

Introduction à la physique statistique et à la physique quantique

Frédéric Hélein, Thierry Lévy

23 avril 2004

Table des matières

1	Thermodynamique classique	11
1.1	La thermodynamique classique	11
1.2	Notions fondamentales	11
1.3	Un exemple de système : le piston	16
1.4	Evolutions isothermes	17
1.5	Evolution adiabatique	19
1.6	Le cycle de Carnot	19
1.7	L'entropie	22
1.8	Le second principe de la thermodynamique	23
1.9	Conclusion	25
2	Physique statistique d'un gaz raréfié	29
2.1	Une description microscopique d'un gaz	29
2.1.1	Un modèle de billard	31
2.1.2	L'espace des phases	33
2.1.3	Rappels sur la mécanique hamiltonienne	34
2.2	Le chaos ou comment les lois du hasard interviennent dans un système déterministe	36
2.3	Premier usage des probabilités : les gaz dilués	41
2.4	Théorie cinétique du gaz parfait	45
2.5	La distribution de Maxwell	48
2.6	Interprétation statistique de l'irréversibilité	49
2.6.1	Le paradoxe de Loschmidt	50
2.6.2	Fluctuations de la densité d'un gaz	52
2.6.3	Une première approche statistique de l'entropie	56
2.6.4	Le paradoxe de Gibbs	60
2.7	Conclusion	62

3	Mécanique statistique avec interactions entre les particules	65
3.1	L'ensemble micro-canonique	65
3.1.1	Le choix de l'espace des phases	65
3.1.2	Représentation d'un état par une mesure	67
3.1.3	La construction d'une mesure décrivant l'état d'équilibre	68
3.1.4	Existe-il d'autres mesures décrivant l'équilibre?	70
3.2	L'ergodicité	71
3.2.1	Flots ergodiques	71
3.2.2	Autres applications de l'hypothèse ergodique	74
3.3	Applications du formalisme micro-canonique	77
3.3.1	Dynamique d'une évolution quasi-statique	77
3.3.2	Le gaz dilué à l'équilibre	79
3.4	Entropie microcanonique	82
3.4.1	Entropie statistique sur des ensembles finis	82
3.4.2	Entropie statistique d'un état macroscopique	84
3.4.3	Entropie microcanonique	87
3.5	Le formalisme canonique	90
3.5.1	L'ensemble canonique	90
3.5.2	Identification de β	92
3.5.3	Répartition de l'énergie	93
4	Interactions thermodynamiques entre la lumière et la matière	97
4.1	Préliminaire : l'équipartition de l'énergie	97
4.2	Le problème du rayonnement du corps noir	99
4.2.1	La pensée thermodynamique de Planck	100
4.2.2	Position du problème	102
4.2.3	Interpolation entre les lois de Rayleigh–Jeans et de Wien	104
4.2.4	Comprendre la loi de Planck	106
4.3	L'effet photoélectrique	108
4.4	Loi de déplacement de Wien et loi de Stefan-Boltzmann	109
4.5	Une approche combinatoire de la loi de Planck	111
4.6	Prologue	114
4.6.1	Généralités	115
4.6.2	La jauge de Coulomb	116
4.6.3	Formulation variationnelle des équations de Maxwell dans le vide	118

4.6.4	La lumière comme oscillation du champ électromagnétique	119
4.6.5	Les ondes électromagnétiques à l'intérieur d'une cavité cubique	121
5	Vers la quantification des états d'énergie de la matière	127
5.1	Quelques éléments de relativité	128
5.2	Retour aux quanta de lumières	136
5.3	Le spectre des atomes	137
5.4	L'équation d'Hamilton–Jacobi	141
5.5	Les débuts de la « mécanique ondulatoire »	149
5.6	L'équation de Schrödinger	153
6	L'équation de Schrödinger et son interprétation	157
6.1	Du paquet d'onde à l'interprétation probabiliste de l'équation de Schrödinger	159
6.1.1	Le paquet d'onde	159
6.1.2	Une loi de conservation	161
6.1.3	Le dogme de l'« école de Copenhague »	164
6.1.4	Le principe de superposition	165
6.2	Observables	167
6.2.1	Une parenthèse mathématique : opérateurs auto-adjoints hermitiens	169
6.3	Le théorème d'Ehrenfest et ses conséquences	174
6.3.1	Le théorème d'Ehrenfest	174
6.3.2	Interprétation et exemples d'applications	176
6.3.3	Un principe de correspondance	177
7	Applications de la mécanique quantique	179
7.1	L'effet tunnel	179
7.2	L'oscillateur harmonique	181
7.2.1	Première méthode	181
7.2.2	Deuxième méthode	182
7.3	Ensembles complets d'observables qui commutent	186
7.4	Le moment cinétique	186
7.5	L'atome d'hydrogène	191

8	Le point de vue de Heisenberg	195
8.1	La mécanique des matrices	195
8.2	Les relations d'incertitudes d'Heisenberg	195
8.3	Reformulation des principes de la mécanique quantique	196
8.4	Le spin	197
8.4.1	Fonctions propres pour \hat{L}_z	197
8.4.2	Retour sur le moment cinétique angulaire	198
8.4.3	Le spin d'un électron	200

Introduction

En apparence l'essentiel des phénomènes physiques dont nous avons une expérience directe dans la vie de tous les jours obéissent à des lois qui ont été formulées il y a plus de trois siècles par Galileo GALILÉE et Isaac NEWTON (entre autres...) pour la mécanique et au XIXème siècle par Michael FARADAY et James Clerck MAXWELL (encore entre autres...) pour l'électromagnétisme. Toute personne relativement familière avec ces lois (comme vous) voit son intuition confortée par des expériences quotidiennes, que ce soit en conduisant sa voiture ou en observant la chute des pommes (mis à part quelques phénomènes un peu plus déroutant comme par exemple le mouvement d'une toupie). Une des étapes les plus sensationnelles de l'histoire de la science fut la découverte au début du XXème siècle qu'à l'échelle microscopique la matière se comporte d'une manière profondément différente de ce que l'on pouvait prévoir à l'aide de la physique classique. Et les physiciens ont été forcés d'admettre qu'il n'était pas possible de décrire ce qui se passe à l'échelle atomique en des termes intelligibles pour nous. En revanche ils ont pu construire des modèles mathématiques (mécanique quantique, théorie quantique des champs) permettant des prédictions d'une précision stupéfiante, mais malheureusement incapables de donner une image de ce qui se passe propre à satisfaire notre intuition. Ce choc dans le monde de la physique a suscité de nombreux débats : si certains physiciens (autour de Niels BOHR) ont prôné une attitude pragmatique, à savoir que puisque le modèle mathématique donne des réponses absolument conformes à l'expérience il n'y a pas lieu de chercher plus loin, d'autres scientifiques en revanche (autour d'Albert EINSTEIN) ont objecté qu'un certain nombre de principes qui sont postulés dans la mécanique quantique, relativement durs à avaler (comme la probabilité de présence ou la réduction du paquet d'onde), ne sont que des concepts provisoires, en attendant l'élaboration d'un modèle satisfaisant d'avantage notre besoin de comprendre. Ajoutons que la physique quantique

n'est pas uniquement la physique du « très petit » et qu'aucun phénomène n'échappe en principe à ses lois, hormis la description de situations extrêmes comme dans les trous noirs ou à l'échelle de Planck (10^{-33} cm) qui nécessiterait l'utilisation d'une théorie unifiant la mécanique quantique et la théorie de la relativité générale (mais cette théorie n'existe pas encore). Simplement la plupart des phénomènes macroscopiques mettent en jeu un nombre gigantesque d'atomes, ce qui fait qu'une partie des lois de la physique est gommée à plus grande échelle, et ce qu'il en reste est ce que nous nommons mécanique classique. Toutefois certains phénomènes particuliers, dont la mise en oeuvre nécessite une technologie sophistiquée, comme les laser ou la supraconductivité, obéissent à des lois quantiques non classiques, même à l'échelle macroscopique.

Nous allons essayer dans ce cours d'exposer les concepts physiques et mathématiques qui permettent actuellement d'appréhender cette théorie en suivant un ordre historique. Deux raisons justifient ce choix, qui n'est pas habituel dans les textes d'introduction à la mécanique quantique, sans doute en partie parce que le chemin historique est toujours plus compliqué qu'un exposé synthétique et axiomatique, mais aussi parce que il nécessite un plus grand bagage mathématique. La première raison est simplement que le développement historique de la mécanique quantique est passionnant. La deuxième raison est qu'une présentation plus synthétique ou privilégiant un certain point de vue pourrait donner l'impression que la théorie est achevée, qu'elle est une construction qui découlerait d'axiomes qui ont été gravés dans le marbre par quelques physiciens dans les années 1920. Rien ne serait plus faux : aucune théorie physique en général n'est définitive et a fortiori la mécanique quantique, où bien des points restent totalement opaques. Certes les axiomes de la mécanique quantiques des années 1920 sont et seront toujours en accord avec la plupart des expériences que l'on peut imaginer aujourd'hui. Mais comme nous l'avons évoqué plus haut et comme nous le verrons en détails plus tard, la théorie actuelle ne peut pas satisfaire notre intuition. Plus grave encore est le fait que l'on ne sait toujours pas comment unifier la physique quantique avec la relativité générale. Enfin une autre difficulté est que la théorie quantique des champs, qui décrit avec une précision extraordinaire les interactions entre particules, repose sur des calculs qui n'ont aucune base mathématique (même si les techniques mathématique qui ont été développées autour de cette théorie, comme par exemple la théorie de la renormalisation, sont des constructions subtiles et puissantes). Il semble que l'on n'ait toujours pas bien compris quelles devaient être les bases physiques et mathématiques

de cette théorie. Ainsi toutes ces considérations conduisent à penser que nos connaissances actuelles ne nous permettent pas de saisir réellement en quoi consiste la physique quantique. En attendant il est donc important de comprendre la théorie actuelle comme le résultat d'un long cheminement, d'après discussions, de changement de perspectives. Ce chemin n'est toujours pas terminé et il y aura probablement d'autres révolutions avant d'arriver (espérons-le) à une théorie complète. Et il est probable que beaucoup d'idées reléguées aux oubliettes seront remises à l'honneur (sous une forme sans doute modifiée) au cours des révolutions à venir, comme c'est l'habitude en physique, plus encore qu'en mathématiques.

