme  $\#R_t(\mathcal{S}_p) = e^{t(S(p)+o(1))}$ .

L'entropie vérifie

$$0 \leq S(p) \leq \ln n$$

et en particulier :

- Entropie nulle  $S(p) = 0$  si et seulement si  $p(x) = \delta_{x=k}$  pour un certain  $k$ , i.e.  $p$  est une mesure presque déterministe.
- Entropie maximale

$$S(p) = \ln n \quad (1.5.10)$$

si et seulement si  $p(x) = \frac{1}{n}, \forall x$ , i.e.  $p$  est la mesure uniforme.

*Démonstration.* On a

$$\begin{aligned} p(x_{[0,t-1]}) &\stackrel{(1.5.4)}{=} p(x_0) p(x_1) \dots p(x_{t-1}) = \prod_{t'=0}^{t-1} p(x_{t'}) \\ &= \exp \left( \sum_{t'=0}^{t-1} \ln p(x_{t'}) \right) = \exp \left( - \sum_{t'=0}^{t-1} \ln \left( \frac{1}{p(x_{t'})} \right) \right) = e^{-tS_t} \end{aligned}$$

avec

$$S_t = \frac{1}{t} \sum_{t'=0}^{t-1} \ln \frac{1}{p(x_{t'})}.$$

D'après la loi des grands nombres (1.4.1), appliquée ici à la fonction  $f(x) = \ln \frac{1}{p(x)}$ , la loi  $p_{S_t}$  de  $S_t$  converge vers un Dirac vers la moyenne  $S$

$$p_{S_t} \xrightarrow[t \rightarrow \infty]{*} \delta_S \quad (1.5.11)$$

avec

$$S = \mathbb{E} \left( \ln \frac{1}{p(x)} \right) \stackrel{(1.1.5)}{=} \sum_x p(x) \ln \frac{1}{p(x)}.$$

Soit  $\mathcal{S}'_p$  l'ensemble des suites typiques pour la mesure  $p$  et la fonction  $f(x) = \ln \frac{1}{p(x)}$  seulement c'est à dire vérifiant (1.5.2). On a donc  $\mathcal{S}_p \subset \mathcal{S}'_p$ . On admet que  $\mathcal{S}_p = \mathcal{S}'_p$ . D'après le théorème central limite, on a  $\lim_{t \rightarrow \infty} p(R_t(\mathcal{S}_p)) = 1$  c'est à dire (1.5.6).

Par définition de  $\mathcal{S}_p$  on a donc (1.5.7) et par conséquence (1.5.9) car

$$\#R_t(\mathcal{S}_p) e^{-t(S(p)+o(1))} \sim 1.$$

Comme  $0 \leq p(x) \leq 1$  on déduit que  $S(p) \geq 0$  et  $S(p) = 0 \Rightarrow p(x) = 0$  ou  $p(x) = 1$ , donc  $p = \delta_{x=k}$ . On a  $\#R_t(\mathcal{S}_p) \leq n^t$  donc  $S(p) \leq \ln n$ . Et pour  $p(x) = \frac{1}{n}, \forall x$ , alors  $S(p) = \ln n$ .  $\square$

### 1.5.3 (\*) Entropie pour un espace $X$ infini

On considère comme en section 1.2, une **mesure de probabilité**  $p(x) dx$  sur un espace  $X$ , de coordonnées  $(x_1, \dots, x_d)$ .

Pour introduire le concept d'entropie, il faut introduire **une autre mesure**  $m = m(x) dx$  **lisse** (pas forcément une mesure de probabilité, mais  $m \in C^\infty(X)$ ), et on suppose  $\text{supp} p \subset \text{supp} m$ . Il y a deux raisons pour introduire cette autre mesure  $m$  :

1. la densité  $p(x)$  a l'unité inverse d'un volume  $dx$ , donc  $\ln \frac{1}{p(x)}$  dans (1.5.8) n'a pas de sens (on ne peut prendre que le log d'un nombre sans dimension). A la place on va considérer  $\ln \frac{m(x)}{p(x)}$  qui est sans dimension.
2. Comme l'espace  $X$  est infini, le cardinal  $\#R_t(\mathcal{S}_p)$  dans (1.5.9) n'a pas de sens, on va le remplacer par le volume  $\text{Vol}_m(R_t(\mathcal{S}_p))$  pour la mesure  $m$ .

On reproduit les définitions et propriétés de la section 1.5.2. La définition 1.5.6 de **suite typique** ne change pas.

Pour une suite restreinte  $x_{[0,t-1]} = (R_t x) = (x_0, x_1, \dots, x_{t-1})$ , comme en (1.5.3), on définit comme en (1.5.4)

$$m(x_{[0,t-1]}) \stackrel{(1.3.1)}{=} m(x_0) m(x_1) \dots m(x_{t-1}), \quad (1.5.12)$$

et pour remplacer le cardinal  $\#R_t(\mathcal{S}_p)$  dans (1.5.9), on écrit

$$\text{Vol}_m(R_t(\mathcal{S}_p)) := \int_{R_t(\mathcal{S}_p)} m(x_{[0,t-1]}) dx_{[0,t-1]}.$$

est le volume de l'ensemble  $R_t(\mathcal{S}_p)$  pour la mesure  $m$ . De même pour remplacer (1.5.5), on écrit

$$p(R_t(\mathcal{S}_p)) = \int_{R_t(\mathcal{S}_p)} p(x_{[0,t-1]}) dx_{[0,t-1]},$$

qui est la probabilité de l'ensemble  $R_t(\mathcal{S}_p)$ .

Le Théorème 1.5.8 ne change pas :

**Théorème 1.5.10.** *Cover (1999, thm 3.1.2 p.51)* Soit  $p \in \mathcal{P}(X)$  une loi de probabilité et  $\mathcal{S}_p$  les suites typiques associées.

$$\lim_{t \rightarrow \infty} p(R_t(\mathcal{S}_p)) = 1, \quad (1.5.13)$$

*c'est à dire que une longue suite empirique est une suite typique presque sûrement.*

et aussi le même Théorème que 1.5.9 mais modifié :

**Théorème 1.5.11.** *Cover (1999, thm 15.7.1 p.475) "Shanon-MacMullen-Bricman"* Soit  $p \in \mathcal{P}(X)$  une loi de probabilité et  $m = m(x) dx$  une mesure sur  $X$ . Alors pour toute suite typique  $(x_t)_t \in \mathcal{S}_p$ ,

$$S_m(p) = \lim_{t \rightarrow \infty} \frac{1}{t} \ln \left( \frac{m(x_{[0,t-1]})}{p(x_{[0,t-1]})} \right) \quad (1.5.14)$$

avec

$$S_m(p) = \int_X p(x) \ln \left( \frac{m(x)}{p(x)} \right) dx, \quad (1.5.15)$$

appelée **entropie de la mesure  $p$  relative à la mesure  $m$** . Eq.(1.5.14) signifie que pour  $t \gg 1$  les suites typiques ont des probabilités "presque" identiques :  $p(x_{[0,t-1]}) = e^{-t(S(p)+o(1))} m(x_{[0,t-1]})$ .

Conséquences :

$$S_m(p) = \lim_{t \rightarrow \infty} \frac{1}{t} \ln (\text{Vol}_m(R_t(\mathcal{S}_p))) \quad (1.5.16)$$

qui signifie que le volume de suites typiques croit comme  $\text{Vol}_m(R_t(\mathcal{S}_p)) = e^{t(S(p)+o(1))}$ .  
L'entropie vérifie

$$S_m(p) \leq \ln \text{Vol}_m(X)$$

et en particulier :

- Entropie maximale  $S(p) = \ln \text{Vol}_m(X)$  si et seulement si  $p(x) = \frac{1}{\text{Vol}_m(X)} m(x)$ ,  $\forall x$  est la mesure uniforme pour  $m$ .
- Pour une mesure de Dirac  $p(x) = \delta_{x_0}$  avec  $x_0 \in X$ , on a  $S(\delta_{x_0}) = -\infty$ .
- Voir ci-dessous une discussion sur les conditions pour avoir  $S_m(p) \geq 0$ .

*Démonstration.* On reprend la preuve page 32.

On a

$$\begin{aligned} p(x_{[0,t-1]}) dx_{[0,t-1]} &= \prod_{t'=0}^{t-1} p(x_{t'}) dx_{t'} = \prod_{t'=0}^{t-1} \frac{p(x_{t'})}{m(x_{t'})} m(x_{t'}) dx_{t'} \\ &= \exp\left(\sum_{t'=0}^{t-1} \ln \frac{p(x_{t'})}{m(x_{t'})}\right) m(x_{[0,t-1]}) dx_{[0,t-1]} = e^{-tS_t} m(x_{[0,t-1]}) dx_{[0,t-1]} \end{aligned}$$

avec

$$S_t = \frac{1}{t} \sum_{t'=0}^{t-1} \ln \frac{p(x_{t'})}{m(x_{t'})}.$$

D'après la loi des grands nombres (1.4.1), appliquée ici à la fonction  $f(x) = \ln \frac{m(x)}{p(x)}$ , la loi  $p_{S_t}$  de  $S_t$  converge vers un Dirac vers la moyenne  $S$

$$p_{S_t} \xrightarrow[t \rightarrow \infty]{*} \delta_S \quad (1.5.17)$$

avec

$$S = \mathbb{E}\left(\ln \frac{m(x)}{p(x)}\right) \stackrel{(1.2.5)}{=} \int_X p(x) \ln \frac{m(x)}{p(x)} dx.$$

Soit  $\mathcal{S}'_p$  l'ensemble des suites typiques pour la mesure  $p$  et la fonction  $f(x) = \ln \frac{m(x)}{p(x)}$  seulement c'est à dire vérifiant (1.5.2). On a donc  $\mathcal{S}_p \subset \mathcal{S}'_p$ . On admet que  $\mathcal{S}_p = \mathcal{S}'_p$ . D'après le théorème central limite, on a  $\lim_{t \rightarrow \infty} p(R_t(\mathcal{S}_p)) = 1$  c'est à dire (1.5.13).

Par définition de  $\mathcal{S}_p$  on a donc (1.5.14) et par conséquence (1.5.16) car

$$\text{Vol}_m(R_t(\mathcal{S}_p)) e^{-t(S(p)+o(1))} \sim 1.$$

On a  $\text{Vol}_m(R_t(\mathcal{S}_p)) \leq (\text{Vol}_m(X))^t$  donc  $S(p) \leq \ln \text{Vol}_m(X)$ . Et pour  $p(x) = \frac{1}{\text{Vol}_m(X)} m(x)$ ,  $\forall x$ , alors  $S(p) = \ln \text{Vol}_m(X)$ .

Pour une mesure de Dirac régularisée  $\delta_{x_0}^\epsilon(x)$  avec  $\epsilon > 0$ , on a  $\lim_{\epsilon \rightarrow 0} \frac{m(x)}{\delta_{x_0}^\epsilon(x)} = 0$  donc  $S(\delta_{x_0}) = \lim_{\epsilon \rightarrow 0} \ln\left(\frac{m(x)}{\delta_{x_0}^\epsilon(x)}\right) = -\infty$ .  $\square$

### 1.5.3.1 Discussion sur la positivité de l'entropie $S_m(p) \geq 0$ .

Supposons que

$$p(x) \leq m(x), \forall x \in X, \quad (1.5.18)$$

Alors d'après (1.5.15), on a

$$S_m(p) \geq 0.$$

Si  $\Omega \subset X$  est un domaine tel que  $\forall x \notin \Omega, p(x) = 0$ , c'est à dire que "la particule étudiée est dans  $\Omega$ " alors

$$\text{Vol}_m(\Omega) = \int_\Omega m(x) dx \geq \int_\Omega p(x) dx = 1,$$

donc “une particule ne peut pas occuper un  $m$ -volume inférieur à 1”. On appelle donc l’hypothèse (1.5.18), “**principe d’incertitude**”.

Si de plus  $p(x) = \frac{1}{\text{Vol}_m(\Omega)} m(x)$ , c’est à dire que la mesure de probabilité est  $m$ -uniforme sur  $\Omega$  alors

$$S_m(p) \stackrel{(1.5.15)}{=} \ln(\text{Vol}_m(\Omega))$$

et  $S_m(p) = 0$  si  $\text{Vol}_m(\Omega) = 1$ .

*Remarque 1.5.12.* En mécanique quantique,  $m = m(x) dx$  sera la mesure de Liouville sur l’espace des phases (espace cotangent). L’hypothèse (1.5.18) est une conséquence du principe d’incertitude en mécanique ondulatoire. On aura  $S_m(p) = 0$  à température nulle, cela est appelé le principe de Nerst, ou 3eme principe, car le système est dans un état quantique.

## 1.6 Dynamique de Markov probabiliste et état d’équilibre

### 1.6.1 Introduction

Le modèle de dynamique probabiliste de Markov présenté dans cette Section modélise l’évolution d’une loi de probabilité au cours du temps. Ce modèle est à la base de la plupart des modèles en physique, biologie, etc. Il est à la base de l’**algorithme de Google rank** (pour classer les pages web), de l’algorithme de monte-carlo, de l’**argument d’Einstein de la balance détaillée** en 1905 pour expliquer la distribution de Planck du corps noir d’un gaz de photon à l’équilibre thermique.

On présente ici le modèle de Markov sur **un espace  $X$  fini et avec un temps discret**  $t \in \mathbb{N}$ . Dans un chapitre ultérieur, on présentera une généralisation de ce modèle pour un espace  $X$  infini et temps  $t \in \mathbb{R}$ .

#### Références :

- Chapitre 4 de ce cours **Faure (2016)**.

### 1.6.2 Dynamique de Markov sur un espace fini $X$ avec temps discret

**Définition 1.6.1.** Soit  $X = \{1, 2, \dots, n\}$  l'espace des évènements que l'on suppose fini. Une **matrice stochastique**  $M$  est une matrice  $M = (M_{x',x})_{x',x \in X} \in \mathbb{R}^{n \times n}$  réelle de taille  $n \times n$  à coefficients  $M_{x',x} \in [0, 1]$  appelés **probabilités de transition de  $x$  vers  $x'$** , vérifiant

$$\forall x \in X = \{1, \dots, n\}, \quad \sum_{x'=1}^n M_{x',x} = 1, \quad (1.6.1)$$

(i.e. la somme sur chaque colonne est 1). A l'instant  $t = 0$ , on considère une mesure de probabilité  $p_{t=0} \in \mathcal{P}(X)$  sur  $X$ , i.e.  $\sum_x p_0(x) = 1$ . La **loi d'évolution de Markov** est que la mesure de probabilité  $p_t \in \mathcal{P}(X)$  à la date  $t \in \mathbb{N}$  est donnée pour tout  $t \geq 1$  par

$$p_t = Mp_{t-1}. \quad (\text{donc } p_t = M^t p_0) \quad (1.6.2)$$

**Exercice 1.6.2.** (\*) Montrer que la condition (1.6.1) est nécessaire et suffisante pour que : si  $p_{t-1} \in \mathcal{P}(X)$  est une mesure de probabilité alors  $p_{t+1} \in \mathcal{P}(X)$  est aussi une mesure de probabilité et que cette condition s'écrit aussi

$$M^T \mathbf{1} = \mathbf{1},$$

avec la matrice transposée  $M^T$  et  $\mathbf{1} = (1, 1, \dots, 1)$ .

**Solution 1.6.3.** En effet

$$\begin{aligned} & \forall p \in \mathcal{P}(X), \quad Mp \in \mathcal{P}(X) \\ \Leftrightarrow & \forall p \in \mathbb{R}^n, \quad \sum_{x'} (Mp)(x') = \sum_x p(x) \\ \Leftrightarrow & \forall p \in \mathbb{R}^n, \quad \langle \mathbf{1} | Mp \rangle = \langle \mathbf{1} | p \rangle \\ \Leftrightarrow & \forall p \in \mathbb{R}^n, \quad \langle M^T \mathbf{1} | p \rangle = \langle \mathbf{1} | p \rangle \\ \Leftrightarrow & M^T \mathbf{1} = \mathbf{1} \\ \Leftrightarrow & \sum_{x'} (M^T)_{x,x'} = 1 \Leftrightarrow \sum_{x'} M_{x',x} = 1. \end{aligned}$$

*Remarque 1.6.4.* Du fait de son importance en science, l'équation d'évolution de Markov (1.6.2) s'appelle aussi "équation maitresse", ou "équation de Fokker Planck discrète", ou "équation de Kolmogorov",...

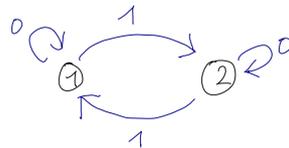
**Question 1.6.5.** Décrire le comportement de la mesure de probabilité  $p_t$  pour  $t \rightarrow \infty$  en fonction de  $p_0$  et  $M$ .

*Remarque 1.6.6.* La formule matricielle  $p_t = Mp_{t-1}$  donne  $p_t = M^t p_0$  en fonction de  $p_0$  qui est un produit de matrices donnant une expression très compliquée et incalculable en général :

$$p_t(x') = \sum_{x_{t-1}} \sum_{x_{t-2}} \dots \sum_{x_0} M_{x',x_{t-1}} M_{x_{t-1},x_{t-2}} \dots M_{x_1,x_0} p_0(x_0)$$

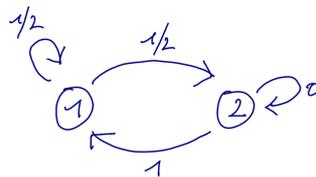
qui est une somme sur toutes les suites  $x_0 \rightarrow x_1 \rightarrow x_2 \dots \rightarrow x_{t-1}$  aussi appelés “chemins”. Pour cette raison cette expression est aussi appelé “**intégrale de chemins de Feynman**”. En fait, pour calculer  $p_t = M^t p_0$  il est préférable de diagonaliser  $M$  au préalable :  $M = PDP^{-1}$  avec  $D = \text{Diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$ , ainsi  $M^t = AD^t A^{-1}$  où  $D^t = \text{Diag}(\lambda_1^t, \lambda_2^t, \dots, \lambda_n^t)$ . De plus pour  $t \rightarrow \infty$  cette expression se simplifie si des valeurs propres ont un module plus petit. C'est que nous allons voir.

**Exemple 1.6.7.** Pour  $X = \{1, 2\}$ , la matrice  $M = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$  représentée par le graphe suivant appelé **graphe de Markov associé** correspond à une dynamique déterministe.



Dans ce cas on a explicitement  $M^t = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$  si  $t$  impair et  $M^t = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$  si  $t$  pair, donc si au départ,  $p_0 = (a, b)$  alors  $p_t = (a, b)$  si  $t$  est pair et  $p_t = (b, a)$  si  $t$  est impair. Il n'a pas été utile de diagonaliser  $M$  dans cet exemple très particulier que l'on considère comme “déterministe”. Un exemple encore plus simple serait  $M = \text{Id}$  qui correspond à aucun mouvement.

**Exemple 1.6.8.** Pour  $X = \{1, 2\}$ , la matrice  $M = \begin{pmatrix} \frac{1}{2} & 0 \\ \frac{1}{2} & 1 \end{pmatrix}$  représentée par le graphe suivant.



On peut deviner que

$$p_t \rightarrow p_{\text{equil.}} \tag{1.6.3}$$

converge vers la mesure de probabilité  $p_{\text{equil.}} = \begin{pmatrix} \frac{2}{3} \\ \frac{1}{3} \end{pmatrix}$  appelée **mesure d'équilibre** pour  $t \rightarrow \infty$ , indépendamment de  $p_0$ .

Remarque 1.6.9. Une différence cruciale entre cet exemple et le précédent est qu'ici on a

$$M^2 = \begin{pmatrix} \frac{3}{4} & \frac{1}{2} \\ \frac{1}{4} & \frac{1}{2} \end{pmatrix},$$

qui est une matrice à coefficients tous positifs et ce qui signifie que à partir du temps  $t = 2$ , il y a transition possible de  $x$  à  $x'$  pour tous  $x, x'$ . On dit que la matrice  $M$  est **mélangeante**.

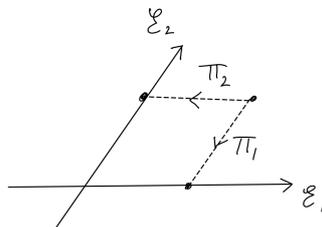
Pour montrer rigoureusement (1.6.3), il faut diagonaliser<sup>3</sup> la matrice  $M$ , i.e. la décomposer selon ses espaces propres  $\mathcal{E}_1, \mathcal{E}_2$  (si  $M$  est diagonalisable) :

$$M = \sum_{k=1}^n \lambda_k \Pi_k$$

où  $\lambda_k \in \mathbb{C}$  sont les valeurs propres de  $M$  et

$$\Pi_k = |U_k\rangle \langle V_k|. : \mathbb{R}^n \rightarrow \mathbb{R}^n$$

le projecteur spectral associé sur l'espace propre  $\mathcal{E}_k$ , i.e.  $\Pi_k^2 = \Pi_k$  et  $\Pi_j \Pi_k = 0$  si  $j \neq k$ , avec  $U_k$  vecteur propre à droite, i.e.  $MU_k = \lambda_k U_k$  et  $V_k$  vecteur propre à gauche i.e.  $V_k M = \lambda_k V_k$ .



Dans cet exemple on calcule<sup>4</sup>

$$M = \lambda_1 \Pi_1 + \lambda_2 \Pi_2$$

avec

$$\lambda_1 = 1, \quad \Pi_1 = |U_1\rangle \langle V_1|., \quad U_1 = \begin{pmatrix} \frac{2}{3} \\ \frac{1}{3} \end{pmatrix}, \quad V_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

3. Dans le cas général une matrice n'est pas toujours diagonalisable, il peut y avoir des blocs de Jordan, cf Appendice.

4. Diagonaliser la matrice  $M = ADA^{-1}$  avec  $D$  diagonale. Pour cela, sur **xcas** en ligne : <https://www-fourier.ujf-grenoble.fr/~parisse/xcasfr.html>, on écrit `M:=[[1/2,0],[1/2,1]]`; `D:=egvl(M)`; `A:=egv(M)`; `inv(A)`; On vérifiera que `simplify(A*D*inv(A))`; redonne bien la matrice

$M$ . On note  $A = (U_1, U_2, U_3)$  vecteurs propres en colonnes définis à une constante près,  $A^{-1} = \begin{pmatrix} V_1 \\ V_2 \\ V_3 \end{pmatrix}$

vecteurs propres en ligne.

$$\lambda_2 = -\frac{1}{2}, \quad \Pi_2 = |U_2\rangle\langle V_2|, \quad U_2 = \begin{pmatrix} -\frac{1}{3} \\ \frac{1}{3} \end{pmatrix}, \quad V_2 = \begin{pmatrix} -1 \\ 2 \end{pmatrix}.$$

On a alors

$$M^t = (\lambda_1 \Pi_1 + \lambda_2 \Pi_2)^t = \lambda_1^t \Pi_1 + \lambda_2^t \Pi_2$$

où

$$|\lambda_2|^t = \frac{1}{2^t} = e^{-t \ln 2} = e^{-t/\tau}$$

qui tend exponentiellement vite vers 0 avec le temps caractéristique  $\tau = \frac{1}{\ln 2} \approx 1.44$ . On a

$$\begin{aligned} p_t &= M^t p_0 = \Pi_1 p_0 + O(e^{-t/\tau}) = U_1 \underbrace{\langle V_1 | p_0 \rangle}_1 + O(e^{-t/\tau}) \\ &= U_1 + O(e^{-t/\tau}) \end{aligned}$$

Il y a donc convergence exponentielle de toute mesure de probabilité initiale  $p_0$  vers la **mesure d'équilibre**

$$p_{\text{equil.}} := U_1 = \begin{pmatrix} \frac{2}{3} \\ \frac{1}{3} \end{pmatrix}.$$

*Remarque 1.6.10. “paradoxe sur l'entropie”.* Si au départ on part de l'état uniforme,  $p_{\text{unif}} = \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \end{pmatrix}$ , on vient de montrer qu'il évolue vers l'état d'équilibre  $p_{\text{equil.}} = \begin{pmatrix} \frac{2}{3} \\ \frac{1}{3} \end{pmatrix}$ . Or d'après la formule (1.5.8) l'entropie de départ est  $S(p_{\text{unif}}) = \ln 2 \approx 0.69$  supérieure à l'entropie finale  $S(p_{\text{equil.}}) = \frac{2}{3} \ln \left(\frac{2}{3}\right)^{-1} + \frac{1}{3} \ln \left(\frac{1}{3}\right)^{-1} \approx 0.63$ . Paradoxalement, cette entropie décroît avec le temps. On verra au théorème 1.6.16 appelé “deuxième principe de la thermodynamique” que l'expression correcte de l'entropie à utiliser est plutôt l'**entropie relative**  $S_{p_{\text{equil.}}}(p(t)) = p_1(t) \ln \left(\frac{2/3}{p_1(t)}\right) + p_2(t) \ln \left(\frac{1/3}{p_2(t)}\right)$  (contenant les composantes de la mesure d'équilibre) qui elle augmente avec le temps.

Le résultat suivant montre que le phénomène de convergence vers une mesure d'équilibre observé dans l'exemple précédent se produit généralement pour une matrice mélangeante.

**Définition 1.6.11.** Une matrice stochastique  $M$  est **mélangeante** si

$$\exists t \geq 1, \forall x', x \in X, \quad (M^t)_{x',x} > 0.$$

Cela signifie que il existe un temps  $t$  auquel il est possible de transiter entre toute paire d'états  $x, x'$ .

**Théorème 1.6.12.** “Perron-Frobenius” (1907), Si  $M$  est une matrice stochastique et mélangeante alors ses valeurs propres  $(\lambda_k)_{k \in \{1,2,\dots,n\}}$  vérifient

$$|\lambda_n| \leq |\lambda_{n-1}| \leq \dots \leq |\lambda_2| < \lambda_1 = 1$$

i.e. la valeur propre dominante est  $\lambda_1 = 1$  et les suivantes sont de module strictement inférieur à 1. De plus le projecteur spectral associé

$$\Pi_1 = |p_{\text{equil.}}\rangle \langle \mathbf{1}|\cdot\rangle$$

avec  $p_{\text{equil.}}$  une mesure de probabilité vérifiant  $p_{\text{equil.}}(x) > 0, \forall x$ , et appelée **mesure d'équilibre**. Par conséquent pour toute mesure initiale  $p_0 \in \mathcal{P}(X)$ , évolue selon

$$p_t = M^t p_0 = p_{\text{equil.}} + O(e^{-t/\tau}), \quad \forall \tau > \frac{1}{\ln |\lambda_2|^{-1}} > 0. \quad (1.6.4)$$

Il y a donc **convergence exponentielle** de toute mesure de probabilité initiale  $p_0$  vers la mesure d'équilibre  $p_{\text{equil.}}$  avec un temps caractéristique  $\tau$ .

*Remarque 1.6.13.* Ce modèle précédent est à la base de la physique statistique. On pense (mais cela n'est pas encore rigoureusement démontré) que le “processus de thermalisation” est un processus de convergence local vers un état d'équilibre. Cela sera étudié dans le chapitre 3.

*Démonstration.* (\*) (ref : Brin [Brin & Stuck \(2002, p.78\)](#), Davies [Davies & Davies \(1996, p.355\)](#)). On note

$$S_1^+ := \left\{ u \in \mathbb{R}^m, \sum_{j=1}^m u_j = 1 \text{ et } \forall j, u_j \geq 0 \right\} = \mathcal{P}(X).$$

qui est le « simplex des mesures de probabilités ». Soit

$$f : \begin{cases} \mathcal{P}(X) & \rightarrow \mathcal{P}(X) \\ u & \rightarrow [Mu] := \frac{Mu}{\|Mu\|_1} \end{cases}$$

$f$  laisse  $\mathcal{P}(X)$  invariant donc d'après le **théorème de Brouwer**  $f$  admet un point fixe :

$$\exists u \in \mathcal{P}(X), \quad f(u) = u$$

Cela implique qu'il existe  $\lambda_1 > 0$  tel que

$$Mu = \lambda_1 u$$

Si  $M$  est mélangeante alors  $u \in \text{Int}(\mathcal{P}(X))$  donc  $u_x > 0, \forall x \in X$ . Soit

$$\tilde{M} := \frac{1}{\lambda} (\text{Diag}(u))^{-1} \circ M \circ \text{Diag}(u)$$

alors

$$\tilde{M}\mathbf{1} = \mathbf{1}$$

donc  $\tilde{M}^\dagger$  préserve  $\mathcal{P}(X)$ . D'après le **théorème de Brouwer**

$$\exists w \in \mathcal{P}(X), \quad w_j > 0, \forall j, \quad \text{t.q. } \tilde{M}^\dagger w = w$$

Soit

$$\mathcal{D} := (\mathcal{P}(X) - w) \subset S_0 := \left\{ u \in \mathbb{R}^m, \quad \sum_{j=1}^m u_j = 0 \right\}.$$

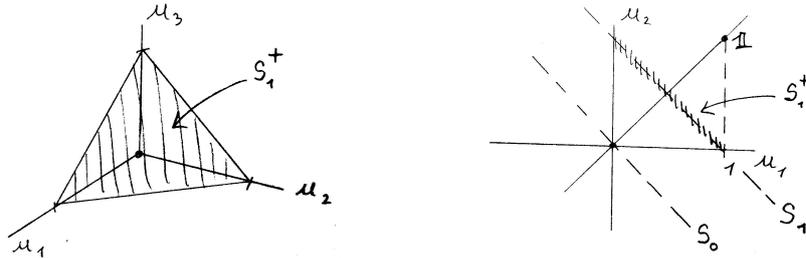
On a

$$\tilde{M}^*(\mathcal{D}) \subset \mathcal{D}$$

donc le spectre de  $\tilde{M}^\dagger_{/S_0}$  est contenu dans le disque unité. Si de plus  $\tilde{M}$  est mélangeante alors il existe  $t \geq 1$  tel que

$$\left(\tilde{M}^\dagger\right)^t(\mathcal{D}) \subset \text{Int}(\mathcal{D})$$

donc  $\left(\tilde{M}^\dagger\right)^t_{/S_0}$  est contractante et donc son spectre est strictement inclu dans le disque unité.



□

### 1.6.3 Augmentation de l'entropie : "2ème principe"

Le **deuxième principe de la thermodynamique** est très célèbre car il affirme l'irréversibilité des phénomènes physique, l'existence d'une **"flèche du temps"** et énonce que l'entropie d'un système physique isolé augmente. C'est cependant un sujet très controversé car il est souvent difficile (et souvent hors de portée) d'obtenir un énoncé satisfaisant et précis pour des dynamiques issues de modèles physique.

Il est intéressant de savoir que ce deuxième principe a une formulation précise dans le cadre des dynamiques de Markov étudiées dans cette section. En particulier en 1872, Boltzmann a tenté de démontrer ce 2ème principe dans le cadre de dynamique probabiliste de particules (modèle de Boltzmann) et obtenu ce que l'on appelle le **théorème H**.

Pour cela on doit utiliser l'entropie  $S_m(p)$  d'une mesure de probabilité  $p$  relative à une mesure  $m$ , comme en (1.5.15) plutôt que l'entropie  $S(p)$ , Eq.(1.5.8). En effet, si on considère Eq.(1.5.8), d'après le théorème de Perron Frobenius (1.6.4), la mesure uniforme

$p_{\text{unif}} = \frac{1}{n}$  d'entropie  $S(p_{\text{unif}})$  maximale va évoluer vers la mesure  $p_{\text{equil}}$  qui a une entropie plus faible et donc l'entropie  $S(p)$  diminue. Par contre, on va démontrer que l'entropie relative  $S_m(p)$  augmente toujours, c'est donc la bonne définition à considérer en général.

On rappelle la définition suivante :

**Définition 1.6.14.** Si  $p \in \mathcal{P}(X)$  est une mesure de probabilité sur  $X$  et  $m$  une mesure sur  $X$  (i.e. on demande seulement  $m(x) \geq 0$ ), l'**entropie relative** de  $p$  par rapport à  $m$  est

$$S_m(p) := \sum_{x \in X} p(x) \ln \left( \frac{m(x)}{p(x)} \right). \quad (1.6.5)$$

*Remarque 1.6.15.* L'entropie relative  $S_m(p)$  est aussi appelée **Kullback–Leibler divergence** par ces auteurs qui l'ont introduite en 1951.

**Théorème 1.6.16.** *Cover (1999, p.34) "Augmentation de l'entropie relative", "2ème principe de la thermodynamique", "Théorème H".* Si  $M$  est une matrice stochastique, pour toute mesure de probabilité  $p \in \mathcal{P}(X)$  et toute mesure  $m$  sur  $X$ , évoluant selon la dynamique de Markov, respectivement  $p_{t+1} = Mp_t$  et  $m_{t+1} = Mm_t$ , on a

$$S_{m_{t+1}}(p_{t+1}) \geq S_{m_t}(p_t). \quad (1.6.6)$$

La preuve est donnée ci-dessous en section 1.6.3.1.

**Corollaire 1.6.17.** *En particulier si  $m = cp_{\text{equil}}$  est proportionnelle à la mesure d'équilibre avec une constante  $c = \sum_x m(x) = \text{Vol}_m(X) > 0$ , alors  $m_t = m$  n'évolue pas, et d'après (1.6.6),  $S_m(p_t)$  **augmente** avec le temps  $t$ . De plus si la matrice  $M$  est mélangeante, alors  $p_t \rightarrow p_{\text{equil}}$  et donc*

$$S_m(p_t) \xrightarrow{t \rightarrow +\infty} S_m(p_{\text{equil}}) = \ln(\text{Vol}_m(X)),$$

*i.e. l'entropie de  $p_t$  augmente avec le temps et converge vers sa valeur maximale  $S_m(p_{\text{equil}}) = \ln(\text{Vol}_m(X))$ .*

### 1.6.3.1 (\*) Preuve du théorème 1.6.16 "2ème principe", suivant Cover (1999, p.34).

Pour cela on va utiliser quelques résultats intermédiaires importants.

**Théorème 1.6.18.** “Inégalité de Jensen” Cover (1999, p.25). Si  $f : \mathbb{R} \rightarrow \mathbb{R}$  est une fonction convexe, i.e.  $f''(s) \geq 0, \forall s \in \mathbb{R}$ , et  $p \in \mathcal{P}(X)$  une mesure de probabilité alors

$$\mathbb{E}_p(f) \geq f(\mathbb{E}_p(x)). \quad (1.6.7)$$

Si  $f$  est strictement convexe, i.e.  $f''(s) > 0, \forall s \in \mathbb{R}$ , et  $\mathbb{E}_p(f) = f(\mathbb{E}_p(x))$  alors la loi est déterministe :  $p = \delta_\mu$  avec  $\mu \in \mathbb{R}$ .

*Démonstration.* Cover (1999, p.25) On le vérifie d’abord pour une mesure  $p = \alpha\delta_{x_0} + (1 - \alpha)\delta_{x_1}$  supportée en deux points, avec  $\alpha \in [0, 1]$ . Dans ce cas  $\mathbb{E}_p(f) = \alpha f(x_0) + (1 - \alpha)f(x_1)$  et  $f(\mathbb{E}_p(x)) = f(\alpha x_0 + (1 - \alpha)x_1)$  donc l’inégalité de Jensen (1.6.7) correspond à l’inégalité de convexité de la fonction  $f$  :

$$\alpha f(x_0) + (1 - \alpha)f(x_1) \geq f(\alpha x_0 + (1 - \alpha)x_1).$$

Par récurrence sur le nombre de point, on déduit l’inégalité de Jensen pour une mesure supportée par un nombre fini arbitraire de points et par passage à la limite, pour toute mesure  $p$ .  $\square$

**Théorème 1.6.19.** “Inégalité de Log-Sum” Cover (1999, p.29). Pour des nombres quelconques  $a_j, b_j \geq 0, j \in \{1, \dots, n\}$ , on a

$$\sum_{i=1}^n a_i \ln \left( \frac{a_i}{b_i} \right) \geq \left( \sum_{i=1}^n a_i \right) \ln \left( \frac{\sum_{i=1}^n a_i}{\sum_{i=1}^n b_i} \right), \quad (1.6.8)$$

et égalité si et seulement si  $\exists c, \forall i, \frac{a_i}{b_i} = c$ .

*Démonstration.* Cover (1999, p.25). Supposons  $a_i, b_i > 0$  et considérons la fonction  $f : s \in \mathbb{R}^+ \rightarrow f(s) = s \ln s$ . On a  $f'(s) = \ln s + 1, f''(s) = \frac{1}{s} > 0$  donc  $f$  est strictement convexe. Soit

$$\alpha_i = \frac{b_i}{\sum_j b_j}, \quad t_i = \frac{a_i}{b_i}.$$

Alors  $\sum_i \alpha_i = 1$  et

$$\sum_i \alpha_i t_i = \sum_i \frac{b_i}{\sum_j b_j} \left( \frac{a_i}{b_i} \right) = \frac{\sum_i a_i}{\sum_j b_j}$$

et l'inégalité de Jensen (1.6.7) pour la mesure de probabilité  $p = \sum_i \alpha_i \delta_{t_i}$  donne

$$\begin{aligned} \sum_i \alpha_i f(t_i) &\geq f\left(\sum_i \alpha_i t_i\right) \\ &\Leftrightarrow \sum_i \frac{b_i}{\sum_j b_j} \left(\frac{a_i}{b_i}\right) \ln\left(\frac{a_i}{b_i}\right) \geq \left(\frac{\sum_i a_i}{\sum_k b_k}\right) \ln\left(\frac{\sum_i a_i}{\sum_k b_k}\right) \\ &\Leftrightarrow \sum_i a_i \ln\left(\frac{a_i}{b_i}\right) \geq \left(\sum_i a_i\right) \ln\left(\frac{\sum_i a_i}{\sum_k b_k}\right) \end{aligned}$$

□

On définit

$$\text{Vol}_m(X) := \sum_{x \in X} m(x). \quad (1.6.9)$$

**Théorème 1.6.20.** *“Inégalité de Gibbs”.* Pour toute mesure de probabilité  $p$  et mesure  $m$ ,

$$S_m(p) \leq \ln(\text{Vol}_m(X)), \quad (1.6.10)$$

Il y a égalité si et seulement si  $p(x) = \frac{1}{\text{Vol}_m(X)} m(x), \forall x$ .