C'est pourquoi nous commencerons par deux questions anciennes, qui sont à l'origine de cette aventure : les interactions thermodynamiques entre la matière et la lumière et le spectre (discret) de la lumière émise par les atomes.

Dans la conception de ce cours, nous avons bien entendu largement utilisé les références bibliographiques qui sont mentionnées au cours du texte. Mais nous nous sommes aussi beaucoup inspirés du livre formidable sur les probabilités de Jacques HARTHONG¹, à la fois comme modèle de style d'exposition et aussi en lui empruntant un certain nombre de discussions sur l'ergodicité et le calcul de Planck. Deuxièmement la construction même de ce cours, en tentant de développer les idées suivant un parcours proche du cheminement historique, doit beaucoup à de nombreuses discussions avec Joseph KOUNEIHHER ainsi qu'à un certain nombre de notes qu'il m'a communiquées.

¹cf. <http://moire4.u-strasbg.fr/bouquins/proba/tabmat1.htm>

Chapitre 1

Thermodynamique classique

1.1 La thermodynamique classique

1.2 Notions fondamentales

Nous rappelons ici brièvement quelques notions de thermodynamique. Pour comprendre quantitativement les liens entre les notions de volume, pression, température, énergie, chaleur, etc., il est nécessaire de considérer dans un premier temps des situations simplifiées et idéalisées comme celle qui suit.

Un système physique — Isolons mentalement une certaine quantité de matière contenue dans une portion de l'espace du reste du monde. Nous pouvons lui associer certaines grandeurs :

- la quantité la plus concrète est *le volume occupé* V
- *la pression* p , égale au quotient de la force exercée par le gaz sur une portion de la paroi par l'aire de cette paroi. Cette définition suppose que ce quotient soit le même partout (ce qui n'est pas le cas par exemple autour d'une aile d'avion, sans quoi l'avion tomberait) et donc que le gaz soit au repos.
- *la température* T : c'est déjà une notion plus subtile, bien que nous en ayons une perception immédiate. D'abord, tout comme la pression, on ne peut définir la température d'un système que s'il est à l'équilibre (Ce qui signifie que l'on a laissé reposer le système suffisamment longtemps pour qu'il « n'évolue plus », le temps de relaxation pouvant varier entre une fraction de seconde et des heures suivant les situa-

tions). La possibilité théorique de définir la température est déduite de l'observation expérimentale du *principe zéro de la thermodynamique*. Celui-ci s'énonce : « deux systèmes, en équilibre thermique avec un même troisième, restent en équilibre entre eux ». C'est dire que la relation d'équivalence $A \sim B$ entre des états physiques A et B , signifiant que « A est en équilibre thermique avec B », qui est bien évidemment réflexive et symétrique, est aussi transitive, et est donc une relation d'équivalence. Alors chaque classe d'équivalence définit une température¹. Ce principe conduit à la définition d'une fonction température T sur l'ensemble des systèmes physiques qui n'est bien évidemment pas unique, puisque, pour tout homéomorphisme croissant ϕ de l'image de T dans vers un ouvert de \mathbb{R} , $\phi \circ T$ est encore une fonction température².

– *la quantité de molécules N* . On l'exprime soit en nombre de molécules N , soit en nombre de moles N/N_A , où $N_A = 6,022\,136\,7 \cdot 10^{23}$ est le nombre d'Avogadro.. Le nombre de molécules joue un rôle important lorsqu'on considère un mélange de différentes espèces chimiques (par exemple de l'hydrogène H_2 avec de l'oxygène O_2 et de l'eau H_2O). On doit alors comptabiliser dans l'énergie totale une énergie chimique (et des potentiels chimiques) dépendant de ces nombres.

Exercice 1 Soit \mathcal{U} un ensemble dénombrable d'éléments que l'on notera A, B, \dots . Supposons que \mathcal{U} soit muni des relations \sim, \succeq et \succ , satisfaisant les propriétés (i) et (ii) énoncées dans la note en bas de page 1. Montrer qu'il existe une fonction $T : \mathcal{U} \rightarrow \mathbb{R}$ telle que $A \succeq B \iff T(A) \geq T(B)$.

Toutes ces grandeurs s'appellent *grandeurs d'état*. Nous les avons défini à peu près comme on aurait pu le faire au milieu du XIXème siècle, sur la base d'un certain nombre de principes issus de l'expérience, qu'il serait difficile de déduire logiquement de la mécanique. Nous verrons ultérieurement, avec la

¹Il faudrait compléter ce principe en introduisant la relation $A \succ B$ signifiant « A est plus chaud que B et n'est pas en équilibre thermique avec B » et la relation $A \succeq B$ signifiant alors que « $A \succ B$ ou $A \sim B$ ». Alors en postulant que (i) \succeq est une relation d'ordre totale sur l'ensemble des systèmes physiques et que (ii) $\forall A, B$ tels que $A \succ B$ il existe C tel que $A \succ C \succ B$, on peut démontrer l'existence d'une fonction température (cf. Exercice 1).

²On peut exiger en plus que lorsque T varie (de façon quasi-statique), par exemple le volume V du système varie en étant une fonction différentiable de T et même que $\frac{dV}{dT}$ soit le plus proche possible d'une constante : c'est exactement le principe du thermomètre à alcool ou à mercure, reliant de façon affine la dilatation d'un fluide à sa température.

mécanique statistique, une approche bien plus profonde de ces notions, notamment en ce qui concerne la température. Notons que parmi ces grandeurs d'état, certaines sont dites *extensives* et d'autres *intensives*. Les grandeurs extensives (volume, nombre de molécules) s'ajoutent lorsque l'on superpose deux systèmes en équilibre thermiques et chimiques : par exemple si j'assemble deux systèmes identiques et en équilibre en un seul, le volume du nouveau système est la somme des volumes de chacune des parties. Les grandeurs intensives (température, pression, potentiel chimique) ne s'ajoutent pas : par exemple si j'assemble deux systèmes à la même température, le résultat est un système qui a encore la même température.

L'énergie interne. Enfin nous aurions pu également incorporer dans la liste des grandeurs d'état *l'énergie interne* U : c'est une notion historiquement issue de la mécanique où elle est définie comme la somme de l'énergie potentielle et de l'énergie cinétique. Pour un mouvement sans frottement, l'énergie totale est conservée : c'est ce qui rend la notion d'énergie si importante. Au milieu du XIX^{ème} siècle, les travaux de N.L.S. CARNOT, J.P. JOULE et R.J. MAYER permirent de comprendre que la chaleur était aussi une forme particulière d'énergie, accomplissant ainsi un premier pas vers une unification de la mécanique et de la thermique. Notons qu'à la différence des grandeurs d'état comme V , p , T , l'énergie totale est un concept relativement abstrait (ou transcendant). En particulier l'énergie totale ne se mesure pas directement, mais est évaluée à partir d'un bilan des échanges d'énergie. En revanche, l'énergie totale peut être en général calculée à partir des autres variables d'état : elle est une fonction $U = U(T, p, V, \dots)$. Enfin, contrairement à la température ou à la pression, il est possible de définir l'énergie totale d'un système, même si celui-ci est en dehors de l'équilibre.