*Démonstration.* On utilise  $\sum_{x \in X} p(x) = 1$  et on écrit

$$\begin{aligned} -S_m(p) &\stackrel{(1.6.5)}{=} \sum_{x \in X} p(x) \ln\left(\frac{p(x)}{m(x)}\right) \\ &\stackrel{(1.6.8)}{\geq} \left(\sum_{x \in X} p(x)\right) \ln\left(\frac{\sum_{x \in X} p(x)}{\sum_{x \in X} m(x)}\right) = -\ln\left(\sum_{x \in X} m(x)\right) = -\ln(\text{Vol}_m(X)). \end{aligned}$$

Il y a égalité si et seulement si  $\frac{p(x)}{m(x)} = c, \forall x$  avec  $c = \text{Vol}_m(X)$ . □

**Quelques définitions :** Si  $Y = X_1 \times X_2$ , et  $p_Y(x_1, x_2) \in \mathcal{P}(Y)$  mesure de probabilité,  $m_Y$  mesure sur  $Y$ , on définit la **mesure induite**  $p_{X_1}$  sur  $X_1$  par

$$p_{X_1}(x_1) = \sum_{x_2 \in X_2} p_Y(x_1, x_2), \quad (1.6.11)$$

Ainsi  $\sum_{x_1} p_{X_1}(x_1) = 1$ . Pour  $x_1 \in X_1$  fixé, on définit la **mesure conditionnelle**  $p_{X_2, x_1}$  sur  $X_2$ ,

$$p_{X_2, x_1}(x_2) := \frac{p_Y(x_1, x_2)}{p_{X_1}(x_1)}. \quad (1.6.12)$$

Ainsi  $\sum_{x_2} p_{X_2, x_1}(x_2) = 1$ . On définit l'entropie relative conditionnelle

$$S_{m_Y, X_2}(p_Y) = \sum_{x_1 \in X_1} p_{X_1}(x_1) S_{m_{X_2, x_1}}(p_{X_2, x_1}) \quad (1.6.13)$$

qui est une moyenne des entropie relative

$$S_{m_{X_2, x_1}}(p_{X_2, x_1}) \stackrel{(1.6.5)}{=} \sum_{x_2 \in X_2} p_{X_2, x_1}(x_2) \ln \left( \frac{m_{X_2, x_1}(x_2)}{p_{X_2, x_1}(x_2)} \right). \quad (1.6.14)$$

**Théorème 1.6.21.** *Cover (1999, p.23) "Chain rule for relative entropy". Avec les définitions précédentes*

$$S_{m_Y}(p_Y) = S_{m_{X_1}}(p_{X_1}) + S_{m_Y, X_2}(p_Y). \quad (1.6.15)$$

*Démonstration.* On a

$$\begin{aligned} S_{m_Y}(p_Y) &\stackrel{(1.6.5)}{=} \sum_{x_1} \sum_{x_2} p(x_1, x_2) \ln \left( \frac{m_Y(x_1, x_2)}{p(x_1, x_2)} \right) \\ &\stackrel{(1.6.12)}{=} \sum_{x_1} \sum_{x_2} p_{X_1}(x_1) p_{X_2, x_1}(x_2) \ln \left( \frac{m_{X_1}(x_1) m_{X_2, x_1}(x_2)}{p_{X_1}(x_1) p_{X_2, x_1}(x_2)} \right) \\ &\stackrel{(1.6.11)}{=} \sum_{x_1} p_{X_1}(x_1) \ln \left( \frac{m_{X_1}(x_1)}{p_{X_1}(x_1)} \right) + \sum_{x_1} \sum_{x_2} p_{X_1}(x_1) p_{X_2, x_1}(x_2) \ln \left( \frac{m_{X_2, x_1}(x_2)}{p_{X_2, x_1}(x_2)} \right) \\ &\stackrel{(1.6.5, 1.6.13)}{=} S_{m_{X_1}}(p_{X_1}) + S_{m_Y, X_2}(p). \end{aligned}$$

□

**Corollaire 1.6.22.** *Avec les définitions précédentes, on a*

$$S_{m_Y, X_2}(p_Y) \leq 0, \quad (1.6.16)$$

*avec égalité si et seulement si  $p_{X_2, x_1}(x_2) = \frac{1}{\text{Vol}_{m_{X_2, x_1}}(X_2)} m_{X_2, x_1}(x_2)$ ,  $\forall x_2 \in X_2$ .*

*Démonstration.* On a

$$S_{m_Y, X_2}(p_Y) \stackrel{(1.6.13)}{=} \sum_{x_1 \in X_1} p_{X_1}(x_1) S_{m_{X_2, x_1}}(p_{X_2, x_1}) \quad (1.6.17)$$

Or

$$S_{m_{X_2, x_1}}(p_{X_2, x_1}) \stackrel{(1.6.10)}{\leq} \ln(\text{Vol}_{m_{X_2, x_1}}(X_2)),$$

$$\begin{aligned}
\text{Vol}_{m_{X_2, x_1}}(X_2) &\stackrel{(1.6.9)}{=} \sum_{x_2} m_{X_2, x_1}(x_2) \stackrel{(1.6.12)}{=} \frac{\sum_{x_2} m_Y(x_1, x_2)}{m_{X_1}(x_1)} \\
&\stackrel{(1.6.11)}{=} \frac{\sum_{x_2} m_Y(x_1, x_2)}{\sum_{x_2} m_Y(x_1, x_2)} = 1.
\end{aligned}$$

Donc  $S_{m_{X_2, x_1}}(p_{X_2, x_1}) \leq 0$  et  $S_{m_Y, X_2}(p_Y) \stackrel{(1.6.17)}{\leq} 0$ .  $\square$

**Preuve du théorème 1.6.16 “2ème principe” :** Si  $M$  est une matrice stochastique,  $p \in \mathcal{P}(X)$  et  $m$  une mesure sur  $X$ , évoluant selon la dynamique de Markov  $p_{t+1} = Mp_t$ ,  $m_{t+1} = Mm_t$ . On considère l'espace produit  $Y = X_1 \times X_2$  qui est deux copies de l'espace  $X_1 = X_2 = X$ , contenant les couples possibles  $(x_1(t), x_2(t+1))$  pour les temps  $t, t+1$ . Sur cet espace  $Y = X_1 \times X_2$  on a les mesures induites  $p_Y, m_Y$  : la probabilité d'un couple  $(x_1, x_2)$  est donnée par la dynamique de Markov

$$p_Y(x_1, x_2) := M_{x_2, x_1} p(x_1), \quad m_Y(x_1, x_2) := M_{x_2, x_1} m(x_1). \quad (1.6.18)$$

Concernant ces mesures, l'entropie relative est

$$\begin{aligned}
S_{m_Y}(p_Y) &\stackrel{(1.6.15)}{=} S_{m_{X_1}}(p_{X_1}) + S_{m_Y, X_2}(p_Y) \\
&\stackrel{(1.6.15)}{=} S_{m_{X_2}}(p_{X_2}) + S_{m_Y, X_1}(p_Y)
\end{aligned} \quad (1.6.19)$$

où  $p_{X_1} = p_t$ ,  $p_{X_2} = p_{t+1}$  donc

$$S_{m_{t+1}}(p_{t+1}) \stackrel{(1.6.19)}{=} S_{m_t}(p_t) + S_{m_Y, X_2}(p_Y) - S_{m_Y, X_1}(p_Y). \quad (1.6.20)$$

Or

$$p_{X_1}(x_1) \stackrel{(1.6.11)}{=} \sum_{x_2 \in X_2} p_Y(x_1, x_2) \stackrel{(1.6.18)}{=} \sum_{x_2 \in X_2} M_{x_2, x_1} p(x_1) \stackrel{(1.6.1)}{=} p(x_1)$$

donc

$$p_{X_2, x_1}(x_2) \stackrel{(1.6.12)}{=} \frac{p_Y(x_1, x_2)}{p_{X_1}(x_1)} \stackrel{(1.6.18)}{=} \frac{M_{x_2, x_1} p(x_1)}{p(x_1)} = M_{x_2, x_1}$$

et de même  $m_{X_2, x_1}(x_2) = M_{x_2, x_1} = p_{X_2, x_1}(x_2)$ . Donc  $\ln \left( \frac{m_{X_2, x_1}(x_2)}{p_{X_2, x_1}(x_2)} \right) = \ln 1 = 0$  et

$$\begin{aligned}
S_{m_{X_2, x_1}}(p_{X_2, x_1}) &\stackrel{(1.6.14)}{=} 0. \\
S_{m_Y, X_2}(p_Y) &\stackrel{(1.6.13)}{=} 0.
\end{aligned} \quad (1.6.21)$$

Par ailleurs

$$S_{m_Y, X_1}(p_Y) \stackrel{(1.6.16)}{\leq} 0$$

donc

$$S_{m_{t+1}}(p_{t+1}) \stackrel{(1.6.20, 1.6.21)}{=} S_{m_t}(p_t) - S_{m_Y, X_1}(p_Y) \geq S_{m_t}(p_t).$$



Deuxième partie

Systemes sans interactions



# Chapitre 2

## Mécanique ondulatoire et limite corpusculaire (Hamilton)

### 2.1 Introduction

On commence par quelques rappels de mécanique.

- En **1685**, **Newton** établit la loi d'évolution d'une particule de position  $q(t) = (q_1(t), q_2(t), q_3(t)) \in \mathbb{R}^3$ , masse  $m$  :

$$m \frac{\partial^2 q_j(t)}{\partial t^2} = F_j(q) = -\frac{\partial U}{\partial q_j}, \quad j = 1, 2, 3, \quad (2.1.1)$$

où  $U(q)$  est la fonction énergie potentielle et  $F(q) = (F_1(q), F_2(q), F_3(q)) \in \mathbb{R}^3$  le vecteur force qui en dérive. L'expression de (2.1.1) peut paraître étrange pour deux raisons : elle est du deuxième ordre en temps et le terme de droite n'est pas quelconque : c'est le gradient d'une fonction  $U$ .

- En **1834**, **Hamilton** reformule l'équation de Newton (2.1.1) en introduisant une nouvelle variable  $p(t) = (p_1(t), p_2(t), p_3(t)) \in \mathbb{R}^3$  appelée impulsion, et la fonction énergie totale (ou Hamiltonien)

$$H(q, p) = \frac{\|p\|^2}{2m} + U(q) \quad (2.1.2)$$

où le premier terme est appelé énergie cinétique avec  $\|p\|^2 := p_1^2 + p_2^2 + p_3^2$ . L'équation de Newton (2.1.1) s'écrit

$$(2.1.1) \Leftrightarrow \begin{cases} \frac{\partial q_j}{\partial t} = \frac{\partial H}{\partial p_j} \\ \frac{\partial p_j}{\partial t} = -\frac{\partial H}{\partial q_j} \end{cases}, \quad j = 1, 2, 3 \quad (2.1.3)$$

appelée **équations de Hamilton**. En effet  $\frac{\partial q_j}{\partial t} = \frac{\partial H}{\partial p_j} = \frac{p_j}{m}$  donne  $p_j = m \frac{\partial q_j}{\partial t}$  et donc  $\frac{\partial p_j}{\partial t} = -\frac{\partial H}{\partial q_j} = -\frac{\partial U}{\partial q_j}$  donne  $m \frac{\partial^2 q_j(t)}{\partial t^2} = -\frac{\partial U}{\partial q_j}$  qui est l'équation de Newton.

Ces “équations de Hamilton” sont aussi un peu étranges car elles ont un aspect “antisymétriques” par rapport aux termes  $q, p$ . La raison de cela était déjà connu à l’époque de Hamilton : il a utilisé ces équations pour décrire l’aspect corpusculaire de la lumière (i.e. les rayons lumineux) comme cas limite de la théorie ondulatoire de la lumière dans la limite des petites longueurs d’ondes (en effet la longueur d’onde de la lumière  $\lambda \approx 1\mu m \ll 1m$  est très petite à l’échelle humaine). Plus généralement on verra en dans la section 2.2 que :

**Proposition 2.1.1.** *toute équation ondulatoire linéaire fait émerger dans la limite des petites longueurs d’ondes une dynamique corpusculaire régit par les équations de Hamilton avec une certaine fonction  $H$  qui dépend de l’équation ondulatoire de départ.*

- En 1920, Schrödinger montre que les équations de Hamilton de la mécanique (et donc l’équation de Newton) sont aussi une approximation (ou émergence) d’une mécanique ondulatoire sous-jacente pour la matière, régit par “l’équation de Schrödinger”. Maintenant on comprends que l’aspect étrange des équations de la mécanique classique, indiquent une réalité ondulatoire sous-jacente.

Il y a deux propriétés importantes des équations de mouvement de Hamilton qui seront utiles dans la suite.

**Théorème 2.1.2. “Conservation de l’énergie”.** *Pour une fonction de Hamilton quelconque  $H(q, p)$ , les trajectoires  $q(t), p(t)$  solutions de (2.1.3) conservent l’énergie définie par  $E(t) = H(q(t), p(t))$ , i.e.  $\frac{dE}{dt} = 0$ .*

*Démonstration.* On écrit

$$\begin{aligned} \frac{dE}{dt} &= \sum_j \frac{\partial H}{\partial q_j} \frac{\partial q_j}{\partial t} + \frac{\partial H}{\partial p_j} \frac{\partial p_j}{\partial t} \\ &\stackrel{(2.1.3)}{=} \sum_j \left( \frac{\partial H}{\partial q_j} \right) \left( \frac{\partial H}{\partial p_j} \right) + \left( \frac{\partial H}{\partial p_j} \right) \left( -\frac{\partial H}{\partial q_j} \right) = 0. \end{aligned}$$

□

**Théorème 2.1.3. “Conservation de la mesure de Liouville dans l’espace des phases”.** *Le champ de vecteur  $X = \left( \frac{\partial q_j}{\partial t}, \frac{\partial p_j}{\partial t} \right)_{j=1,2,3}$  défini par (2.1.3) dans l’espace des phases  $(q, p)$  préserve le volume  $dqdp$  (appelée mesure de Liouville), i.e. vérifie  $\text{div} X = 0$ .*

*Démonstration.* On écrit

$$\text{div} X = \sum_j \frac{\partial X_{q_j}}{\partial q_j} + \frac{\partial X_{p_j}}{\partial p_j} = \sum_j \frac{\partial}{\partial q_j} \left( \frac{\partial H}{\partial p_j} \right) + \frac{\partial}{\partial p_j} \left( -\frac{\partial H}{\partial q_j} \right) = 0.$$

□

## 2.2 Mécanique ondulatoire et limite semi-classique de Hamilton

Il y a un fait général et important qui explique la forme particulière des équations de Newton et Hamilton ci-dessus et qui suggère l'origine ondulatoire de la matière, découverte par Schrödinger en 1920.

Considérons une “dynamique ondulatoire” pour une fonction  $\psi(q, t)$  de la position  $q$  et du temps  $t$ , régie par une équation linéaire que l'on écrit sous la forme

$$i\hbar \frac{\partial \psi}{\partial t} = \hat{H}\psi \quad (2.2.1)$$

où  $\hat{H}$  est un opérateur différentiel (qui a l'unité de  $\hbar/t$  qui est une énergie). On verra des exemples importants plus bas.

*Remarque 2.2.1.* Dans les cas qui nous intéressent, l'opérateur  $\hat{H}$  est auto-adjoint, c'est à dire  $\hat{H}^\dagger = \hat{H}$ . Cela implique que ses valeurs propres sont réelles que l'on notera donc  $E_1 \leq E_2 \leq \dots$  ( $E_j$  a l'unité d'énergie) et que les espaces propres sont orthogonaux entre eux.

*Remarque 2.2.2.* La présence du facteur  $i$  et  $\hbar$  dans (2.2.1) est une convention. Il serait peut être plus naturel d'écrire  $\frac{\partial \psi}{\partial t} = \hat{\Omega}\psi$  avec  $\hat{\Omega} = -\frac{i}{\hbar}\hat{H}$  qui est anti auto-adjoint, i.e.  $\hat{\Omega}^\dagger = -\hat{\Omega}$  et qui a l'unité d'une fréquence (comme  $1/t$ ). Ses valeurs propres sont des fréquences  $-i\frac{E_j}{\hbar} = i\omega_j$ , imaginaires pures (on retrouve le facteur  $i$ ).

**Proposition 2.2.3.** *Le comportement des ondes  $\psi(q, t)$ , solutions de l'équation (2.2.1), dans la limite des petites longueurs d'ondes a un “aspect corpusculaire”, au sens où  $\psi$  se décompose comme superposition de petits paquets d'ondes de la forme*

$$\varphi_{q,p}(q') = e^{ikq'} e^{-\frac{1}{2}\left(\frac{\|q'-q\|}{\sigma}\right)^2},$$

avec une largeur  $\sigma > 0$  position  $q$ , vecteur d'onde  $k = \frac{1}{\hbar}p \in \mathbb{R}^3$  et chaque paquet d'onde évolue comme  $\varphi_{q(t),p(t)}$  avec des paramètres  $q(t), p(t)$  solutions des équations de Hamilton (2.1.3) avec une certaine fonction  $H(q, p)$  déterminée par l'opérateur  $\hat{H}$  de l'équation ondulatoire, en remplaçant les opérateur  $-i\frac{\partial}{\partial q_j}$  par la variable impulsion  $\frac{p_j}{\hbar} = k_j$  autrement dit

$$\frac{\partial}{\partial q_j} \rightarrow ik_j = i\frac{p_j}{\hbar} \quad (2.2.2)$$

appelé **principe de correspondance**.

*Remarque 2.2.4.* Pour une formulation plus précise de cette proposition, voir le cours [Faure \(2020\)](#). Il est important de comprendre que  $\hbar$  ne joue aucun rôle et sert seulement de “petit

paramètre”. Il faut observer que pour une onde plane (ou mode de Fourier)  $\varphi_k(q) = e^{ik \cdot q} = e^{i(\frac{p}{\hbar}) \cdot q}$ ,  $k \in \mathbb{R}^3$ ,  $q \in \mathbb{R}^3$ , on a  $\frac{\partial \varphi_k}{\partial q_j} = (ik_j) \varphi_k$ , ce qui explique la correspondance (2.2.2). La proposition est valable dans la limite des petites longueurs d’ondes  $\lambda \ll L$  devant la taille  $L$  caractéristique du problème, donc  $\|k\| \gg \frac{1}{L}$  ou  $\|p\| \gg \frac{\hbar}{L} \Leftrightarrow \hbar \ll L \|p\|$ . Il faut retenir que l’impulsion  $p = \hbar k$  introduite par Hamilton est donc similaire au vecteur d’onde  $k$ , c’est une “fréquence spatiale”. En géométrie,  $p$  s’appelle un “**vecteur cotangent**”.

Donnons ici deux exemples importants. Un premier exemple bien connu qui est l’équation ondulatoire de Schrödinger et l’autre plus ancien, mais moins connu qui est l’équation des ondes. On verra d’autres exemples avec l’équation de Schrödinger relativiste pour les électrons.

**Exemple 2.2.5. “L’équation de Schrödinger”.** En 1920, Schrödinger montre que la matière est modélisée par des ondes  $\psi(q, t)$  appelée ondes de matière ou ondes quantiques et solutions de l’équation d’évolution (2.2.1) avec l’opérateur

$$\hat{H} = -\frac{\hbar^2}{2m} \Delta + U(q)$$

où  $\Delta = \sum_{j=1}^3 \frac{\partial^2}{\partial q_j^2}$  est l’opérateur Laplacien. D’après le principe de correspondance (2.2.2), on obtient la fonction

$$\begin{aligned} H(q, p) &= -\frac{\hbar^2}{2m} \left( \sum_{j=1}^3 \left( i \frac{p_j}{\hbar} \right)^2 \right) + U(q) \\ &= \frac{1}{2m} \left( \sum_{j=1}^3 p_j^2 \right) + U(q) = \frac{\|p\|^2}{2m} + U(q) \end{aligned}$$

qui est bien la fonction de Hamilton de la mécanique (2.1.2).

**Exemple 2.2.6. “L’équation d’onde”.** Dans le cas simple d’une onde scalaire i.e.  $\psi(q, t) \in \mathbb{R}$  évoluant dans un milieu d’indice variable, l’équation d’onde s’écrit

$$\frac{\partial^2 \psi}{\partial t^2} + (c(q))^2 \Delta \psi = 0, \quad (2.2.3)$$

avec une fonction  $c(q) > 0$  qui correspond à l’indice du milieu. On peut la ré-écrire<sup>1</sup> sous la forme (2.2.1) avec l’opérateur

$$\hat{H}_{\pm} = \pm i \hbar c(q) \sqrt{-\Delta}.$$

---

1. L’opérateur  $\sqrt{-\Delta}$  est défini par son action sur la base de Fourier  $\varphi_k(q) = e^{ik \cdot q}$ ,  $k \in \mathbb{R}^3$ ,  $q \in \mathbb{R}^3$ , observant que  $\Delta \varphi_k = -\|k\|^2 \varphi_k$ , par

$$\sqrt{-\Delta} \varphi_k = \|k\| \varphi_k.$$

On vérifie l’équation (2.2.1),  $\frac{\partial \psi}{\partial t} = -\frac{i}{\hbar} \hat{H}_{\pm} \psi$  avec  $\hat{H}_{\pm} = \pm i \hbar c(q) \sqrt{-\Delta}$  donne  $\hat{H}_{\pm}^2 = \hbar^2 (c(q))^2 \Delta$  et donc  $\frac{\partial^2 \psi}{\partial t^2} = \left( -\frac{i}{\hbar} \hat{H}_{\pm} \right)^2 \psi = - (c(q))^2 \Delta \psi$  qui est l’équation d’onde (2.2.3).

D'après le principe de correspondance (2.2.2), on obtient la fonction de Hamilton de "l'optique géométrique" :

$$H_{\pm}(q, p) = \pm i\hbar c(q) \|p\|. \quad (2.2.4)$$

Ici  $\hbar$  ne joue aucun rôle et on pourrait fixer  $\hbar = 1$ . La fonction  $H(q, p)$  a donc l'unité de  $1/t$ , c'est une fréquence que l'on note parfois  $\omega(q, p)$  et que l'on appelle **relation de dispersion**, et  $\frac{\partial q}{\partial t} = \frac{\partial H}{\partial p} = \frac{\partial \omega}{\partial p}$  s'appelle la **vitesse de groupe**.

Par exemple utilisant les équations de Hamilton, on peut déduire la loi de la réfraction en optique :

**Exercice 2.2.7. (\*) Réfraction : Loi de Snell (1621)-Descartes (1637) (trouvée avant par Ibn Sahl (983) à Bagdad).** Supposons que dans l'espace  $q = (q_1, q_2) \in \mathbb{R}^2$ , il y a une interface en  $q_1 = 0$ , de sorte que pour  $q_1 < 0$  on a  $c(q) = c$  et pour  $q_1 > 0$  on a  $c(q) = c'$  avec  $c' < c$ . Déduire de (2.2.4), de l'équation de mouvement de Hamilton (2.1.3) et de la conservation de l'énergie que : dans chaque domaine un rayon (lumineux ou sonore) se déplace en ligne droite et à l'interface on a

$$\frac{1}{c} \sin \theta = \frac{1}{c'} \sin \theta' \quad (2.2.5)$$

où  $\theta$  (respect.  $\theta'$ ) est l'angle du rayon avec la normale de l'interface. Voir schéma.

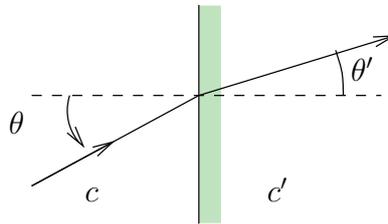


FIGURE 2.2.1 – Si  $c(q) = c, c'$  prend deux valeurs différentes de part et d'autre d'une interface alors le rayon est dévié d'après la loi de Snell-Descartes (2.2.5).

*Remarque 2.2.8.* En optique, on définit l'indice du milieu par  $n(q) = \frac{c_{\text{vide}}}{c(q)}$  où  $c_{\text{vide}}$  est la vitesse de la lumière dans le vide.

**Solution :** L'énergie est  $H(q) = c(q_1) \|k\|$  avec le vecteur d'onde  $k = \frac{p}{\hbar}$  (ou poser  $\hbar = 1$ ). L'équation de mouvement (2.1.3) donne  $\partial_t k_2 = -\partial_{q_2} H = 0$  donc  $k'_2 = k_2$  à l'interface. La conservation de l'énergie à l'interface donne

$$c \|k\| = c' \|k'\|$$

On a

$$\sin \theta = \frac{k_2}{\|k\|}, \quad \sin \theta' = \frac{k'_2}{\|k'\|}.$$

On déduit

$$\sin \theta' = \frac{k_2'}{\|k'\|} = \frac{k_2}{\frac{c}{c'} \|k\|} = \frac{c'}{c} \sin \theta$$

et donc (2.2.5).

## 2.3 Formule de Weyl semi-classique

Pour un opérateur donné  $\hat{H}$ , on souhaite estimer la position des valeurs propres  $E_1 \leq E_2 \leq \dots$  qui sont des niveaux d'énergie. Rappelons d'abord la signification physique de des vecteurs (fonctions) propres  $\varphi_j$  associés par  $\hat{H}\varphi_j = E_j\varphi_j$ .

**Proposition 2.3.1. “Ondes stationnaires”.** Si  $\hat{H}\varphi = E\varphi$  avec une fonction propre  $\varphi : q \rightarrow \varphi(q)$  et valeur propre (énergie)  $E$ , alors la fonction

$$\psi(q, t) = e^{-i\frac{E}{\hbar}t} \varphi(q)$$

est solution de l'équation d'évolution (2.2.1),  $i\hbar\partial_t\psi = \hat{H}\psi$ , appelée **solution stationnaire** car le module  $|\psi(q, t)| = |\varphi(q)|$  est indépendant du temps.

*Démonstration.* On calcule  $\partial_t\psi = (-i\frac{E}{\hbar})\psi$  donc  $i\hbar\partial_t\psi = E\psi = \hat{H}\psi$ . □

*Remarque 2.3.2.* Les vecteurs propres d'un opérateur  $\hat{H}$  sont en général impossible à calculer exactement. On sait parfois montrer leur existence seulement. Expérimentalement, on les observe approximativement en régime stationnaire, c'est à dire en excitant le système par un signal d'amplitude temporelle  $e^{i\omega t} = e^{-iEt/\hbar}$  appelé signal monochromatique (ce peut être une onde lumineuse).

L'intérêt de la **formule de Weyl** ci-dessous ou formule du “comptage d'états ondulatoires” est de donner approximativement la position des niveaux d'énergie  $E_j$  et leur densité moyenne. (Voir par exemple ce **cours de mécanique quantique**).

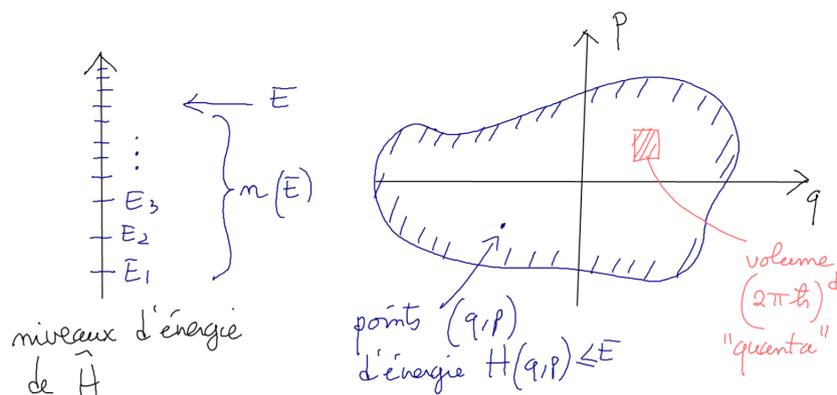
**Théorème 2.3.3. “Formule de Weyl semi-classique”.** Si un opérateur (Hamiltonien quantique)  $\hat{H}$  est défini à partir d'un Hamiltonien classique  $H(q, p)$  où  $q, p \in \mathbb{R}^d$  sont les variables de position et impulsion pour  $d$  degrés de liberté et dont l'expression dépend du paramètre  $\hbar$  (la **constante de Planck**) et si on note

$$\hat{H}\varphi_j = E_j\varphi_j$$

avec  $E_j$  niveaux d'énergie (valeurs propres) de  $\hat{H}$  alors le nombre de valeurs propres  $E_j$  inférieures à  $E$  donnée, noté  $n(E) := \#\{j \mid E_j \leq E\}$  est

$$n(E) = \frac{\text{Vol}(\{(q, p) \text{ t.q. } H(q, p) \leq E\})}{(2\pi\hbar)^d} + o(\hbar^{-d}) \quad (2.3.1)$$

où le terme d'erreur devient négligeable dans la limite où  $\hbar$  est petit devant les grandeurs du problème. Dans le terme de droite, il apparaît le volume dans l'espace des phases, occupé par les points  $(q, p)$  d'énergie inférieure à  $E$ .



*Remarque 2.3.4.*

- En termes plus simples, cette formule montre que à cause du **principe d'incertitude**, un état quantique occupe le volume  $(2\pi\hbar)^d$  appelé « quanta » dans l'espace des phases  $(q, p)$ .
- On utilisera cette formule de la façon suivante. Par dérivation, on peut déduire une approximation de la **densité d'états**  $\rho(E) := \left(\frac{dn}{dE}\right)(E)$  et pour un petit intervalle d'énergie de largeur  $\Delta E$  autour de  $E$ , déduire le nombre d'états  $\Delta n = \rho(E) \Delta E$ . Finalement en supposant tous ces états équiprobables (ce sera "l'hypothèse ergodique"), on déduira une expression de l'**entropie**  $S(E) = k \ln(\Delta n)$  (avec la **constante de Boltzmann**  $k$ ).

On précise la signification du volume  $\text{Vol}(\cdot)$  dans la formule de Weyl (2.3.1). Cela sera utile pour quelques exemples qui suivront.

### 2.3.1 Rappel de géométrie sur le volume

Par définition, l'espace  $\mathbb{R}^n$  contenant des points  $x = (x_1, x_2, \dots, x_n)$  caractérisés par  $n$  variables indépendantes. On dit que sa dimension est  $\dim \mathbb{R}^n = n$ . La mesure de l'espace occupé par certains points s'appelle un **volume** en général, mais dans les cas particulier de la dimension 1 on appelle cela **longueur**, et en dimension 2 on l'appelle **surface**. Dans la suite il cependant plus simplement d'utiliser le terme volume pour toutes les dimensions.

**Sphères :** Pour  $R > 0$  fixé, l'ensemble

$$S^{n-1} := \{x \text{ t.q. } x_1^2 + x_2^2 + \dots + x_n^2 = R^2\}$$

est un ensemble de dimension  $n - 1$  (car il n'y a plus que  $n - 1$  variables indépendantes à cause de la contrainte) appelé la « **sphère**  $S^{n-1}$  ». Par exemple  $S^1$  est le cercle de rayon  $R$  dans  $\mathbb{R}^2$ .  $S^2$  est la sphère habituelle dans  $\mathbb{R}^3$ .  $S^0$  est constitué des deux points  $\{-R, +R\} \subset \mathbb{R}$ . On montre que le « volume » de  $S^{n-1}$  est  $\text{Vol}(S^{n-1}) = C_{n-1}R^{n-1}$  avec  $C_{n-1} \geq 0$  qui ne dépend pas de  $R$ . Par exemple  $\text{Vol}(S^1) = 2\pi R$  est aussi appelé périmètre du cercle.  $\text{Vol}(S^2) = 4\pi R^2$ ,  $\text{Vol}(S^0) = 0$ , etc.

**Boules :** L'ensemble

$$B^n := \{x \text{ t.q. } x_1^2 + x_2^2 + \dots + x_n^2 \leq R^2\} \quad (2.3.2)$$

est un ensemble de dimension  $n$  appelée la « **boule fermée**  $B^n$  » que l'on appelle aussi intérieur de la sphère  $S^{n-1}$ . Par exemple  $B^2$  est en fait le disque de rayon  $R$  dans  $\mathbb{R}^2$ .  $B^1$  est le segment  $[-R, +R] \subset \mathbb{R}$ . On montre que le « volume » de  $B^n$  est  $\text{Vol}(B^n) = C'_n R^n$  avec  $C'_n > 0$  qui ne dépend pas de  $R$ . Par exemple  $\text{Vol}(B^2) = \pi R^2$  est aussi appelé surface du disque.  $\text{Vol}(B^1) = 2R$ ,  $\text{Vol}(B^3) = \frac{4}{3}\pi R^3$ , etc. Pour  $n \gg 1$ , on a l'**équivalence**

$$\frac{1}{n} \ln C'_n \sim \frac{1}{2} \ln \left( \frac{2\pi e}{n} \right), \quad (2.3.3)$$

autrement dit  $C'_n \asymp \left( \frac{2\pi e}{n} \right)^{n/2}$ .

**Cubes :** Une situation plus simple est l'ensemble

$$C^n := \{x \text{ t.q. } 0 \leq x_i \leq R, \forall i = 1 \dots n\}$$

qui est un espace de dimension  $n$  appelé **cube** de côté  $R$ . Son volume est simplement le produit des côtés  $\text{Vol}(C^n) = R^n$ . Par exemple  $C^2$  est en fait un carré de surface  $\text{Vol}(C^2) = R^2$ .  $C^1$  est un segment de longueur  $\text{Vol}(C^1) = R$ .

**Produit cartésien d'ensembles :** Considérons maintenant un cas intermédiaire. Par exemple si on considère l'ensemble de points dans  $\mathbb{R}^4$

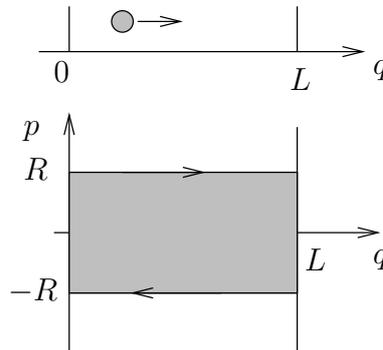
$$E := \{x \text{ t.q. } 0 \leq x_1 \leq L, 0 \leq x_2 \leq L, x_3^2 + x_4^2 = R^2\},$$

on observe que les conditions sur les variables  $(x_1, x_2)$  correspondent à un cube  $C^2$  (un carré) que les conditions sur les variables  $(x_3, x_4)$  correspondent à une sphère  $S^{2-1} = S^1$  (un cercle). C'est deux conditions sont indépendantes, par conséquent  $E = C^2 \times S^1$  est le **produit cartésien** d'un carré  $C^2$  de dimension 2 et d'une sphère  $S^1$  de dimension 1. Ainsi  $\text{Vol}(E) = \text{Vol}(C^2) \text{Vol}(S^1) = L^2 2\pi R$ .

### 2.3.2 Exemple : une particule libre dans une boîte uni-dimensionnelle.

Nous traitons un cas simple pour illustrer la formule de Weyl (2.3.1). On considère un segment de longueur  $L$ . La position d'une particule libre de masse  $m$  est  $q(t) \in [0, L]$ . Particule libre signifie qu'il n'y a pas de force extérieure, sauf au moment du rebond sur les parois en  $q = 0$  et  $q = L$ . Pour  $q \in ]0, L[$ , son énergie totale est donc

$$H(q, p) \stackrel{(2.1.2)}{=} \frac{p^2}{2m}.$$



On se fixe une énergie  $E \geq 0$ . On a

$$H(q, p) \leq E \Leftrightarrow |p| \leq R = \sqrt{2mE}$$

Le volume recherché est un rectangle

$$\begin{aligned} \text{Vol}(\{(q, p) \text{ t.q. } H(q, p) \leq E\}) &= \text{Vol}(\{q \in [0, L]\}) \times \text{Vol}(\{p \in [-R, R]\}) = L \times 2R \\ &= 2L\sqrt{2mE} \end{aligned}$$

Donc la formule de Weyl donne

$$n(E) \stackrel{(2.3.1)}{=} \frac{1}{2\pi\hbar} 2L\sqrt{2mE} + o(\hbar^{-1}) = \frac{1}{\pi\hbar} L\sqrt{2mE} + o(\hbar^{-1}).$$

*Remarque 2.3.5.* On déduit inversement que l'énergie  $E_n$  du niveau  $n$  est approximativement

$$E_n = \frac{1}{2m} \left( \frac{\pi \hbar n}{L} \right)^2 + o(1).$$

On peut comparer à la solution exacte  $E_n^{(\text{exact})} = \frac{1}{2m} \left( \frac{\pi \hbar n}{L} \right)^2$  qui s'obtient en résolvant l'équation de Schrödinger stationnaire, [Faure \(2014a\)](#). On observe qu'il n'y a pas de terme d'erreur.