A présent nous pouvons commencer à récolter les fruits du travail des thermodynamiciens : sur des systèmes parfois extrêmement complexes, nous avons défini un petit nombre de grandeurs d'état (à partir du moment où les systèmes sont **à l'équilibre**) et nous pouvons faire des prédictions en utilisant des relations qui ne font intervenir que ces grandeurs ! Considérons un exemple simple : un gaz monoatomique contenu dans une enceinte fermée, à l'équilibre. Les grandeurs d'état sont V , N , p et T . L'état du système semble a priori dépendre de quatre variables. Mais ces grandeurs sont liées par une relation d'état (entre V , N , p et T). Par ailleurs U peut être calculée à partir des autres variables. Par exemple si nous avons un gaz parfait, alors l'équa-

tion d'état (R. BOYLE, E. MARIOTTE, L.J. GAY-LUSSAC, L. BOLTZMANN) est

$$pV = NkT,$$

où $k = \frac{R}{N_A} = 1,380\,658 \cdot 10^{-23} JK^{-1}$ est la constante de Boltzmann. Et l'énergie totale s'exprime en fonction de la température par

$$U = \frac{3}{2}NkT.$$

Ainsi l'état macroscopique de ce système dépend de 3 paramètres. Nous pourrions alors choisir trois grandeurs parmi V , N , p et T et tout exprimer en fonction de ces trois grandeurs. Mais quelles grandeurs choisir ? La pratique des calculs de thermodynamique s'apparente souvent à une sorte de gymnastique différentielle, consistant à dériver une variable par rapport à une autre, tout en gelant une troisième variable, etc., d'où il résulte rapidement une certaine confusion. Afin de clarifier cela nous allons représenter chaque état A d'un système à l'équilibre thermodynamique par un point m dans une variété différentiable \mathcal{M} , de dimension 3, sorte d'espace de configuration thermodynamique³. Pourquoi une variété différentiable ? parce que beaucoup de raisonnements faits en thermodynamique supposent que l'on laisse ou fait évoluer un système de façon suffisamment lente pour qu'à chaque instant il soit à l'équilibre et qu'on puisse ainsi définir sa température et sa pression. Donc on peut à chaque instant associer au système un point dans la variété \mathcal{M} et modéliser l'évolution par une courbe dans \mathcal{M} . On parle alors d'une *évolution réversible* ou *quasi-statique*. On peut alors étudier les variations infinitésimales des grandeurs d'état et leurs relations⁴. Ainsi chaque grandeur d'état est manipulée comme une fonction différentiable sur \mathcal{M} .

Afin de quantifier les échanges d'énergie d'un système avec l'extérieur sous forme de chaleur ou de travail, nous devons adjoindre aux 4 fonctions différentiables qui satisfont les relations d'états écrites plus haut deux 1-formes différentielles sur \mathcal{M} :

- la *quantité de chaleur absorbée* par le système q
- la *quantité de travail apportée* par l'extérieur au système w .

³ici \mathcal{M} peut être choisie difféomorphe à \mathbb{R}^3 , ou bien être défini comme étant la sous-variété de $\{(V, N, p, T) \in \mathbb{R}^4\}$ définie par la relation d'état.

⁴Il s'agit là bien sûr d'une situation idéale, mais que l'on approche expérimentalement

Alors si un système évolue d'une façon *réversible* ou *quasi-statique*, et si nous représentons cette évolution par un chemin orienté γ sur \mathcal{M} allant d'un état initial m_0 à un état final m_1 , la quantité de chaleur (respectivement de travail) absorbée par le système pour évoluer de m_0 à m_1 est $\int_{\gamma} q$ (respectivement $\int_{\gamma} w$). Ces deux 1-formes ne sont pas fermées en général. En revanche nous

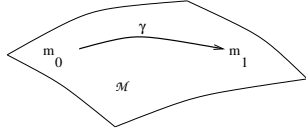


FIG. 1.1 – Une évolution quasi-statique sur la variété des configurations thermodynamiques

avons :

$$dU = q + w.$$

Remarquons que l'on aurait pu prendre cette relation comme définition de l'énergie (à une constante près).

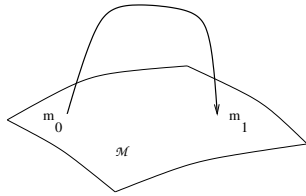


FIG. 1.2 – Une évolution **non** quasi-statique joignant deux points sur la variété \mathcal{M}

A l'opposé une évolution non réversible ne peut pas être représentée par une courbe sur \mathcal{M} , car en général le système passe par des états en dehors de l'équilibre thermodynamique. Ainsi on ne peut définir les états de départ et d'arrivée dans \mathcal{M} .

1.3 Un exemple de système : le piston

Imaginons un gaz contenu dans une enceinte formée par un tube cylindrique fermé à un bout par une paroi fixe et à l'autre bout par un piston sans masse qui glisse sans frottement (figure 1.3). Dans cette expérience on remarque que (sauf s'il y a des fuites) le nombre de molécules N est fixé, égal à une valeur N_0 . On peut donc remplacer la variété \mathcal{M} par

$$\mathcal{M}_{N_0} := \{m \in \mathcal{M} / N(m) = N_0\}.$$

Envisageons un déplacement infiniment petit dh du piston (h désigne ici la

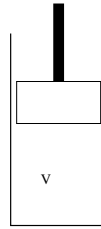


FIG. 1.3 – Un cylindre fermé par un piston

hauteur du piston). Nous supposons qu'en même temps une force d'intensité **très légèrement inférieure** à $p\Sigma$ (où Σ est l'aire du piston) est exercée de l'extérieur sur le piston. Nous pouvons alors conclure en vertu du principe d'action et de réaction que le gaz a lui aussi exercé une force de même intensité sur le piston et donc le gaz fournit un travail pratiquement égal à $p\Sigma dh = p dV$. Donc le travail apporté par l'extérieur est :

$$w = -p dV.$$

L'expression de q en revanche n'est dictée par aucune considération mécanique. Tout ce que l'on peut faire est utiliser deux grandeurs en guise de coordonnées locales sur \mathcal{M} et décomposer q linéairement en fonction des différentielles de ces grandeurs. Par exemple si on prend comme grandeurs (T, V) (ou bien respectivement (T, p)), on a

$$q = C_V dT + a dV \quad (\text{ou bien respectivement } = C_p dT + b dp),$$

où C_V est la capacité thermique à volume constant et C_p est la capacité thermique à pression constante.

1.4 Evolutions isothermes

Dilatation isotherme quasi-statique — Comme premier exemple d'application de ce qui précède, imaginons que notre cylindre, qui contient le gaz, est plongé dans une grosse marmite d'eau bouillante à $T_0 := 373\text{K}$ (100C) (figure 1.4). Cette quantité d'eau maintient le gaz dans le cylindre à la température

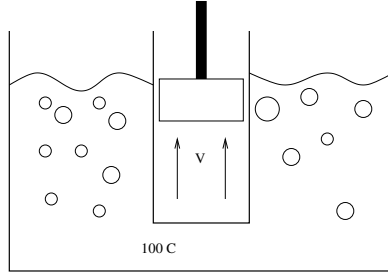


FIG. 1.4 – Le même système dans un thermostat à 100 C ...

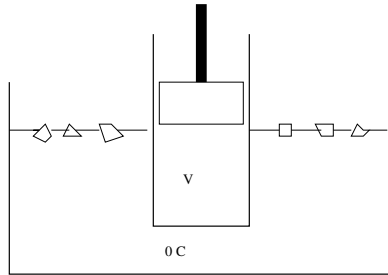


FIG. 1.5 – ... ou dans un thermostat à 0 C

constante T_0 , elle joue le rôle de *thermostat*. Initialement une certaine force (d'intensité égale à $p\Sigma$, où Σ est l'aire du piston) est exercée de l'extérieur sur le piston, de façon à comprimer le gaz. Le système est à l'état initial m_0 de grandeurs d'état (p_0, V_0, T_0) . Ensuite nous réduisons très légèrement la force exercée sur le piston, de façon à ce que le gaz puisse se détendre, tout en restant à l'équilibre. On continue ainsi jusqu'à un certain état m_1 de grandeurs d'état (p_1, V_1, T_1) . Le chemin γ entre m_0 et m_1 a pour effet d'augmenter V , de diminuer p , tout en restant à la température constante

$T = T_0 = T_1$. Le travail fourni à l'extérieur par le gaz est, pour un gaz parfait,

$$-W = \int_{\gamma} -w = \int_{\gamma} p dV = \int_{\gamma} \frac{NkT dV}{V} = NkT_0 \int_{V_0}^{V_1} \frac{dV}{V} = NkT_0 \log \frac{V_1}{V_0} > 0.$$

Le processus inverse, qui consiste à faire revenir le système à son état d'origine en poussant le piston, nécessite qu'une machine extérieure fournisse un travail égal à $NkT_0 \log \frac{V_1}{V_0}$. Notons que, dans le cas du gaz parfait, U est fonction uniquement de T et donc ne varie pas ici. Donc $dU = w + q = 0$. Par conséquent le travail fourni par le gaz lors de la dilatation est « alimenté » en énergie par la chaleur du thermostat et lorsque le gaz se contracte le travail fourni par l'extérieur est converti en chaleur intégralement transmise au thermostat.

Dilatation isotherme non quasi-statique — A présent nous reprenons le dispositif expérimental précédent mais nous n'exerçons plus de force sur le piston : comme le piston est supposé avoir une masse négligeable, le gaz n'exerce pratiquement aucune force et le travail fourni par le gaz est pratiquement nul. On observera aussi une détente beaucoup plus rapide que dans le cas précédent : il s'agit d'une dilatation non quasi-statique. Nous laissons monter ainsi librement le piston jusqu'à ce qu'il atteigne la même hauteur que dans l'expérience précédente, puis nous attendons un certain temps, pour que le système revienne à un état d'équilibre. Alors l'état final du système doit être le même que dans l'expérience quasi-statique : la température à l'intérieur du cylindre doit s'équilibrer avec celle du thermostat, c'est être à $T_1 = T_0$, le volume, le nombre de molécules sont les mêmes, et donc la pression p_1 , liée à T_1 , V_1 et N_1 par la relation d'état est aussi identique. Le bilan énergétique est donc le même. Et pourtant le système n'a pas fourni de travail. L'explication est que le gaz dans le cylindre a puisé une quantité de chaleur dans le thermostat qui est plus petite que dans l'expérience précédente : le caractère violent de la détente a créé une agitation des molécules dans le cylindre qui a contribué à échauffer le gaz. Enfin, contrairement à l'expérience quasi-statique, le système n'est pas passé par des états d'équilibre. Donc il n'est pas possible de définir toutes les grandeurs d'état (en particulier T) et on ne peut pas modéliser cette expérience par un chemin dans \mathcal{M} .

1.5 Evolution adiabatique

Une évolution est adiabatique si elle s'effectue sans échange de chaleur avec l'extérieur. Noter que'une telle évolution peut parfois être très rapide. Dans l'exemple du piston, on doit remplacer le thermostat dans lequel baignait le piston par une enveloppe isolante.

Il est possible de caractériser les courbes γ dans \mathcal{M}_{N_0} qui modélisent une évolution **réversible et adiabatique** par une équation différentielle. En effet il s'agit de la condition $q|_\gamma = 0$. Et étant donné que \mathcal{M}_{N_0} est de dimension 2, en chaque point $m \in \mathcal{M}_{N_0}$ cette condition prescrit totalement la droite tangent à γ . Voyons ce que cela donne dans le cas d'un gaz parfait. Des relations

$$dU = q + w \quad \text{et} \quad q|_\gamma = 0,$$

on tire, sachant que $U = \frac{3}{2}NkT$ et $w = -pdV$, que

$$\frac{3}{2}Nk dT|_\gamma = -p dV|_\gamma.$$

Il y a ici trois grandeurs d'état susceptibles de varier : T , V et p . Nous pouvons éliminer p en fonction de T et V , à l'aide de la relation d'état $pV = NkT$. Cela donne

$$\frac{3}{2}Nk dT|_\gamma = -\frac{NkT}{V} dV|_\gamma \quad \iff \quad \frac{3}{2} \frac{dT|_\gamma}{T} = -\frac{dV|_\gamma}{V}.$$

En intégrant nous obtenons

$$T^{3/2}V = \text{Constante}.$$

1.6 Le cycle de Carnot

Nous avons déjà remarqué que q est une 1-forme non exacte sur \mathcal{M} . En particulier, cela signifie que, si γ est une courbe *fermée* dans \mathcal{M} décrivant une évolution réversible, le travail reçu par le système au cours d'un tel cycle est

$$W = \int_\gamma w = \int_D dw = \int_D -dp \wedge dV \neq 0,$$

(où D est l'immersion d'un disque orienté dans \mathcal{M} dont le bord est γ) en vertu de la formule de Stokes. La quantité de chaleur reçue est, quand à

elle, l'opposée de cette quantité. Nous percevons donc qu'un système peut transformer de la chaleur en travail ou l'inverse au cours d'un cycle, tout en revenant à son état initial à la fin. C'est le principe de la machine thermodynamique (moteur à vapeur, réfrigérateur, etc.). Une situation où le bilan peut être calculée relativement simplement a été imaginée par N.L.S. CARNOT. Cela consiste en un chemin constitué d'évolutions réversibles alternativement isothermiques ou adiabatiques :

- Chemin γ_0 — Détente réversible isotherme $(p_0, V_0, T_0) \longrightarrow (p_1, V_1, T_0)$
- Chemin γ_1 — Détente réversible adiabatique $(p_1, V_1, T_0) \longrightarrow (p_2, V_2, T_2)$
- Chemin γ_2 — Compression réversible isotherme $(p_2, V_2, T_2) \longrightarrow (p_3, V_3, T_2)$
- Chemin γ_3 — Compression réversible adiabatique $(p_3, V_3, T_2) \longrightarrow (p_0, V_0, T_0)$.

On peut imaginer par exemple que le chemin γ_0 se fait dans un thermostat à $T_0 \simeq 373 \text{ K}$ (100 C) (plongé dans l'eau bouillante), puis que, durant le chemin γ_1 , on laisse le piston se détendre jusqu'à ce que la température descende à $T_2 \simeq 273 \text{ K}$ (0 C). Le chemin γ_2 se fait alors dans un thermostat rempli de glace à 0 C, le chemin γ_3 est une recompression jusqu'à ce qu'on retrouve la température de 100 C. La quantité de chaleur Q reçue par le système se

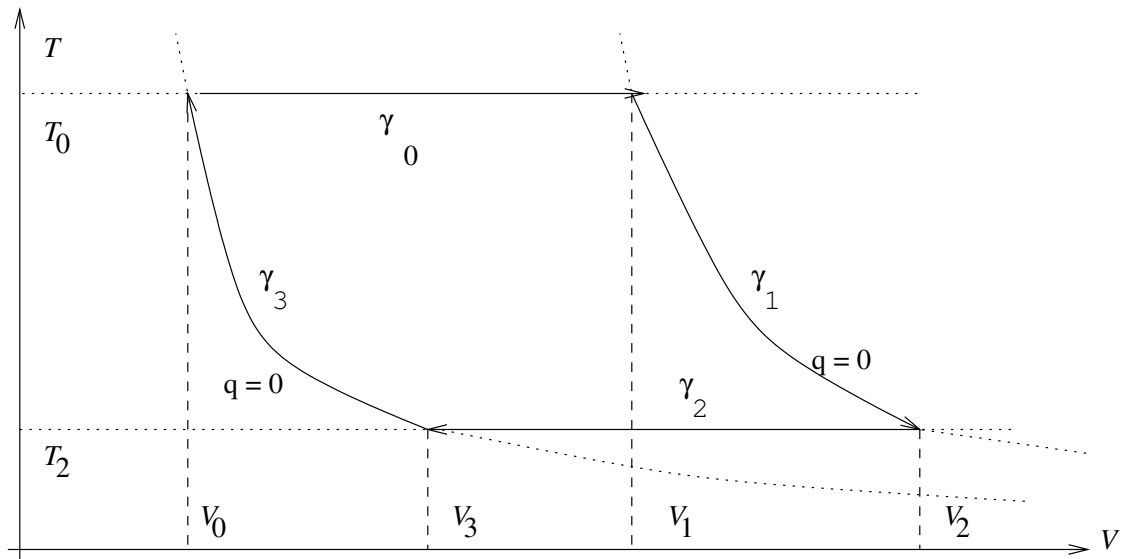


FIG. 1.6 – Représentation schématique du cycle de Carnot

décompose selon $Q = Q_0 + Q_1 + Q_2 + Q_3$, où chaque Q_a est la quantité reçue pendant le chemin γ_a . De même on décompose la quantité de travail

absorbées par le système $W = W_0 + W_1 + W_2 + W_3$. Ces quantités peuvent

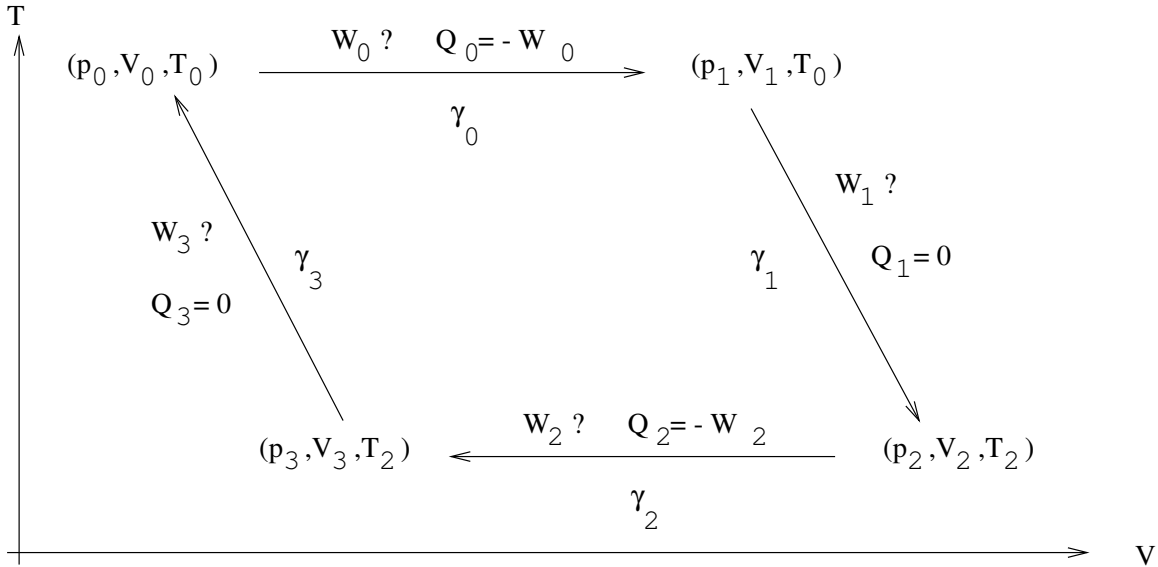


FIG. 1.7 – Le bilan énergétique du cycle de Carnot

être calculées à partir des considérations précédentes :

- Détente $(p_0, V_0, T_0) \longrightarrow (p_1, V_1, T_0)$: $W_0 = -NkT_0 \log \frac{V_1}{V_0}$ (voir précédemment) et comme T est constant et on a affaire à un gaz parfait, U est constant, donc $Q_0 = NkT_0 \log \frac{V_1}{V_0}$
- Détente $(p_1, V_1, T_0) \longrightarrow (p_2, V_2, T_2)$: $Q_1 = 0$ et donc $W_1 = U(p_2, V_2, T_2) - U(p_1, V_1, T_0) = \frac{3}{2}Nk(T_2 - T_0)$.
- Compression $(p_2, V_2, T_2) \longrightarrow (p_3, V_3, T_2)$: $W_2 = -NkT_2 \log \frac{V_3}{V_2}$ et $Q_2 = NkT_2 \log \frac{V_3}{V_2}$.
- Compression $(p_3, V_3, T_2) \longrightarrow (p_0, V_0, T_0)$: $Q_3 = 0$ et $W_3 = \frac{3}{2}Nk(T_0 - T_2)$.