## 2.4 Modèle du gaz parfait : $N$ particules libres identiques dans une boîte 3D

*Références :* [Diu et al. \(1989\)](#) p.83. [Couture & Zitoun \(1992\)](#) p.360.

### 2.4.1 Calcul de la densité d'états

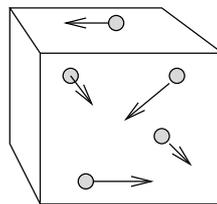
On considère un espace  $\Omega \subset \mathbb{R}^3$  de volume fini  $V = \text{Vol}(\Omega)$  qui contient  $N$  particules ponctuelles.

La position de la particule  $j$  est notée  $\vec{q}^{(j)} = (q_1^{(j)}, q_2^{(j)}, q_3^{(j)}) \in \mathbb{R}^3$ . Son impulsion est notée  $\vec{p}^{(j)} = (p_1^{(j)}, p_2^{(j)}, p_3^{(j)}) \in \mathbb{R}^3$ . Ainsi l'état du système  $(q, p)$ , i.e. configuration des particules, appelé **micro-état** est caractérisé par la position de toutes les particules :

$$q = (\vec{q}^{(1)}, \vec{q}^{(2)}, \dots, \vec{q}^{(N)}) \in \mathbb{R}^{3N}$$

et leur impulsions :

$$p = (\vec{p}^{(1)}, \vec{p}^{(2)}, \dots, \vec{p}^{(N)}) \in \mathbb{R}^{3N}.$$



On suppose les particules libres, c'est à dire que leur énergie est seulement l'énergie cinétique :

$$H(q, p) = \sum_{j=1}^N \frac{1}{2m} \|\vec{p}^{(j)}\|^2.$$

## 2.4. MODÈLE DU GAZ PARFAIT : $N$ PARTICULES LIBRES IDENTIQUES DANS UNE BOITE 3D

Pour appliquer la formule de Weyl (2.3.1), on se fixe une énergie  $E \geq 0$ . On a

$$\begin{aligned} H(q, p) \leq E &\Leftrightarrow \sum_{j=1}^N \|\vec{p}^{(j)}\|^2 \leq 2mE = R^2 \\ &\Leftrightarrow q \in \underbrace{\Omega \times \dots \times \Omega}_N \text{ et } p \in B^{3N}(R) \end{aligned}$$

car d'après (2.3.2), on reconnaît l'équation d'une boule  $B^{3N}(R)$  de rayon  $R$  dont le volume est donnée en (2.3.3). On suppose les **particules identiques** et donc indiscernables. Cela signifie que la permutation  $j \longleftrightarrow j'$  de deux particules quelconques ne change pas l'état et qu'il faut donc réduire le volume de l'espace de phase par un facteur  $N!$  qui est le nombre de **permutations**. La **formule de Stirling** donne  $\frac{1}{n} \ln(n!) \sim \ln\left(\frac{n}{e}\right)$  si  $n \gg 1$ , avec  $e = \exp(1)$  soit

$$n! \asymp \left(\frac{n}{e}\right)^n.$$

Alors

$$\text{Vol}(\{(q, p \text{ t.q. } H(q, p) \leq E)\}) = \frac{1}{N!} (\text{Vol}(\Omega))^N \text{Vol}(B^{3N}(R)) \stackrel{(2.3.3)}{=} \frac{1}{N!} V^N C_{3N} R^N.$$

La formule de Weyl donne alors

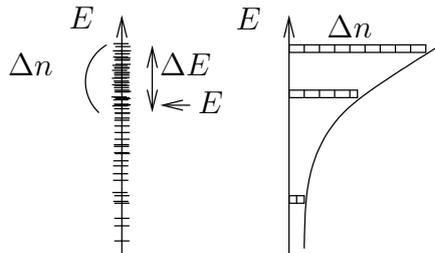
$$\begin{aligned} n(E) &\stackrel{(2.3.1)}{=} \frac{\text{Vol}(\{(q, p \text{ t.q. } H(q, p) \leq E)\})}{(2\pi\hbar)^{3N}} \\ &= \frac{V^N C_{3N} R^N}{N! (2\pi\hbar)^{3N}} \asymp \left( \frac{Ve}{3N(2\pi\hbar)^3} \left( \frac{2\pi e 2mE}{3N} \right)^{3/2} \right)^N \\ &= \left( \frac{Ve^{5/2}}{N} \left( \frac{mE}{3\pi\hbar^2 N} \right)^{3/2} \right)^N \end{aligned} \quad (2.4.1)$$

On utilisera cette formule de la façon suivante. Par dérivation, on déduit une approximation de la **densité d'états**

$$\rho(E) := \left( \frac{dn}{dE} \right) (E) = n(E) \left( \frac{3N}{2} \right) \frac{1}{E}. \quad (2.4.2)$$

Pour un petit intervalle d'énergie de largeur  $\Delta E$  autour de  $E$  donnés, on déduit le nombre d'états  $\Delta n$  par

$$\frac{\Delta n}{\Delta E} = \rho(E) \stackrel{(2.4.2)}{=} n(E) \left( \frac{3N}{2} \right) \frac{1}{E} \quad (2.4.3)$$





# Chapitre 3

## Modèle de Boltzmann Gibbs de l'équilibre thermodynamique

Dans ce chapitre on considère un ensemble physique de plusieurs particules comme le modèle du gaz parfait étudié en section 2.4. On va faire une hypothèse de “chaos”, appelé hypothèse ergodique et déduire des propriétés importantes sur l'état d'équilibre et sur le comportement de deux systèmes qui sont en contact.

A partir de maintenant les notions ne sont plus très rigoureuses. On va se baser plus sur l'intuition physique et sur quelques modèles physiques particuliers.

### 3.1 Système thermodynamique et hypothèse ergodique

Dans cette première section on donne une définition un peu générale de système thermodynamique. Le gaz parfait sera un exemple étudié dans la suite.

**Définition 3.1.1.** Un **système thermodynamique** est un modèle de dynamique en physique contenant **un grand nombre  $N$  particules** (ou plus généralement  $N$  degrés de liberté) qui interagissent entre elles avec des **forces à courte portée**. On considère **toutes les grandeurs physiques conservées** comme  $(E, V, N)$  : l'énergie totale  $E$  (car si le système est isolé, l'énergie est conservée en mécanique Hamiltonienne), le volume  $V$  occupé par le système, le nombre de particules  $N$ . On fait l'**hypothèse ergodique** qui suppose que à part ces quantités conservées comme  $(E, V, N)$ , tous les autres paramètres évoluent de façon "totalement chaotique", i.e. imprévisible et plus précisément selon une **dynamique de Markov mélangeante** qui fait que toutes les configurations possibles appelées **micro-état** (par exemple position et impulsion précise des particules) deviennent **équi-probables**. C'est l'**état d'équilibre thermodynamique** (aussi appelé **ensemble micro-canonique**, d'après Gibbs 1902).

L'entropie de Shannon de cet état d'équilibre contenant  $\Delta n$  micro-états est

$$S_{\text{Shannon}} \stackrel{(1.5.10)}{=} \ln(\Delta n)$$

et l'**entropie de Boltzmann** est définie par (la **constante de Boltzmann** fixée en 1966 est  $k = 1.380649 \cdot 10^{-23} \text{ J/K}$ )

$$S(E, V, N) := k S_{\text{Shannon}} = k \ln(\Delta n) \quad (3.1.1)$$

et **caractérise le système thermodynamique à l'échelle macroscopique**. D'après la localité des interactions, la fonction entropie est supposée **extensive**, c'est à dire que pour tout facteur de proportionnalité  $\lambda > 0$ ,

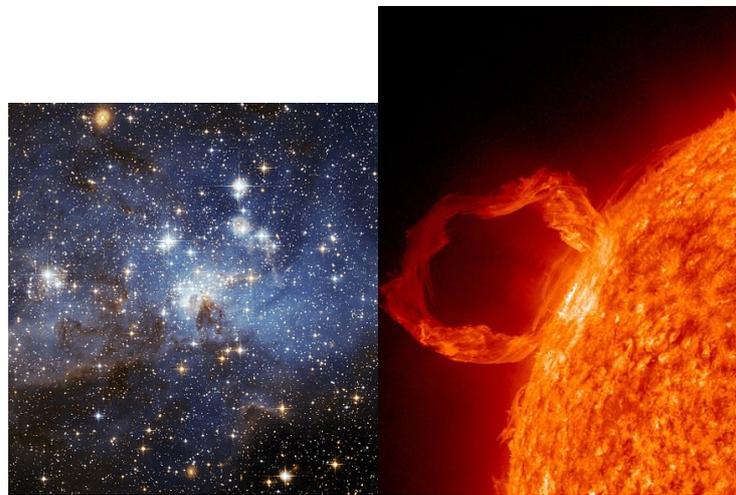
$$S(\lambda E, \lambda V, \lambda N) = \lambda S(E, V, N). \quad (3.1.2)$$

Finalement on suppose que la fonction  $S(E, V, N)$  **est concave**, i.e.  $\frac{\partial^2 S}{\partial E^2} < 0$ . Cela sera nécessaire pour déduire la stabilité du système thermodynamique, cf Prop. 3.3.3.

### 3.1.1 Exemples et contre-exemples de systèmes thermodynamiques

- Le modèle le plus simple est le **gaz parfait** présenté plus haut :  $N$  particules libres ponctuelles (ou petites sphères) qui subissent des chocs. Du fait de ces chocs on imagine que la dynamique totale est chaotique et donc il y a mélange du système et convergence vers un unique état d'équilibre qui est la mesure de Liouville, donc la mesure uniforme sur les  $\Delta n$  états, conformément à l'hypothèse ergodique. Cependant cette hypothèse de "chaos" est démontrée des modèles très simplifiés comme le billard de Sinaï mais reste **une grande conjecture en science** pourtant à la base de la physique statistique. Voir cette [vidéo avec 1 bille](#) et [vidéo avec 1e6 billes](#).

- Un autre modèle plus sophistiqué est un ensemble de  $N$  particules mais avec des **forces entre particules de courte portée**, décrite par une énergie potentielle  $U(|\vec{q}^{(j)} - \vec{q}^{(k)}|)$  où  $\vec{q}^{(j)} \in \mathbb{R}^3$  est la position de la particule  $j$  et  $U(\cdot)$  peut être une énergie effective entre molécules comme celui de **Lennard-Jones**. Pour ce modèle on ne sait pas calculer l'état d'équilibre de façon exacte. Il apparaît des phénomènes de **transition de phases** étudiée au chapitre suivant comme la condensation des particules en des **phases liquide** ou **solide** (cristallin) qui peuvent coexister avec le gaz. On verra des méthodes approximative pour expliquer cela. Les phases gaz ou liquide où les atomes peuvent se déplacer sont appelés **fluide** au contraire de l'état cristallin. Voir vidéo. Par exemple un **plasma** est un fluide de noyaux atomiques avec des électrons libres.



- Attention dans les systèmes décrit précédemment, gaz, liquide ou solide, l'état d'équilibre est atteints que sur une petite échelle de la taille de quelques micromètres appelée **échelle mésoscopique**, car grande devant l'échelle microscopique des particules et petite devant l'échelle humaine macroscopique. Ainsi un système thermodynamique est toujours de taille mésoscopique. **A l'échelle macroscopique, il faut considérer les fluides comme un ensemble de plusieurs systèmes thermodynamiques qui interagissent entre eux.** On étudiera ces interactions dans la suite.
- D'autres modèles importants en physique sont les **ondes de vibration d'un solide cristallin** (appelés phonons, bien que les phonons soient précisément la quantification de ces ondes). A cause des non linéarité, il y a des interactions entre les paquets d'ondes (appelés quasi-particules) et on peut supposer que ces comportements sont chaotiques, justifiant l'hypothèse ergodique et donc le fait de considérer ces ondes comme un système thermodynamique.
- Dans un solide, les moments magnétiques des particules, appelés spins interagissent entre eux. Un modèle simplifié est de considérer des états de **spins sur un réseau fixé**. Ces spins interagissent avec leur proches voisins. De même, supposant

quelques interactions générant du chaos, on peut considérer ce modèle comme un système thermodynamique. Son étude montre l'existence de transitions de phase, par exemple des états où les spins sont collectivement parallèle, donnant une **aimantation** au matériau. On étudiera le modèle d'Ising du ferromagnétisme qui est un exemple simple de tel modèle. (Pour avoir la thermalization il faut supposer que les **ondes de spins** ont des comportements non linéaires?)

- Un **système dynamique simple n'est pas un système thermodynamique** (exemple une planète considérée comme un point ou un pendule). On ne le décrit pas par une mesure de probabilité d'équilibre qui n'évolue pas mais par son état dynamique qui évolue d'après les lois de Newton par exemple. Par contre un tel système dynamique interagit éventuellement avec des systèmes thermodynamiques, par exemple le pendule bouge dans un fluide comme l'air ou l'eau. Il sera important d'étudier les interactions entre ces différents systèmes et observer les phénomènes de **fluctuation et dissipation**.

*Remarque 3.1.2. “Sur les quantités conservées”.* Plus généralement, on peut avoir un système thermodynamique avec  $p$  familles différentes de particules où le nombre de particule  $N_j$  dans chaque famille est conservé  $j = 1, \dots, p$ . On considérera l'ensemble de ces quantités conservées  $N_1, \dots, N_p$ . Par exemple en chimie, avec des réactions comme  $2H_2O \Leftrightarrow 2H_2 + O_2$ , le nombre de molécules  $N_{H_2O}$  n'est pas conservé mais le nombre d'atomes  $N_H, N_O$  est conservé. Dans le cas des photons, le nombre de photons  $N_\gamma$  n'est pas une quantité conservée, car un photon peut être absorbé et/ou émit par la matière.

En mécanique des fluides on peut avoir un déplacement du système et considérer l'**impulsion totale** qui est une grandeur conservée, ou le moment cinétique total.

*Remarque 3.1.3. “Sur la constante de Boltzmann  $k$ ”.* On verra que  $S_{\text{Shannon}}(E) \approx N$  qui est de l'ordre du **nombre d'Avogadro**  $\mathcal{N}_{\text{Av.}} = 6.02 \cdot 10^{23}$  pour un gaz à l'échelle humaine. Pour cette raison on multiplie par  $k = \frac{1}{\mathcal{N}_{\text{Av.}}} R$  dans la définition (3.1.1) de l'entropie de Boltzmann, où  $R = 8.31 J/K$  est une grandeur à l'échelle humaine. L'entropie de Shannon  $S_{\text{Shannon}}$  n'a pas d'unité, mais l'entropie de Boltzmann  $S$  a l'unité  $J/K$ .

*Remarque 3.1.4. “Sur l'entropie à basse énergie”.* Pour un système à basse énergie il y a un seul état fondamental,  $\Delta n = 1$  donc l'entropie est nulle  $S = k \ln \Delta n = 0$ . Cette propriété s'appelle le **3eme principe de la thermodynamique** ou principe de Nernst (1906).

### 3.1.2 Equations d'états

Pour cela voici des définitions générales qui vont être utiles pour étudier plus loin, en proposition 3.3.3, la relaxation thermique vers l'état d'équilibre.

**Définition 3.1.5.** Pour une fonction entropie  $S(E, V, N)$  donnée, on définit la **température**  $T$ , la **pression**  $p$ , le **potentiel chimique**  $\mu$  par

$$\frac{\partial S}{\partial E} = \frac{1}{T}, \quad \frac{\partial S}{\partial V} = \frac{p}{T}, \quad \frac{\partial S}{\partial N} = -\frac{\mu}{T},$$

appelée **équations d'état**.

*Remarque 3.1.6.* Noter que ce sont des fonctions  $T(E, V, N)$ ,  $p(E, V, N)$ ,  $\mu(E, V, N)$  et que l'unité de la température  $T$  est en  $K$ , la pression  $p$  est en  $\frac{J}{m^3} \equiv Pa$ , le potentiel chimique  $\mu$  est en  $J$ .

*Remarque 3.1.7.* D'après la définition générale de la différentielle d'une fonction

$$dS = \left( \frac{\partial S}{\partial E} \right) dE + \left( \frac{\partial S}{\partial V} \right) dV + \left( \frac{\partial S}{\partial N} \right) dN$$

on peut écrire la définition de  $T, p, \mu$  sous la forme

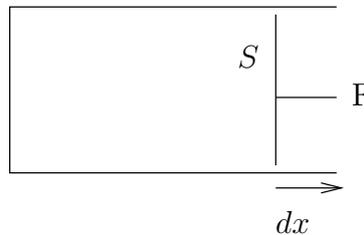
$$dS = \frac{1}{T} dE + \frac{p}{T} dV - \frac{\mu}{T} dN \quad (3.1.3)$$

ou

$$dE = T dS - p dV + \mu dN.$$

*Remarque 3.1.8.* On peut justifier dès à présent que la pression  $p$  correspond à la **pression mécanique**  $p_{\text{mec.}} = \frac{F}{S}$  qui est une force par surface de la façon suivante. Considérons que une paroi de surface  $S$  de l'enceinte du gaz peut se déplacer de  $dx$  et que le déplacement est **isentropique**, i.e. à entropie constante, alors  $0 = dS = \frac{1}{T} (dE + p dV)$  donc  $\frac{dE}{dV} = -p$ . Alors le changement de volume est  $dV = S dx$ . La force est  $F = -\frac{dE}{dx}$  et donc

$$p_{\text{mec.}} = \frac{F}{S} = \frac{1}{S} \left( -\frac{dE}{dx} \right) = -\frac{dE}{dV} = p.$$



*Remarque 3.1.9. “sur le signe de la pression  $p$ ”* : si  $p < 0 \Leftrightarrow \left(\frac{\partial S}{\partial V}\right) < 0$ , on aurait que  $S$  augmente en diminuant  $V$ . Par l'évolution naturelle (2eme principe,  $S$  augmente) le système aurait tendance à s'effondrer jusqu'à retrouver une pression positive (ou pas). Ce n'est donc pas une situation stable. Par contre si  $p > 0$ , la système a tendance à s'étendre, mais en pratique il est limité par des parois (gaz dans une enceinte) ou des forces de rappels comme la force gravitationnelle à longue portée qui limite la taille d'une étoile ou d'une atmosphère planétaire.

Pour la même raison  $T > 0$  en général, car si  $T < 0$ , alors le système a tendance à donner de l'énergie pour augmenter l'entropie, jusqu'à ce que  $T > 0$ . C'est le cas dans un système de spins, cf TD.

*Remarque 3.1.10. “sur le potentiel chimique  $\mu$ ”*. (ref : [Diu et al. \(1989, p.168\)](#)). On va expliquer que l'on a  $\mu < 0$  pour un gaz de particules sans interactions ou avec des interactions attractives, mais on peut avoir  $\mu > 0$  dans un système où il y a des forces suffisamment répulsives entre particules, comme entre les électrons par leur charges ou par l'[effet d'exclusion de Pauli](#).

En effet, si on ajoute un certain nombre  $\Delta N > 0$  de particules immobiles au système (c'est à dire avec énergie cinétique nulle), alors cela induit une variation d'entropie positive  $\Delta S > 0$  car cela ajoute des degrés de liberté donc augmente le nombre de micro-états  $\Delta n$ . Si il n'y a pas d'interactions entre particules (i.e. pas d'énergie potentielle  $\Delta E_p = 0$ ), alors l'énergie totale est inchangée  $\Delta E = \Delta E_c + \Delta E_p = 0$ . Si au contraire il y a des interactions répulsives entre les particules, i.e.  $\Delta E_p > 0$ , alors il y a une variation positive de l'énergie totale  $\Delta E > 0$  et si il y a des interactions attractives, i.e.  $\Delta E_p < 0$ , alors  $\Delta E < 0$ . Dans tous les cas, si  $T > 0$ , on obtient

$$\Delta S = \left(\frac{\partial S}{\partial E}\right) \Delta E + \left(\frac{\partial S}{\partial N}\right) \Delta N = \frac{1}{T} \Delta E - \frac{\mu}{T} \Delta N$$

Donc

$$\mu = \frac{1}{\Delta N} (\Delta E - T \Delta S)$$

peut être positif si  $\Delta E > T \Delta S > 0$ , i.e. des forces suffisamment répulsives mais  $\mu < 0$  si  $\Delta E = 0$  ou  $\Delta E < 0$ .

**Proposition 3.1.11. “Fonctions intensives”**. Pour tout facteur de proportionnalité  $\lambda > 0$ ,

$$\begin{aligned} T(\lambda E, \lambda V, \lambda N) &= T(E, V, N), \\ p(\lambda E, \lambda V, \lambda N) &= p(E, V, N), \quad \mu(\lambda E, \lambda V, \lambda N) = \mu(E, V, N). \end{aligned}$$

*Démonstration.* Par exemple pour la température,

$$\frac{1}{T(\lambda E, \lambda V, \lambda N)} = \frac{\partial S(\lambda E, \lambda V, \lambda N)}{\partial(\lambda E)} \stackrel{(3.1.2)}{=} \frac{\lambda \partial S(E, V, N)}{\lambda \partial E} = \frac{1}{T(E, V, N)}.$$

□

**Proposition 3.1.12.** “*Relation d'Euler*”. On a

$$S = \frac{1}{T}E + \frac{p}{T}V - \frac{\mu}{T}N$$

qui s'écrit aussi

$$E = TS - pV + \mu N.$$

à comparer avec (3.1.3).

*Remarque 3.1.13.* Cela montre que deux équations d'états déterminent la 3eme.

*Démonstration.* D'après l'extensivité  $S(\lambda E, \lambda V, \lambda N) \stackrel{(3.2.2)}{=} \lambda S(E, V, N)$ ,

$$\frac{d}{d\lambda} S(\lambda E, \lambda V, \lambda N) = \frac{d}{d\lambda} (\lambda S(E, V, N)) = S(E, V, N)$$

En  $\lambda = 1$  cela donne

$$\begin{aligned} \left(\frac{\partial S}{\partial E}\right) E + \left(\frac{\partial S}{\partial V}\right) V + \left(\frac{\partial S}{\partial N}\right) N &= S \\ \Leftrightarrow \frac{1}{T}E + \frac{p}{T}V - \frac{\mu}{T}N &= S \end{aligned}$$

□

*Remarque 3.1.14.* En mathématiques, la *relation d'Euler* s'applique plus généralement, aux fonctions homogènes.

## 3.2 Exemple du gaz parfait : entropie et équation d'états

On va vérifier que le gaz parfait est un exemple de système thermodynamique comme défini en 3.1.1 et calculer précisément sa fonction entropie  $S(E, V, N)$  pour déduire les équations d'états.

On calcule l'entropie de Boltzmann (3.1.1) du gaz parfait à partir de la densité d'états (2.4.2). On verra l'utilité dans la suite.

$$\begin{aligned} S(E, V, N) &\stackrel{(3.1.1)}{=} k \ln(\Delta n) \stackrel{(2.4.3)}{=} k \ln \left( n(E) \left( \frac{3N}{2} \right) \frac{\Delta E}{E} \right) \\ &= k \ln(n(E)) + k \ln \left( \left( \frac{3N}{2} \right) \frac{\Delta E}{E} \right) \end{aligned}$$

Avec (2.4.1) on obtient que  $S \approx N$  et donc le deuxième terme est négligeable et

**Proposition 3.2.1.** “Entropie du gaz parfait”, “formule de Sachur Tetrode”.  
*Diu et al. (1989, p.197).*

$$S(E, V, N) \stackrel{(2.4.1)}{=} kN \ln \left( \frac{V e^{5/2}}{N} \left( \frac{mE}{3\pi\hbar^2 N} \right)^{3/2} \right). \quad (3.2.1)$$

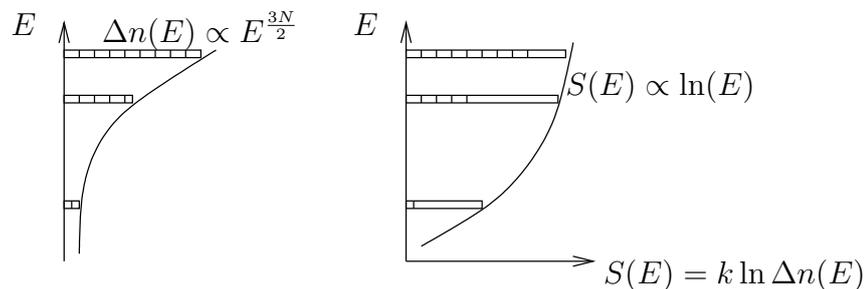
Une conséquence importante de cette formule est le comportement de  $S$  si on considère un système identique plus grand (par exemple en doublant l'énergie, le volume et le nombre de particules).

**Proposition 3.2.2.** “Extensivité de l'entropie”. Pour tout facteur de proportionnalité  $\lambda > 0$ ,

$$S(\lambda E, \lambda V, \lambda N) = \lambda S(E, V, N). \quad (3.2.2)$$

*Remarque 3.2.3.* Le facteur  $N!$  venant de l'indiscernabilité des particules est essentiel pour déduire cette propriété d'extensivité. Sinon il apparaît le “paradoxe de Gibbs”.

Voici un schéma représentant pour le modèle du gaz parfait, le nombre de micro-états  $\Delta n(E)$  à l'énergie  $E \geq 0$  dans un petit intervalle  $\Delta E$  fixé d'énergie.



**Concavité de l'entropie :** On a vu que pour le gaz parfait

$$\Delta n(E) \propto E^{\frac{3N}{2}}$$

augmente très vite avec l'énergie et par conséquent, l'entropie se comporte comme

$$S(E) = k \ln(\Delta n) \approx k \frac{3N}{2} \ln E \propto \ln E$$

Ses dérivées sont

$$\begin{aligned} \frac{1}{T} &= \frac{\partial S}{\partial E} \propto \frac{1}{E} > 0 \\ \frac{\partial^2 S}{\partial E^2} &\propto -\frac{1}{E^2} < 0 \end{aligned} \quad (3.2.3)$$

donc  $S(E)$  est **une fonction concave**<sup>1</sup>. Ce sera important dans la suite pour établir la stabilité des systèmes thermodynamiques.

Dans le cas du gaz parfait, en utilisant l'expression de l'entropie (3.2.1), on peut expliciter les équations d'état.

**Proposition 3.2.4. “Equations d'états du gaz parfait”.** Pour le gaz parfait on a les trois équations d'état

$$\frac{E}{N} = \frac{3}{2}kT \quad (3.2.4)$$

signifiant que l'énergie moyenne par particule est proportionnelle à la température.

$$pV = NkT \quad (3.2.5)$$

qui montre que  $\frac{V}{N} = \frac{1}{p}kT$ , i.e. la pression est inversement proportionnel au volume moyen  $V/N$  par particule.

$$e^{-\frac{\mu}{kT}} = \frac{V}{N\lambda_T^3} \quad (3.2.6)$$

avec

$$\lambda_T = \left( \frac{2\pi\hbar^2}{mkT} \right)^{1/2}$$

appelée *longueur d'onde thermique de de Broglie*.

*Démonstration.* On calcule

$$\frac{1}{T} = \frac{\partial S}{\partial E} \stackrel{(3.2.1)}{=} kN \frac{3}{2} \frac{1}{E}$$

donnant (3.2.4). De même

$$\frac{p}{T} = \frac{\partial S}{\partial V} \stackrel{(3.2.1)}{=} kN \frac{1}{V}.$$

donnant (3.2.5). Finalement

$$\begin{aligned} -\frac{\mu}{T} &= \frac{\partial S}{\partial N} \stackrel{(3.2.1)}{=} k \ln \left( \frac{V e^{5/2}}{N} \left( \frac{mE}{3\pi\hbar^2 N} \right)^{3/2} \right) - kN \frac{5}{2} \frac{1}{N} \\ &= k \ln \left( \frac{V}{N} \left( \frac{mE}{3\pi\hbar^2 N} \right)^{3/2} \right) \stackrel{(3.2.4)}{=} k \ln \left( \frac{V}{N} \left( \frac{mkT}{2\pi\hbar^2} \right)^{3/2} \right) = k \ln \left( \frac{V}{N} \lambda_T^3 \right) \end{aligned}$$

donnant (3.2.6). □

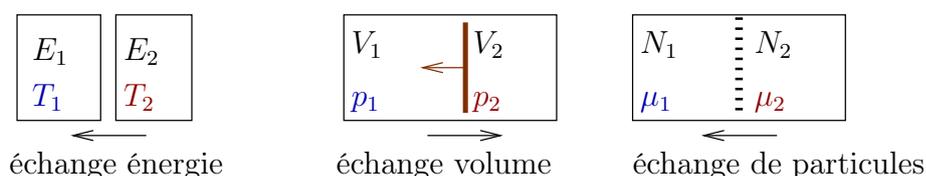
1. De même on a vu pour le gaz parfait que concernant le volume et la pression,

$$\begin{aligned} \Delta n(V) &\propto V^N, & S(V) &= k \ln(\Delta n) \propto \ln V \\ \frac{p}{T} &= \frac{\partial S}{\partial V} \propto \frac{1}{V} > 0, & \frac{\partial^2 S}{\partial V^2} &\propto -\frac{1}{V^2} < 0. \end{aligned}$$

### 3.3 Échanges entre deux systèmes thermodynamiques

Dans la nature, il y a de nombreux exemples de systèmes thermodynamiques qui interagissent entre eux. On souhaite décrire maintenant ces interactions et évolutions mutuelles. Il y a aussi des systèmes que l'on ne peut pas considérer comme thermodynamiques (i.e. qui ne sont pas dans un état d'équilibre thermodynamique).

*Remarque 3.3.1.* Pour simplifier on considère une unique famille de particules, mais cela s'étend à plusieurs famille qui coexistent.



#### 3.3.1 Relaxation vers l'état d'équilibre

Dans la suite on va supposer la situation suivante.

**Définition 3.3.2.** Deux systèmes thermodynamiques 1 et 2 sont en contact si chacun est à l'équilibre thermodynamique, caractérisé par leur fonction entropie  $S_j(E_j, V_j, N_j)$  respectives  $j = 1, 2$ , mais que cet équilibre peut varier du à un échange d'énergie, volume, ou particules entre ces deux systèmes :

- Si la paroi est non rigide, elle permet d'échanger de l'énergie (via des vibrations) mais on suppose

$$E = E_1 + E_2 \text{ est constant.}$$

- Si la paroi peut se déplacer, elle permet l'échange de volume mais

$$V = V_1 + V_2 \text{ est constant.}$$

- Si la paroi est poreuse (trous), elle permet l'échange de particules mais

$$N = N_1 + N_2 \text{ est constant.}$$

Donc les valeurs  $E_1, V_1, N_1$  peuvent varier au cours du temps. Dans cette section, on souhaite décrire leur évolution et tout d'abord l'état d'équilibre qui est la situation la plus probable des systèmes 1 et 2.

**Proposition 3.3.3.** *Si deux systèmes thermodynamiques 1, 2 sont en contact alors l'état d'équilibre, i.e. situation la plus probable, est donnée par les conditions*

$$T_1 = T_2, \quad p_1 = p_2, \quad \mu_1 = \mu_2.$$

- L'existence de cet état d'équilibre nécessite que  $S(E, V, N)$  est concave, i.e.  $\frac{\partial^2 S}{\partial E^2} < 0$ , etc.
- Les fluctuations autour de cet état d'équilibre sont très petites, de l'ordre de  $\frac{1}{\sqrt{N}} \ll 1$ .
- Toute situation initiale hors équilibre évolue vers cet état d'équilibre. Par exemple
  - Si au départ  $E_1 > E_1^{\text{equil.}}$  (et donc  $T_1 > T_2$ ) alors le système évolue vers l'état d'équilibre avec transfert d'énergie de 1 vers 2.
  - Si au départ  $V_1 > V_1^{\text{equil.}}$  (et donc  $p_1 > p_2$ ) alors le système évolue vers l'état d'équilibre avec transfert de volume de 1 vers 2.
  - Si au départ  $N_1 > N_1^{\text{equil.}}$  (et donc  $\mu_1 > \mu_2$ ) alors le système évolue vers l'état d'équilibre avec transfert de particules de 1 vers 2.

*Démonstration.* Pour simplifier on considère seulement un échange d'énergie avec

$$E = E_1 + E_2 \tag{3.3.1}$$

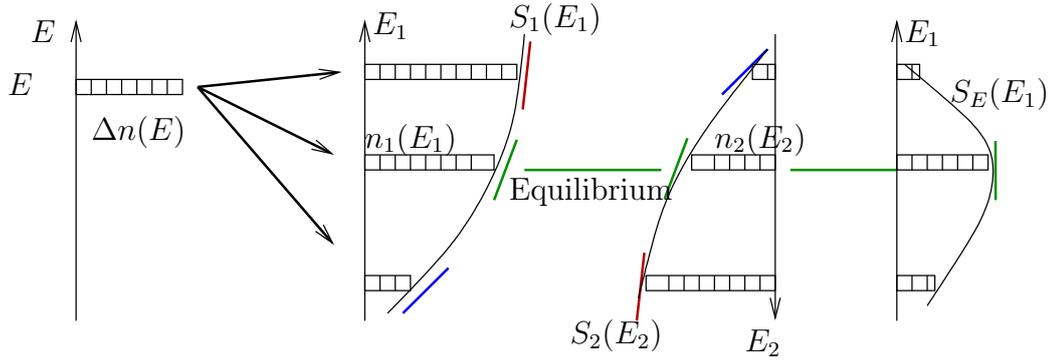
fixé. Et l'on cherche la mesure de probabilité induite  $p_{E_1}(E_1)$  sur la variable  $E_1$ . On décompose l'espace des micro-états du système total selon leur énergie  $E_1$  :  $\Delta n(E) = \sum_{E_1} n(E_1)$ , où  $n(E_1)$  est le nombre de micro-états du système total, à  $E_1$  fixé. D'après l'hypothèse ergodique que tous les micro-états sont équiprobables et d'après la loi de probabilité induite (1.3.2), on a la proportionnalité

$$p_{E_1}(E_1) \propto n(E_1)$$

Notons  $n_j(E_j)$  le nombre de micro-états du système  $j = 1, 2$  à l'énergie  $E_j$ . Pour des raisons combinatoires, du fait de la faible interaction entre les deux systèmes, on a approximativement

$$n(E_1) \approx n_1(E_1) n_2(E_2).$$

□



*Démonstration.* Ainsi, utilisant la définition de l'entropie du système  $j$  :

$$S_j(E_j) = k \ln n_j(E_j) \quad (3.3.2)$$

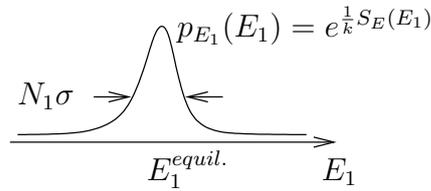
i.e.  $n_j(E_j) = e^{\frac{1}{k}S_j(E_j)}$ , on obtient

$$\begin{aligned} p_{E_1}(E_1) \propto n(E_1) &\approx n_1(E_1)n_2(E_2) \stackrel{(3.3.2)}{=} e^{\frac{1}{k}(S_1(E_1)+S_2(E_2))} \\ &\stackrel{(3.3.1)}{=} e^{\frac{1}{k}(S_1(E_1)+S_2(E-E_1))} = e^{\frac{1}{k}S_E(E_1)} \end{aligned}$$

avec

$$S_E(E_1) := S_1(E_1) + S_2(E - E_1).$$

□



*Démonstration.* On déduit que à la état d'équilibre  $E_1 = E_1^{\text{equil}}$ , la probabilité  $p_{E_1}(E_1^{\text{equil}})$  est maximale et vérifie donc

$$\begin{aligned} \frac{\partial p_{E_1}(E_1^{\text{equil}})}{\partial E_1} = 0 &\Leftrightarrow \frac{\partial S_E(E_1^{\text{equil}})}{\partial E_1} = 0 \Leftrightarrow \frac{\partial S_1}{\partial E_1}(E_1) - \frac{\partial S_2}{\partial E_2}(E - E_1) = 0 \\ &\Leftrightarrow T_1 = T_2. \end{aligned}$$

De plus pour  $E_1$  proche du maximum  $E_1^{\text{equil}}$ , par un développement de Taylor

$$\frac{1}{k}S_E(E_1) = \frac{1}{k}S_E(E_1^{\text{equil}}) + \frac{1}{2}(E_1 - E_1^{\text{equil}})^2 \frac{1}{k} \left( \frac{\partial^2 S_E}{\partial E_1^2} \right) + \dots$$

où  $\left(\frac{\partial^2 S_E}{\partial E_1^2}\right) < 0$  d'après la concavité de l'entropie (3.2.3). On se rappelle que  $E_j, S_j \propto N_j$  car sont extensives. En faisant le changement de variable  $E_1 \rightarrow x = \frac{E_1 - E_1^{\text{equil}}}{N_1}$  qui représente l'écart à l'équilibre, on a

$$p_{E_1}(E_1) \propto e^{-\frac{1}{2}\left(\frac{x}{\sigma}\right)^2}$$

avec un écart type  $\sigma > 0$  qui mesure les fluctuations, donné par

$$\begin{aligned} -\left(\frac{x}{\sigma}\right)^2 &= \left(E_1 - E_1^{\text{equil}}\right)^2 \frac{1}{k} \left(\frac{\partial^2 S_E}{\partial E_1^2}\right) \\ \Leftrightarrow \sigma &= \frac{1}{N_1} \left(-\frac{1}{k} \left(\frac{\partial^2 S_E}{\partial E_1^2}\right)\right)^{-1/2} \propto \frac{1}{N_1} \left(\frac{N}{N_1^2}\right)^{-1/2} \propto \frac{1}{N^{1/2}}. \end{aligned}$$

Avec  $N = N_1 + N_2$ . Ainsi pour le système thermodynamique où on a supposé  $N \gg 1$ , les fluctuations sont relativement très petites. Pour la dernière remarque, regarder la figure □

### 3.3.2 Mesure de probabilité de Gibbs-Boltzmann d'un micro-état

Dans cette section, on considère toujours deux systèmes 1 et 2 en contacts, mais on va supposer que le système deux est beaucoup plus grand, appelé "réservoir" :

**Définition 3.3.4.** On dira que le système 2 est un **réservoir** ou **thermostat** pour le système 1 si c'est un système thermodynamique et

$$E_2 \gg E_1 \text{ et } V_2 \gg V_1 \text{ et } N_2 \gg N_1.$$

*Remarque 3.3.5.* Dans la suite, **le système 1 n'est pas forcément un système thermodynamique**, ce peut être une seule particule (ou plusieurs), ou un spin unique (ou plusieurs), on suppose seulement que l'interaction avec le système 2 est négligeable au sens où elle ne perturbe pas la dynamique de 1. Par contre **on supposera que le système 2 est un système thermodynamique** comme précisé dans la définition 3.1.1, en plus que **c'est un réservoir** pour le système 1, selon la définition 3.3.4.

On a le résultat fondamental suivant, qui est à la base de la "physique statistique à l'équilibre".

**Proposition 3.3.6.** *Pour deux systèmes 1,2 en contact, supposant que le système 2 est un réservoir pour le système 1, alors*

$$T_2, p_2, \mu_2 \text{ sont constants,}$$

*La mesure de probabilité sur un micro-état  $m_1$  du système 1 d'énergie  $E_1(m_1)$ , volume  $V_1(m_1)$  et nombre de particules  $N_1(m_1)$ , appelée "mesure de probabilité de Boltzmann Gibbs" est*

$$p(m_1) = \frac{1}{Z} e^{-\frac{1}{kT_2}(E_1(m_1) + p_2 V_1(m_1) - \mu_2 N_1(m_1))} \quad (3.3.3)$$

*avec une constante de normalisation  $Z > 0$  telle que  $\sum_{m_1} p(m_1) = 1$ .*

*Remarque 3.3.7.* Si il y a seulement échange d'énergie, la mesure est  $p(m_1) = \frac{1}{Z} e^{-\frac{1}{kT_2} E_1(m_1)}$  et appelée aussi "ensemble canonique". Si il y a aussi échange de particules (paroi poreuse ou autre) alors la mesure de probabilité (3.3.3) est aussi appelée "ensemble grand canonique".

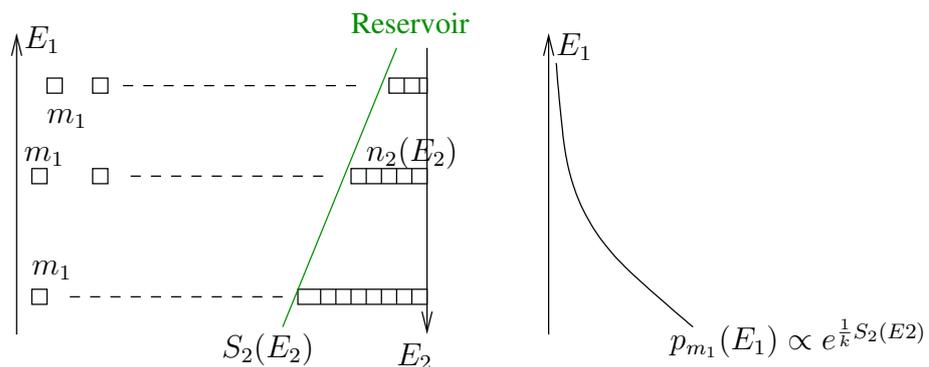
La constante de normalisation  $Z = Z(T_2, p_2, \mu_2)$  qui dépend des paramètres s'appelle la **fonction de partition**.

*Remarque 3.3.8.* Comme  $T_2, p_2 > 0$ , on observe que comme conséquence de (3.3.3), le **micro-état  $m_1$  le plus probable** est celui avec  $E_1(m_1), V_1(m_1)$  minimaux et  $N_1(m_1)$  minimal si  $\mu_2 < 0$ .

*Démonstration.* Considérons seulement un échange d'énergie pour simplifier. D'après le schéma,

$$p(m_1) \propto n_2(E_2) \stackrel{(3.3.2)}{=} e^{\frac{1}{k} S_2(E_2)} \stackrel{(3.3.1)}{=} e^{\frac{1}{k} S_2(E - E_1)}.$$

□



*Démonstration.* Avec  $E_1 = E_1(m_1)$ . Or par un développement de Taylor

$$S_2(E - E_1) = S_2(E) - E_1 \left( \frac{\partial S_2}{\partial E_2}(E - E_1) \right) + O \left( E_1^2 \frac{\partial^2 S_2}{\partial E_2^2} \right)$$

Le deuxième terme est  $O\left(N_1 \frac{N_2}{N_2}\right) = O(N_1)$  et le dernier terme est  $O\left(N_1^2 \frac{N_2}{N_2^2}\right) = O\left(\frac{N_1^2}{N_2}\right)$  donc relativement négligeable si et seulement si

$$\frac{N_1^2}{N_2} \ll N_1 \Leftrightarrow N_1 \ll N_2$$

ce qui est garanti par l'hypothèse que 2 est un réservoir pour 1. Autrement dit la fonction  $S_2(E_2)$  est une droite affine. Or  $\frac{1}{T_2} = \frac{\partial S_2}{\partial E_2}$  donc

$$S_2(E - E_1) \approx S_2(E) - E_1 \left( \frac{\partial S_2}{\partial E_2} \right) = S_2(E) - E_1 \frac{1}{T_2}$$

le premier terme est constant donc

$$p(m_1) \propto e^{\frac{1}{k} \left( -E_1 \frac{1}{T_2} \right)},$$

soit

$$p(m_1) = \frac{1}{Z} e^{-\frac{1}{kT_2} E_1(m_1)}.$$

Plus généralement, si il y a aussi un échange de volume et particules, d'après les définitions  $\frac{p_2}{T_2} = \frac{\partial S_2}{\partial V_2}$ ,  $-\frac{\mu_2}{T_2} = \frac{\partial S_2}{\partial N_2}$  on obtient

$$\begin{aligned} S_2(E - E_1) &\approx S_2(E) - E_1 \left( \frac{\partial S_2}{\partial E_2} \right) - V_1 \left( \frac{\partial S_2}{\partial V_2} \right) - N_1 \left( \frac{\partial S_2}{\partial N_2} \right) \\ &= S_2(E) - E_1 \frac{1}{T_2} - V_1 \frac{p_2}{T_2} - N_1 \left( -\frac{\mu_2}{T_2} \right) \end{aligned}$$

donnant (3.3.3). □

*Remarque 3.3.9.* Dans la preuve précédente, l'entropie  $S_1(E_1)$  ne joue plus aucun rôle puisque on ne considère que des micro-états individuels du système 1.

*Remarque 3.3.10.* (\*) Plus généralement, si  $Q = (Q_1, Q_2, \dots, Q_n)$  sont des quantités extensives conservées dans un système thermodynamique, on définit

$$\kappa_j := \frac{\partial S}{\partial Q_j}$$

qui donne  $\kappa = (\kappa_1, \kappa_2, \dots, \kappa_n)$ , quantités intensives et la mesure de probabilité d'un micro-état

$$p(m_1) = \frac{1}{Z} e^{-\frac{1}{k} \sum_j Q_j^{(1)}(m_1) \kappa_j^{(2)}} = \frac{1}{Z} e^{-\frac{1}{k} \langle Q^{(1)}(m_1) | \kappa^{(2)} \rangle}$$

avec  $Q_j^{(1)}(m_1)$  : quantité  $Q_j$  du système 1 dans le micro-état  $m_1$  et  $\kappa_j^{(2)} = \frac{\partial S^{(2)}}{\partial Q_j^{(2)}}$  quantité intensive du système 2 supposé réservoir, donc  $\kappa_j^{(2)}$  est constante.

*Remarque 3.3.11.* Pour deux états  $m', m$  on a  $\frac{p(m')}{p(m)} = e^{-\frac{E_1(m') - E_1(m)}{kT_2}}$  qui montre que le rapport des probabilités décroît avec l'écart en énergie  $\Delta E = E_1(m') - E_1(m)$  comparativement aux « fluctuations d'énergie thermique »  $kT_2$ . Par exemple on retient que les deux micro-états ont à peu près la même probabilité d'apparaître si

$$p(m) \simeq p(m') \quad \Leftrightarrow \quad |E_1(m') - E_1(m)| \ll kT_2.$$

### 3.4 Exemples de systèmes physique en contact avec un réservoir

Pour illustrer les résultats précédents, on présente quelques exemples de systèmes physique en contact avec un réservoir d'énergie, de volume ou de particules. Ces systèmes physiques ne sont pas forcément des systèmes thermodynamique, voir remarque 3.3.5.

#### 3.4.1 Moment magnétique dans un champ magnétique.

Comme exemple d'application de la loi de Boltzmann, on considère le moment magnétique d'un atome individuel dans un champ magnétique extérieur, en contact avec un environnement à la température  $T$  (thermostat), permettant des échanges d'énergie. On cherche la loi de probabilité sur le moment magnétique et sa moyenne  $\langle \vec{M} \rangle$  qui est l'aimantation.

##### 3.4.1.1 Modèle 1 : moment magnétique classique (Langevin)

Le moment magnétique est un vecteur  $\vec{M}$  de longueur fixée  $\|\vec{M}\| = \mu_B$  fixée. Son énergie dans un champ magnétique  $\vec{B}$  fixé est

$$E(\vec{M}) = -\vec{M} \cdot \vec{B} = -\mu_B B \cos \theta$$

où  $\theta = \left( \widehat{\vec{M}, \vec{B}} \right) \in [0, \pi]$  est l'angle entre  $\vec{M}$  et  $\vec{B}$  et  $\mu_B = \frac{e\hbar}{2m_e} = 0,927 \cdot 10^{-23} J/T$  (**magnéton de Bohr**). Le micro-état du système 1 est ici le vecteur  $\vec{M}$  caractérisé par sa direction  $(\theta, \varphi)$  en coordonnées sphériques (par rapport à la direction de  $\vec{B}$ ). La loi de Boltzmann dit que sa mesure de probabilité est

$$p(\vec{M}) d\theta d\varphi = \frac{1}{Z} e^{-\frac{E(\vec{M})}{kT}} \sin \theta d\theta d\varphi$$

Attention, la mesure de Boltzmann est la densité par rapport à la mesure uniforme  $\sin \theta d\theta d\varphi$  sur les directions. On déduit sa moyenne  $\langle \vec{M} \rangle$  qui est l'aimantation par (calcul détaillé en

TD)

$$\langle \vec{M} \rangle = \int \vec{M} p(\vec{M}) d\theta d\varphi = \dots = \mu_B \left( \coth x - \frac{1}{x} \right) \frac{\vec{B}}{\|\vec{B}\|}$$

avec  $x = \frac{\mu_B B}{kT}$ . On observe un état aimanté à basse température et non aimanté à haute température, appelé “paramagnétisme”.

En particulier à  $T$  fixé, l’aimantation s’annule si  $B \rightarrow 0$  et la susceptibilité magnétique est

$$\chi := \left( \frac{d\langle M \rangle}{dB} \right)_{B=0} = \dots = \frac{\mu^2}{2kT}$$

### 3.4.1.2 Modèle 2 : moment magnétique quantique (Langevin)

On modélise le moment magnétique de chaque atome selon un axe donnée, par deux états possible  $m = \pm \frac{1}{2}$  (spin 1/2). L’aimantation d’un atome selon la direction de  $\vec{B}$  est  $M_m = g\mu_B m$  avec  $g = 2$  (facteur de Landé). L’énergie de l’atome dans le champ magnétique extérieur  $\vec{B}$  est alors :

$$E(m) = -M_m B$$

Le micro-état du système 1 est ici caractérisé par la variable discrète  $m = \pm \frac{1}{2}$ . La loi de Boltzmann dit que sa mesure de probabilité est

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

On déduit sa moyenne  $\langle \vec{M} \rangle$  qui est l’aimantation par (calcul détaillé en TD)

$$\langle M \rangle = \sum_{m=\pm \frac{1}{2}} M_m p(m) = \dots = \mu_B \tanh x$$

avec  $x = \frac{\mu_B B}{kT}$ . En particulier à  $T$  fixé, l’aimantation s’annule si  $B \rightarrow 0$  et la susceptibilité magnétique est

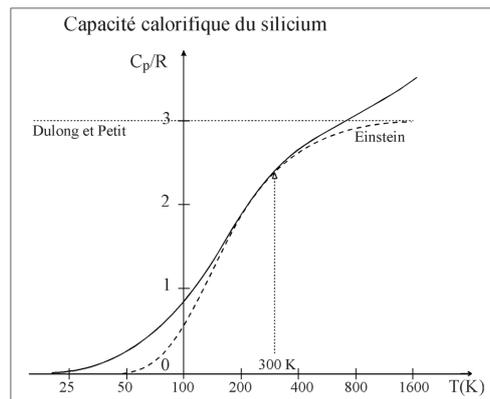
$$\chi := \left( \frac{d\langle M \rangle}{dB} \right)_{B=0} = \dots = \frac{\mu^2}{kT}$$

### 3.4.2 Capacité thermique d’un matériau solide

Dans cet exemple d’application de la loi de Boltzman-Gibbs, on cherche à modéliser la vibration des atomes dans un matériau solide et de déduire les échanges d’énergie entre le matériau et son environnement à la température  $T$  (thermostat), précisément calculer la capacité calorifique molaire

$$C_{\text{mole}}(T) := \frac{d\langle E_{\text{mole}} \rangle(T)}{dT}$$

qui mesure la variation d’énergie interne pour une mole d’atomes, lorsque la température externe varie.



On va considérer 3 modèles différents, chacun étant plus précis que le précédent. Ces modèles ont joué un grand rôle historique. Plus de détails en TD.

Dans chaque modèle on décrit les vibrations au premier ordre, c'est à dire à l'ordre linéaire avec des termes quadratiques dans le Hamiltonien (i.e. oscillateurs harmoniques).

### 3.4.2.1 Modèle 1 : particules individuelles classiques :

Dans le modèle 1, considère chaque atome **indépendant** et décrit par un **oscillateur harmonique en mécanique classique** : à 1 dimension cela s'écrit :

$$E(x, p) = \underbrace{\frac{1}{2m} p^2}_{E_c} + \underbrace{\frac{1}{2} K x^2}_{E_p}.$$

En appliquant la **loi de Boltzmann** qui affirme que la mesure de probabilité sur l'espace des phases est  $p(x, p) dx dp = \frac{1}{Z} e^{-\frac{E(x, p)}{kT}} dx dp$ , on calcule que les valeurs moyennes des énergie et obtient

$$\langle E_c \rangle = \int E_c(x, p) p(x, p) dx dp = \dots = \frac{1}{2} kT$$

$$\langle E_p \rangle = \int E_p(x, p) p(x, p) dx dp = \dots = \frac{1}{2} kT$$

où le calcul (...) utilise une formule d'intégrale Gaussienne. Il est remarquable que le résultat ne dépend pas ni de  $m$ , ni de  $K$ , donc **indépendant du matériau**. Cela s'appelle le « théorème d'équipartition de l'énergie » qui montre en général que **l'énergie moyenne est  $\frac{1}{2} kT$  par terme quadratique dans le Hamiltonien**. Ainsi pour un atome à 3 dimensions (somme de 3 oscillateurs), obtient que l'énergie moyenne est  $\langle E \rangle = 3 \times (\frac{1}{2} kT + \frac{1}{2} kT) = 3kT$  et pour une mole d'atomes  $N = \mathcal{N}_a$ , on déduit que l'énergie interne est proportionnelle à la température :

$$\langle E_{\text{mole}} \rangle = \mathcal{N}_a 3kT = 3RT,$$

la constante de proportionnalité, appelée capacité calorifique molaire est une constante indépendante du matériau :

$$C_{\text{mole}}(T) := \frac{d \langle E_{\text{mole}} \rangle}{dT} = 3R \approx 25 J/K.$$

### 3.4. EXEMPLES DE SYSTÈMES PHYSIQUE EN CONTACT AVEC UN RÉSERVOIR<sup>81</sup>

Ce résultat important appelé **loi de dulong-petit** (1819) était connu de façon expérimentale depuis longtemps. Cependant ce modèle ne correspond pas aux mesures expérimentales à basse température.

#### 3.4.2.2 Modèle 2 : particules individuelles quantiques (Einstein 1907) :

Le modèle 2, appelé modèle d'Einstein (1907) considère chaque atome **indépendant** et décrit par un **oscillateur harmonique en mécanique quantique**. Ainsi les états classique  $(x, p)$  du modèle précédent sont remplacés par les états  $\varphi_n$  d'énergie  $E_n = \hbar\omega (n + \frac{1}{2})$  de l'oscillateur harmonique avec  $\omega = \sqrt{\frac{K}{m}}$  et  $n \in \mathbb{N}$ . En appliquant la loi de Boltzmann qui affirme que la mesure de probabilité sur les états  $\varphi_n$  discrets est  $p(n) = \frac{1}{Z} e^{-\frac{E_n}{kT}}$ , on calcule que les valeurs moyennes des énergie et obtient (calculs détaillés en TD)

$$\langle E \rangle = \sum_{n \in \mathbb{N}} E_n p(n) = \dots = \frac{\hbar\omega}{2} \left( \frac{1 + e^{-\frac{\Theta_E}{T}}}{1 - e^{-\frac{\Theta_E}{T}}} \right)$$

avec  $\Theta_E := \frac{\hbar\omega}{k}$  appelée température d'Einstein. On déduit

$$C_{\text{mole}}(T) := \frac{d\langle E_{\text{mole}} \rangle}{dT} = 3R^2 \frac{\left(\frac{\Theta_E}{2T}\right)^2}{\sinh^2\left(\frac{\Theta_E}{2T}\right)}$$

En particulier  $C_{\text{mole}}(T) \approx 3R$  si  $T \gg \Theta_E$  et  $C_{\text{mole}}(T) \approx 0$  si  $T \ll \Theta_E$ . L'intuition de ce dernier résultat est si  $T \ll \Theta_E$ , les oscillateurs sont **coincés dans l'état fondamental quantique** et donc n'absorbent pas d'énergie, même si l'environnement est à température positive. Ainsi le diamant matériau dur, donc  $K$  grand,  $\omega$  grand,  $\Theta_E \approx 2000K$  grand est dans ce régime à température ambiante.

#### 3.4.2.3 Modèle 3 : ondes collectives quantiques (phonons) (Debye 1912)

Le modèle 3, appelé modèle de Debye (1912) considère que les **atomes oscillent mais sont aussi couplés entre eux** et encore décrit par un **Hamiltonien quadratique en mécanique classique** puis en **mécanique quantique**. C'est donc un modèle de **champs libres quantiques**.

- Dans un premier temps on évalue les modes classiques de vibration, chacun étant formellement assimilé à un **oscillateur harmonique classique**.
- Dans un deuxième temps on modélise chacun de ces oscillateurs par un modèle quantique, un **oscillateur harmonique quantique**, ce qui revient à "quantifier les champs".
- Dans un troisième temps, on applique la loi de Boltzmann à tous les micro-états obtenus pour déduire la capacité calorifique  $C_{\text{mole}}(T)$ .

La différence essentielle avec le modèle d'Einstein est qu'il n'y a pas de gap spectral près de l'état fondamental à cause des ondes de grande longueur d'onde et donc la capacité

calorifique ne s'annule pas très vite à basse température comme dans le modèle d'Einstein. On trouve plutôt : si  $T \ll \Theta_E$  alors (à 1D)

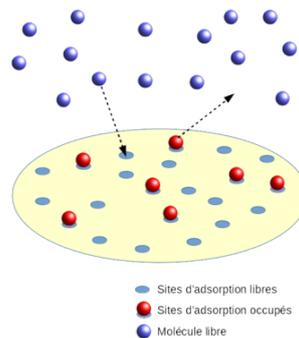
$$C_{\text{mole}}(T) \approx \frac{\pi}{3} R \left( \frac{T}{\Theta_E} \right),$$

qui correspond aux expériences à basses températures, mieux que le modèle d'Einstein.

### 3.4.3 Adsorption de particules

Références : [Diu \*et al.\* \(1989, p.677\)](#), [Balian \(2007, Tome1, p174\)](#), [adsorption sur wikipedia](#).

L'adsorption est le phénomène qu'une surface peut capturer (et/ou libérer) de façon aléatoire des particules d'un gaz ou d'un liquide. Par exemple l'adsorption avec du **charbon actif** est un traitement efficace pour enlever la matière organique, les phénols, les hydrocarbures saturés qui sont des molécules insolubles difficilement attaquables par l'ozone, les pesticides, les métaux lourds, les agents tensio-actifs...



Dans le **modèle de Langmuir** 1918, on suppose qu'un gaz est en contact avec une surface. Le gaz est modélisé par le modèle du gaz parfait de  $N$  particules à la température  $T$ , pression  $p$ , potentiel chimique  $\mu$  et considéré comme un réservoir d'énergie et de particules. La surface est modélisée par un ensemble de sites indépendants, chaque site peut contenir  $m = 0$  particule ou  $m = 1$  particule dans un état d'énergie  $(-\epsilon) < 0$  fixée.

L'objectif est d'exprimer  $P_m(p, T)$  qui est la probabilité pour qu'un site soit dans l'état  $m = 0, 1$ , en fonction des paramètres  $p, T$  du gaz que l'on peut contrôler.

— Dans un premier temps, d'après la **loi de Boltzmann de l'ensemble grand canonique**, la probabilité que le site est dans l'état  $m = 0$  ou  $m = 1$  est

$$P_m = \frac{1}{Z} \exp \left( -\frac{1}{kT} (E_m - \mu N_m) \right)$$

où  $T$  est la température externe,  $\mu$  le potentiel chimique externe,  $E_m$  l'énergie du site,  $N_m$  le nombre de particules sur le site. Or dans ce modèle,

$$\text{si } m = 0, \text{ alors } E_0 = 0, N_0 = 0,$$

### 3.4. EXEMPLES DE SYSTÈMES PHYSIQUE EN CONTACT AVEC UN RÉSERVOIR 83

si  $m = 1$ , alors  $E_1 = -\epsilon$ ,  $N_1 = 1$ ,

ainsi

$$P_{m=0} = \frac{1}{Z}, \quad P_{m=1} = \frac{1}{Z} \exp\left(-\frac{1}{kT}(-\epsilon - \mu)\right) = \frac{1}{Z} \exp\left(\frac{1}{kT}(\epsilon + \mu)\right)$$

D'après la normalisation  $1 = P_0 + P_1$ , on déduit que  $Z = 1 + \exp\left(\frac{1}{kT}(\epsilon + \mu)\right)$  donc

$$P_{m=1} = \frac{e^{\frac{\epsilon+\mu}{kT}}}{1 + e^{\frac{\epsilon+\mu}{kT}}} = \frac{1}{1 + e^{-\frac{\epsilon+\mu}{kT}}} \quad (3.4.1)$$

— Dans un deuxième temps on utilise le modèle du gaz parfait pour exprimer  $\mu$  à partir de  $p, T$  :

$$e^{-\frac{\mu}{kT}} \stackrel{(3.2.6)}{=} \frac{V}{N \left(\frac{2\pi\hbar^2}{mkT}\right)^{3/2}} \stackrel{(3.2.5)}{=} \frac{kT}{p \left(\frac{2\pi\hbar^2}{mkT}\right)^{3/2}} = C \frac{T^{5/2}}{p}$$

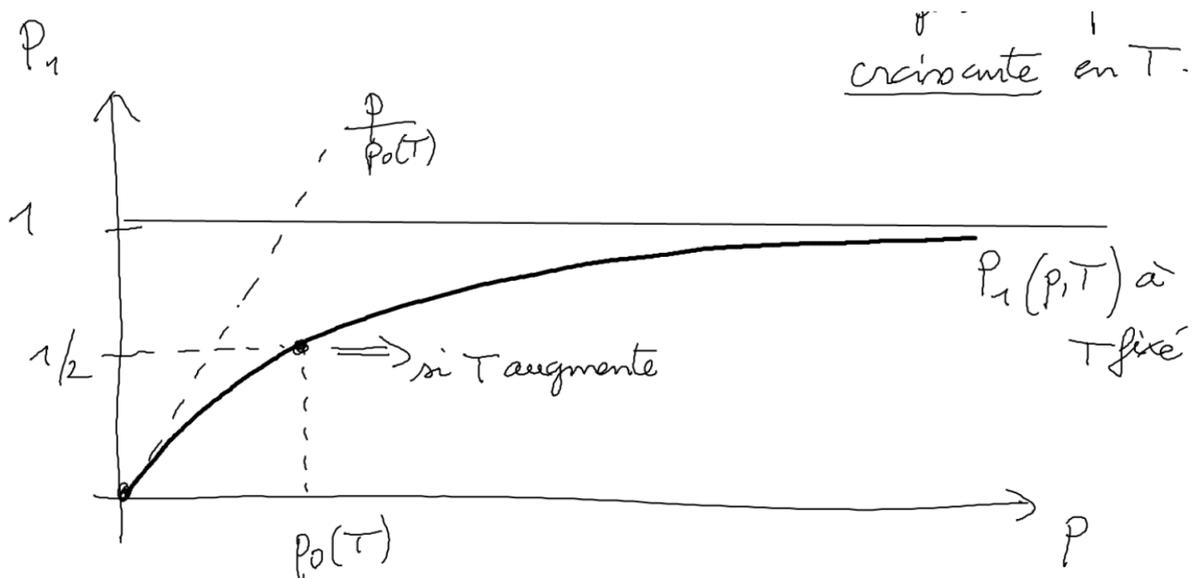
avec  $C_N$  qui ne dépend pas de  $p, T$ . On déduit que

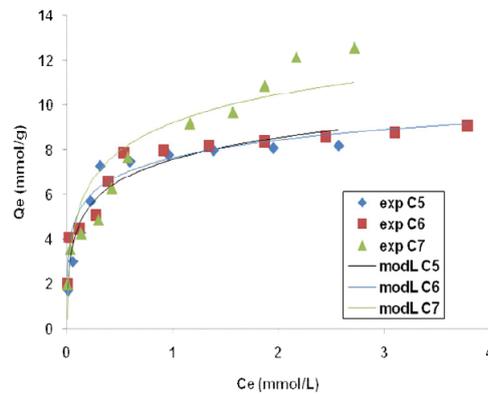
$$P_{m=1}(p, T) \stackrel{(3.4.1)}{=} \frac{1}{1 + e^{-\frac{\epsilon}{kT}} C \frac{T^{5/2}}{p}} = \frac{1}{1 + \frac{p_0(T)}{p}}$$

avec

$$p_0(T) := C e^{-\frac{\epsilon}{kT}} T^{5/2}$$

qui ne dépend pas de  $p$  et qui est une fonction croissante de  $T$ . A  $T$  fixé, les courbes  $p \rightarrow P_{m=1}(p, T)$  sont appelées **isothermes de Langmuir**.





La deuxième figure avec des données expérimentales est de [Morvan Clotaire Nko'o Abuiboto et al](#), Cameroun.

### 3.5 Le principe variationnel

On a vu ci-dessus que la distribution de Boltzmann-Gibbs apparaît de façon naturelle et universelle lorsqu'un système physique est en contact avec un réservoir thermodynamique avec lequel il échange de l'énergie et éventuellement du volume et des particules. Ce réservoir est caractérisé par sa température  $T$ , et éventuellement sa pression  $P$  et son potentiel chimique  $\mu$ .

Dans cette section on va donner une caractérisation mathématique de cette **distribution de Boltzmann-Gibbs comme minimisant une fonction**  $F_T(p)$  appelée **énergie libre** sur l'espace des mesures de probabilité  $p$ . Si de plus, il y a échange de volume et/ou particules, l'énergie libre  $F_T$  est remplacé de façon analogue par l'enthalpie libre  $G_{T,P}$  ou le grand potentiel  $J_{T,\mu}$  ou  $\Omega_{T,P,\mu}$ .

On verra dans les chapitres suivants que cette caractérisation est utile pour les calculs. En particulier elle permettra de mettre en œuvre une méthode d'approximation appelée **méthode variationnelle** qui est très utile pour les problèmes compliqués, comme l'étude des transitions de phase.

**Proposition 3.5.1. “Le principe variationnel”.** *Considérons un système physique dont les micro-états sont indicés par  $i$ , ayant des énergies  $(E_i)_i$ . Supposons que ce système est en contact (échange d'énergie) avec un réservoir de température  $T$ . Pour une distribution de probabilité  $p = (p_i)_i$  quelconque sur les états  $i$  (i.e. vérifiant  $\sum_i p_i = 1$  et  $p_i \in [0, 1]$ ), on définit l'énergie libre*

$$F_T(p) := \langle E \rangle_p - TS(p) \quad : \text{énergie libre} \quad (3.5.1)$$

avec l'entropie de Boltzmann (d'après (3.1.1, 1.5.8)) :  $S(p) = kS_{\text{Schannon}} = -k \sum_{i=1}^n p_i \ln p_i$  et l'énergie moyenne  $\langle E \rangle_p := \sum_i p_i E_i$ . La fonction  $p \rightarrow F_T(p)$  est **minimale** pour la mesure de **Boltzmann-Gibbs** (3.3.3)

$$p_i = \frac{1}{Z} e^{-\frac{1}{kT} E_i},$$

et sa valeur au minimum est

$$F_T(p) = -kT \ln Z.$$

**Proposition 3.5.2. “Le principe variationnel plus général”.** *Plus généralement, notant le volume  $V_i$ , le nombre de particules  $N_i$  du micro-état  $i$ , supposons que le système est en contact avec un réservoir de volume et/ou de particules de pression  $P$ , de potentiel chimique  $\mu$ . Pour une mesure de probabilité  $p = (p_i)_i$  sur les micro-états, on définit les fonctionnelles appelées **potentiel thermodynamique***

$$G_{T,P}(p) = \langle E \rangle_p + P \langle V \rangle_p - TS(p) \quad : \text{enthalpie libre}$$

$$J_{T,\mu}(p) = \langle E \rangle_p - \mu \langle N \rangle_p - TS(p) \quad : \text{grand potentiel}$$

$$\Omega_{T,P,\mu}(p) = \langle E \rangle_p + P \langle V \rangle_p - \mu \langle N \rangle_p - TS(p) \quad : \text{grand potentiel}$$

avec le volume moyen  $\langle V \rangle_p := \sum_i p_i V_i$ , le nombre moyen de particules  $\langle N \rangle_p = \sum_i p_i N_i$ . Alors ces potentiels thermodynamiques sont minimaux respectivement pour les mesures d'équilibres de **Boltzmann-Gibbs** (3.3.3) :  $p_i = \frac{1}{Z} e^{-\frac{E_i + PV_i}{kT}}$ ,  $p_i = \frac{1}{Z} e^{-\frac{E_i - \mu N_i}{kT}}$ ,  $p_i = \frac{1}{Z} e^{-\frac{1}{kT}(E_i + PV_i - \mu N_i)}$ , avec une valeur au minimum égale à  $-kT \ln Z$ .

*Remarque 3.5.3.* La constante de normalisation  $Z(T, P, \mu)$  obtenue dépend des paramètres  $T, P, \mu$  et s'appelle aussi **fonction de partition**.

*Remarque 3.5.4.* Par exemple en **Thermo-chimie**, un système chimique peut échanger de l'énergie et du volume (il évolue à température  $T$  et pression  $p$  fixés) et donc l'équilibre chimique correspond au minimum de l'enthalpie libre  $G$ .

*Démonstration.* (\*) (en TD) On va tout d'abord démontrer le principe variationnel dans un cadre plus simple avec des variables sans dimensions et déduire ensuite les propositions précédentes. On considère  $n$  états d'énergie  $\epsilon_1 \leq \epsilon_2 \dots \leq \epsilon_n$  et une mesure de probabilité quelconque sur ces états notée  $p = (p_1, p_2 \dots p_n)$  avec  $p_i \geq 0$  et  $\sum_{i=1}^n p_i = 1$ . L'énergie moyenne pour cette mesure est notée  $\langle E \rangle_p := \sum_i p_i \epsilon_i$  et l'entropie de la mesure est  $S(p) = -\sum_{i=1}^n p_i \ln p_i$ . On introduit la fonctionnelle suivante sur les mesures de probabilités

$$F(p) := \langle E \rangle_p - S(p) \quad (3.5.2)$$

appelée **énergie libre**.

**Lemme 3.5.5. “Le principe variationnel”.** *La fonctionnelle  $p \rightarrow F(p)$ , en (3.5.2), est minimale pour la mesure de Boltzmann*

$$p_i = \frac{1}{Z} e^{-\epsilon_i}, \forall i$$

avec  $Z = \sum_i e^{-\epsilon_i}$  constante de normalisation, et la valeur minimale est

$$F = -\ln Z.$$

*Démonstration.* Considérons des nombres  $p_i \in \mathbb{R}$  sans contrainte et posons  $\sigma = \sum_i p_i$ . On calcule les différentielles

$$d\sigma = \sum_i dp_i$$

et

$$dF = d\langle E \rangle - dS = \sum_i \left( \epsilon_i + \ln p_i + \frac{p_i}{p_i} \right) dp_i \quad (3.5.3)$$

Pour résoudre  $\frac{\partial F}{\partial p_i} = 0, \forall i$ , i.e. trouver un point critique de  $F$ , avec la contrainte  $\sigma = 1$ , on doit avoir  $\lambda \in \mathbb{R}$ , appelé **multiplicateur de Lagrange** tel que (voir schéma)

$$dF = \lambda d\sigma.$$

Cela donne

$$\begin{aligned} \exists \lambda \in \mathbb{R}, \forall i, \quad \epsilon_i + \ln p_i + 1 &= \lambda, \\ \Leftrightarrow \exists \lambda \in \mathbb{R}, \forall i, \quad p_i &= \exp(\lambda - 1 - \epsilon_i) = \frac{1}{Z} e^{-\epsilon_i} \end{aligned} \quad (3.5.4)$$

avec  $Z = e^{1-\lambda}$  donné par la contrainte  $\sum_i p_i = 1$  donc  $Z = \sum_i e^{-\epsilon_i}$ . On doit ensuite vérifier que cette solution  $p = (p_i)_i$  est bien un minimum de  $F$ . D'après (3.5.3), on a

$$\frac{\partial F}{\partial p_i} = \epsilon_i + \ln p_i + 1, \quad \frac{\partial^2 F}{\partial p_j \partial p_i} = \frac{1}{p_i} \delta_{i=j},$$

On observe que la matrice  $\left(\frac{\partial^2 F}{\partial p_i \partial p_i}\right)_{i,j}$  est diagonale, définie positive, donc  $F$  est convexe. C'est bien un minimum. Pour cette solution (3.5.4), on calcule la valeur de  $F$  :

$$F(p) \stackrel{(3.5.2)}{=} \sum_i \epsilon_i p_i + p_i \ln p_i = \frac{1}{Z} \sum_i \epsilon_i e^{-\epsilon_i} + e^{-\epsilon_i} (-\ln Z + \epsilon_i) = -\frac{1}{Z} \sum_i e^{-\epsilon_i} \ln Z = -\ln Z.$$

□

On considère maintenant l'expression (3.5.1) de l'énergie libre

$$\begin{aligned} F_T(p) &= \langle E \rangle_p - TS(p) = \sum_i E_i p_i + T k p_i \ln p_i \\ &= kT \left( \sum_i \left( \frac{E_i}{kT} \right) p_i + p_i \ln p_i \right) \end{aligned}$$

On se ramène donc au problème précédent, en posant  $\epsilon_i = \frac{E_i}{kT}$  sans dimension. On a obtenu la mesure d'équilibre

$$p_i = \frac{1}{Z} e^{-\epsilon_i} = \frac{1}{Z} e^{-\frac{E_i}{kT}}.$$

De même

$$\begin{aligned} G_{T,P}(p) &= \langle E \rangle_p + P \langle V \rangle_p - TS(p) \\ &= \sum_i E_i p_i + P p_i V_i + T k p_i \ln p_i \\ &= kT \left( \sum_i \left( \frac{E_i + P V_i}{kT} \right) p_i + p_i \ln p_i \right) \end{aligned}$$

et posant  $\epsilon_i = \frac{E_i + P V_i}{kT}$  on a la mesure d'équilibre

$$p_i = \frac{1}{Z} e^{-\epsilon_i} = \frac{1}{Z} e^{-\frac{E_i + P V_i}{kT}},$$

etc.

□



# Chapitre 4

## Statistiques quantiques

Dans ce chapitre on décrit les “statistiques de Fermi-Dirac et de Bose-Einstein” qui permettent de décrire les particules élémentaires observées dans la nature et par conséquent les propriétés des éléments composés de particules élémentaires comme les noyaux atomiques, les atomes, les molécules, les matériaux etc.