On voit tout de suite que $Q_1 + Q_3 = 0$ et $W_1 + W_3 = 0$ et donc les étapes adiabatiques ne jouent aucun rôle dans le bilan de Q et W . En revanche, il est possible de comparer W_0 et W_2 comme suit.

Sachant que $T_0^{3/2}V_1 = T_2^{3/2}V_2$ (conservation de $T^{3/2}V$ le long de la détente adiabatique γ_1), et $T_2^{3/2}V_3 = T_0^{3/2}V_0$ (compression adiabatique γ_2) on déduit

que

$$V_2 = \left(\frac{T_0}{T_2}\right)^{3/2} V_1 \quad \text{et} \quad V_3 = \left(\frac{T_0}{T_2}\right)^{3/2} V_0 \quad \implies \quad \frac{V_3}{V_2} = \left(\frac{V_1}{V_0}\right)^{-1}.$$

Donc

$$W_2 = -NkT_2 \log \frac{V_3}{V_2} = NkT_2 \log \frac{V_1}{V_0} = -\frac{T_2}{T_0} W_0.$$

Donc

$$W = W_0 \left(1 - \frac{T_2}{T_0}\right) \quad \text{et} \quad Q = -W_0 \left(1 - \frac{T_2}{T_0}\right).$$

Ainsi, si $T_0 > T_2$ et $V_2 > V_0$ (ce qui correspond au scénario que nous avons imaginé), nous avons $W_0 < 0$ et donc $W < 0$ et $Q > 0$: le système a fourni du travail à l'extérieur, en l'échange d'une quantité de chaleur. Mais si on parcourt le cycle en sens inverse, c'est l'inverse.

1.7 L'entropie

La relation trouvée entre W_2 et W_0 entraîne la même relation entre Q_2 et Q_0 , c'est à dire $Q_2 = -\frac{T_2}{T_0} Q_0$ ou encore

$$\int_{\gamma_0} \frac{q}{T} + \int_{\gamma_2} \frac{q}{T} = \frac{Q_0}{T_0} + \frac{Q_2}{T_2} = 0.$$

Et comme q s'annule sur les processus adiabatiques γ_1 et γ_3 ,

$$\int_{\gamma} \frac{q}{T} = 0.$$

Cette identité peut permettre de démontrer que la 1-forme $\frac{q}{T}$ est exacte. En effet, elle est valable pour tous les cycles de Carnot (correspondant à des choix de températures T_0 et T_2 et de volumes V_0 et V_1 arbitraires) donc en particulier pour des cycles infiniment petits. Il est donc possible de démontrer que cette identité est valable sur tous les cycles *réversibles* et d'en déduire que

$$d\left(\frac{q}{T}\right) = 0.$$

(*Exercice* : le vérifier !) Mais surtout on *vérifie expérimentalement* que cette loi s'applique pour tous les processus réversibles concernant des gaz. Cela entraîne ne qu'il existe une fonction $S : \mathcal{M} \rightarrow \mathbb{R}$ telle que

$$dS = \frac{q}{T} \quad \text{sur } \mathcal{M}.$$

Cette fonction, définie ici à une constante d'intégration près, s'appelle **l'entropie** du système.

Dans le cas particulier des gaz parfaits, nous allons vérifier directement cela par le calcul :

$$\frac{q}{T} = \frac{dU - w}{T} = \frac{3}{2} \frac{Nk dT}{T} + \frac{p dV}{T},$$

et en substituant $p = \frac{NkT}{V}$ dans la deuxième expression :

$$\frac{q}{T} = \frac{3}{2} \frac{Nk dT}{T} + \frac{Nk dV}{V} = Nk d(\log(T^{3/2}V)).$$

L'entropie est donc

$$\begin{aligned} S(p, V, T, N) &= S(p_0, V_0, T_0, N) + Nk \log \frac{T^{3/2}V}{T_0^{3/2}V_0} \\ &= S(p_0, V_0, T_0, N) + Nk \log \left(\frac{T^{5/2}}{V} \frac{V_0}{T_0^{5/2}} \right), \end{aligned}$$

où l'on a utilisé pour déduire la deuxième expression de la première la substitution $V = \frac{NkT}{p}$.

Nous remarquons que, à cause du facteur N devant le logarithme, l'entropie est une grandeur extensive : elle croît dans les mêmes proportions que la quantité de matière. Ainsi l'entropie d'un système constitué de plusieurs sous-systèmes couplés est la somme des entropies de chacun des sous-systèmes.

1.8 Le second principe de la thermodynamique

Tout ce qui précède concernait les évolutions réversibles ou quasi-statiques. Mais chacun sait qu'un système du type de celui considéré précédemment

subit des forces de frottement (inévitables) et s'il est laissé à lui-même, il finira toujours par ne plus être animé d'aucun mouvement, toute l'énergie mécanique ayant été « dégradée » en chaleur. Une avancée importante réalisée au XIX^{ème} siècle a été de comprendre cela et d'être capable de quantifier cette perte.

Cela n'apparaît pas immédiatement dans le calcul des échanges de chaleur et de travail que nous avons effectué pour le cycle de Carnot. En effet lors d'un cycle complet, le système part d'un état d'équilibre et revient au même état d'équilibre. Donc comme l'entropie est une fonction des grandeurs d'états, elle doit être la même au début et à la fin. Cela est vrai même si, au lieu de faire un cycle de Carnot réversible, le système évolue de façon non quasi-statique⁵, il suffit d'attendre à la fin que le système revienne à l'équilibre pour que l'on retrouve la même valeur que la valeur initiale pour l'entropie, à nouveau puisque l'entropie est une fonction des grandeurs d'état. Donc pour observer l'irréversibilité, c'est à dire un phénomène de dégradation de l'énergie, on ne doit pas raisonner sur le système constitué par le piston seul, mais sur un système plus grand, constitué par le piston (P) plus un système environnant (E)⁶. Ainsi par exemple, durant la phase γ_0 , le travail reçu par (E) est $W_0(E) = -W_0(P)_{irr} = \int_{\gamma_0} p_{irr} dV$. Mais comme le processus n'est pas exactement réversible, la pression p_{irr} réellement exercée par le piston est plus petite que la pression calculée pour un processus parfaitement réversible, donc

$$W_0(E) = \int_{\gamma_0} p_{irr} dV < \int_{\gamma_0} p dV = -W_0(P)_{rev}.$$

A l'opposé, pendant la phase γ_2 , de compression isotherme, la pression p_{irr} à exercer sur le piston est supérieure à celle pour un processus réversible : $p_{irr} > p$, donc le gain⁷ de travail de (E) est (en tenant compte ici du fait que $dV|_{\gamma_2} < 0$)

$$W_2(E) = \int_{\gamma_2} p_{irr} dV < \int_{\gamma_2} p dV = -W_2(P)_{rev}.$$

⁵par exemple en « sautant » de (p_0, V_0, T_0) à (p_1, V_1, T_0) (c'est à dire par une détente isotherme *non réversible*), puis de (p_1, V_1, T_0) à (p_2, V_2, T_2) (par une dilatation adiabatique *non réversible*), puis de (p_2, V_2, T_2) à (p_3, V_3, T_3) et enfin de (p_3, V_3, T_3) à (p_0, V_0, T_0)

⁶comprenant les thermostats plus par exemple des poids que l'on accrocherait à des hauteurs différentes pour stocker le travail sous forme d'énergie potentielle

⁷négalif ici, puisqu'il s'agit d'une perte

Pour ces deux processus isothermes, on a donc

$$W_0(E) + W_0(P)_{rev} < 0 \quad \text{et} \quad W_2(E) + W_2(P)_{rev} < 0.$$

On aurait pu arriver au même résultat en considérant les flux de chaleurs : pendant la phase γ_0 , l'agitation thermique dans le piston crée de la chaleur qui n'est pas puisée dans le thermostat et donc (E) perd moins d'énergie que pour un processus réversible, donc a un gain $Q_0(E)_{irr} = -Q_0(P)_{irr} > -Q_0(P)_{rev} = W_0(P)_{rev}$. Et de même, pendant la compression γ_2 une partie du travail effectué par (E) est dissipée dans le piston, ce qui fait que le thermostat soit absorber une quantité de chaleur supérieure au cas réversible. Donc à nouveau $Q_2(E)_{irr} = -Q_2(P)_{irr} > -Q_2(P)_{rev} = W_2(P)_{rev}$.