### 4.1 Introduction

Référence : [Particule data book du CERN](#) ou [Review of particle physics](#) téléchargeable. Voici un aperçu des particules élémentaires selon le **modèle standard** (1970) (en conformité avec les expériences) et un aperçu des particules composées de particules élémentaires.

#### 4.1.1 Champ et particules

On utilisera le terme de « **champs** » plutôt que particule, car on verra que la notion de **particule correspond à des états d’excitations quantique du champ**. Par exemple dans l’univers, il y a un unique “champ électron” et toutes les “particules électrons” de l’univers sont des excitations quantiques de ce champ. De même il y a un unique “champ photon” et les particules photons sont des excitations quantiques de ce champ.

#### 4.1.2 Fermions/Bosons

On donnera plus loin une définition plus précise des fermions et des bosons, qui concerne une propriété du **comportement collectif** des particules et donc une propriété du champ. Ainsi un champ donné est de type fermion ou boson. Par exemple les électrons sont des Fermions et les photons sont des Bosons.

En terme simple :

— Un champ est de type **Fermion** si il n’y a pas plus qu’une particule (excitation) dans un état physique donné.

Par exemple : les électrons sont des Fermions : il peut y avoir 0 ou 1 électron

dans l'état  $1s, +$  qui est l'état orbital de basse énergie, spin  $+\frac{1}{2}$  autour du noyau hydrogène.

— Pour un champ de type **Boson**, il n'y a pas cette contrainte.

Par exemple les photons sont des Bosons : il peut y avoir un nombre arbitraire de photons dans un mode électromagnétique donné, comme une onde plane  $\varphi_{\vec{k},+1}(\vec{x}) = e^{i\vec{k}\cdot\vec{x}}$  de vecteur d'onde  $\vec{k}$  et hélicité  $+1$  (i.e. polarisation circulaire  $+1$ ) dans un faisceau laser.

### 4.1.3 Liste des champs existants d'après le **modèle standard** (1970)

On va distinguer les champs de matière et les champs de Jauge. De façon général **les champs de matière sont des Fermions** et **les champs de Jauge sont des Bosons**.

#### 4.1.3.1 Champs de matière (Fermions)

Voici la liste des champs de matière élémentaires :

- Les **leptons** (électron  $e$ , muon  $\mu$ , tau  $\tau$  et neutrinos associés  $\nu_e, \nu_\mu, \nu_\tau$ ).
- les **baryons** (les quarks  $u, d, c, s, t, b$ ).

Pour chaque champ (par exemple le champ d'électrons  $e$ ), il existe le champ d'**antimatière** noté  $\bar{e}$  (aussi appelé **positron**, noté  $e^+$ ).

Dans la matière ordinaire autour de nous, il y a seulement les champs de matière  $e, u, d$ .

A cause du **confinement de couleur**, les champs de quarks ne se manifestent pas sous forme de particule élémentaire mais sous forme composée appelé **hadron**, par exemple par groupe de deux, appelés **méson** ex :  $\pi^+ = \{u, \bar{d}\}$ , ou par groupe de trois, ex : **proton**  $p^+ = \{u, u, d\}$  ou **neutron**  $n = \{u, d, d\}$ .

*Remarque 4.1.1.* (\*) Ci-dessus, on distingue trois générations de champs de matière :

1. Première génération : les leptons  $e, \nu_e$  et baryons  $u, d$ .
2. Deuxième génération : les leptons  $\mu, \nu_\mu$  et baryons  $c, s$
3. Troisième génération : les leptons  $\tau, \nu_\tau$  et baryons  $t, b$

Les champs de 2eme et 3eme génération ont des masse plus élevées et donc se désintègrent en champs de première génération, cf discussion plus bas. Ainsi la matière ordinaire est formée de champs de 1ere génération.

#### 4.1.3.2 Champs de jauge (Bosons)

Les champs de matières sont couplés à des « **champs de Jauge** », que l'on appelle aussi interactions fondamentales :

- Les **photons**  $\gamma$ , (champ ou force électromagnétique)
- Les **bosons** faibles  $W^\pm, Z_0$  (champ ou force nucléaire faible)
- Les **gluons**  $g$  (champ ou force nucléaire forte)

Les leptons sont couplés au champs de photons  $\gamma$  et de bosons faibles  $W^\pm, Z_0$  mais les **leptons ne sont pas couplés aux champs de gluons**  $g$ . Les baryons sont couplés aux trois champs de Jauge.

**Exemple 4.1.2.** Un **proton** est constitué de trois quarks  $p = (u, u, d)$  couplés à des gluons et plus faiblement à des photons et bosons faibles. Un **atome d'hydrogène** est constitué d'un proton et d'un électron couplés par l'intermédiaire du champ de photon.

*Remarque 4.1.3.* (\*) Le **champ de gravitation** est aussi une théorie de Jauge qui concerne la métrique de l'espace-temps. **Tous les champs de matière et de Jauge sont couplés au champ métrique.** Ainsi l'énergie et la masse déforme la géométrie de l'espace-temps via sa métrique et inversement la trajectoire d'un photon (lumière), d'une particule massive ou d'une planète est déviée (suit une géodésique) dans l'espace-temps. Cela explique les trajectoires elliptiques des planètes autour du soleil (dans le cadre de la théorie d'Einstein). Mais il n'y a pas de formulation quantique de la gravitation connue ni même d'expérience suggérant cela. L'excitation quantique du champ de gravitation est donc très hypothétique, mais ce serait un boson de spin  $s = \pm 2$ , appelé **graviton**.

*Remarque 4.1.4.* (\*) La théorie de Jauge aussi appelée **théorie géométrique de Yang-Mills** est une formulation géométrique où les champs de matière sont des sections de fibrés vectoriels complexes et les champs de Jauge sont des connections de ces fibrés, respectivement de rang 1, 2, 3. Cependant le cadre mathématique de la version quantique n'est pas établie.

*Remarque 4.1.5.* (\*) Le **champ de Higgs** est un champ scalaire (i.e. à valeur complexe) de type Boson. Sa présence permet d'expliquer la masse des champs de matière et des champ de Jauge des bosons faibles, par un mécanisme appelé "brisure de symétrie". Quelle est sa signification géométrique en termes d'espaces fibrés ?

#### 4.1.4 Grandeurs caractéristiques de chaque champ

Chaque champ élémentaires décrit ci-dessus, et donc chaque particule élémentaire associée a des caractéristiques que nous décrivons maintenant, outre la propriété Fermion/-Boson déjà définit ci-dessus.

##### 4.1.4.1 Masse et spin

Toutes les particules élémentaires possèdent une **masse**  $m \geq 0$  et un moment angulaire appelé **spin** ou **moment angulaire intrinsèque** caractérisé par son intensité  $s \in \{0, \frac{1}{2}, 1, \frac{3}{2}, 2, \frac{5}{2}, \dots\}$  si la masse est  $m > 0$  (et si la masse est  $m = 0$ , le spin appelé hélicité est caractérisé par  $s \in \{0, \pm\frac{1}{2}, \pm 1, \pm\frac{3}{2}, \dots\}$ ). Il apparaît que **les Fermions ont un spin demi-entier**  $s \in \{\frac{1}{2}, \frac{3}{2}, \frac{5}{2}, \dots\}$  et **les Bosons ont un spin entier**  $s \in \{0, 1, 2, \dots\}$  (idem dans le cas  $m = 0$ ). Cette correspondance s'appelle « le **théorème spin-statistique** ».

*Remarque 4.1.6.* (\*) Les paramètres de masse  $m \geq 0$  et de spin  $s$  sont associés à chaque champ élémentaire et sont des invariants (représentations irréductibles) du groupe de symétrie de Poincaré qui est le groupe de symétrie de l'espace temps, voir [Sternberg \(1994, p.148\)](#).

#### 4.1.4.2 Charges

Les charges sont des quantités associées à chaque particules élémentaire

- La charge électrique qui peut être  $q = 0, \pm\frac{1}{3}, \pm\frac{2}{3}, \pm 1$  (en unité<sup>1</sup> de  $q = 1, 602176634 \times 10^{-19}C$ , appelée **charge élémentaire**). Cette charge est associée aux interactions avec le champ de Jauge  $\gamma$ .
- Le **nombre leptonique**  $L_l = N_l - N_{\bar{l}}$  où  $l$  est une certaine famille de **leptons** (exemple électron)  $N_l$  est le nombre de leptons et  $N_{\bar{l}}$  est le nombre d'anti-leptons. Voir exemples ci-dessous.
- Le **nombre baryonique**  $B = \sum_q \frac{1}{3} (N_q - N_{\bar{q}})$  concernant la conservation du nombre de quarks.

*Remarque 4.1.7.* pour le photon et autres bosons de Jauge il n'y a pas de règle de conservation sur leur nombre.

#### 4.1.5 Lois de conservation

Les champs évoluent d'après l'équation de Schrödinger

$$i \frac{\partial \psi}{\partial t} = \hat{H} \psi$$

où  $\psi$  représente tous les champs quantiques et  $\hat{H}$  le Hamiltonien quantique du modèle standard. Dans ce Hamiltonien il y a des couplages entre les champs de matière et les champs de Jauge et par conséquent il y a des transitions possibles, les particules élémentaires peuvent apparaître/disparaître au cours du temps. Il y a seulement quelques contraintes globales appelées **lois de conservations** et souvent associées à des symétries du modèle :

- Il y a **conservation de l'impulsion totale**  $\vec{p} \in \mathbb{R}^3$  et **énergie totale**  $E$  d'un système de particules (cela découle de l'invariance par translation dans l'espace-temps de  $\hat{H}$ ). Ne pas oublier de plus que pour chaque particule on a la relation

$$E^2 - \|\vec{p}\|^2 c^2 = m^2 c^4.$$

- Il y a **conservation du moment angulaire total** (qui découle de l'invariance par rotation dans l'espace temps de  $\hat{H}$ ).
- Conservation de la charge électrique totale  $Q$ , des nombres leptoniques  $L_l$ , du nombre baryonique  $B$ .

---

1. Nouvelle valeur à partir du 20 mai 2019

### 4.1.6 Désintégrations. Particules stables et instables

Une **réaction de désintégration** est de la forme

$$A \rightarrow B$$

où la masse totale de  $A$  (énergie au repos) est supérieure à la masse totale de  $B$  :  $M_A \geq M_B$  se sorte que pour satisfaire la conservation de l'énergie totale, les composés  $B$  auront éventuellement de l'énergie cinétique qui est toujours positive. Une telle réaction est possible si elle ne contredit pas une des règles de conservation ci-dessus et donc se produira avec une probabilité plus ou moins grande donc un certain temps moyen  $\tau$ .

Ainsi la plupart des particules élémentaires et composés sont instables et se désintègrent en particules ou composés stables (qui eux n'ont pas la possibilité de se désintégrer). Une particule instable  $A$  est souvent appelé **résonance**.

#### Exemples :

- Par exemple le **muon** se désintègre ainsi

$$\mu^- \rightarrow e^- + \bar{\nu}_e + \nu_\mu$$

en  $\tau \sim 2 \times 10^{-6} s$  **ref**. La charge  $Q = -1$ , les nombres leptoniques  $L_\mu = 1$  et  $L_e = 0$  sont conservés. Les masses<sup>2</sup> sont  $m_{\mu^-} = 0.11u$ ,  $m_{e^-} = 0.5 \times 10^{-3}u$ ,  $m_{\bar{\nu}_e}, m_{\nu_\mu} < 2.4 \times 10^{-7}u$ , ainsi  $m_{\mu^-} > m_{e^-} + m_{\bar{\nu}_e} + m_{\nu_\mu}$ , la réaction n'est pas interdite par les règles de conservation.

- Le neutron se désintègre ainsi

$$n \rightarrow p^+ + e^- + \bar{\nu}_e$$

en  $\tau \sim 879s$  **ref**. La charge  $Q = 0$ , les nombres leptoniques  $L_e = 0$ , nombre baryonique  $B = 1$  sont conservés. Les masses sont  $m_n = 1.0087u$ ,  $m_{p^+} = 1.0073u$ ,  $m_{e^-} = 0.0005u$ ,  $m_{\bar{\nu}_e} < 2.4 \times 10^{-7}u$  ainsi  $m_n > m_{p^+} + m_{e^-} + m_{\bar{\nu}_e}$ , la réaction n'est pas interdite par les règles de conservation.

- Voici la liste des masses et de toutes les réactions de désintégrations connues **Particule data book du CERN**.
- Les particules plus lourdes ont plus de possibilités de désintégrations. Par exemple le lepton  $\tau$  (tau)

$$\tau^- \rightarrow \mu^- + \bar{\nu}_\mu + \nu_\tau$$

$$\tau^- \rightarrow e^- + \bar{\nu}_e + \nu_\tau$$

- Au final, **les seules particules élémentaires stables** sont électrons  $e^-$ , **neutrinos**  $\nu$  (mais les **neutrinos oscillent**), photons  $\gamma$ .

---

2. en **unité de masse atomique unifiée**  $u = N_A^{-1} \times 10^{-3} \text{kg}$

### 4.1.7 Compositions de particules

Toute la richesse et la difficulté en physique est de comprendre les interactions entre les particules élémentaires, (i.e. excitations quantiques des champs).

- Certaines particules élémentaires s’assemblent pour former des particules composées très variées, par exemple les **protons** et **neutrons** sont des assemblées de quarks et gluons. Le proton est stable mais le neutron seul est instable.
- Les **noyaux atomiques** sont des assemblages de protons et neutrons qui sont stables si il y a un nombre adéquat de protons et neutrons situés dans la **vallée de stabilité**.
- Les atomes sont des assemblages de noyaux et électrons, les molécules sont des assemblages d’atomes, etc.. (les cristaux, cellules vivantes, fluides sont des assemblages de molécules).

Dans la nature, il y a différents régimes : sur Terre la température est relativement faible et seuls les particules et composés stables se manifestent. Les noyaux atomiques sont stables (sauf certains noyaux instables appelés radioactifs). Dans les étoiles, la température est plus élevée, des réactions nucléaires sont possibles. En laboratoires, pour obtenir des énergies cinétiques élevés, les expériences se font dans les **accélérateurs de particules** et par collisions.

*Remarque 4.1.8.* Tous les éléments atomiques présents sur Terre ont été produits par des réactions nucléaires dans des étoiles d’anciennes générations (notre Soleil serait déjà de la **3ème génération**, fille de l’étoile **Coatlicue**). Avant de le savoir, avant le XVIIIème siècle, certaines personnes cherchaient à effectuer ces transmutations par “**Alchimie**”.

#### 4.1.7.1 Propriétés des compositions de particules

Pour les compositions, **les propriétés s’ajoutent** : énergie, moment angulaire, charges etc.

**Exemple 4.1.9.** Le proton est constitué de 3 quarks  $u, u, d$  de spin respectifs  $s = \frac{1}{2}, \frac{1}{2}, -\frac{1}{2}$  et charges électriques  $s = \frac{2}{3}, \frac{2}{3}, -\frac{1}{3}$ , donc son spin est  $s = +\frac{1}{2}$  et sa charge électrique est  $q = +1$ . De même le neutron  $n = (u, d, d)$  a un spin  $s = \frac{1}{2}$  et une charge  $q = 0$ .

Attention que l’énergie du composé est l’énergie des constituants plus l’énergie de liaison qui est négative en générale.

**Exemple 4.1.10.** L’énergie du noyau d’**Hélium 4** constitué de  $(p, p, n, n)$  est<sup>3</sup>  $E_{He_4} = 4.002u$  inférieure à  $4 \times E_p = 4 \times 1.007u$  donc il ne peut se désintégrer, il est stable (ainsi les neutrons qu’ils contient sont devenus stables).

Concernant les propriétés statistiques qui nous intéressent dans ce cours, en première approximation (i.e. à faible dilution et petit nombre de degrés de liberté), un composé de **nombre impair de Fermions a les propriétés d’un Fermion** et un composé de **nombre pair de Fermions a les propriétés d’un Boson**.

---

3.  $u = 931.494 \text{Mev}/c^2$  appelé **unité de masse atomique**.

**Exemple 4.1.11.** Le proton et le neutron sont chacun composés de gluons (Bosons) et de trois quarks (Fermions), ce sont donc des Fermions. Le **noyau He4** est composé de 2 protons et 2 neutrons (Fermions), c'est donc un Boson. Le **noyau He3** est composé de 2 protons et 1 neutron, c'est donc un Fermion.

## 4.2 Définition des Fermions et Bosons

On donne une **définition des Fermions et des Bosons** dans le cadre de la mécanique quantique en définissant l'espace de Fock Fermionique et l'espace de Fock Bosonique des états possibles  $(\psi_m)_m$  en terme d'états à une particule  $(\varphi_n)_{n \geq 1}$ .

Considérons une particule élémentaire Fermion ou Boson d'une famille donnée (par exemple un électron ou un photon). En mécanique quantique cette particule est décrite par un vecteur  $\psi \in \mathcal{H}_1$  (sa fonction d'onde) dans un espace de Hilbert  $\mathcal{H}_1$  (indice 1 pour signifier que l'on considère  $N = 1$  particule). On note  $(\varphi_n)_{n \geq 0}$  **une base orthonormée quelconque de cet espace  $\mathcal{H}_1$  à une particule**. Les états  $(\varphi_n)_n$  ne sont pas forcément des états stationnaires. Par exemple on peut penser à  $\varphi_n$  comme un paquet d'onde en  $(q, p)$ , caractérisant la position  $q$ , impulsion  $p$ , et aussi le spin d'une particule. Autre exemple,  $(\varphi_n)_n$  peut être la base des états orbitaux de l'électron dans un atome d'hydrogène (et son spin).

On considère maintenant **plusieurs de particules d'une même famille** (ex : des électrons). Elles ont toutes les même caractéristiques physique : même masse, charge, spin etc.. on dit qu'elles sont **identiques** et **indiscernables**. Conformément à l'expérience, on postule :

**Définition 4.2.1.** Dans le cas des **Bosons**, chaque état possible  $m$  d'un ensemble de plusieurs particules est caractérisé par le **nombre de particules  $N_n \geq 0$  dans chaque état individuel  $\varphi_n$**  avec  $n \geq 0$ . Ainsi un état possible est noté

$$m = (N_0, N_1, N_2, \dots), \quad \text{avec } N_n \geq 0.$$

Le nombre total de particules dans cet état  $m$  est donc

$$N(m) = \sum_{n \geq 0} N_n$$

Ces états  $m$  forment une base orthonormée de l'espace quantique de tous les états possibles, appelé **espace de Fock Bosonique  $\mathcal{F}$** . Autrement dit, un état quelconque de plusieurs particules est une combinaison linéaire de ces états de base  $m$ . Voir figure [4.2.1](#).

**Définition 4.2.2.** Dans le cas des **Fermions**, chaque état possible  $m$  d'un ensemble de plusieurs particules est caractérisé par le **nombre de particules**  $N_n \geq 0$  dans **chaque état individuel**  $\varphi_n$  avec  $n \geq 0$ , mais avec la **contrainte supplémentaire** que  $N_n \leq 1$ . Ainsi un état possible est noté

$$m = (N_0, N_1, N_2, \dots), \quad \text{avec } 0 \leq N_n \leq 1.$$

Le nombre total de particules dans cet état  $m$  est donc

$$N(m) = \sum_{n \geq 0} N_n$$

Ces états  $m$  forment une base orthonormée de l'espace quantique de tous les états possibles, appelé **espace de Fock Fermionique**  $\mathcal{F}$ . Autrement dit, un état quelconque de plusieurs particules est une combinaison linéaire de ces états de base  $m$ . Voir figure 4.2.1.

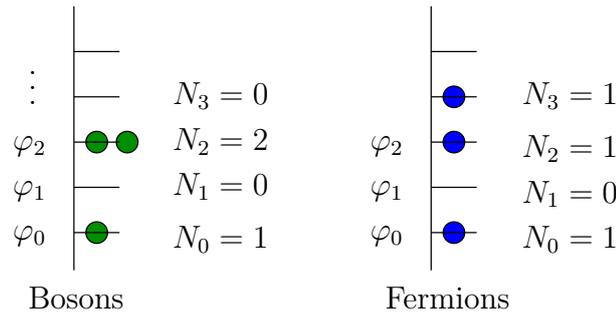


FIGURE 4.2.1 – On note  $N_n$  le nombre de particules dans l'état  $\varphi_n$ .  $N = \sum_n N_n$  est le nombre total. Exemples d'un état  $m = (N_0, N_1, N_2, N_3, \dots)$  à  $N = 3$  particules. Dans le cas des Fermions on a la contrainte  $0 \leq N_n \leq 1$ .

*Remarque 4.2.3.* On note  $\mathcal{F}_N$  l'espace de Fock à nombre fixé  $N \geq 0$  de particules. Ainsi

$$\mathcal{F} = \mathcal{F}_0 \oplus \mathcal{F}_1 \oplus \mathcal{F}_2 \dots \quad (4.2.1)$$

en particulier  $\mathcal{F}_1 = \mathcal{H}_1$  est l'espace à une particule et  $\mathcal{F}_0$  est un espace à zéro particule, contenant un seul état  $m = (0, 0, \dots)$  appelé état du **vide quantique**. Cette décomposition (4.2.1) signifie qu'un état quantique général du système  $\psi \in \mathcal{F}$  peut être une combinaison linéaire d'états (i.e. superposition) appartenant à chaque espace  $\mathcal{F}_N$ . L'état  $\psi$  n'a pas un nombre déterminé de particules.

Comme expliqué plus haut, dans le modèle standard, il y a plusieurs familles de particules : électrons, muons, etc. Il y a un espace de Fock  $\mathcal{F}^{(f)}$  pour chaque famille  $f$  et l'espace de Fock total de toutes les familles est

$$\mathcal{F} = \bigoplus_f \mathcal{F}^{(f)}$$

Noter que les interactions sont responsables de transitions entre ces différentes composantes.

### 4.2.1 (\*) Statistiques quantiques

Des solutions et plus d'explications dans cette référence : [Cours de MQ, chap. 5.3](#)

Considérons une famille donnée de particules (par exemples les électrons qui sont des Fermions ou les photons qui sont des Bosons). On a noté  $\mathcal{F}_1 = \mathcal{H}_1$  l'espace (quantique) à une particule et  $\mathcal{F}_N$  "l'espace de Fock" quantique à  $N$  particules. On note  $(\varphi_n)_{n \geq 0}$  une base arbitraire de  $\mathcal{H}_1$ .

1. Si la dimension de l'espace à une particule est fini :  $\dim \mathcal{H}_1 = d$ , calculer  $\dim \mathcal{F}_N$  pour les Bosons et pour les Fermions.
2. En mécanique quantique, on donne la définition suivante des états possible des **bosons**. Ce sont des combinaisons linéaires des états suivants

$$\psi_m = \hat{S}(\varphi_{n_1} \otimes \varphi_{n_2} \cdots \otimes \varphi_{n_N})$$

avec  $\varphi_{n_j} \in \mathcal{H}_1$  et  $\hat{S}$  est l'opérateur de symétrisation

$$\hat{S} := \frac{1}{N!} \sum_{\sigma \in S_N} \hat{\sigma}$$

où  $S_N$  est le **groupe des permutations** de  $N$  éléments et  $\hat{\sigma}$  est l'opérateur qui permute les vecteurs du produit tensoriel, ex :  $\hat{\sigma}(\varphi_1 \otimes \varphi_2) = \varphi_2 \otimes \varphi_1$ . De même les états possible des **Fermions** sont les combinaisons linéaires des états suivant

$$\psi_m = \hat{A}(\varphi_{n_1} \otimes \varphi_{n_2} \cdots \otimes \varphi_{n_N})$$

avec  $\varphi_{n_j} \in \mathcal{H}_1$  et  $\hat{A}$  est l'opérateur de anti-symétrisation

$$\hat{A} := \frac{1}{N!} \sum_{\sigma \in S_N} \epsilon(\sigma) \hat{\sigma}$$

avec  $\epsilon(\sigma) = \pm 1$  la **signature de la permutation**. Montrer que ces définitions sont équivalentes aux précédentes. Pour cela, montrer qu'une base de  $\mathcal{F}_N$  des bosons est

$$\psi_m = \hat{S}(\varphi_{n_1} \otimes \varphi_{n_2} \cdots \otimes \varphi_{n_N})$$

avec  $n_1 \leq n_2 \leq \dots \leq n_N$  et est caractérisé par  $m = (N_0, N_1, N_2, \dots)$  avec  $N_n$  que l'on exprimera. Montrer qu'une base de  $\mathcal{F}_N$  des Fermions est

$$\psi_m = \hat{A}(\varphi_{n_1} \otimes \varphi_{n_2} \dots \otimes \varphi_{n_N})$$

avec  $n_1 < n_2 < \dots < n_N$  et est caractérisé par  $m = (N_0, N_1, N_2, \dots)$  avec  $N_n$  que l'on exprimera.

3. Montrer que pour  $\sigma \in S_N$ ,  $\hat{\sigma}\hat{S} = \hat{S}$  et  $\hat{\sigma}\hat{A} = \hat{A}$ .
4. Dans l'espace  $\mathcal{H}^{\otimes N} := \mathcal{H}_1 \otimes \mathcal{H}_1 \otimes \dots \otimes \mathcal{H}_1$ , montrer que  $\hat{S}^2 = \hat{S}$ ,  $\hat{S}^\dagger = \hat{S}$  et donc  $\hat{S}$  est un projecteur orthogonal sur le sous espace de Fock Bosonique  $\mathcal{F}_N$ . De même montrer que  $\hat{A}$  est un projecteur orthogonal sur le sous espace de Fock Fermionique.

### 4.3 Distribution de Bose-Einstein et de Fermi-Dirac

référence : [Diu et al. \(1989, p.670\)](#).

On considère maintenant le **cas particulier de particules identiques et indépendantes** (i.e. sans interaction mutuelle) qui sont des Bosons ou des Fermions. On suppose que  $(\varphi_n)_{n \geq 0}$  **sont les états propres d'énergie pour une particule seule** et on note  $\epsilon_n$  l'énergie de  $\varphi_n$ , i.e.

$$\hat{H}\varphi_n = \epsilon_n\varphi_n$$

avec  $\epsilon_0 \leq \epsilon_1 \leq \epsilon_2 \leq \dots$

On suppose que ce système échange de l'énergie et des particules avec son environnement, un réservoir de température  $T$  et potentiel chimique  $\mu$ . On note  $m = (N_0, N_1, N_2, \dots)$  un état du système à plusieurs particules comme expliqué ci-dessus.

Pour la suite, il est utile de remarquer le résultat suivant qui est assez intuitif.

**Lemme 4.3.1.** *Pour un système de **particules identiques et indépendantes**, (Fermions ou Bosons), la mesure de probabilité  $P(m)$  de Boltzmann-Gibbs d'un microétat  $m = (N_0, N_1, N_2, \dots)$  du système, d'énergie  $E(m) = \sum_{n \geq 0} N_n \epsilon_n$  et de nombre de particules  $N(m) = \sum_{n \geq 0} N_n$ , se factorise (i.e. événements indépendants)*

$$\begin{aligned} P(m) &= \frac{1}{Z} \exp(-\beta(E(m) - \mu N(m))) \\ &= \prod_{n \geq 0} p_n(N_n) \end{aligned}$$

où  $p_n(N_n)$  est la probabilité pour que le niveau  $n$  possède  $N_n$  particules, aussi donné par la mesure de Boltzmann Gibbs

$$p_n(N_n) = \frac{1}{Z_n} e^{-\beta(N_n \epsilon_n - \mu N_n)}.$$

*Démonstration.* Considérons un état à plusieurs particules noté  $m = (N_0, N_1, N_2, \dots)$ . D'après la loi de Boltzmann-Gibbs, la probabilité d'apparition de l'état  $m$  est

$$P(m) = \frac{1}{Z} \exp(-\beta(E(m) - \mu N(m)))$$

Dans cette expression,

$$\begin{aligned} \exp(-\beta(E(m) - \mu N(m))) &= \exp\left(-\beta\left(\sum_{n \geq 0} N_n \epsilon_n - \mu N_n\right)\right) \\ &= \prod_{n \geq 0} \exp(-\beta N_n (\epsilon_n - \mu)) \end{aligned}$$

Donc  $P(m)$  se factorise. □

**Proposition 4.3.2.** *Considérons un système de **particules identiques et indépendantes**. Supposons que le système échange de l'énergie et des particules avec un environnement à la température  $T$  (on pose  $\beta = \frac{1}{kT}$ ) et potentiel chimique  $\mu$ ,*

- Dans le cas de **Bosons**, si  $\mu < \epsilon_n, \forall n$ , le nombre moyen de particules dans un état à une particule d'énergie  $\epsilon_n$  est

$$\langle N_n \rangle = \frac{1}{e^{\beta(\epsilon_n - \mu)} - 1} \quad (4.3.1)$$

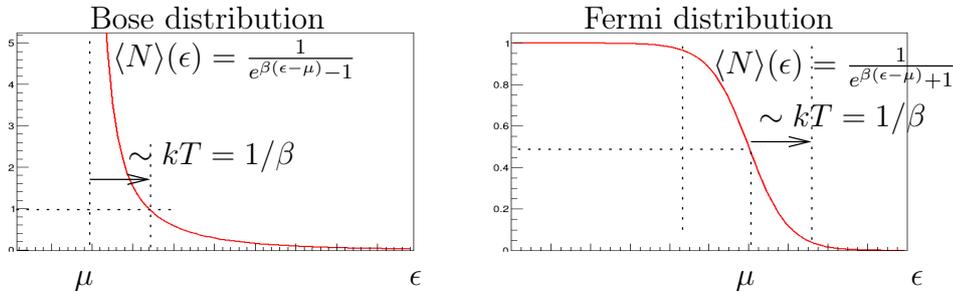
*appelée distribution de Bose-Einstein.*

- Dans le cas de **Fermions**, le nombre moyen de particules dans un état à une particule d'énergie  $\epsilon_n$  est

$$\langle N_n \rangle = \frac{1}{e^{\beta(\epsilon_n - \mu)} + 1} \quad (4.3.2)$$

*appelée distribution de Fermi-Dirac.*

*Remarque 4.3.3.* Les distributions dépendent de l'état  $n$  à travers son énergie  $\epsilon_n$ . On trace donc les fonctions  $\langle N \rangle(\epsilon) = (\exp(\beta(\epsilon - \mu)) \pm 1)^{-1}$  avec  $\beta, \mu$  fixés :



*Remarque 4.3.4.* Dans les deux cas, pour  $\beta(\epsilon_n - \mu) \gg 1$ , on a

$$\langle N_n \rangle \approx e^{-\beta(\epsilon_n - \mu)},$$

qui correspond à la décroissance exponentielle observée à la droite des figures. L'effet statistique Boson/Fermion n'a pas d'importance dans ce régime à "faible densité".

*Démonstration.* Faisons la preuve de la proposition 4.3.2. pour un niveau  $n \geq 0$  donné,

$$\begin{aligned} \langle N_n \rangle &= \sum_{N_n} p_n(N_n) N_n \\ &= \frac{1}{Z_n} \sum_{N_n \geq 0} e^{-\beta N_n(\epsilon_n - \mu)} N_n \end{aligned} \quad (4.3.3)$$

avec

$$Z_n = \sum_{N_n \geq 0} e^{-\beta N_n(\epsilon_n - \mu)}. \quad (4.3.4)$$

Dans le cas des Bosons, la somme est sur  $N_n \in \mathbb{N}$ . Cette somme est convergente seulement si  $\epsilon_n - \mu > 0, \forall n$ . On calcule (suite géométrique)

$$Z_n(\mu) \stackrel{(4.3.4)}{=} \sum_{N_n \geq 0} (e^{-\beta(\epsilon_n - \mu)})^{N_n} \stackrel{(A.1.1)}{=} \frac{1}{1 - e^{-\beta(\epsilon_n - \mu)}}$$

et

$$\frac{\partial Z_n}{\partial \mu} = \sum_{N_n \geq 0} \beta N_n e^{-\beta N_n(\epsilon_n - \mu)} = \frac{\beta e^{-\beta(\epsilon_n - \mu)}}{(1 - e^{-\beta(\epsilon_n - \mu)})^2}$$

donc

$$\begin{aligned} \langle N_n \rangle &\stackrel{(4.3.3)}{=} (1 - e^{-\beta(\epsilon_n - \mu)}) \frac{e^{-\beta(\epsilon_n - \mu)}}{(1 - e^{-\beta(\epsilon_n - \mu)})^2} \\ &= \frac{e^{-\beta(\epsilon_n - \mu)}}{(1 - e^{-\beta(\epsilon_n - \mu)})} = \frac{1}{e^{\beta(\epsilon_n - \mu)} - 1} \end{aligned}$$

Dans le cas des Fermions, la somme est sur  $N_n \in \{0, 1\}$ . On calcule

$$Z_n \stackrel{(4.3.4)}{=} 1 + e^{-\beta(\epsilon_n - \mu)},$$

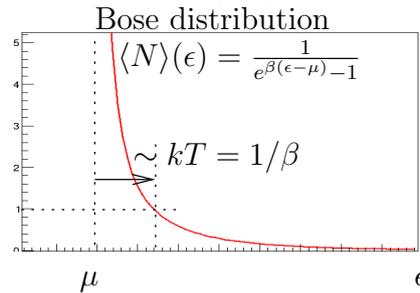
$$\begin{aligned} \langle N_n \rangle &\stackrel{(4.3.3)}{=} \frac{1}{1 + e^{-\beta(\epsilon_n - \mu)}} (0 + e^{-\beta(\epsilon_n - \mu)}) \\ &= \frac{1}{e^{\beta(\epsilon_n - \mu)} + 1} \end{aligned}$$

□

## 4.4 Interprétations et applications de la distribution de Bose-Einstein

Dans cette section on interprète la formule (4.3.1) donnant que le nombre moyen de bosons dans un état à une particule d'énergie  $\epsilon_n$  est

$$\langle N_n \rangle = \frac{1}{e^{\beta(\epsilon_n - \mu)} - 1}.$$



où  $\beta = \frac{1}{kT}$  est relié à la température extérieure et  $\mu$  est le potentiel chimique extérieur. On discute surtout d'une application importante qui est la densité d'énergie des photons appelée la loi de Planck (1900).

### 4.4.1 Densité d'énergie des photons : la loi de Planck (1900).

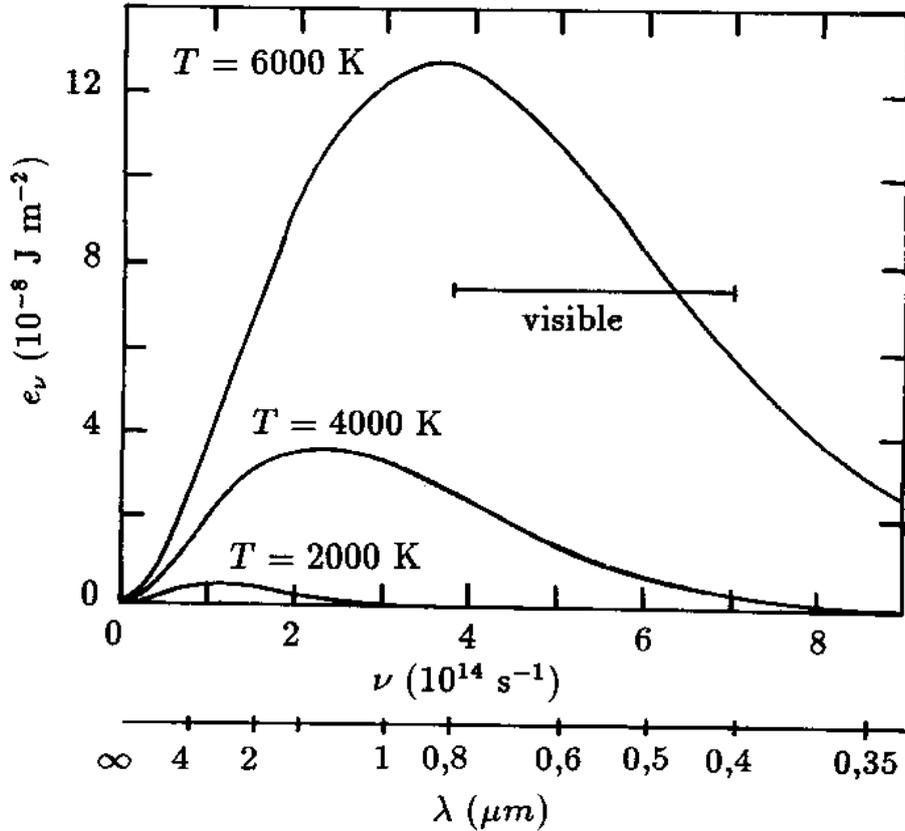
référence : [Diu et al. \(1989, p.825\)](#).

Dans cette section nous allons considérer un "gaz de photon à l'équilibre thermique, à la température  $T$ ". Les photons sont les excitations quantiques du champ électromagnétique. Comme expliqué dans l'introduction, c'est un champ de Jauge et par conséquent les photons sont des bosons.

Considérons un volume  $V$  fixé, qui contient un gaz de photons à l'équilibre thermodynamique à la température  $T$  fixée. Cela signifie que ces photons sont en contact avec de la matière qui a la température  $T$  (car il n'y a pas d'interaction directe entre les photons). Noter qu'il n'y a pas de loi de conservation sur le nombre de photons. Donc **il n'y a pas de potentiel chimique pour les photons**. On pose  $\mu = 0$  dans la formule (4.3.1).

**Proposition 4.4.1.** *La densité d'énergie en fréquence  $\nu$ , par unité de volume d'un gaz de photons à la température  $T$  est*

$$u(\nu)d\nu = \frac{8\pi h\nu^3 d\nu}{c^3 (e^{h\nu/kT} - 1)} : \text{loi de Planck}$$



*Démonstration.* (\*) voir TD. Considérons une enceinte de volume  $V$ . Il contient des modes électromagnétiques. On peut estimer leur densité i.e. le nombre de modes ondulatoires par intervalle de fréquence  $\frac{dn}{d\nu}$  par la loi de Weyl de la façon suivante. Pour un mode électromagnétique, la fréquence  $\nu = \frac{\omega}{2\pi}$  est reliée au vecteur d'onde  $\vec{k}$  par la relation de dispersion

$$\omega = c |\vec{k}|.$$

D'après la loi de Weyl, et n'oubliant pas que chaque mode spatial a deux états de polarisation possible, le nombre de modes de fréquence inférieure à  $\omega$  est

$$n = \frac{2 \times V \times \text{Vol} \left( \left\{ \vec{k}, \left( c |\vec{k}| \right) \leq \omega \right\} \right)}{(2\pi)^3} = \frac{2V \frac{4}{3}\pi \left( \frac{\omega}{c} \right)^3}{(2\pi)^3} = \frac{V 8\pi \omega^3}{3 (2\pi)^3 c^3}$$

La densité est donc

$$\frac{dn}{d\nu} = 2\pi \frac{dn}{d\omega} = 2\pi \frac{V 8\pi \omega^2}{(2\pi)^3 c^3} = \frac{2V \omega^2}{\pi c^3}$$

On rappelle que d'après la mécanique quantique, chaque mode électromagnétique est peuplé de photons et l'énergie d'un photon est  $E = \hbar\omega = h\nu$ , son impulsion est  $\vec{p} = \hbar\vec{k}$ . Ces photons sont des bosons, donc le nombre moyen dans chaque mode  $\frac{d\langle N \rangle}{dn}$  est donné par

la distribution de Bose-Einstein (4.3.1). Et l'énergie s'obtient en multipliant l'énergie d'un photon par le nombre de photons donc l'énergie par mode est  $\frac{d\langle N \rangle}{dn} E$ .

On cherche  $u(\nu)$  qui est la densité d'énergie en fréquence  $\nu$  et par unité de volume. On écrit donc

$$\begin{aligned} u(\nu) &= \frac{1}{V} \frac{d\langle N \rangle}{d\nu} E = \frac{1}{V} \left( \frac{d\langle N \rangle}{dn} \right) \left( \frac{dn}{d\nu} \right) E \\ &= \frac{1}{V} \left( \frac{1}{e^{\beta E} - 1} \right) \left( \frac{2V\omega^2}{\pi c^3} \right) (\hbar\omega) = \left( \frac{1}{e^{\beta E} - 1} \right) \left( \frac{2\hbar\omega^3}{\pi c^3} \right) \\ &= \frac{8\pi h\nu^3}{c^3 \left( e^{\frac{h\nu}{kT}} - 1 \right)}. \end{aligned}$$

□

#### 4.4.2 Conséquences et applications de la loi de Planck

- A la surface du Soleil, le magma a la température  $T = 6000^0K$ . Il émet un rayonnement de photons qui suit la loi de Planck. Pour cette raison la couleur du Soleil est plutôt jaune. Le **Diagramme Hertzsprung-Russel** relie la couleur des étoiles (leur température) avec leur âge, et leur masse. En hiver on observe **Bételgeuse** qui est rouge avec une température de surface  $T = 3600K$ , **Rigel** qui est bleue avec  $T = 10\,000K$ . L'été, **Antares** qui est rouge et **Deneb** qui est bleue.
- Dans un four, on peut avoir  $T = 600^0K$ . Cf Diu p826-917 **Diu et al. (1989)** pour une barre de fer. Le forgeron a des tables de couleurs, lui donnant la température, à partir de la couleur observée. (Zitoun p204 **Couture & Zitoun (1992)**).
- Le rayonnement fossile de l'univers suit la loi de Planck pour  $T = 2,725K \pm 0.002$ . Mais ce rayonnement n'est plus à l'équilibre thermique. Il était à l'équilibre thermique en étant couplé au plasma contenu dans l'univers avec une température  $T = 3000K$ , lorsque l'univers avait 300 000 ans. Ensuite, à cause de l'expansion de l'espace, le rayonnement s'est découplé de la matière et se propage de façon libre, en conservant sa distribution initiale, mais ses longueurs d'ondes ont été étirée par un facteur  $\times 1000$ .

#### 4.4.3 Propriétés de la distribution de Planck (voir figure)

- **Limites des basses fréquences** : on a pour  $h\nu \ll kT$  :

$$u_{cl}(\nu) \simeq 8\pi \frac{kT}{c^3} \nu^2$$

c'est la Formule de **Rayleigh-Jeans**. Ce résultat s'obtient aussi en considérant le champs électromagnétique classique, qui est composé d'oscillateurs indépendants (champs  $\mathcal{E}$  et  $\mathcal{B}$ ), et en utilisant le théorème d'équipartition de l'énergie, qui montre que il y a une énergie moyenne  $kT$  par oscillateur. Alors  $du_{cl} = kT.dn$ . ( $dn$  est le nombre de oscillateurs, ci-dessus).

- **Limite des hautes fréquences** : on a pour  $h\nu \gg kT$  :

$$u(\nu) \simeq 8\pi \frac{h}{c^3} \nu^3 e^{-\frac{h\nu}{kT}}$$

c'est la loi de **Wien (1893)** qui a trouvé cette décroissance exponentielle ; mais il lui manquait la constante fondamentale  $h$  pour préciser cette décroissance. C'est une manifestation des photons en tant que particules. En effet, ce résultat se retrouve en utilisant la loi de Boltzmann, et en tenant compte de la quantification de l'énergie par le photon :  $du = (h\nu) dn e^{-E/(kT)}$ .

- **Fréquence du maximum de  $u(\nu)$**  : on calcule (voir TD)

$$\nu_{max} \simeq 2,821 \frac{kT}{h}$$

la position du maximum est linéaire avec  $T$  : c'est la loi **du déplacement de Wien** déduit de l'étude empirique des courbes expérimentales.

- **Energie totale** : pour un gaz de photons à l'équilibre, à la température  $T$  dans une enceinte de volume  $V$ , l'énergie totale est (voir TD) :

$$\begin{aligned} E(T, V) &= V \int_0^{+\infty} u(\nu) d\nu \\ &= \frac{\pi^2}{15} \frac{k^4}{h^3 c^3} V T^4 \\ &= V \sigma T^4 \end{aligned}$$

avec

$$\sigma = \frac{\pi^2}{15} \frac{k^4}{h^3 c^3} = 5.67 \cdot 10^{-8} J.m^{-3}.K^{-4}$$

c'est la loi de **Stephan-Boltzmann**. Stephan (1879) avait observé la relation empirique  $E = V\sigma T^4$ , et Boltzmann (1884) l'a obtenue par des arguments de thermodynamique (voir Bransden p.5).

#### 4.4.4 Autres remarques

- Nous avons dérivé la loi de Planck, en supposant l'existence du photon, qui provient de la quantification du champ électromagnétique (i.e. spectre discret d'énergie  $N(h\nu)$  pour chaque mode de fréquence  $\nu$ ). Si la constante de Planck tendait vers 0, ce qui signifie que la quantification disparaît, et que le champ est classique, alors l'énergie totale ci-dessus  $E(T, V)$  diverge. La quantification du champ est donc nécessaire pour expliquer l'existence d'un équilibre thermique entre matière et rayonnement. Autrement dit, le fait que le champ électromagnétique possède un nombre infini de degrés de liberté fait que, s'il était classique, il absorberait toute l'énergie de la matière si celle-ci a un nombre fini de degrés de liberté. Si le champ doit être quantifié, la matière doit aussi être quantifiée. Autrement dit le modèle classique de la physique possédant des champs n'est pas acceptable car il ne permet pas l'existence d'états d'équilibres en physique statistique.

- Historique : Cette formule a été trouvée par **Rayleigh** dans le domaine des **basses basse fréquences**, par théorème d'équipartition  $\frac{1}{2}kT$  appliquée au modes du champ classique. Puis **Wien (1893)** pour la partie **hautes fréquences**, à partir de la distribution de Boltzmann (c'est une manifestation des photons en tant que particule). Puis **Planck en 1901** a trouvé la formule ci-dessus, en introduisant l'idée de quantification de l'énergie électromagnétique. **Einstein en 1905** introduit la notion de photon.
- Autre application de la loi de Bose : **He<sup>4</sup> superfluide** : propriétés du **condensat** (cf ci-dessous). La loi de Bose-Einstein se manifeste aussi dans les modèles de vibration d'un cristal: modèle de Einstein et de Debye.

#### 4.4.5 La condensation de Bose-Einstein

Référence : [Diu et al. \(1989\)](#).

Considérons maintenant une autre conséquence de la distribution de Bose Einstein (4.3.1) appelée la **condensation de Bose Einstein**. Elle se manifeste dans des expériences récentes de condensats d'atomes froids, et correspond récemment à plusieurs prix Nobels de physique. (Elle a aussi un lien avec la superfluidité, la supraconductivité, les faisceaux laser).

**Système physique** : Le système physique considéré est un ensemble d'**atomes identiques à très basse température** ( $10^{-7}K$  par **refroidissement laser** et **évaporatif**), appelés "atomes froids", contenant un nombre pair de protons et de neutrons. Par conséquent un nombre pair d'électrons. Ces atomes ont donc un spin entier et sont **des bosons**, comme discuté en section 4.1.7.1. Par exemple l'Helium 4 (2 protons et 2 neutrons). A la différence des photons, les atomes sont de la matière et il y a des règles de conservation du nombre d'atome lors des interactions. Dans les expériences (à basse énergie) le **nombre total d'atome  $N$  est une quantité conservée**.  $N \approx 10^3$  à  $10^6$ . La densité est assez faible pour pouvoir négliger les interactions mutuelles en première approximation.

**Modèle** : On va décrire un **gaz de  $N$  atomes identiques qui sont des bosons** de masse  $m > 0$  de spin  $s$  entier, à l'équilibre thermique, caractérisé par une **température  $T \geq 0$**  et un **potentiel chimique  $\mu$** .

On rappelle que le terme "gaz" signifie que dans ce modèle il n'y a pas d'interaction entre les particules. Par ailleurs  $\mu < \epsilon_0$  où  $\epsilon_0$  est le niveau d'énergie le plus bas pour un atome seul. On va supposer  $N \gg 1$  très grand et la densité de particules  $\rho = \frac{N}{V}$  fixée. On va noter  $N_n$  le nombre moyen d'atomes dans le niveau d'énergie  $\epsilon_n$  à une particule. On va s'intéresser à la proportion  $N_n/N$  dans la limite  $N \gg 1$ .

On va aussi considérer des atomes libres dans une boîte cubique de côté  $L$ , donc de volume  $V = L^3$ . Ainsi, notant  $x = (x_1, x_2, x_3)$  avec  $x_j \in [0, L]$ ,  $j = 1, 2, 3$ , les niveaux d'énergie à une particule sont  $\varphi_n(x) = \prod_{j=1}^3 \sin\left(\pi n_j \frac{x_j}{L}\right)$  avec  $n = (n_1, n_2, n_3) \in (\mathbb{N} \setminus \{0\})^3$

d'énergie  $\epsilon_n = \frac{1}{2m} \left(\frac{\pi\hbar}{L}\right)^2 |n|^2$ . Ainsi l'écart entre l'état fondamental et le premier niveau d'énergie est  $\Delta_1 = \epsilon_{1,0,0} - \epsilon_0 = \frac{\pi^2}{2} \left(\frac{\hbar^2}{mL^2}\right)$ .

On rappelle (TD2) la loi de Weyl

$$n(\epsilon) = \frac{(2s+1)}{(2\pi\hbar)^3} \text{Vol} \left\{ (x, p) \text{ t.q. } \frac{|p|^2}{2m} \leq \epsilon \right\} = \frac{(2s+1)}{(2\pi\hbar)^3} V \frac{4}{3} \pi (2m\epsilon)^{\frac{3}{2}}$$

où  $s$  est le spin (entier) de chaque atome, et qui donne la densité moyenne d'états à 1 particule dans la boîte (on remplace  $\epsilon$  par  $(\epsilon - \epsilon_0)$  car  $n \gg 1$ )

$$\frac{dn}{d\epsilon} = C (\epsilon - \epsilon_0)^{1/2}$$

avec

$$C = \frac{(2s+1)}{2^{1/2}\pi^2} \left(\frac{mL^2}{\hbar^2}\right)^{3/2} = \frac{(2s+1)\pi}{4\Delta_1^{3/2}}$$

En effet

**Proposition 4.4.2.** (Einstein 1925). On définit la **température de Bose**  $T_B$  par

$$kT_B = \left(\frac{N}{CI}\right)^{2/3},$$

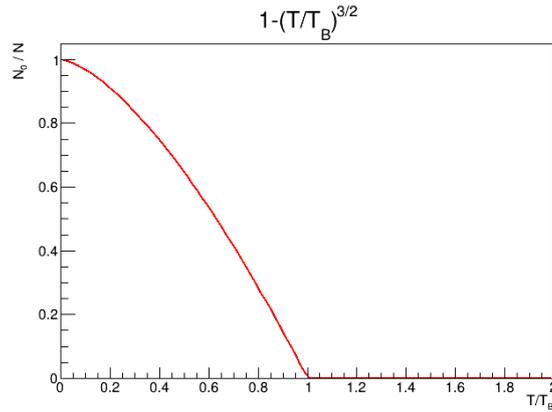
avec la constante numérique ( $\zeta$  est la **fonction zeta de Riemann**)

$$I = \int_0^{+\infty} \frac{1}{e^x - 1} x^{1/2} dx = \frac{\sqrt{\pi}}{2} \zeta\left(\frac{3}{2}\right) \approx 2,31\dots$$

Dans la limite  $N \gg 1$ , la proportion d'atomes  $N_0/N$  dans l'état fondamental d'énergie  $\epsilon_0$  est nul pour  $T \geq T_B$ , mais pour  $T < T_B$ , il y a "**condensation dans l'état fondamental**" :

$$\frac{N_0}{N} = 1 - \left(\frac{T}{T_B}\right)^{3/2},$$

Pour les autres niveaux d'énergie  $\epsilon_n > \epsilon_0$ , la proportion  $N_n/N$  est toujours nulle.



*Remarque 4.4.3.* On peut écrire (comme dans la plupart des livres) :

$$kT_B = \left( \frac{N}{CI} \right)^{2/3} = \frac{2\pi\hbar^2}{m} \left( \rho \frac{1}{(2s+1)\zeta\left(\frac{3}{2}\right)} \right)^{2/3}.$$

*Démonstration.* Dans la preuve, on considère la limite  $N \gg 1$ . Posons  $\Delta_0 = \epsilon_0 - \mu > 0$  et  $\Delta_1 = \epsilon_1 - \epsilon_0 \geq 0$ . On va traiter différents cas séparément.

1. On supposera

$$\Delta_0 \ll \Delta_1 \ll kT.$$

On a  $N_n = \frac{1}{e^{(\epsilon_n - \mu)/(kT)} - 1}$  pour chaque niveau  $n \geq 0$ . On a  $N_0 = \frac{1}{e^{\Delta_0/kT} - 1} \approx \frac{kT}{\Delta_0} \gg 1$ . Alors pour les niveaux supérieurs, i.e.  $|n| > 0$ , on a  $\epsilon_n \geq \epsilon_0 + \Delta_1$ ,

$$N_n = \frac{1}{e^{(\epsilon_n - \mu)/(kT)} - 1} < \frac{1}{e^{(\Delta_1 + \Delta_0)/(kT)} - 1} \approx \frac{kT}{(\Delta_1 + \Delta_0)} \approx \frac{kT}{\Delta_1} \ll \frac{kT}{\Delta_0} \approx N_0.$$

Donc  $N_n/N \ll N_0/N \leq 1$  et  $N_n/N$  peut être considérée comme nulle. On applique la loi de Weyl à partir des niveaux  $|n| \geq 1$ , et en posant  $x = (\epsilon - \mu)/(kT)$ ,

$$\begin{aligned} N &\approx N_0 + \int_{\epsilon_1}^{+\infty} \frac{1}{e^{(\epsilon - \mu)/(kT)} - 1} C (\epsilon - \epsilon_0)^{1/2} d\epsilon \\ &= N_0 + C (kT)^{3/2} \int_{(\Delta_1 + \Delta_0)/(kT)}^{+\infty} \frac{1}{e^x - 1} (x - \Delta_0/(kT))^{1/2} dx \\ &\approx N_0 + C (kT)^{3/2} I \end{aligned}$$

Ainsi

$$\frac{N_0}{N} \approx 1 - \frac{CI}{N} (kT)^{3/2}.$$

A la limite où il n'y a plus de condensation  $N_0/N = 0$ , la température de transition  $T_B$  est donnée par

$$0 = 1 - \frac{C(kT_B)^{3/2}}{N} I \Leftrightarrow kT_B = \left( \frac{N}{CI} \right)^{2/3}$$

et on peut écrire

$$\frac{N_0}{N} \approx 1 - \left( \frac{T}{T_B} \right)^{3/2}.$$

2. Si

$$kT \ll \Delta_1,$$

alors  $\frac{N_0}{N} \approx 1$ .

3. Si

$$\Delta_1 \ll \Delta_0 \ll kT$$

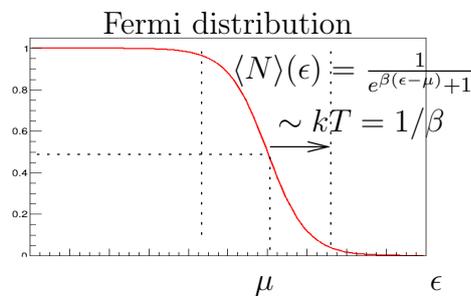
alors  $\frac{N_0}{N} \approx 0$ .

□

## 4.5 Interprétations et applications de la distribution de Fermi-Dirac

Dans cette section on interprète la formule (4.3.2) donnant que le nombre moyen de fermions dans un état à une particule d'énergie  $\epsilon_n$  est

$$\langle N_n \rangle = \frac{1}{e^{\beta(\epsilon_n - \mu)} + 1},$$



où  $\beta = \frac{1}{kT}$  est relié à la température extérieure et  $\mu$  est le potentiel chimique extérieur. On discute surtout d'une application importante qui est la distribution des électrons dans les métaux et semi-conducteurs.

### 4.5.1 Interprétations

Il est habituel de définir

$$\theta_F := \frac{\mu}{k}$$

appelée **température de Fermi**. Voici son interprétation

- Si  $kT \ll \mu \Leftrightarrow T \ll \theta_F$  : La distribution de Fermi devient une "marche d'escalier" positionnée à  $\epsilon = \mu$  appelé le **niveau de Fermi**. Cela signifie que les niveaux d'énergie  $\epsilon_n < \mu$  contiennent  $\langle N_n \rangle = 1$  particule et les niveaux d'énergie  $\epsilon_n > \mu$  contiennent  $\langle N_n \rangle = 0$  particules. **Seules les particules proches du niveau de Fermi jouent un rôle**, car ce sont les seules qui peuvent changer éventuellement d'état par un petit changement d'énergie. On parle de **gaz dégénéré de Fermions**.
- Pour  $kT \gtrsim \mu \Leftrightarrow T \gtrsim \theta_F$ , la distribution de Fermi est plus douce. Si  $\epsilon_n < \mu$ , alors  $\langle N_n \rangle > 0.5$  et le niveau est plus souvent rempli que vide, sinon c'est le contraire. On appelle un **trou** l'absence de particule. La distribution des trous est donc  $\langle \text{trou}_n \rangle = 1 - \langle N_n \rangle$ .

### 4.5.2 Application de la distribution de Fermi : électrons dans un métal avec potentiel constant

Dans le **modèle de Sommerfeld des métaux**, on considère seulement les électrons de valence (voir cours de physique du solide). On considère le métal comme une "boite" contenant des électrons libres. L'énergie d'une particule individuelle est donc  $E = \frac{p^2}{2m}$ .

**Proposition 4.5.1.** *Si le nombre de particules  $N$  et le volume  $V$  sont fixés, la densité de particules (d'électrons libres)  $n = \frac{N}{V}$  est relié au potentiel chimique  $\mu$ , à  $T = 0$  par*

$$\mu = \frac{\hbar^2}{2m} (3\pi^2 n)^{2/3} \quad (4.5.1)$$

*Démonstration.* Cette relation s'obtient directement à partir de la formule de Weyl du comptage d'états (voir TD2) dans l'espace des phases  $(x, k)$  où  $x$  est la position dans un volume  $V$  et  $k \in \mathbb{R}^3$  est le vecteur d'onde. L'impulsion<sup>4</sup> est  $p = \hbar k \in \mathbb{R}^3$ . L'énergie d'une particule libre est son énergie cinétique  $E = \frac{|p|^2}{2m} = \frac{\hbar^2}{2m} |k^2|$ . A  $T = 0$ , tous les niveaux d'énergie jusqu'à  $E = \mu$  sont occupés par les  $N$  particules. La condition que l'énergie est inférieure à  $\mu$ , s'écrit

$$E \leq \mu \Leftrightarrow \frac{\hbar^2}{2m} |k^2| \leq \mu \Leftrightarrow |k| \leq R_\mu = \left( \frac{2m\mu}{\hbar^2} \right)^{1/2}$$

qui s'interprète comme le fait que le vecteur d'onde est dans une sphère de rayon  $R_\mu = \left( \frac{2m\mu}{\hbar^2} \right)^{1/2}$ , appelée sphère de Fermi.

4. on rappelle que en physique l'énergie est reliée à une fréquence par  $E = \hbar\omega$  et l'impulsion à une fréquence spatiale ou vecteur d'onde par  $p = \hbar k$ . C'est la définition de la constante de Planck.

La loi de Weyl donne que le nombre d'état  $N$  d'énergie plus petite que  $\mu$  est (le préfacteur 2 est pour compter deux états de spins différents)

$$\begin{aligned} N(E) &= \frac{2}{(2\pi)^3} \text{Vol} \{(x, k) \text{ t.q. } E \leq \mu\} = \frac{2}{(2\pi\hbar)^3} V \left( \frac{4}{3} \pi R_\mu^3 \right) \\ &= \frac{V}{3\pi^2} \left( \frac{2m\mu}{\hbar^2} \right)^{3/2} \end{aligned}$$

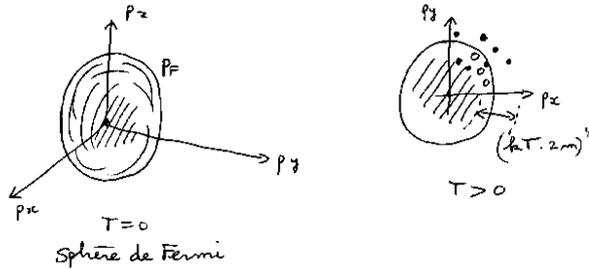
Donnant

$$\mu = \frac{\hbar^2}{2m} \left( 3\pi^2 \frac{N}{V} \right)^{2/3}.$$

□

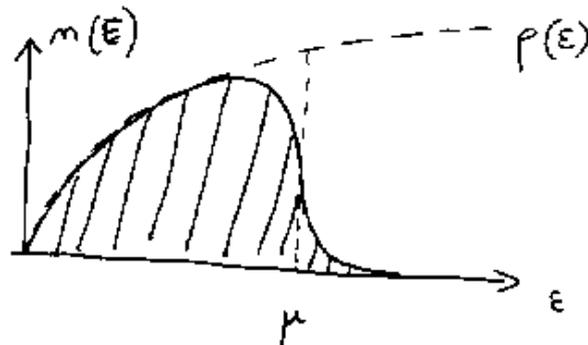
#### 4.5.2.1 Remarques :

- Comme  $E = \frac{p^2}{2m}$ , à  $T = 0$ , dans l'espace des impulsions, les états occupés sont ceux d'impulsion  $|p| < p_F = \sqrt{2m\mu}$ .  $p_F$  s'appelle **l'impulsion de Fermi**. Ces impulsions forment **la sphère de Fermi**. A la surface de cette sphère (**surface de Fermi**), les électrons ont une grande vitesse  $V_F = p_F/m \simeq 10^6 m/s$  pour les électrons du Sodium!



- Si on note  $N$  le nombre de particules considérées, et  $n$  le nombre d'états quantiques individuels, alors pour connaître la densité de particules par intervalle d'énergie  $dN/d\varepsilon$ , il faut multiplier la fonction d'occupation de Fermi  $\langle N_n \rangle = dN/dn$  par la densité de niveaux à une particule  $\rho(\varepsilon) = dn/d\varepsilon$

$$\frac{dN}{d\varepsilon} = \frac{dN}{dn} \frac{dn}{d\varepsilon} = \langle N \rangle (\varepsilon) \rho(\varepsilon)$$



#### 4.5.2.2 Application numérique : le Sodium :

cf Diu p838, Guinnier p96.

Dans le sodium, il y a un électron libre par atome (car c'est un alcalin. En fait il y a 11 électrons par atome, et 10 d'entre eux sont liés, car ils sont dans des bandes internes. voir cours de phys. du solide) donc la densité est

$$n = \frac{N}{V} = \frac{MN\mathcal{N}}{V\mathcal{N}M} = \frac{M_{vol}\mathcal{N}}{M_{mol}} = \dots = 2.62 \cdot 10^{22} \text{ e}^- / \text{cm}^3$$

où  $\mathcal{N}$  est le nombre d'avogadro. On a

$$m_e = 9.3110^{-31} \text{ kg}$$

D'après (4.5.1),

$$\mu = 3eV$$

donc

$$\theta_F = \mu/k = 35000K$$

à température ambiante,  $T_{metal} \ll \theta_F$ . La vitesse de Fermi, (i.e. vitesse des électrons au niveau de Fermi) est  $V_F = \sqrt{\frac{2k\theta_F}{m}} = 10^6 \text{ m/s}$ .

#### 4.5.2.3 Remarques :

- Ces résultats sont très différents de l'idée que à  $T = 0$  les vitesses sont nulles.
- Le principe d'exclusion de Pauli explique en particulier que l'on ne pénètre pas le plancher, car les électrons du plancher et de la chaussure s'excluent !
- À température ordinaire, pour un métal, on a  $T \ll \theta_F$  : donc les électrons forment un gaz dégénéré.
- Pour  $\text{He}^3$  liquide :  $m_{\text{HE3}}$  est  $10^4$  plus lourd que  $m_e$  donc  $\theta_F = 2,7K$  et  $\mu = 10^{-4}eV$ .

- Pour les noyaux nucléaires :  $\theta_F$  très élevé, car  $m$  lourd mais la densité  $n$  est très élevée. Donc  $T \ll \theta_F$ . Les nucléons forment donc un gaz dégénéré.
- Pour les étoiles à neutrons : gaz de neutron à volume  $\infty$ , voir TD.
- Pour les étoiles naines blanches :  $T = 10^7 K \ll \theta_F$ , cf Diu p794, voir TD.
- (Notion de pression quantique , cf Diu P785, p659)

### 4.5.3 Autres applications (\*)

(voir Balian [Balian \(2007\)](#))

- Effet Thermo-ionique (Balian, Feynmann) : l'émission d'électrons est contrôlée par  $\mu$ . **Voir TD**
- Effet Volta : contact entre deux métaux avec  $\mu$  initiaux différents  $\mu_1, \mu_2$ . Il y a égalisation de  $\mu$  par échange d'électrons, donc l'apparition d'une différence de potentiel électrique  $V = (\mu_1 - \mu_2) / e$ .
- Effet photo-électrique : possible si  $E_\gamma = h\nu > (U - \mu)$ . avec  $U$  : potentiel de la barrière. cf Feynmann

Troisième partie  
Systèmes en interactions



# Chapitre 5

## Phases et transitions de phases

### 5.1 Introduction

#### 5.1.1 Phases de la matière

Dans la nature, la matière est constituée de particules, d'atomes et molécules qui s'organisent selon différentes “**phases**”, dépendant des conditions du milieu :

- **Gaz** : si les molécules sont “libres”. Cela signifie que chaque molécule a une énergie suffisante pour pouvoir s'échapper (se libérer) de l'attraction exercée par les autres.
- **Liquide** si les molécules subissent des interactions attractives entre elles, et ont une énergie pas suffisante pour se libérer, mais suffisante pour se déplacer par rapport aux autres. Gaz et liquide sont appelés “**fluides**”.
- **Solide** si les molécules subissent des interactions attractives entre elles mais n'ont pas assez d'énergie pour se déplacer et sont donc figées près d'une certaine position moyenne. Dans la plupart des cas la configuration des molécules est périodique, appelée **crystal**.
- **Plasma** : si les atomes sont ionisés (i.e. noyaux atomiques seuls), les électrons sont libres et il y a aussi un gaz de photons. L'ensemble du plasma est neutre électriquement. Le soleil et autres étoiles du même type sont un plasma.
- **Condensat de Bose** : des atomes de spin entier (bosons) dans un état de superposition quantique.
- **Supraconductivité** : dans un matériau, les électrons s'assemblent en “**paires de Cooper**” qui circulent sans dissipation. Chaque paire est un boson, et ces bosons sont dans un état de condensat.
- En **cosmologie**, on imagine que les champs élémentaires du modèle standard de la physique des particules sont en fait des phases de champs plus fondamentaux, apparus après le big bang, lors du refroidissement de l'univers. Voir **Phase Transitions in the Early Universe** .
- En **économie** ou **sociologie**, attention ce ne sont pas des sciences exactes, il y a des tentatives d'appliquer des analyses de physique statistique. Ex : livre de Oded Galor qui évoque que la société humaine qui a évolué depuis des milliers d'années est

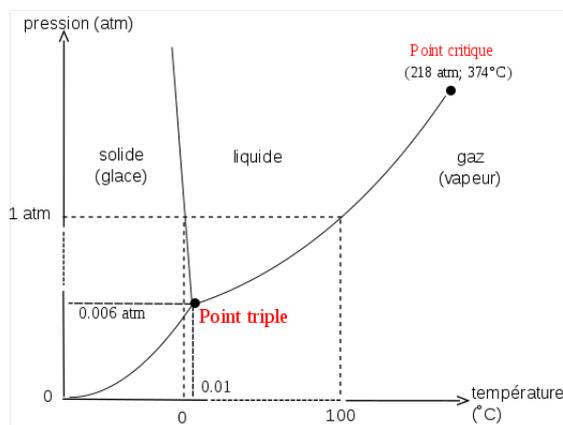
actuellement dans une “transition de phase”, mesurée par la richesse par individus. Article de BRUCE M. BOGHOSIAN, POUR LA SCIENCE N° 507 / Janvier 2020 / p.61, qui évoque que les différentes classes sociales sont comme des phases qui coexistent, aussi mesurée par la richesse par individus.

Dans la formation de l'état solide, les atomes peuvent se trier. Par exemple dans le manteau terrestre, à haute températures, il y a du **magma** qui contient des liquides et gaz dissous. Avec le temps les atomes se trient et s'assemblent pour former par exemple :

- De la **silice**  $\text{SiO}_2$  formant 60% de la croûte terrestre. La température de fusion de la silice est  $T_{\text{fusion}} = 1713\text{C}^\circ$ , la température d'ébullition est  $T_{\text{fusion}} = 2700\text{C}^\circ$  (il fait  $3500\text{C}^\circ$  à la limite supérieure du noyau), et à température ambiante, la silice se cristallise sous forme **quartz** ou se présente sous la forme du **verre** qui est un solide amorphe (structure désordonnée).
- Le **calcaire** ou **carbonate de calcium**  $\text{CaCO}_3$ . A température ambiante, le calcaire se cristallise sous forme **Calcite** qui est un cristal stable, ou Aragonite, cristal instable ( $\tau = 10^8$  ans), bio-synthétisé (**nacre** des coquillages), ou **Vaterite**, etc..
- Les **argiles**, sous forme de feuillets complexes, (début de la vie) ?

### 5.1.2 Transitions de phase

On appelle **transition de phase**, lorsque la matière change de phase. Il peut y avoir coexistence de phases, par exemple gaz et liquide, liquide et solide etc., voici le diagramme de phase de l'eau,  $\text{H}_2\text{O}$  :



L'huile d'olive a une température de solidification, fusion à  $T = 14\text{C}^\circ$  :



## 5.2 Modèle d'Ising du ferromagnétisme

Dans cette section on présente un modèle simple d'interactions entre moments magnétiques disposés sur un réseau en dimension 1 ou 2 ou plus, appelé le **modèle d'Ising**. Ce modèle est un des modèles les plus simples pour étudier les transitions de phase. Il a joué (et continue de jouer) un rôle historique important en science. Voir [Friedli & Velenik \(2017\)](#).

Lorsque la température varie, il y a une variation des propriétés d'aimantation du système, avec une phase aimantée à basse température vers une phase non aimantée à haute température. Cette variation de l'aimantation est continue avec la température mais non dérivable en dimension supérieure ou égale à 2 (sur un réseau infini), où l'aimantation moyenne est strictement nulle si la température est supérieure à une certaine température  $T_c$  appelée **température critique**.

### 5.2.1 Définition du modèle

Soit  $N \geq 1$  entier qui est la taille du réseau. On considère le réseau de <sup>1</sup>  $\Lambda = (\mathbb{Z}/(N\mathbb{Z}))^2$  c'est à dire de taille  $N \times N$  et périodique. Chaque site est noté  $X = (x, y) \in \Lambda$  et possède une variable  $f_X = \pm 1$  qui modélise un moment magnétique ou spin up ou down. Une configuration notée <sup>2</sup>  $f = (f_X)_{X \in \Lambda} \in \{-1, +1\}^\Lambda$  est un choix  $f_X = \pm 1$  pour chaque point  $X \in \Lambda$  du réseau.

*Remarque 5.2.1.* il y a donc  $N^2$  sites et chaque site  $X \in \Lambda$  a deux états possibles  $f_X = \pm 1$ , donc il y a  $2^{N^2}$  configurations possibles. Par exemple pour  $N = 5$ , cela donne  $2^{5^2} \approx 33 \cdot 10^6$  configurations possibles.

Soit  $B \in \mathbb{R}$  donné qui correspond à un champ magnétique extérieur. On note  $X \sim Y$  si  $X, Y$  sont proches voisins (chaque site a 4 proches voisins). L'**énergie** de la configuration  $f$  est définie par

$$E(f) := \sum_X \sum_{Y \text{ tq } X \sim Y} (-f_X \cdot f_Y) + \sum_X B f_X \quad (5.2.1)$$

L'**aimantation de la configuration**  $f$  est définie par

$$M(f) := \frac{1}{N^2} \sum_X f_X \in [-1, +1] \quad (5.2.2)$$

*Remarque 5.2.2.* Pour deux sites voisins,  $X \sim Y$ , si  $f_X = f_Y$  alors leur énergie mutuelles est  $(-f_X \cdot f_Y) = -1$  et sinon  $(-f_X \cdot f_Y) = +1$ . Ainsi si le champ extérieur  $B = 0$  est nul, alors les deux configurations où tous les spins sont égaux sont les deux configurations d'énergie minimale du système, appelé **état fondamental**, correspondent respectivement à

1. On rappelle que  $\mathbb{Z}/(N\mathbb{Z})$  est l'ensemble des entiers  $\mathbb{Z}$  modulo  $N$ . Ainsi  $\mathbb{Z}/(N\mathbb{Z}) = \{0, 1, \dots, N-1\}$  et avec la règle que  $(N-1) + 1 = 0$ .

2. Rappel : en mathématiques, si  $A, B$  sont deux ensembles, on note  $B^A$  l'ensemble des applications  $f : A \rightarrow B$ .

1.  $f_X = +1, \forall X \in \Lambda$ , donnant  $E(f) = \sum_X \sum_Y t_{XY} (-1) = -4N^2$  et  $M_f = +1$ ,
2.  $f_X = -1, \forall X \in \Lambda$ , donnant  $E(f) = -4N^2$  et  $M_f = -1$ .

On suppose que le système est en **contact avec un réservoir à la température  $T$  fixée** (i.e. il échange de l'énergie avec ce réservoir d'énergie). On notera

$$\beta = \frac{1}{kT}.$$

D'après la loi de Boltzmann Gibbs, la **probabilité d'une configuration  $f$**  est

$$p(f) = \frac{1}{Z} e^{-\beta E(f)}. \quad (5.2.3)$$

Ainsi l'**aimantation moyenne** est

$$\langle M \rangle_p = \sum_{f \in \{-1, +1\}^\Lambda} M(f) p(f). \quad (5.2.4)$$

*Remarque 5.2.3.* L'**objectif** est de **calculer cette aimantation moyenne**  $\langle M \rangle_p$  en fonction de la température  $\beta$  et éventuellement du champ extérieur  $B$ . En fait ce calcul exact est impossible en dimension plus grande que 2 et très difficile en dimension 2 (mais facile en dimension 1). La difficulté vient des termes de "couplage"  $f_X \cdot f_Y$  dans l'énergie (5.2.1). Cette énergie apparaît dans l'exponentielle de la probabilité (5.2.3).