En conclusion, en supposant pour simplifier que l'entropie de (P) et (E) ne varie pas pendant les phases adiabatique γ_1 et γ_3 , l'augmentation de l'entropie de (E) est :

$$\delta S(E)_{irr} = \frac{\delta Q_0(E)_{irr}}{T_0} + \frac{\delta Q_2(E)_{irr}}{T_2} > -\frac{\delta Q_0(P)_{rev}}{T_0} - \frac{\delta Q_2(P)_{rev}}{T_2} = -\delta S(P)_{rev} = 0.$$

A présent nous savons que $\delta S(P)_{irr} = \delta S(P)_{rev}$ (puisque $S(P)$ est une fonction d'état), donc

$$\delta S(P + E)_{irr} = \delta S(P)_{rev} + \delta S(E)_{irr} > 0.$$

Enonçons le second principe de la thermodynamique :

Pour un système isolé en évolution au cours du temps, l'entropie ne décroît jamais. Pour un système à l'équilibre, l'entropie est constante.

Un corollaire des deux cas envisagés dans ce principe est qu'un système va évoluer jusqu'à ce que son entropie ne puisse plus augmenter, c'est à dire est *maximale*. Noter qu'il existe bien d'autres formulations plus ou moins équivalentes de ce principe (en terme de mouvement perpétuels ou en des termes bien plus abstraits). Nous ne nous étendrons pas là-dessus, car le chapitre suivant va apporter un éclairage différent sur l'entropie.

1.9 Conclusion

Le second principe de la thermodynamique et les propriétés physiques qu'il contient (l'irréversibilité) sont le premier concept dans l'histoire de la phy-

sique moderne qui entre en conflit avec le postulat du déterminisme, étroitement liée à l'idée de réversibilité. Citons à ce sujet P.-S. de LAPLACE :

Une intelligence qui, pour un instant donné, connaît trait toutes les forces dont la nature est animée, et la situation respective des êtres qui la composent, si d'ailleurs elle était assez vaste pour soumettre ces données à l'analyse, embrasserait dans la même formule les mouvements des plus grands corps de l'univers et ceux du plus léger atome : rien ne serait incertain pour elle, et l'avenir comme le passé serait présent à ses yeux. (1819).

Afin de concilier la réversibilité et le second principe de la thermodynamique une idée a été proposée au milieu du XIX^{ème} siècle (notamment par R. CLAUSIUS et Sir J. J. THOMSON) : supposons que la matière soit constituée d'atomes (une notion pas très claire à l'époque). Supposons de plus que ces atomes bougent d'une manière totalement désordonnée, et qu'ils se trouvent dans des positions aléatoires et bougent dans des directions et avec des vitesses aléatoires. Alors on pourrait imaginer que l'énergie interne d'un gaz (ou d'un système plus complexe) soit la somme des énergies cinétiques de chacun de ces atomes, que la température mesure le degré d'agitation des atomes et qu'enfin la pression⁸ résulte des chocs des atomes entrant en collision avec une paroi. Cette hypothèse va s'avérer être très fructueuse. Elle sera mise en oeuvre d'abord par J.C. MAXWELL, qui retrouvera ainsi les propriétés du gaz parfait, puis par L. BOLTZMANN, qui comprendra réellement⁹ ce qu'est l'entropie et quel est le lien entre l'entropie et le désordre. Cela nécessite l'introduction des statistiques et des probabilités dans la physique, objet du prochain chapitre.

Enfin il faut souligner l'originalité des raisonnements faits en thermodynamique, par rapport à la façon de penser traditionnelle de la mécanique. En effet, il n'est pas possible en thermodynamique de décrire avec exactitude ce qui passe dans un gaz (par exemple en modélisant l'évolution par une équation différentielle) ou a fortiori dans un système plus complexe. En revanche, l'identification de grandeurs d'état qui, même si elles ne sont pas observables à tout instant (en particulier quand le système n'est pas à l'équilibre), sont mesurables au début ou à la fin d'une expérience lorsque le système est à l'équilibre, permet de faire de nombreuses prédictions. C'est un peu une phy-

⁸une explication déjà proposée par D. Bernoulli en 1736

⁹et pendant longtemps restera le seul à comprendre

sique de la « boîte noire » : on observe avant et après une expérience et on essaie de tirer des conclusions à partir d'un bilan sur certaines grandeurs. Ce type de raisonnement est à garder en tête lorsqu'on aborde la mécanique quantique : nous allons chercher à comprendre des expériences dans lesquelles on peut mesurer des grandeurs avant et après, mais il nous sera impossible d'avoir une idée correspondant à un schéma « classique » de ce qui se passe durant l'expérience.

Référence : W. Greiner, L. Neise, H. Stöcker, *Thermodynamique et mécanique statistique*, Springer.

Chapitre 2

Physique statistique d'un gaz raréfié

2.1 Une description microscopique d'un gaz

Nous allons imaginer un gaz comme une assemblée de N particules *newtoniennes*, c'est à dire dont le mouvement obéit à la dynamique newtonienne. Pour simplifier nous supposons que ces particules sont identiques entre elles (en particulier elles ont toutes la même masse m). La position à chaque instant t du centre de masse de ces particules est donnée par les valeurs des N fonctions $(\gamma_i)_{i=1}^N$ en t . Si le gaz est supposé être contenu dans une enceinte fermée, chaque fonction γ_i est astreinte à prendre des valeurs dans un ouvert borné Ω de \mathbb{R}^3 . Nous supposons que la frontière $\partial\Omega$ de Ω est régulière (de classe \mathcal{C}^2). Parmi les modèles *dynamiques*, les plus simples consistent à supposer que chaque particule décrit un mouvement de translation rectiligne uniforme en dehors de chocs. Cela signifie que si la distance entre deux particules différentes est plus grande qu'une certaine valeur $2r$ (où r est le « rayon » de la particule), ces deux particules n'interagissent pas. Donc γ_i est une fonction continue de t , affine par morceau et non dérivable aux instants des chocs. Nous supposons les chocs élastiques, c'est à dire :

- pour un choc entre deux particules, l'énergie cinétique et la quantité de mouvement sont conservées. Ainsi, en numérotant les deux particules

qui se rencontrent par i et j , nous avons les relations :

$$\begin{aligned} m\dot{\gamma}_i(t_c^-) + m\dot{\gamma}_j(t_c^-) &= m\dot{\gamma}_i(t_c^+) + m\dot{\gamma}_j(t_c^+) \\ \frac{m}{2}|\dot{\gamma}_i(t_c^-)|^2 + \frac{m}{2}|\dot{\gamma}_j(t_c^-)|^2 &= \frac{m}{2}|\dot{\gamma}_i(t_c^+)|^2 + \frac{m}{2}|\dot{\gamma}_j(t_c^+)|^2, \end{aligned}$$

où t_c est l'instant du choc et on note $f(t_c^-) := \lim_{t \uparrow t_c} f(t)$ et $f(t_c^+) := \lim_{t \downarrow t_c} f(t)$ les valeurs limites prises par une fonction à gauche et à droite respectivement de t_c . Notant $v_k^\pm := \dot{\gamma}_k(t_c^\pm)$, les solutions de ce système sont de la forme

$$v_i^+ = \frac{1}{2}(v_i^- + v_j^-) + \frac{1}{2}|v_i^- - v_j^-|n \quad \text{et} \quad v_j^+ = \frac{1}{2}(v_i^- + v_j^-) - \frac{1}{2}|v_i^- - v_j^-|n, \quad (2.1)$$

où n est un vecteur quelconque de norme 1. (*Exercice.*)

- pour un choc d'une particule contre la paroi $\partial\Omega$ l'énergie cinétique est conservée¹ et la variation de la quantité de mouvement est égale à la force exercée sur la paroi sur la particule. En numérotant i la particule incidente et en notant $F_{\partial\Omega i}\delta(t - t_c)$ la force exercée² par la paroi sur la particule incidente, on a

$$m\dot{\gamma}_i(t_c^+) = F_{\partial\Omega i} + m\dot{\gamma}_i(t_c^-) \quad \text{et} \quad \frac{m}{2}|\dot{\gamma}_i(t_c^-)|^2 = \frac{m}{2}|\dot{\gamma}_i(t_c^+)|^2,$$

avec, en plus, les conditions que $\langle \dot{\gamma}_i(t_c^-), \nu \rangle > 0$ et $\langle \dot{\gamma}_i(t_c^+), \nu \rangle < 0$, où ν est le vecteur normal sortant à $\partial\Omega$ au point d'impact $\gamma_i(t_c)$. D'où nous déduisons ici, notant toujours $v_i^\pm := \dot{\gamma}_i(t_c^\pm)$, que

$$v_i^+ = |v_i^-|n \quad \text{et} \quad F_{\partial\Omega i} = |v_i^-|n - v_i^-, \quad (2.2)$$

où n est un vecteur quelconque de norme 1, tel que $\langle n, \nu \rangle < 0$. (*Exercice.*)

¹Une hypothèse simplificatrice est faite ici : en réalité les molécules qui constituent la paroi sont animées de mouvements de vibration thermique et des échanges d'énergie et de quantité de mouvement ont lieu entre les molécules du gaz et celles de la paroi. Ces échanges contribuent à l'équilibre thermodynamique entre les deux.