Dans la suite on va présenter :

1. une méthode numérique de l'aimantation, appelée méthode de Monte-Carlo
2. une méthode approximative qui consiste à calculer l'aimantation pour une expression de  $p(f)$  simplifiée qui ne présente pas de terme de couplage, appelée méthode du champ moyen,
3. un calcul exact de la température critique  $\beta_c^{\text{exact}} = \frac{1}{4} \ln(1 + \sqrt{2}) \approx 0.22$  donnée par la **méthode de Kramers-Wannier** (1941) qui est un argument de dualité. Il faut savoir que L. Onsager a obtenu en 1944 l'expression exacte de l'énergie libre pour le modèle d'Ising sur  $\mathbb{Z}^2$ , **Friedli & Velenik** (2017, p.132), **Yvan Velenik. Le modèle d'Ising. DEA. 2009. cel-00392289.**

La figure 5.2.1 présente un résultat numérique pour  $B = 0$  (en fait  $B > 0$  très petit, proche de 0 dans le calcul) :

A température assez basse (ou  $\beta \gg 1$ ), on observe un état collectif où les moments magnétiques ont une aimantation totale moyenne non nulle (ici positive), donc une aimantation du matériau. Au dessus d'une certaine température appelée "transition" ou température critique, ici  $\beta < \beta_c \approx 0.2$ , l'aimantation totale moyenne est nulle. On va aussi étudier l'effet d'un champ magnétique extérieur  $B$  sur l'aimantation du matériau.

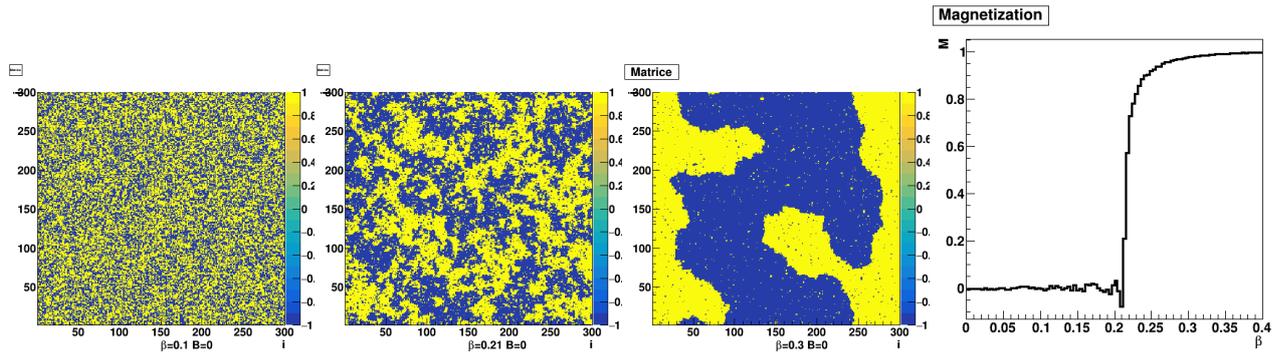


FIGURE 5.2.1 – Résultat numérique de la méthode de Monte-Carlo pour le modèle d’Ising. Configurations prises au hasard pour les paramètres  $\beta = 0.1$  (désordre, aimantation nulle),  $\beta = 0.21$  (transition de phase) et  $\beta = 0.3$  (ordre magnétique, aimantation non nulle). Aimantation  $M = \frac{1}{N^2} \sum_X f_X$  en fonction de  $\beta$ .

## 5.2.2 Dynamique probabiliste de Markov d’un réseau de spins (Monte Carlo)

Dans cette partie on présente un modèle de dynamique du réseau de spin défini précédemment, qui évolue vers l’état d’équilibre de Boltzmann. Cet algorithme permet de calculer numériquement l’aimantation moyenne et d’observer la transition de phase.

### 5.2.2.1 Algorithme de Monte-Carlo

On considère le modèle suivant de dynamique probabiliste des configurations  $f(t)$  du réseau en fonction du temps  $t \in \mathbb{N}$  discret. La configuration  $f(t+1)$  à la date  $t+1$  s’obtient à partir de la configuration  $f(t)$  à la date  $t$  par les règles suivantes. On note  $f = f(t)$ .

1. on choisit un site  $X \in \Lambda$  au hasard (avec probabilité uniforme, donc chaque site  $X$  a la probabilité  $\frac{1}{N^2}$  d’apparaître). On note  $f'$  la configuration identique à  $f$  sauf que le spin du site  $X$  est retournée :

$$f'_Y = \begin{cases} f_Y & \text{si } Y \neq X \\ -f_X & \text{si } Y = X \end{cases}$$

2. Si  $E(f') < E(f)$  alors on pose  $f(t+1) = f'$ .
3. Si  $E(f') \geq E(f)$  alors on choisit  $f(t+1) = f'$  avec la probabilité  $e^{-\beta(E(f')-E(f))}$  et on reste sur l’état  $f$  avec la probabilité complémentaire  $1 - e^{-\beta(E(f')-E(f))}$ .

### 5.2.2.2 Dynamique de Markov

On peut reformuler cet algorithme comme étant une dynamique probabiliste de Markov comme définie en section 1.6.2. Ici l’espace des états est l’espace des configurations possibles  $\{-1, +1\}^\Lambda$  et la matrice  $M = (M_{f',f})_{f',f}$  de transitions  $f \rightarrow f'$  a pour éléments de matrices :

1. si  $f$  et  $f'$  ont au moins deux spins différents alors  $M_{f',f} = 0$
2. sinon, si  $f$  et  $f'$  ont tous les spins identiques sauf en un site  $X$  où  $f_X = -f'_X$  alors
  - Si  $E(f') < E(f)$  alors  $M_{f',f} = \frac{1}{N^2}$ .
  - Si  $E(f') \geq E(f)$  alors  $M_{f',f} = \frac{1}{N^2} e^{-\beta(E(f')-E(f))}$  et à  $M_{f',f}$  on ajoute  $\frac{1}{N^2} (1 - e^{-\beta(E(f')-E(f))})$ .
 La somme de ces deux contributions est  $\frac{1}{N^2}$ .

On vérifie que cette matrice  $M = (M_{f',f})_{f',f}$  est stochastique car à  $f$  fixé,

$$\sum_{f'} M_{f',f} = \sum_{f' \text{ cas 2}} \frac{1}{N^2} = \sum_{X \in \Lambda} \frac{1}{N^2} = 1$$

On démontre en TD que cette matrice est mélangeante et que l'état d'équilibre est la mesure de probabilité de Boltzmann-Gibbs (en fait on a choisit la dynamique pour cela) :

$$p(f) = \frac{1}{Z} e^{-\beta E(f)}.$$

Ainsi d'après le théorème de Perron Frobenius, le système évolue vers cet état d'équilibre avec un écart qui décroît exponentiellement vite, i.e. comme  $e^{-t/\tau}$  avec un temps caractéristique  $\tau(\beta)$  qui dépend de la température.

### 5.2.3 Méthode variationnelle du champ moyen dans le modèle d'Ising (Curie-Weiss)

Référence : [Balian \(2007, Tome 1, p.427, p264,p301\)](#), [Friedli & Velenik \(2017\)](#).

On considère à nouveau le modèle d'Ising (5.2.1) mais en généralisant un petit peu. On suppose donnée une fonction de couplage

$$V : X \in \Lambda \rightarrow \mathbb{R}^+.$$

L'énergie d'une configuration est

$$E(f) := \sum_X \sum_{Y, Y \neq X} (-V(X-Y) f_X \cdot f_Y) + \sum_X B f_X \quad (5.2.5)$$

Dans l'exemple (5.2.1), la fonction couplage est  $V(X-Y) = 1$  si  $X \sim Y$  et  $V(X-Y) = 0$  sinon.

On souhaite calculer l'aimantation moyenne (5.2.4) pour la mesure de probabilité de Boltzmann-Gibbs (5.2.3)  $p(f)$ , mais ce calcul est trop difficile (ou impossible) alors on va utiliser une mesure approchée  $p_{x_0}(f)$  plus simple pour lequel le calcul est possible, et que l'on va d'abord déterminer (optimiser) grâce au principe variationnel s'inspirant de la proposition 3.5.1.

### 5.2.3.1 Famille variationnelle et calcul de l'énergie libre

On va définir une famille de mesures de probabilité  $(p_x)_{x \in \mathbb{R}}$  dépendant d'un paramètre  $x \in \mathbb{R}$  donné. Pour une configuration  $f \in \{-1, +1\}^\Lambda$  et un site  $X \in \Lambda$ , on définit la probabilité

$$\tilde{p}_x(f_X) = \frac{1}{z} e^{xf_X} \quad (5.2.6)$$

et une constante de normalisation

$$z = e^x + e^{-x}.$$

On définit ainsi la mesure de probabilité produit

$$p_x(f) = \prod_{X \in \Lambda} \tilde{p}_x(f_X) \quad (5.2.7)$$

qui signifie l'indépendance statistique des sites entre eux. Pour  $x, T$  fixés (rappel  $\beta = \frac{1}{kT}$ ), on calcule l'énergie moyenne  $\langle E \rangle_{p_x}$ , l'entropie  $S(p_x)$  et on déduit l'énergie libre (voir exercice de TD)

$$\begin{aligned} F_T(x) &:= \langle E \rangle_{p_x} - TS(p_x) \\ &= \dots = N \left( -v \tanh^2(x) + B \tanh x + kT(x \tanh x - \ln 2 - \ln \cosh x) \right) \end{aligned}$$

avec

$$v = \sum_{X, X \neq 0} V(X)$$

et  $N$  est le nombre total de sites.

### 5.2.3.2 Recherche du minimum de $F_T(x)$

Le principe variationnel suggère de chercher la distribution de probabilité  $p_x$ , donc le paramètre  $x \in \mathbb{R}$ , qui minimise l'énergie libre  $F_T(x)$ .

On écrit

$$\frac{dF_T}{dx}(x_0) = 0,$$

et cela donne  $x_0$  solution de

$$\tanh x_0 = \frac{1}{2v} (kT x_0 + B)$$

qui est une équation transcendante. La résolution graphique montre qu'il y a une température critique

$$T_c := \frac{2v}{k},$$

et

- si  $T > T_c$  alors  $x_0(T) = 0$
- si  $T < T_c$  alors il y a deux solutions, qui sont pour  $T$  proche de  $T_c$ ,

$$x_0(T) = \pm \left( \frac{3k}{2v} (T_c - T) \right)^{1/2} \quad (5.2.8)$$

### 5.2.3.3 Aimantation moyenne

Pour le paramètre  $x_0$  obtenu qui minimise l'énergie libre dans la famille variationnelle, on calcule l'aimantation moyenne (5.2.4) pour la mesure de probabilité  $p_{x_0(T)}$  (voir TD) :

$$\begin{aligned} \langle M \rangle_{p_{x_0(T)}} &= \sum_f M(f) p_{x_0(T)}(f) = \dots = \frac{1}{2v} (kT x_0(T) + B) \\ &\stackrel{(5.2.8)}{=} \pm \sqrt{3} \left(1 - \frac{T}{T_c}\right)^{1/2} + \frac{B}{2v} \end{aligned}$$

où la dernière ligne est pour  $T$  proche de  $T_c$ . Pour  $B = 0$ , on obtient l'allure de la figure 5.2.1. Dans le modèle (5.2.1), on a  $v = 4$  donc  $\beta_c = \frac{1}{kT_c} = \frac{1}{2v} = \frac{1}{8} = 0.125$ .

### 5.2.3.4 Interprétation de l'approximation comme un champ moyen

La famille variationnelle de mesures de probabilité choisie (5.2.7) a la particularité d'être un produit de mesures sur chaque site, signifiant que l'on ne considère pas les corrélations entre sites. D'après la remarque 5.2.3 il y a des termes de corrélations dans la mesure de Boltzmann à cause des termes de couplages dans l'énergie. Ces **effets de corrélations sont donc simplement ignorés** (en fait ils ne sont pas négligeables) dans cette approche variationnelle.

Mais c'est cela qui a permis d'effectuer le calcul. Précisément, dans le calcul de l'énergie libre on avait un terme

$$\left\langle \sum_{X,Y} f_X f_Y V(X-Y) \right\rangle = \sum_{X,Y} \langle f_X \rangle \langle f_Y \rangle V(X-Y) = \sum_X \langle f_X \rangle \underbrace{\left( \sum_Y \langle f_Y \rangle V(X-Y) \right)}_{B_X}$$

qui s'interprète que l'interaction d'un site  $X$  avec les autres sites  $Y$  agit comme un champ magnétique effectif  $B_X$  obtenu par la contribution des moyennes  $\langle f_Y \rangle$  des autres sites voisins. Pour cette raison, cette approximation s'appelle aussi **approximation du champ moyen**.

## 5.2.4 Dualité de Kramers-Wannier pour le modèle d'Ising 2D (1940)

Nous présentons ici l'argument de symétrie de Kramers-Wannier de 1940, qui permet de calculer la valeur exacte de la température critique pour le réseau carré  $\mathbb{Z}^2$  pour  $B = 0$  :

$$\beta_c^{\text{exact}} = \frac{1}{4} \ln \left(1 + \sqrt{2}\right) \approx 0.22..$$

Ensuite en 1944, Onsager a effectué le calcul exact de l'énergie libre de ce même modèle, ce qui permet de déduire toutes les grandeurs thermodynamiques.

Pour le modèle d'Ising dont l'énergie d'une configuration est  $f$

$$E(f) = - \sum_{X \sim Y} f_X f_Y$$

Voici les étapes de l'argument de Kramers-Wannier (voir détails en TD).

— La mesure de probabilité de Boltzmann est

$$p(f) = \frac{1}{Z} e^{-\beta E(f)}$$

avec  $\beta = \frac{1}{kT}$  et la constante de normalisation

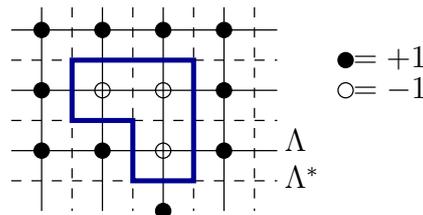
$$Z(\beta) = \sum_f e^{-\beta E(f)}.$$

— La température critique recherchée  $\beta_c$  est une valeur de  $\beta$  donnant des grandeurs thermodynamique singulières (non dérivabilité, divergente, etc) donc pour laquelle **la fonction  $\beta \rightarrow Z(\beta)$  est non analytique** (i.e. n'admet pas une expression en série entière convergente). Plus précisément, par définition une transition de phase est dite :

— **transition de phase du premier ordre** si  $Z(\beta)$  n'est pas continue (pas  $C^0$ ) en  $\beta_c$ .

— **transition de phase du deuxième ordre** si  $Z(\beta)$  est continue mais pas différentiable (pas  $C^1$ ) en  $\beta_c$ . etc

— On considère le réseau carré fini  $\Lambda = (\mathbb{Z}/(N\mathbb{Z}))^2$ . A partir de l'observation qu'une configuration  $f$  (ici points noirs/blancs) est uniquement associée à un ensemble d'arêtes sur le réseau dual noté  $E \in \mathcal{E}(\Lambda^*)$ , formant des courbes fermées (ici en bleu) et représentant **les parois de domaines d'aimantation** :



on déduit deux expressions différentes pour la fonction  $\beta \rightarrow Z(\beta)$ , en utilisant le fait que **le réseau carré est self-dual**.

1. Une série appelée « High Temperature » car convergente dans la limite thermodynamique  $N \rightarrow +\infty$ , pour  $\beta$  assez petit, notée

$$Z^{HT}(\beta) = \lim_{N \rightarrow +\infty} (\cosh \beta)^{|\mathbb{E}(\Lambda)|} 2^{|\Lambda|} \sum_{E \in \mathcal{E}(\Lambda^*)} (\tanh \beta)^{|E|},$$

où  $\mathbb{E}(\Lambda)$  est l'ensemble des arêtes et  $|\mathbb{E}(\Lambda)|$  le cardinal.

2. Une série appelée « Low Temperature » car convergente pour  $\beta$  assez grand est notée

$$Z^{LT}(\beta) = \lim_{N \rightarrow +\infty} e^{\beta |\mathbb{E}(\Lambda)|} \sum_{E \in \mathcal{E}(\Lambda^*)} (e^{-2\beta})^{|E|}.$$

On suppose que ces séries divergent à une certaine valeur  $\beta_c$  que l'on suppose unique. Cela implique

$$\tanh \beta_c = e^{-2\beta_c},$$

donnant

$$\beta_c = \frac{1}{2} \ln(1 + \sqrt{2}) \approx 0.44$$

## 5.3 Modèle BCS : transition de phase en supraconductivité

### 5.3.1 Introduction

Le phénomène de **super conductivité** est un phénomène physique très surprenant : en dessous d'une certaine température  $T_c$  appelée température critique et généralement très basse, certains métaux n'ont plus aucune résistance aux courants électriques (circulation de charges sans aucune dissipation). Ces propriétés font que les matériaux supraconducteurs sont très utilisés en technologie, pour créer de **forts champs magnétique**, pour l'**effet Josephson**, etc.

D'un point de vue microscopique, il apparaît que ces propriétés découlent de la présence d'un **gap en énergie**  $E_0$  au dessus de l'état fondamental du système constitué par le gaz d'électrons pour des température assez basse  $T < T_c$ . Il y a une transition de phase à  $T = T_c$  et l'état supraconducteur disparaît pour  $T > T_c$ . Voir tableau 5.1.

Elément	T° critique(K) $T_c$	gap (K) $\frac{E_0}{k}$	rapport $\frac{E_0}{kT_c}$	Debye (K) $\Theta_D = \frac{E_D}{k}$
Al	1.196	4.06	3.4	394
Zn	0.875	2.8	3.2	234
Sn	3.72	13.02	3.5	170
Pb	7.19	30.917	4.3	88

TABLE 5.1 – Exemples de matériaux supraconducteur, température critique  $T_c$  en Kelvin (K), énergie du gap  $E_0$  qui sépare l'état fondamental des états excités en K, énergie de Debye des vibrations du réseau en K.

Dans cette section, on verra que le modèle BCS prédit le rapport universel  $\frac{E_0}{kT_c} = \frac{2\pi}{e^\gamma} \approx 3.53$  bien observé ici (avec la constante de Euler  $\gamma := \lim_{n \rightarrow \infty} \sum_{k=1}^n \frac{1}{k} - \ln n \approx 0.57$ ). L'existence de ce gap se manifeste dans le fait que la capacité calorifique s'annule très

vite pour  $T \rightarrow 0$  (idem modèle d'Einstein) et dans des expériences d'absorption d'ondes électromagnétiques.

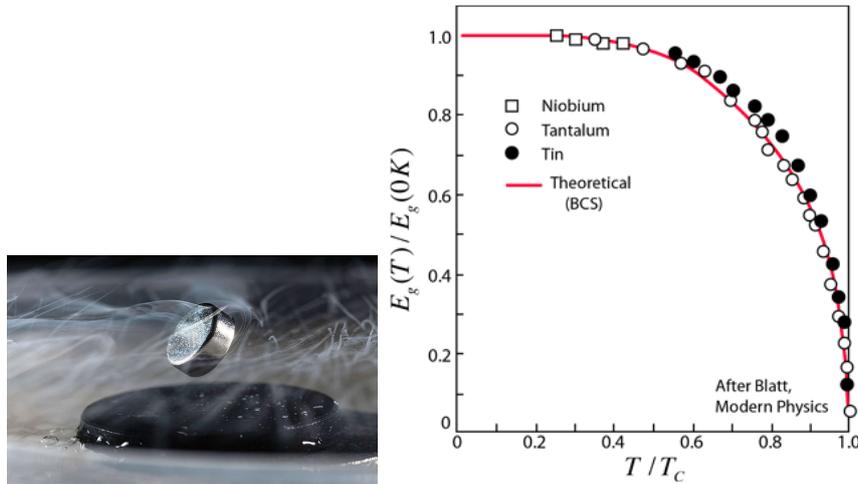


FIGURE 5.3.1 – (a) Lévitation d'un aimant à basse température, il n'y a pas de dissipation grâce à l'état supraconducteur. (b) Gap en énergie dans différents supraconducteurs, comparé à la prédiction BCS  $\Delta(T)$ , eq.(5.3.26).

**Histoire :** (voir [Combescot & Shiau \(2015\)](#), chap.6). L'effet supraconducteur a été découvert expérimentalement en 1911 par [Heike Kamerlingh Onnes](#). Ensuite Meissner et Ochsenfeld 1933 ont observé que le champ magnétique est expulsé du matériau supraconducteur. London 1935, puis Ginzburg et Landau 1950, ont proposé un modèle phénoménologique. En 1950, Fröhlich a suggéré que les électrons peuvent s'apparier en paires suite à une force attractive effective entre eux, via le réseau des atomes.

Le premier modèle microscopique démontrant l'existence d'un gap spectral au-dessus de l'état fondamental a été proposé en 1957 par Bardeen, Cooper, Schrieffer appelé **modèle BCS**. Les calculs ont été améliorés par Bogoliubov 1958, Gorkov 1958.

Dans cette section on présente ce modèle. L'effet Josephson a été prédit en 1962 à partir de la théorie BCS.

*Remarque 5.3.1.* Ce modèle BCS est basé sur l'appariement des électrons en paires d'électrons ayant des impulsions opposées ( $p, -p$ ) et formant un état stable collectif appelé **condensat de Bose**. Ce phénomène d'appariement se retrouve dans d'autres domaines de la physique, par exemple pour expliquer la **superfluidité** de l'He3, ou la stabilité de certains noyaux nucléaires (**appariement de paires de nucléons**), les **condensats d'atomes froids**, etc, voir [Martin & Rothen \(1990\)](#).

Ce qui suit est basé sur le livre [Couture & Zitoun \(1992\)](#), chap 15.3. Voir aussi [Combescot & Shiau \(2015\)](#), chap.6, chap.9.

### 5.3.2 Modèle d'un métal comme gaz d'électrons libres

Comme dans la section 4.5.2, on considère au départ les électrons libres dans un métal comme un gaz de  $N$  électrons libres dans un volume  $V$ . A basse température, on a vu que dans l'espace des impulsions d'un électron  $p \in \mathbb{R}^3$ , les états à un électron occupés sont  $p \in \mathbb{R}^3$  tel que

$$E_{\text{cin}}(p) = \frac{|p|^2}{2m} \leq \mu \quad (5.3.1)$$

l'intérieur de la « sphère de Fermi ». En appliquant la loi de Weyl, on a déduit que le niveau de Fermi  $\mu$  est donné par (4.5.1) :

$$\mu = \frac{\hbar^2}{2m} \left( 3\pi^2 \frac{N}{V} \right)^{2/3}$$

Ainsi l'état fondamental du gaz d'électrons libres est la configuration où tous les états à un électron dans la sphère de Fermi sont occupés par un électron et les états hors de la sphère sont inoccupés.

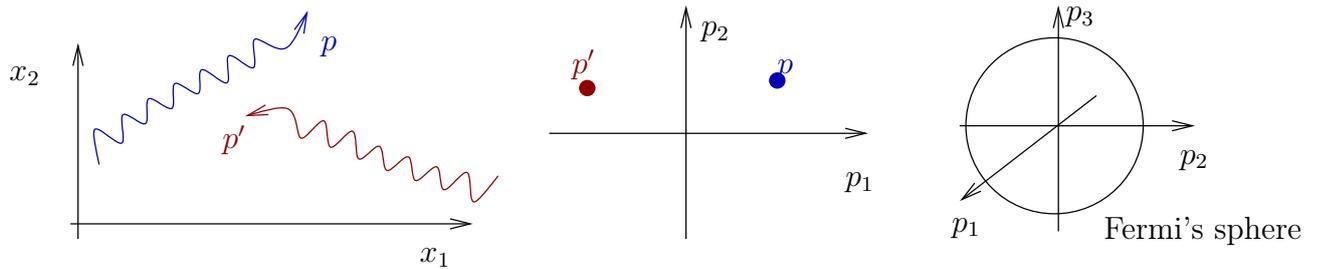


FIGURE 5.3.2 – Représentation de deux ondes planes ou modes de Fourier  $\psi_p(x) = e^{ipx/\hbar}$  dans l'espace réel où les fonctions sont délocalisées et dans l'espace de Fourier (ou espace réciproque) où ce sont des points.

### 5.3.3 Base de l'espace de Fock Fermionique et opérateurs de création/annihilation

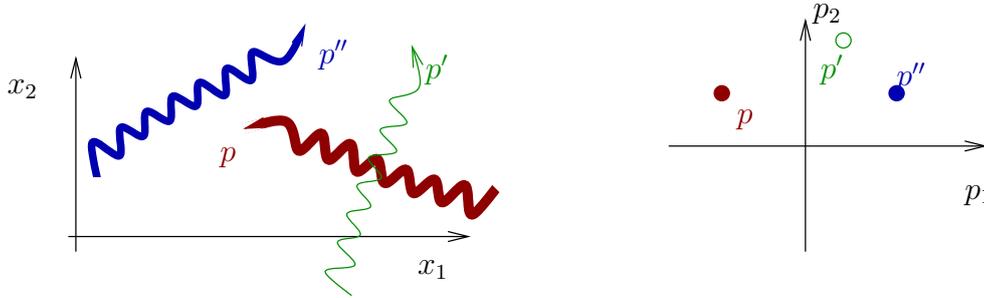
Dans ce paragraphe on revient sur la description de l'espace Fermionique  $\mathcal{F}$  des électrons défini en section 4.2 et on reformule l'état fondamental Faure (2022, chap 5.3) à l'aide d'opérateurs dans cet espace  $\mathcal{F}$ .

Les électrons sont des Fermions, cela signifie que le nombre d'électron dans chaque état d'impulsion  $p \in \mathbb{R}^3$  décrivant une particule est  $N(p) \in \{0, 1\}$ . On a vu qu'une configuration du système à plusieurs particules est caractérisée par la donnée simultanée de toutes les valeurs d'occupation que l'on notera

$$N(\cdot) := (N(p))_{p \in \mathbb{R}^3}.$$

Autrement dit  $N(\cdot)$  est une **fonction**  $N(\cdot) : p \rightarrow N(p)$  aussi appelé **champs classique en physique**.

**Exemple graphique :** On considère 3 ondes plane  $p, p', p''$ , et  $p, p''$  sont occupées par un électron,  $p'$  n'est pas occupé. L'état est donc noté  $N(\cdot) = |1, 0, 1\rangle$  :



En mécanique quantique, d'après le principe de superposition, les différentes configurations possibles  $N(\cdot)$  forment une base orthonormée de l'espace vectoriel quantique  $\mathcal{F}$  des états possibles, appelé **espace de Fock Fermionique**. On notera  $|N(\cdot)\rangle$  le vecteur normalisé de base associé au champs classique  $N(\cdot)$ . Ainsi un vecteur quelconque de  $\mathcal{F}$  est une superposition (ou combinaison linéaire) de ces états de base :

$$\psi = \sum_{N(\cdot)} C_{N(\cdot)} |N(\cdot)\rangle \in \mathcal{F}, \quad C_{N(\cdot)} \in \mathbb{C}.$$

On appelle  $\psi$  un **champs quantique**. Les vecteurs de base  $|N(\cdot)\rangle$  sont des champs quantiques particuliers. En particulier parmi ces vecteurs de base il y a le champs  $N(p) = 0, \forall p$ , où tous les états sont vides. On l'appelle **état du vide**, noté  $|0\rangle$ . Ne pas confondre le vecteur nul 0 avec l'état du vide  $|0\rangle = |0, \dots, 0\rangle$  qui est un vecteur de base de  $\mathcal{F}$  de norme 1.

Dans le modèle des électrons libres et indépendants, l'énergie est formée seulement par l'énergie cinétique et chaque configuration  $|N(\cdot)\rangle$  a une énergie qui est la somme des énergie cinétiques individuelles  $E_{\text{cin}}(p) = \frac{|p|^2}{2m}$  :

$$E_{N(\cdot)} = \int_{\mathbb{R}^3} N(p) E_{\text{cin}}(p) d^3p.$$

On peut décrire cela par un **opérateur énergie cinétique**  $\hat{H}_{\text{cin}}$ . dans cet espace  $\mathcal{F}$  définit par le fait que les vecteurs de base  $|N(\cdot)\rangle$  sont des vecteurs propres, de valeur propre  $E_{N(\cdot)}$  :

$$\hat{H}_{\text{cin}} |N(\cdot)\rangle = E_{N(\cdot)} |N(\cdot)\rangle$$

De même le nombre total d'électron dans l'état  $|N(\cdot)\rangle$  est

$$N_{N(\cdot)} = \int_{\mathbb{R}^3} N(p) d^3p$$

et on peut définir l'opérateur nombre de particules  $\hat{N}$  défini par :

$$\hat{N}|N(\cdot)\rangle = N_{N(\cdot)}|N(\cdot)\rangle$$

Pour une valeur  $\mu \in \mathbb{R}$  donnée du potentiel chimique, on définit alors l'opérateur appelé **Hamiltonien grand canonique**

$$\hat{H}_{GC} := \hat{H}_{\text{cin.}} - \mu\hat{N}$$

et on a

$$\hat{H}_{GC}|N(\cdot)\rangle = \epsilon'_{N(\cdot)}|N(\cdot)\rangle$$

avec la valeur propre

$$\epsilon'_{N(\cdot)} := (E_{N(\cdot)} - \mu N_{N(\cdot)}) = \int_{\mathbb{R}^3} N(p) \epsilon'(p) d^3p,$$

$$\epsilon'(p) := E_{\text{cin}}(p) - \mu \quad (5.3.2)$$

On rappelle que  $N(p) \in \{0, 1\}$ . L'état fondamental de  $\hat{H}_{GC}$ , i.e. le vecteur propre  $|N(\cdot)\rangle$  de valeur propre minimale, est celui pour lequel

$$N(p) = 1 \text{ si } \epsilon'(p) < 0 \Leftrightarrow E_{\text{cin}}(p) < \mu$$

$$N(p) = 0 \text{ si } \epsilon'(p) > 0 \Leftrightarrow E_{\text{cin}}(p) > \mu$$

correspondant donc à la sphère de Fermi (5.3.1) des états occupés.

*Remarque 5.3.2.* Jusqu'à présent, dans ce cours, on n'avait pas insisté sur cette vision d'espace de Fock car jusqu'alors on a considéré des électrons libres et il n'était pas nécessaire de considérer ces combinaisons linéaires et ces opérateurs dans l'espace de Hilbert  $\mathcal{F}$ . Dans la suite, on va considérer des interactions entre les électrons et il deviendra nécessaire d'utiliser ce formalisme pour décrire les états propres du Hamiltonien  $\hat{H}$  avec interactions.

On introduit maintenant d'autres opérateurs qui n'ont pas directement de sens physique mais qui seront utile pour décrire l'opérateur Hamiltonien et les termes d'interaction qu'il contient.

**Définition 5.3.3.** Pour une impulsion  $p \in \mathbb{R}^3$  fixée,

- l'opérateur de création  $a_p^\dagger : \mathcal{F} \rightarrow \mathcal{F}$  est défini par son action sur un vecteur de base (voir schéma ci-dessous)

$$\begin{aligned} \text{si } N(p) = 0, \quad a_p^\dagger |N(\cdot)\rangle &= |N'(\cdot)\rangle \text{ avec } N'(p') = \begin{cases} N(p') & \text{pour } p' \neq p \\ 1 & \text{pour } p' = p \end{cases} \\ \text{si } N(p) = 1, \quad a_p^\dagger |N(\cdot)\rangle &= 0. \end{aligned}$$

Autrement dit, cet opérateur  $a_p^\dagger$  fait apparaître un électron dans l'état  $p$  si il n'y en avait pas et donne le vecteur nul sinon.

- De même on définit l'opérateur d'annihilation  $a_p : \mathcal{F} \rightarrow \mathcal{F}$  par

$$\begin{aligned} \text{si } N(p) = 1, \quad a_p |N(\cdot)\rangle &= |N'(\cdot)\rangle \text{ avec } N'(p') = \begin{cases} N(p') & \text{pour } p' \neq p \\ 0 & \text{pour } p' = p \end{cases} \\ \text{si } N(p) = 0, \quad a_p |N(\cdot)\rangle &= 0. \end{aligned}$$

Autrement dit, cet opérateur  $a_p$  fait disparaître un électron dans l'état  $p$  si il y en avait un et donne le vecteur nul sinon.

**Exemple 5.3.4.** quelques exemples d'action des opérateurs de création sur les vecteurs de base  $|N(\cdot)\rangle = |N_1, N_2\rangle$  avec  $N_1, N_2 \in \{0, 1\}$ . C'est un espace de dimension  $2 \times 2 = 4$  et 0 désigne le vecteur nul.

$ N(\cdot)\rangle$	$a_1^\dagger  N(\cdot)\rangle$	$a_1  N(\cdot)\rangle$	$a_2 a_1^\dagger  N(\cdot)\rangle$	$(a_1^\dagger a_1 + a_2^\dagger a_2)  N(\cdot)\rangle$
$ 0, 1\rangle$	$ 1, 1\rangle$	0	$ 1, 0\rangle$	$1 \times  0, 1\rangle$
$ 1, 0\rangle$	0	$ 0, 0\rangle$	0	$1 \times  1, 0\rangle$
$ 0, 0\rangle$	$ 1, 0\rangle$	0	0	$0 = 0 \times  0, 0\rangle$
$ 1, 1\rangle$	0	$ 0, 1\rangle$	0	$2 \times  1, 1\rangle$

**Lemme 5.3.5.** *L'opérateur  $a_p^\dagger$  est l'adjoint de  $a_p$  c'est à dire que*

$$\langle u | a_p^\dagger v \rangle = \langle a_p u | v \rangle, \forall u, v \in \mathcal{F}.$$

*On a les expressions*

$$\hat{H}_{\text{cin.}} = \int_{\mathbb{R}^3} E_{\text{cin}}(p) a_p^\dagger a_p d^3 p \quad (5.3.3)$$

$$\hat{N} = \int_{\mathbb{R}^3} a_p^\dagger a_p d^3 p$$

$$\hat{H}_{GC} = \int_{\mathbb{R}^3} \epsilon'(p) a_p^\dagger a_p d^3 p$$

*L'état fondamental de  $\hat{H}_{GC}$  dans  $\mathcal{F}$  s'écrit*

$$|\psi_{\text{fond.}}\rangle = \left( \prod_{p \text{ t.q. } E_{\text{cin}}(p) < \mu} a_p^\dagger \right) |0\rangle$$

*où  $|0\rangle$  est l'état  $|N(\cdot)\rangle$  tel que  $N(p) = 0, \forall p$ , appelé **état du vide**.*

*Démonstration.* Il suffit de vérifier ces relations sur les vecteurs de base  $|N(\cdot)\rangle$ . Par exemple on observe que

$$a_p^\dagger a_p |N(\cdot)\rangle = N(p) |N(\cdot)\rangle.$$

□

*Remarque 5.3.6.* Le formalisme qui consiste à utiliser des opérateurs de création/annihilation dans l'espace de Fock s'appelle “**quantum many-body problem**”, “**second quantization**”, “**quantum field theory**” et est à la base de la physique (quantique) depuis les années 1930. Il y a de nombreux ouvrages conseillés, [Martin & Rothen \(1990\)](#), [Peskin & Schroeder \(1995\)](#), [Blaizot & Ripka \(1986\)](#), [Negele & Orland \(1998\)](#), etc.

### 5.3.4 Interactions entre électrons et Hamiltonien B.C.S.

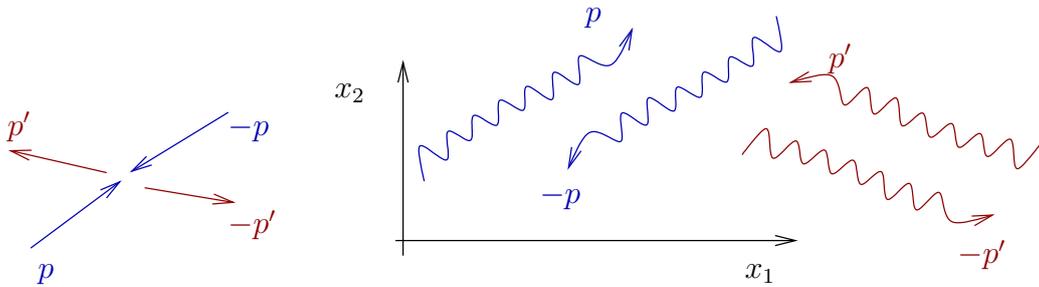
On va s'intéresser maintenant à l'interaction entre deux électrons. On ne mentionne pas le spin bien qu'il soit important. Les paires d'électrons considérées sont dans l'état singulet (spin total nul). On considère une interaction (i.e. collision) entre deux électrons ayant des impulsions respectives  $p_1, p_2 \in \mathbb{R}^3$  avant la collision et  $p'_1, p'_2 \in \mathbb{R}^3$  après la collision.

**Paires de Cooper :** On souhaite que l'impulsion totale soit nulle. D'après la conservation de l'impulsion totale on a

$$0 = p_1 + p_2 = p'_1 + p'_2.$$

Pour simplifier on notera  $p = p_1$  et  $p' = p'_1$ . Ainsi avant collision on a une paire d'électrons d'impulsions  $(p, -p)$  devenant  $(p', -p')$  après collision, avec  $p, p' \in \mathbb{R}^3$  arbitraires. **Ces paires d'électron qui interagissent (ou collisionnent) sont appelées *paire de Cooper*.**

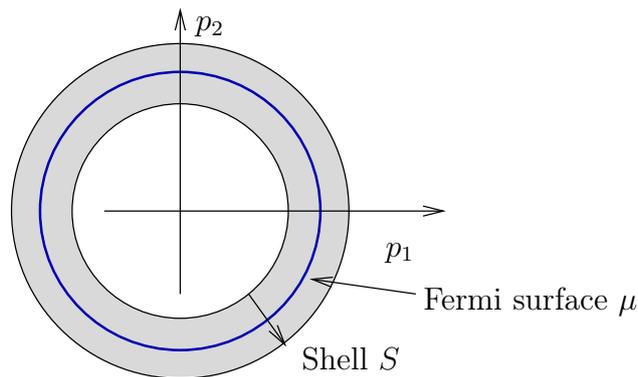
Voici deux représentations schématiques possibles dans l'espace réel :



**Couche d'interaction de DeBye :** L'interaction entre les électrons ne provient pas de la répulsion électrostatique qui est écrantée par les atomes, mais provient d'un couplage avec les déformations du réseau cristallin (ondes de vibrations, phonons). Seuls les électrons près de la surface de Fermi contribuent et sont perturbés par les vibrations du réseau d'atome dont l'énergie est de l'ordre de  $E_D$  (énergie de Debye), on note  $\epsilon'(p) := E_{\text{cin}}(p) - \mu$  l'écart à l'énergie de Fermi comme en (5.3.2), et on va considérer dans la suite les électrons dans cette couche (Shell) notée :

$$S := \{p \in \mathbb{R}^3, \quad |\epsilon'(p)| < E_D\}.$$

Par exemple pour l'étain  $\Theta_D = \frac{E_D}{k} = 200K$ .



Ce couplage induit une force effective attractive entre les électrons (Fröhlich 1954), modélisée dans la suite par l'énergie potentielle

$$-V_0 < 0.$$

**Expression de l'interaction avec des opérateurs :** De la même façon que l'énergie cinétique a été exprimée par  $a_p^\dagger a_p$  dans (5.3.3), il nous sera utile, pour décrire ces processus de transition en terme d'opérateurs, d'utiliser les **opérateurs de création et annihilation de paires de Cooper**, pour  $p \in \mathbb{R}^3$  donné, définis par :

$$b_p^\dagger := a_p^\dagger a_{-p}^\dagger \quad : \text{opérateur création de paire } (p, -p)$$

$$b_p := a_p a_{-p} \quad : \text{opérateur annihilation de paire } (p, -p)$$

**Définition 5.3.7.** Le modèle BCS est le Hamiltonien suivant dans l'espace de Fock  $\mathcal{F}$  :

$$\hat{H}_{BCS} = \hat{H}_{\text{cin.}} + \hat{H}_{\text{inter.}} - \mu \hat{N} \quad (5.3.4)$$

avec le terme d'interaction

$$\hat{H}_{\text{inter.}} := -V_0 \int_{p,p' \in S} b_p^\dagger b_p d^3 p d^3 p' \quad (5.3.5)$$

**Objectif :** A basse température, l'état du gaz d'électrons en interaction correspond à l'état fondamental de l'opérateur  $\hat{H}_{BCS}$ . A plus haute température il est dans la superposition statistique des états d'énergie plus élevés (les états excités)  $E$  selon la mesure de probabilité de Boltzmann Gibbs  $p = \frac{1}{Z} e^{-\beta E}$ . Il faudrait donc trouver ces états propres d'énergie pour ensuite déduire des propriétés thermodynamiques du gaz d'électron en calculant des valeurs moyenne d'observables. La difficulté est que le terme d'interaction  $\hat{H}_{\text{inter.}}$  rend ce calcul très difficile. Dans la section suivante on va plutôt faire un **calcul approché du spectre** en utilisant la "méthode variationnelle de Hartree-Fock-Bogoliubov" ou "**approximation du champ moyen**".

Le calcul exact du spectre de  $\hat{H}_{BCS}$  a été réalisé par Richardson et Gaudin dans les années 60',70' ref : [Combescot & Shiau \(2015, chap.10\)](#).

### 5.3.5 Approximation du champs moyen, méthode variationnelle de Hartree-Fock-Bogoliubov

Dans ce paragraphe on décrit l'approximation variationnelle de façon directe et efficace par la méthode de Gorkov, qui évite le formalisme abstrait de l'espace de Fock, voir [Combescot & Shiau \(2015, chap.9\)](#). Une description équivalente plus formelle sera faite dans une section ultérieure et permettra de mettre en valeur le fonctionnement général de la méthode variationnelle.

### 5.3.5.1 Opérateur approché et apparition d'un gap spectral

Ce sont les termes  $b_{p'}^\dagger, b_p$  dans (5.3.5) qui rendent difficile le calcul du spectre de  $\hat{H}_{BCS}$ . On considère à la place l'opérateur

$$\left(b_{p'}^\dagger b_p\right)^{(\text{approx})} := b_{p'}^\dagger b_p - \left(b_{p'}^\dagger - \langle b_{p'}^\dagger \rangle\right) (b_p - \langle b_p \rangle) \quad (5.3.6)$$

$$= b_{p'}^\dagger \langle b_p \rangle + \langle b_{p'}^\dagger \rangle b_p - \langle b_{p'}^\dagger \rangle \langle b_p \rangle \quad (5.3.7)$$

où  $\langle b_p \rangle, \langle b_{p'}^\dagger \rangle \in \mathbb{C}$  sont des nombres (paramètres) que l'on choisira plus tard en imposant qu'ils approchent au mieux les valeurs moyennes de  $b_p, b_{p'}^\dagger$  dans l'état fondamental. Le terme correctif  $\left(b_{p'}^\dagger - \langle b_{p'}^\dagger \rangle\right) (b_p - \langle b_p \rangle)$  s'interprète comme des "fluctuations quantiques" que l'on soustrait du problème original. Au lieu de  $\hat{H}_{BCS}$ , eq.(5.3.4), on considère l'opérateur d'approximation

$$\hat{H}_{BCS}^{(\text{approx})} := \hat{H}_{\text{cin.}} + \hat{H}_{\text{inter.}}^{(\text{approx})} - \mu \hat{N} \quad (5.3.8)$$

$$\hat{H}_{\text{inter.}}^{(\text{approx})} := -V_0 \int_{p, p' \in S} \left(b_{p'}^\dagger b_p\right)^{(\text{approx})} d^3 p' d^3 p$$

L'avantage de cette approximation est que le spectre de l'opérateur  $\hat{H}_{BCS}^{(\text{approx})}$  dans l'espace de Fock  $\mathcal{F}$  sera calculable explicitement, car il n'y a plus d'opérateurs de transition  $b_{p'}^\dagger b_p$ , mais seulement des opérateurs  $b_p^\dagger$  ou  $b_p$  qui créent/annihile une paire de Cooper. On va pouvoir considérer chaque paire de Cooper individuellement. Introduisons donc l'espace de Fock  $\mathcal{F}_{(p, -p)}$  d'une paire de Cooper.

**Proposition 5.3.8.** *Pour une impulsion  $p \in \mathbb{R}^3$  donnée, on considère la paire de Cooper  $(p, -p)$  est son espace de Fock associé noté  $\mathcal{F}_{(p, -p)}$  **de dimension 4** dont les vecteurs de base sont  $|N(p), N(-p)\rangle$ , avec  $N(p), N(-p) \in \{0, 1\}$ . Cela donne 4 vecteurs de base notés*

$$|1, 0\rangle, |0, 1\rangle, |0, 0\rangle, |1, 1\rangle. \quad (5.3.9)$$

*L'espace de Fock total  $\mathcal{F}$  est le produit tensoriel antisymétrique de tous ces espaces*

$$\begin{aligned} \mathcal{F} &= \mathcal{F}_{(p, -p)} \wedge \mathcal{F}_{(p', -p')} \wedge \dots \\ &= \bigwedge_{p \in D} \mathcal{F}_{(p, -p)}. \end{aligned} \quad (5.3.10)$$

*où  $D = \mathbb{R}^3 / \sim$  est le **demi espace** avec l'identification de parité  $p \sim (-p)$ , afin de ne considérer que les paires.*



**Interprétation :**

- Remarquer que  $\Delta = 0$  correspond au modèle sans interaction.
- Si  $\epsilon'_p = E_{\text{cin}}(p) - \mu > 0 \Leftrightarrow E_{\text{cin}}(p) > \mu$ , cela signifie que  $p$  est hors de la sphère de Fermi.
  - Dans le cas  $\Delta = 0$ , l'état fondamental est  $|0, 0\rangle$  sans électron, comme attendu.
  - Dans le cas  $\Delta \neq 0$ , l'état fondamental est un mélange  $|0, 0\rangle + a|1, 1\rangle$  avec  $a \neq 0$ . (paire est partiellement vide/occupée).
- Si  $\epsilon'_p = E_{\text{cin}}(p) - \mu < 0 \Leftrightarrow E_{\text{cin}}(p) < \mu$ , cela signifie que  $p$  est dans la sphère de Fermi.
  - Dans le cas  $\Delta = 0$ , l'état fondamental est  $|1, 1\rangle$  occupé par des électrons, comme attendu.
  - Dans le cas  $\Delta \neq 0$ , l'état fondamental est un mélange  $|1, 1\rangle + b|0, 0\rangle$  avec  $b \neq 0$ . (paire est partiellement occupée/vide).
- La **règle de conservation du nombre d'électrons** interdit les transitions de l'état fondamental  $E_-$  ayant nombre pair de particules, vers les intermédiaires (5.3.12) à nombre impair de particules. Mais les transition de  $E_-$  vers  $E_+$  sont permises. Pour cette raison, on observe expérimentalement un **gap spectral** qui est l'écart

$$|E_+ - E_-| := 2\delta_p > 2|\Delta|$$

dont l'expression en fonction de la température  $\Delta(T)$  sera obtenu en (5.3.26) et correspondra à la figure 5.3.1 expérimentale.

*Démonstration.* On a

$$\begin{aligned} \hat{H}_{BCS}^{(\text{approx})} &= \hat{H}_{\text{cin.}} - \mu \hat{N} + \hat{H}_{\text{inter.}}^{(\text{approx})} \\ &= \int_{p \in \mathbb{R}^3} \epsilon'_p a_p^\dagger a_p d^3p - V_0 \int_{p, p' \in S} \left( b_{p'}^\dagger \langle b_p \rangle + \langle b_{p'}^\dagger \rangle b_p - \langle b_{p'}^\dagger \rangle \langle b_p \rangle \right) d^3p' d^3p \\ &= \int_{p \in S} \hat{H}_{BCS}^{(\text{approx})}(p) d^3p + C \text{Id} \end{aligned} \quad (5.3.17)$$

avec l'opérateur spécifique à la paire  $(p - p)$  :

$$\hat{H}_{BCS}^{(\text{approx})}(p) = \epsilon'_p a_p^\dagger a_p + \epsilon'_p a_{-p}^\dagger a_{-p} + \Delta b_p^\dagger + \bar{\Delta} b_p \quad (5.3.18)$$

et la constante

$$C = - \int_{p \in S} \bar{\Delta} \langle b_p \rangle d^3p.$$

On calcule,

$$a_p^\dagger a_p |1, 0\rangle = |1, 0\rangle, \quad b_p^\dagger |1, 0\rangle = a_p^\dagger a_{-p}^\dagger |1, 0\rangle = 0, \quad b_p |1, 0\rangle = a_p a_{-p} |1, 0\rangle = 0$$

donc

$$\hat{H}_{BCS}^{(\text{approx})}(p) |1, 0\rangle = \epsilon'_p |1, 0\rangle$$

De même

$$\begin{aligned}\hat{H}_{BCS}^{(\text{approx})}(p)|0,1\rangle &= \epsilon'_p|0,1\rangle, & b_p^\dagger|0,0\rangle &= a_p^\dagger a_{-p}^\dagger|0,0\rangle = |1,1\rangle \\ b_p|0,0\rangle &= a_p a_{-p}|0,0\rangle = 0, & b_p^\dagger|1,1\rangle &= a_p^\dagger a_{-p}^\dagger|1,1\rangle = 0 \\ b_p|1,1\rangle &= a_p a_{-p}|1,1\rangle = |0,0\rangle, & \hat{H}_{BCS}^{(\text{approx})}(p)|0,0\rangle &= \Delta|1,1\rangle \\ \hat{H}_{BCS}^{(\text{approx})}(p)|1,1\rangle &= 2\epsilon'_p|1,1\rangle + \bar{\Delta}|0,0\rangle\end{aligned}$$

On déduit l'expression (5.3.14). Cette matrice contient une sous matrice  $2 \times 2$  qu'il faut diagonaliser. On obtient (5.3.16).  $\square$

### 5.3.5.2 Optimisation

Dans le calcul précédent, les paramètres complexes  $\langle b_p \rangle, \langle b_{p'}^\dagger \rangle \in \mathbb{C}$  sont arbitraires. Pour justifier l'approximation (5.3.6), on impose qu'ils coïncident avec les valeurs moyennes de  $b_p, b_{p'}^\dagger$  pour l'état de Boltzmann Gibbs de l'Hamiltonien approché  $\hat{H}_{BCS}^{(\text{approx})}$ . Avant cela rappelons les propriétés générales suivantes

**Proposition 5.3.10.** *En mécanique quantique, si une grandeur physique  $A$  est représentée par l'opérateur autoadjoint  $\hat{A}$  (appelée observable) qui a des valeurs propres  $a_j$  et vecteurs propres  $\psi_j$  alors pour tout état quantique  $\varphi$ , la probabilité d'observer la valeur  $a_j$  est*

$$p_\varphi(j) = \frac{1}{\|\psi_j\|^2 \|\varphi\|^2} |\langle \psi_j | \varphi \rangle|^2 \quad (5.3.19)$$

et donc la "valeur moyenne quantique" dans l'état  $\varphi$  est

$$\langle A \rangle_\varphi = \sum_j a_j p_\varphi(j) \quad (5.3.20)$$

que l'on peut écrire

$$\langle A \rangle_\varphi = \frac{1}{\|\varphi\|^2} \langle \varphi | \hat{A} \varphi \rangle.$$

*Démonstration.* On a la décomposition spectrale de l'opérateur

$$\hat{A} = \sum_j a_j \frac{|\psi_j\rangle\langle\psi_j|\cdot\rangle}{\|\psi_j\|^2} \quad (5.3.21)$$

donc

$$\begin{aligned}\frac{1}{\|\varphi\|^2} \langle \varphi | \hat{A} \varphi \rangle &\stackrel{(\text{??})}{=} \frac{1}{\|\varphi\|^2} \sum_j \frac{1}{\|\psi_j\|^2} \langle \varphi | a_j \psi_j \rangle \langle \psi_j | \varphi \rangle = \frac{1}{\|\varphi\|^2} \sum_j a_j \frac{1}{\|\psi_j\|^2} |\langle \varphi | \psi_j \rangle|^2 \\ &\stackrel{(5.3.19)}{=} \sum_j p_\varphi(j) a_j \stackrel{(5.3.20)}{=} \langle A \rangle_\varphi\end{aligned}$$

$\square$

**Proposition 5.3.11.** Si  $\hat{H}$  est le Hamiltonien du système avec  $\hat{H}\varphi_j = E_j\varphi_j$  qui sont les vecteurs propres et valeurs propres d'énergie de  $\hat{H}$ , si  $\beta = \frac{1}{kT}$  est la température de l'environnement, la valeur moyenne de la grandeur physique  $A$  pour la mesure de probabilité de Boltzmann Gibbs  $p_{BG}$  est définie par

$$\langle A \rangle_{p_{BG}} := \sum_j \langle A \rangle_{\varphi_j} p_{BG}(j) \quad (5.3.22)$$

où  $\langle A \rangle_{\varphi_j} = \frac{1}{\|\varphi_j\|^2} \langle \varphi_j | \hat{A} \varphi_j \rangle$  est la valeur moyenne quantique dans l'état  $\varphi_j$ , et

$$p_{BG}(j) := \frac{1}{Z} e^{-\beta E_j}. \quad (5.3.23)$$

On peut écrire

$$\langle A \rangle_{p_{BG}} = \frac{1}{Z} \text{Tr} \left( \hat{A} e^{-\beta \hat{H}} \right) \quad (5.3.24)$$

avec la constante de normalisation

$$Z = \text{Tr} \left( e^{-\beta \hat{H}} \right).$$

*Remarque 5.3.12.* Noter que la valeur moyenne (5.3.22) fait intervenir une **moyenne quantique** (qui repose sur du hasard quantique dont l'origine est encore inconnue) et aussi une **moyenne statistique** liée à la température décrite par la loi de Boltzmann et qui provient des échanges d'énergie avec l'environnement chaotique.

*Démonstration.* Comme l'opération Trace ne dépend pas de la base utilisée, on utilise la base des vecteurs propres  $(\varphi_j)_j$ . On calcule<sup>3</sup>

$$\text{Tr} \left( e^{-\beta \hat{H}} \right) = \sum_j \frac{1}{\|\varphi_j\|^2} \langle \varphi_j | e^{-\beta \hat{H}} \varphi_j \rangle = \sum_j e^{-\beta E_j} = Z$$

et

$$\frac{1}{Z} \text{Tr} \left( \hat{A} e^{-\beta \hat{H}} \right) = \frac{1}{Z} \sum_j \frac{1}{\|\varphi_j\|^2} \langle \varphi_j | \hat{A} e^{-\beta \hat{H}} \varphi_j \rangle = \frac{1}{Z} \sum_j \frac{1}{\|\varphi_j\|^2} \langle \varphi_j | \hat{A} \varphi_j \rangle e^{-\beta E_j} \stackrel{(5.3.23)}{=} \sum_j \langle A \rangle_{\varphi_j} p(j) \stackrel{(5.3.22)}{=} \langle A \rangle_{p_{BG}}$$

□

3. Si  $(\varphi_j)_j$  est une base quelconque d'un espace de Hilbert, la base duale  $(\varphi_k^*)_k$  est une base de l'espace dual définie par  $\varphi_k^*(\varphi_j) = \delta_{j=k}$ . La trace d'un opérateur  $\hat{A}$  est donnée par  $\text{Tr}(A) = \sum_j \varphi_j^*(\hat{A}\varphi_j)$ . Par exemple si  $(\varphi_j)_j$  est une base orthogonale d'un espace de Hilbert (pas forcément normée) alors  $\varphi_k^* = \frac{1}{\|\varphi_k\|^2} \langle \varphi_k | \cdot \rangle$  est la base duale et donc  $\text{Tr}(A) = \sum_j \frac{1}{\|\varphi_j\|^2} \langle \varphi_j | \hat{A} \varphi_j \rangle$ .

**Proposition 5.3.13.** Soit  $T$  une température fixée et  $\beta = \frac{1}{kT}$ . Si on impose que les paramètres  $\langle b_p \rangle, \langle b_p^\dagger \rangle \in \mathbb{C}$  sont donnés par les valeurs moyennes pour le Hamiltonien  $\hat{H}_{BCS}^{(\text{approx})}$

$$\langle b_p \rangle = \langle b_p \rangle_{pBG}, \quad \langle b_p^\dagger \rangle = \langle b_p^\dagger \rangle_{pBG}, \quad (5.3.25)$$

Pour une température assez basse  $T < T_c, \beta > \beta_c$  avec la température critique (valable si  $e^{-\frac{1}{V_0g}} \ll 1$ ) donnée par

$$kT_c \approx E_D \frac{2e^\gamma}{\pi} e^{-\frac{1}{V_0g}}$$

avec la **constante de Euler**  $\gamma := \lim_{n \rightarrow +\infty} \left( \sum_{k=1}^n \frac{1}{k} - \ln n \right) = 0.577\dots$ , alors  $\Delta = \Delta(T) > 0$  est solution de l'équation du gap

$$\int_0^{E_D} \frac{\tanh\left(\frac{\beta}{2}\sqrt{\epsilon'^2 + \Delta^2}\right)}{\sqrt{\epsilon'^2 + \Delta^2}} d\epsilon' = \frac{1}{V_0g} \quad (5.3.26)$$

avec  $g$  la densité d'états au niveau de Fermi. A basse température  $T \rightarrow 0$ , la solution est

$$\Delta(0) = 2E_D e^{-\frac{1}{V_0g}}$$

près de la température critique,  $T \xrightarrow{T < T_c} T_c$ ,

$$\Delta(T) \approx 3.06 kT_c \left(1 - \frac{T}{T_c}\right)^{1/2} \quad (5.3.27)$$

Voir figure 5.3.1(b) qui représente la solution  $T \rightarrow \Delta(T)$ . Pour représenter cette courbe, il faut résoudre numériquement l'équation (5.3.26).

**Remarque 5.3.14. "Gap effectif"**. Le fait que le gap spectral dépend de la température est un artefact de la méthode utilisée. En effet pour le problème exact, il y a un gap fixé  $\Delta(0)$  dans le spectre, et des probabilité de transition vers les états excités qui vont dépendre de la température.

**Remarque 5.3.15. "Universalité"**. Une conséquence remarquable relie le gap spectral  $E_0 = 2\Delta(0)$  à la température critique :

$$\frac{E_0}{kT_c} = \frac{2\Delta(0)}{kT_c} = \frac{2\pi}{e^\gamma} \approx 3.53\dots$$

est une constante universelle. Les expérience sont en bon accord avec cette prédiction. Voir tableau 5.1, Couture & Zitoun (1992, p479), Kittel (1987, p.336).

*Remarque 5.3.16. “Effet isotopique”.* Une prédiction de ce modèle, bien vérifié expérimentalement est que si on change de matériau par des atomes isotopes, c’est à dire de même charge donc même propriétés chimique, mais masse  $M$  des noyaux différentes, alors l’énergie de Debye  $E_D$  lié aux fréquences de vibration, varie comme  $E_D \propto M^{-1/2}$  donc  $T_c \propto M^{-1/2}$

*Démonstration.* On part de

$$\langle b_p \rangle_{pBG} \stackrel{(5.3.24)}{=} \frac{1}{Z} \text{Tr} \left( b_p e^{-\beta \hat{H}_{BCS}^{(\text{approx})}} \right)$$

On calcule

$$\begin{aligned} Z &= \text{Tr} \left( e^{-\beta \hat{H}_{BCS}^{(\text{approx})}} \right) = \text{Tr} \left( \exp \left( -\beta \left( \int_{p \in D} \hat{H}_{BCS}^{(\text{approx})}(p) d^3 p + C \text{Id} \right) \right) \right) \\ &= \text{Tr} \left( e^{-\beta C \text{Id}} \right) \prod_{p \in D} \text{Tr} \left( e^{-\beta \hat{H}_{BCS}^{(\text{approx})}(p)} \right) \end{aligned}$$

On a de même

$$\text{Tr} \left( b_p e^{-\beta \hat{H}_{BCS}^{(\text{approx})}} \right) = \text{Tr} \left( e^{-\beta C \text{Id}} \right) \prod_{p' \in D} \text{Tr} \left( b_p e^{-\beta \hat{H}_{BCS}^{(\text{approx})}(p')} \right)$$

où  $b_p$  agit comme Id dans  $\mathcal{F}_{(p', -p')}$  pour  $p' \neq p$ . Donc il reste

$$\langle b_p \rangle_{pBG} = \frac{\text{Tr} \left( b_p e^{-\beta \hat{H}_{BCS}^{(\text{approx})}(p)} \right)}{\text{Tr} \left( e^{-\beta \hat{H}_{BCS}^{(\text{approx})}(p)} \right)}. \quad (5.3.28)$$

On va calculer chacun de ces termes. Comme la trace ne dépend pas de la base, on utilise les valeurs propres (5.3.16) de  $\hat{H}_{BCS}^{(\text{approx})}(p)$  pour calculer

$$\begin{aligned} \text{Tr} \left( e^{-\beta \hat{H}_{BCS}^{(\text{approx})}(p)} \right) &= 2e^{-\beta \epsilon'_p} + e^{-\beta E_+(p)} + e^{-\beta E_-(p)} \\ &= 2e^{-\beta \epsilon'_p} + e^{-\beta \epsilon'_p} e^{-\beta \delta_p} + e^{-\beta \epsilon'_p} e^{\beta \delta_p} \\ &= e^{-\beta \epsilon'_p} \left( 2 + e^{-\beta \delta_p} + e^{\beta \delta_p} \right) = e^{-\beta \epsilon'_p} \left( e^{\frac{\beta}{2} \delta_p} + e^{-\frac{\beta}{2} \delta_p} \right)^2 \end{aligned} \quad (5.3.29)$$

$$= 4e^{-\beta \epsilon'_p} \cosh^2 \left( \frac{\beta}{2} \delta_p \right) \quad (5.3.30)$$

Pour le numérateur de (5.3.28), on observe que

$$b_p e^{-\beta \hat{H}_{BCS}^{(\text{approx})}(p)} \stackrel{(5.3.18)}{=} \left( \frac{1}{-\beta} \right) \frac{\partial}{\partial \Delta} e^{-\beta \hat{H}_{BCS}^{(\text{approx})}(p)}$$

Donc

$$\begin{aligned}
\text{Tr} \left( b_p e^{-\beta \hat{H}_{BCS}^{(\text{approx})}(p)} \right) &= \left( \frac{1}{-\beta} \right) \frac{\partial}{\partial \Delta} \text{Tr} \left( e^{-\beta \hat{H}_{BCS}^{(\text{approx})}(p)} \right) \stackrel{(5.3.29)}{=} \left( \frac{1}{-\beta} \right) \frac{\partial}{\partial \Delta} \left( 4e^{-\beta \epsilon'_p} \cosh^2 \left( \frac{\beta}{2} \delta_p \right) \right) \\
&= \left( \frac{1}{-\beta} \right) 4e^{-\beta \epsilon'_p} \left( \frac{\partial \cosh^2 \left( \frac{\beta}{2} \delta_p \right)}{\partial \delta_p} \right) \left( \frac{\partial \delta_p}{\partial \Delta} \right) \\
&\stackrel{(5.3.16)}{=} \left( \frac{1}{-\beta} \right) 4e^{-\beta \epsilon'_p} \left( 2 \frac{\beta}{2} \cosh \left( \frac{\beta}{2} \delta_p \right) \sinh \left( \frac{\beta}{2} \delta_p \right) \right) \frac{\Delta}{2\delta_p} \\
&= -2e^{-\beta \epsilon'_p} \cosh \left( \frac{\beta}{2} \delta_p \right) \sinh \left( \frac{\beta}{2} \delta_p \right) \frac{\Delta}{\delta_p}
\end{aligned}$$

et

$$\langle b_p \rangle_{pBG} \stackrel{(5.3.28)}{=} - \frac{2e^{-\beta \epsilon'_p} \cosh \left( \frac{\beta}{2} \delta_p \right) \sinh \left( \frac{\beta}{2} \delta_p \right) \frac{\Delta}{\delta_p}}{4e^{-\beta \epsilon'_p} \cosh^2 \left( \frac{\beta}{2} \delta_p \right)} = - \frac{\Delta}{2\delta_p} \tanh \left( \frac{\beta}{2} \delta_p \right). \quad (5.3.31)$$

Ensuite

$$\begin{aligned}
\Delta &\stackrel{(5.3.15)}{=} \int_{p'} -V_0 \langle b_{p'} \rangle dp' \stackrel{(5.3.25)}{=} \int_{p'} -V_0 \langle b_{p'} \rangle_{pBG} dp' \\
&\stackrel{(5.3.31)}{=} - \int_{p'} -V_0 \frac{\Delta}{2\delta_{p'}} \tanh \left( \frac{\beta}{2} \delta_{p'} \right) dp'
\end{aligned}$$

Ainsi  $\Delta$  vérifie l'équation du gap

$$\Delta = \int_{p' \text{ t.q. } |\epsilon'_p| < E_D} V_0 \frac{\Delta}{2\delta_{p'}} \tanh \left( \frac{\beta}{2} \delta_{p'} \right) d^3 p'.$$

Il y a la solution  $\Delta = 0$ . Si  $\Delta \neq 0$  alors en notant  $g$  la densité d'état au niveau de Fermi  $\epsilon = \mu$ , (on néglige ses variation sur l'intervalle d'énergie  $|\epsilon'| < E_D$ )

$$1 = \int_{p' \text{ t.q. } |\epsilon'_p| < E_D} \frac{V_0}{2\delta_{p'}} \tanh \left( \frac{\beta}{2} \delta_{p'} \right) d^3 p' = g \frac{V_0}{2} \int_{-E_D}^{E_D} \frac{1}{\delta_{p'}} \tanh \left( \frac{\beta}{2} \delta_{p'} \right) d\epsilon' \quad (5.3.32)$$

$$= g V_0 \int_0^{E_D} \frac{1}{\delta_{p'}} \tanh \left( \frac{\beta}{2} \delta_{p'} \right) d\epsilon' \quad (5.3.33)$$

avec  $\delta_{p'} \stackrel{(5.3.16)}{=} \sqrt{\epsilon'^2 + \Delta^2}$ . On a obtenu (5.3.26).

**Basse température :** Pour  $\beta \rightarrow \infty$ , c'est à dire  $T \rightarrow 0$ , on a  $\tanh \left( \frac{\beta}{2} \delta_{p'} \right) \rightarrow 1$ , et utilisant que pour  $y > 0$  et  $y \gg 1$ ,

$$\int_0^y (1+x^2)^{-1/2} dx \stackrel{(\text{xcas})}{=} -\ln \left( \sqrt{y^2+1} - y \right) = -\ln \left( y \sqrt{1 + \frac{1}{y^2}} - y \right) \quad (5.3.34)$$

$$= -\ln \left( y \left( 1 + \frac{1}{2y^2} + o(y^{-2}) \right) - y \right) \quad (5.3.35)$$

$$= -\ln \left( \frac{1}{2y} (1 + o(y^{-2})) \right) = \ln(2y) + o(y^{-2}) \quad (5.3.36)$$

et posant  $x = \frac{\epsilon'}{\Delta}$ , on obtient pour  $\Delta \ll E_D$

$$\begin{aligned} 1 &= gV_0 \int_0^{E_D} \frac{1}{\sqrt{\epsilon'^2 + \Delta^2}} d\epsilon' = \frac{gV_0}{\Delta} \int_0^{E_D} \frac{1}{\sqrt{\left(\frac{\epsilon'}{\Delta}\right)^2 + 1}} d\epsilon' \\ &= gV_0 \int_0^{\frac{E_D}{\Delta}} \frac{1}{\sqrt{x^2 + 1}} dx \stackrel{(5.3.34)}{=} gV_0 \left( \ln \left( 2 \frac{E_D}{\Delta} \right) + o \left( \left( \frac{\Delta}{E_D} \right)^2 \right) \right) \end{aligned}$$

donc

$$\Delta(0) = 2E_D e^{-\frac{1}{gV_0}}. \quad (5.3.37)$$

**Température critique :** La valeur critique  $\beta_c$  est donnée par le fait que  $\Delta(\beta_c) = 0$ , donc  $\delta_{p'} \stackrel{(5.3.16)}{=} \epsilon'$ . On utilisera que pour  $y > 0$ ,  $u' = \frac{1}{x}$ ,  $v = \tanh(x)$ ,  $u = \ln x$ ,  $v' = \frac{1}{\cosh^2(x)}$ ,  $\ln(x) \tanh(x) \xrightarrow{x \rightarrow 0} 0$ ,

$$\int_0^y \frac{1}{x} \tanh(x) dx \stackrel{(p.p.)}{=} \ln(y) \tanh(y) - \int_0^y \frac{\ln(x)}{\cosh^2(x)} dx \quad (5.3.38)$$

$$(5.3.39)$$

Et pour  $y \gg 1$ ,  $\tanh(y) \rightarrow 1$  et

$$\int_0^y \frac{\ln(x)}{\cosh^2(x)} dx \rightarrow \ln \left( \frac{\pi}{4e^\gamma} \right)$$

avec  $\gamma := \lim_{n \rightarrow +\infty} \left( \sum_{k=1}^n \frac{1}{k} - \ln n \right) = 0.577..$  la **constante de Euler**. Donc pour  $y \gg 1$ , l'équation

$$\int_0^y \frac{1}{x} \tanh(x) dx = C$$

a la solution

$$\ln(y) = C + \ln \left( \frac{\pi}{4e^\gamma} \right) \Leftrightarrow y \approx \frac{\pi}{4e^\gamma} e^C$$

On pose  $x = \frac{\beta_c}{2} \epsilon'$ . Alors

$$1 \stackrel{(5.3.32)}{=} gV_0 \int_0^{E_D} \frac{1}{\epsilon'} \tanh \left( \frac{\beta_c}{2} \epsilon' \right) d\epsilon' = gV_0 \int_0^{\frac{\beta_c E_D}{2}} \frac{1}{x} \tanh(x) dx$$

Donc dans le cas  $\frac{\beta_c E_D}{2} \gg 1$ , la solution est

$$\begin{aligned} \frac{\beta_c E_D}{2} &\approx \frac{\pi}{4e^\gamma} e^{\frac{1}{gV_0}} \Leftrightarrow \beta_c \approx \frac{\pi}{2e^\gamma E_D} e^{\frac{1}{gV_0}} \\ &\Leftrightarrow kT_c = E_D \frac{2e^\gamma}{\pi} e^{-\frac{1}{gV_0}} \end{aligned}$$

Conséquence

$$\frac{\Delta(0)}{kT_c} \stackrel{(5.3.37)}{=} 2E_D e^{-\frac{1}{gV_0}} \frac{\pi}{2e^\gamma E_D} e^{\frac{1}{gV_0}} = \frac{\pi}{e^\gamma} \approx 1.76..$$

est une constante universelle. Voir Kittel (1987, p.336). Notre hypothèse a été

$$\frac{\beta_c E_D}{2} \gg 1 \Leftrightarrow e^{-\frac{1}{gV_0}} \ll 1.$$

**Exposant critique :** Pour montrer (5.3.27), on affine le calcul précédent, en supposant  $\Delta \ll 1$  et on cherche  $\beta(\Delta)$  par développement limité □

# Annexe A

## Formulaire

### A.1 Série géométrique

**Proposition A.1.1.** *La série*

$$S_n := \sum_{n'=0}^n r^{n'}$$

*s'appelle série géométrique. Si  $r \neq 1$ , on a*

$$S_n = \frac{r^{n+1} - 1}{r - 1},$$

*et si  $r = 1$  alors  $S_n = n + 1$ . En particulier si  $|r| < 1$  alors la série*

$$S_\infty = \sum_{n'=0}^{\infty} r^{n'} = \frac{1}{1 - r}. \quad (\text{A.1.1})$$

*est convergente.*

*Démonstration.* On a  $S_n = r^0 + \dots + r^{n-1} + r^n$  donc

$$rS_n = r^1 + \dots + r^n + r^{n+1} = S_n + r^{n+1} - r^0$$

donc si  $r \neq 1$ ,

$$S_n = \frac{r^{n+1} - r^0}{r - 1}.$$

□



# Bibliographie

- Balian, Roger. 2007. *From microphysics to macrophysics : methods and applications of statistical physics*. Vol. 1,2. Springer Science & Business Media.
- Blaizot, J.P., & Ripka, G. 1986. *Quantum theory of finite systems*.
- Brin, M., & Stuck, G. 2002. *Introduction to Dynamical Systems*. Cambridge University Press.
- Combescot, Monique, & Shiau, Shiue-Yuan. 2015. *Excitons and Cooper pairs : two composite bosons in many-body physics*. Oxford University Press.
- Couture, L., & Zitoun, R. 1992. *Physique statistique*.
- Cover, Thomas M. 1999. *Elements of information theory*. John Wiley & Sons.
- Davies, E.B., & Davies, E.B. 1996. *Spectral theory and differential operators*. Vol. 42. Cambridge Univ Pr.
- Den Hollander, Frank. 2008. *Large deviations*. Vol. 14. American Mathematical Soc.
- Diu, B., Guthmann, C., Lederer, D., & Roulet, B. 1989. *Physique statistique*.
- Faure, F. 2014a. *Cours de Mécanique quantique pour Master M1 de physique*. [link](#).
- Faure, F. 2014b. Introduction au chaos quantique. *In : journées X-UPS*. Éditions de l'école polytechnique [link](#).
- Faure, F. 2016. *Cours de systèmes dynamiques*. [link](#).
- Faure, F. 2020. *Cours d'acoustique musicale. Niveau Licence 3*. [link](#).
- Faure, F. 2022. *Cours de Physique statistique*. [link](#).
- Friedli, Sacha, & Velenik, Yvan. 2017. *Statistical Mechanics of Lattice Systems : A Concrete Mathematical Introduction*. Cambridge University Press.
- Kittel, C. 1987. *Quantum theory of Solids*.

- Martin, Philippe, & Rothen, Francis. 1990. *Problèmes à N-corps et champs quantiques*. Presses Polytechniques et Universitaires Romandes (PPUR).
- Negele, J.W., & Orland, H. 1998. *Quantum many-particle systems*. Perseus Books.
- Peskin, M.E., & Schroeder, D.V. 1995. *An introduction to quantum field theory*. Vol. 94. Westview press.
- Pottier, Noelle. 2000. *Physique statistique hors d'équilibre*.
- Sternberg, S. 1994. *Group theory and physics*. Cambridge University Press.