² δ est ici la *masse de Dirac*, ce n'est pas une fonction, mais une *distribution*, une sorte de fonction généralisée (voir un cours sur les distributions) : $\delta(t - t_c)$ est nul sur tout compact dans $\mathbb{R} \setminus \{t_c\}$ et $\int_{t_c - \varepsilon}^{t_c + \varepsilon} \delta(t - t_c) dt = 1, \forall \varepsilon > 0$.

Dans cette description une certaine incertitude apparaît déjà, puisqu'à chaque choc entre deux particules, les vitesses après le choc ne sont pas complètement déterminées. En particulier le paramètre $n \in S^2$ est ici arbitraire. Il est possible, en faisant des hypothèses sur la structure des particules (atomes ou molécules), de modéliser de ce qui se passe durant le choc par une loi donnant n en fonction des vitesses avant le choc (section efficace).

Notre propos ici sera de montrer que, même en partant d'un modèle microscopique *déterministe*, une modélisation aléatoire s'impose. Nous allons donc considérer un modèle simple qui, même s'il n'est pas forcément réaliste, est en accord avec les règles de conservation d'énergie et de quantité de mouvement énoncées plus haut et est de plus déterministe.

2.1.1 Un modèle de billard

Un modèle simple est le suivant : nous supposons que chacune des particules est une sphère parfaitement dure et élastique de rayon r (comme une boule de billard). Ces particules évoluent dans une enceinte $\Omega' := \{x \in \mathbb{R}^3 / \text{dist}(x, \Omega) < r\}$. A chaque choc, la force (et donc la variation de quantité de mouvement) exercée sur chaque particule est normale à la surface de la sphère au point d'impact. Donc :

- un choc entre les deux particules i et j a lieu lorsque $\lim_{t \rightarrow t_c} |\gamma_i(t) - \gamma_j(t)| = 2r$. Alors nous supposons que la force $F_{ji}\delta(t - t_c)$, exercée par j sur i , est parallèle au vecteur

$$\nu := \frac{\gamma_j(t_c) - \gamma_i(t_c)}{|\gamma_j(t_c) - \gamma_i(t_c)|}.$$

Cela signifie que $v_i^+ - v_i^- = \frac{F_{ji}}{m}$ est parallèle à ν , donc $\exists \lambda \in \mathbb{R}$ tel que

$$v_i^+ - v_i^- = \lambda \nu.$$

En utilisant la conservation de l'énergie on trouve que $\lambda = -\langle v_i^- - v_j^-, \nu \rangle$ et donc

$$\begin{cases} v_i^+ &= v_i^- - \langle v_i^- - v_j^-, \nu \rangle \nu \\ v_j^+ &= v_j^- + \langle v_i^- - v_j^-, \nu \rangle \nu. \end{cases}$$

Remarquer qu'ici, comme $|\gamma_j(t) - \gamma_i(t)| \geq 2r, \forall t$, on a toujours $\langle v_i^- - v_j^-, \nu \rangle \geq 0$. (*Exercice.*)

- un choc entre la particule i et la paroi a lieu lorsque $\text{dist}(\gamma_i(t_c), \partial\Omega') = r$. Alors nous supposons que la force exercée par la paroi sur la particule est parallèle au vecteur normal extérieur ν à $\partial\Omega'$ au point d'impact. Donc v_i^+ s'exprime fonction de v_i^- par la loi

$$v_i^+ = v_i^- - 2\langle v_i^-, \nu \rangle \nu. \quad (2.3)$$

Observer que cette loi est la même que celle de la réflexion d'un rayon lumineux sur un miroir. Noter aussi ici que $\langle v_i^-, \nu \rangle \geq 0$. (*Exercice.*)

Nous avons exclu ici les collisions simultanées de trois particules ou plus : cela peut arriver, mais est extrêmement improbable (en fait la probabilité d'une telle collision est nulle).

Il est possible de réinterpréter les équations précédentes comme celles du mouvement d'un point de coordonnées (x_1, \dots, x_N) dans un ouvert de \mathbb{R}^{3N} . Pour cela, nous observons que la particule sphérique centrée en $x \in \mathbb{R}^3$ est contenue dans Ω' si et seulement si $x \in \Omega$. Ensuite nous considérons

$$\Omega_r^N := \{(x_1, \dots, x_N) \in \Omega^N / \forall i, j \in \mathbb{N} \text{ tels que } 1 \leq i, j \leq N \text{ et } i \neq j, |x_i - x_j| > 2r\}$$

et nous noterons également dans la suite (dans le cas où $r = 0$ et $\Omega' = \Omega$)

$$\Omega_\star^N := \{(x_1, \dots, x_N) \in \Omega^N / \forall i, j \in \mathbb{N} \text{ tels que } 1 \leq i, j \leq N \text{ et } i \neq j, |x_i - x_j| > 0\}.$$

Il est clair que l'on peut associer de façon bijective à tout point $x_\star = (x_1, \dots, x_N) \in \Omega_\star^N$ une configuration de N boules dans Ω' qui ne se touchent pas entre elles et que les points de $\partial\Omega_r^N$ représentent des collisions particule-paroi ou particule-particule. Soit maintenant $\gamma_\star = (\gamma_1, \dots, \gamma_N)$ décrivant le mouvement de ces N particules. Chaque collision s'interprète comme une réflexion de γ_\star sur $\partial\Omega_r^N$. Cela est clair pour les collisions contre la paroi $\partial\Omega'$. Dans le cas d'une collision entre deux particules i et j , le vecteur normal unitaire à $\partial\Omega_r^N$ en $x_\star = \gamma_\star(t_c)$ est

$$\begin{aligned} \nu_N &= \frac{-1}{\sqrt{2}|x_i - x_j|} \sum_{\alpha=1}^3 (x_i^\alpha - x_j^\alpha) \frac{\partial}{\partial x_i^\alpha} + (x_j^\alpha - x_i^\alpha) \frac{\partial}{\partial x_j^\alpha} \\ &= \frac{-1}{\sqrt{2}} \sum_{\alpha=1}^3 \nu^\alpha \left(\frac{\partial}{\partial x_i^\alpha} - \frac{\partial}{\partial x_j^\alpha} \right). \end{aligned}$$

Et si nous demandons que γ_* se réfléchisse sur $\partial\Omega_r^N$, à savoir $v_*^+ = v_*^- - 2\langle v_*^-, \nu^N \rangle \nu^N$, nous obtenons

$$\begin{cases} v_k^+ = v_k^- & \text{si } k \neq i, j \\ v_i^+ = v_i^- - \langle v_i^- - v_j^-, \nu \rangle \nu \\ v_j^+ = v_j^- + \langle v_i^- - v_j^-, \nu \rangle \nu, \end{cases}$$

comme précédemment. (*Exercice.*) Nous en déduisons que :

Le mouvement de N particules assimilées à des sphères dures parfaitement élastiques de même rayon et même masse dans une enceinte fermée est décrit par le mouvement d'un point dans un ouvert de dimension $3N$ se réfléchissant sur le bord de cet ouvert selon la loi (2.3).

2.1.2 L'espace des phases

Une autre propriété de la dynamique de ces N particules est qu'il s'agit d'un *mouvement hamiltonien* (singulier dans le cas du modèle de billard à cause des chocs). En effet, premièrement toute l'évolution du système au cours du temps (passé et futur) peut être retrouvée en connaissant les positions et vitesses de chaque particule à un instant donné. Plus précisément le système peut être décrit par un champ de vecteur (singulier) ξ dans l'*espace des phases*

$$T^*\Omega_r^N := \{(x_*, p_*) / x_* \in \Omega_r^N, p_* \in (\mathbb{R}^{3N})^*\},$$

où $p_* := (p_1, \dots, p_N)$ est la donnée des impulsions $p_i = mv_i$. Cet espace est le fibré cotangent à Ω_r^N et est muni de la forme symplectique canonique

$$\omega := \sum_{i=1}^N \sum_{\alpha=1}^3 dp_{i,\alpha} \wedge dx_i^\alpha.$$

Ainsi pour tout $t \in \mathbb{R}$, nous pouvons considérer l'application³ $\Phi_t : T^*\Omega_r^N \longrightarrow T^*\Omega_r^N$ qui, à tout point $(x_*, p_*) \in T^*\Omega_r^N$ associe les positions et impulsions des particules dont les positions et impulsions à l'instant initial 0 étaient (x_*, p_*)

³Dans la définition qui suit, on exclut à nouveau les collisions à plus de trois particules ; cela signifie qu'en réalité Φ_t est défini uniquement sur un sous-ensemble dense de $T^*\Omega_r^N$, celui des conditions initiales qui précisément ne donnent pas lieu à ces collisions.

et que l'on a laissé évoluer pendant le temps t . La famille des homéomorphismes $(\Phi_t)_{t \in \mathbb{R}}$ peut être considérée comme le flot du champ de vecteur singulier ξ . Deuxièmement ce champ de vecteur singulier est *hamiltonien*.

2.1.3 Rappels sur la mécanique hamiltonienne

Rappelons que, si \mathcal{N} est une variété différentielle de dimension n , la variété fibré cotangent $T^*\mathcal{N}$ est munie de la forme symplectique canonique $\omega = \sum_{\alpha=1}^n dp_\alpha \wedge dx^\alpha$. A toute fonction $H \in \mathcal{C}^2(\mathcal{N}, \mathbb{R})$, on associe l'unique champ de vecteur ξ_H sur $T^*\mathcal{N}$, défini par

$$\forall V \in T_{(q,p)}(T^*\mathcal{N}), \quad \omega(\xi_H, V) + dH(V) = 0.$$

Notant $\xi_H \lrcorner \omega$ la 1-forme définie par $\forall V \in T_{(q,p)}(T^*\mathcal{N}), \xi_H \lrcorner \omega(V) := \omega(\xi_H, V)$, cette relation peut encore s'écrire

$$\xi_H \lrcorner \omega + dH = 0 \quad \text{ou bien} \quad \xi_H = \sum_{\alpha=1}^n \frac{\partial H}{\partial p_\alpha} \frac{\partial}{\partial x^\alpha} - \frac{\partial H}{\partial x^\alpha} \frac{\partial}{\partial p_\alpha}.$$

On dit alors que ξ_H est un *champ de vecteur hamiltonien*. L'équation d'évolution $\dot{\phi} = \xi_H(\phi)$ est le système des équations d'Hamilton. Dans beaucoup de cas, H est obtenu à partir d'une densité lagrangienne $L : T\mathcal{N} \rightarrow \mathbb{R}$ par la transformée de Legendre : alors la transformation

$$\begin{aligned} T\mathcal{N} &\longrightarrow T^*\mathcal{N} \\ (x^\alpha, v^\alpha) &\longmapsto (x^\alpha, p_\alpha) = \left(x^\alpha, \frac{\partial L}{\partial v^\alpha}(x, v)\right) \end{aligned}$$

est un difféomorphisme et $H(x, p) + L(x, v) = p_\alpha v^\alpha, \forall x, v, p$ tel que $p_\alpha = \frac{\partial L}{\partial v^\alpha}(x, v)$. De plus $t \mapsto \gamma(t)$ est un point critique⁴ de $\int_{\mathbb{R}} L(\gamma(t), \dot{\gamma}(t)) dt$ si et seulement si $t \mapsto (\gamma(t), \frac{\partial L}{\partial v}(\gamma(t), \dot{\gamma}(t)))$ est solution des équations de Hamilton.

Une propriété importante des équations de Hamilton est que le flot $e^{t\xi_H}$ du champ vecteur ξ_H préserve la forme symplectique. Autrement dit, si Σ est une portion de surface dans $T^*\mathcal{N}$ (l'image du plongement d'un ouvert de

⁴et donc une solution des équations d'Euler-Lagrange $\frac{d}{dt} \left(\frac{\partial L}{\partial v^\alpha}(\gamma(t), \dot{\gamma}(t)) \right) = \frac{\partial L}{\partial x^\alpha}(\gamma(t), \dot{\gamma}(t))$

\mathbb{R}^2), alors $\int_{\Sigma} \omega = \int_{e^{t\xi_H}(\Sigma)} \omega$. Pour montrer cela on remarque d'abord que $\int_{e^{t\xi_H}(\Sigma)} \omega = \int_{\Sigma} (e^{t\xi_H})^* \omega$. Ensuite on démontre que

$$Lie_{\xi_H} \omega := \lim_{t \rightarrow 0} \frac{(e^{t\xi_H})^* \omega - \omega}{t} = 0,$$

qui entraîne bien le résultat. Cette dernière identité est immédiate à partir du moment où l'on dispose de la formule

$$Lie_{\xi_H} \omega = d(\xi_H \lrcorner \omega) + \xi_H \lrcorner (d\omega),$$

où \lrcorner est le produit intérieur d'une forme par un vecteur. En effet, sachant que $\xi_H \lrcorner \omega = -dH$ et $d\omega = 0$, on a tout de suite $Lie_{\xi_H} \omega = d(-dH) + \xi_H \lrcorner (d\omega) = 0$.

Ici, on peut invoquer deux raisons pour justifier le fait que le champ singulier ξ est « hamiltonien » (bien qu'il ne soit pas régulier). Une première raison est que le mouvement de réflexion d'une particule est variationnel (principe de Hénou). Une deuxième raison, plus satisfaisante sur le plan mathématique, consiste à régulariser le champ ξ , c'est à dire montrer qu'il est limite d'une famille de champs de vecteur hamiltoniens. Nous pouvons pour cela considérer l'hamiltonien sur $T^*\Omega_{\star}^N$:

$$H_{\varepsilon}(x_*, p_*) = \sum_{i=1}^N \sum_{\alpha=1}^3 \frac{|p_{i,\alpha}|^2}{2m} + \frac{\varepsilon^2}{\text{dist}(x, \partial\Omega_{\star}^N)^2},$$

où $\varepsilon > 0$ est censé tendre vers 0.

Exercice 2 Calculer $\xi_{\varepsilon} := \xi_{H_{\varepsilon}}$. Montrer que lorsque $\varepsilon \rightarrow 0$, $\xi_{H_{\varepsilon}}$ tend ξ uniformément sur tout compact de $T^*\Omega_{\star}^N$.

Faire une étude près du bord dans le cas où l'on remplace Ω_{\star}^N par le demi-espace \mathbb{R}_+^n : montrer que, dans $T^*\mathbb{R}_+^n$, les trajectoires hamiltoniennes de ξ_{ε} se projettent sur \mathbb{R}_+^n en des branches d'hyperboles et que, lorsque $\varepsilon \rightarrow 0$, ces trajectoires convergent vers celles de ξ .

A partir de cette étude, on peut donc conclure que la famille des homéomorphismes $(\Phi_t)_{t \in \mathbb{R}}$ préserve la forme symplectique ω .

Le crochet de Poisson — Si $F, H : T^*\mathcal{N} \longrightarrow \mathbb{R}$ sont deux fonctions de classe \mathcal{C}^k (pour $k \in \mathbb{N}$ tel que $k \geq 1$), nous définissons le crochet de Poisson de F et H comme étant la fonction $\{F, H\} : T^*\mathcal{N} \longrightarrow \mathbb{R}$ définie par

$$\{F, H\} := \omega(\xi_H, \xi_F) = \sum_{\alpha=1}^n \frac{\partial H}{\partial p_\alpha} \frac{\partial F}{\partial x^\alpha} - \frac{\partial H}{\partial x^\alpha} \frac{\partial F}{\partial p_\alpha}.$$

Il est immédiat que $\{F, H\}$ est de classe \mathcal{C}^{k-1} . Cette quantité apparaît naturellement lorsque l'on considère une solution (x, p) des équations de Hamilton pour $H : \frac{d}{dt}(x(t), p(t)) = \xi_H(x(t), p(t))$, et que l'on calcule le taux de variation de $F(x(t), p(t))$ au cours du temps, c'est à dire

$$\frac{d}{dt}(F(x(t), p(t))) = dF_{(x(t), p(t))}(\xi_H(x(t), p(t))).$$

Un calcul très simple (*Exercice*) donne

$$\frac{d}{dt}(F(x(t), p(t))) = \{F, H\}(x(t), p(t)).$$

Deux propriétés fondamentales du crochet de Poisson sont qu'il est antisymétrique :

$$\{H, F\} + \{F, H\} = 0$$

et qu'il satisfait l'identité de Jacobi :

$$\{F, \{G, H\}\} + \{G, \{H, F\}\} + \{H, \{F, G\}\} = 0.$$

Dans le cas où le problème variationnel est invariant au cours du temps, c'est à dire où le lagrangien ou l'hamiltonien sont indépendants du temps, l'antisymétrie du crochet de Poisson implique que, si $\frac{d}{dt}(x(t), p(t)) = \xi_H(x(t), p(t))$,

$$\frac{d}{dt}H(x(t), p(t)) = \{H, H\}(x(t), p(t)) = 0,$$

c'est à dire la conservation de l'énergie.

2.2 Le chaos ou comment les lois du hasard interviennent dans un système déterministe

Dans les deux paragraphes qui précèdent, nous avons présenté un modèle microscopique d'un gaz qui entre parfaitement dans le cadre de la mécanique

classique, c'est-à-dire dont la solution, du point de vue mathématique, se ramène à la résolution d'une équation différentielle. Dans l'esprit du texte de Laplace cité à la fin du premier chapitre, on peut dire que la connaissance des conditions initiales du système que nous étudions doit permettre de calculer complètement son évolution.

Nous allons voir dans ce paragraphe que ceci est impossible pour deux raisons principales. Nous verrons aussi comment il faut modifier notre approche mathématique pour pouvoir faire des prédictions pertinentes.

La première raison est très simple : l'équation différentielle qu'il nous faut résoudre est une équation du premier ordre en $6N$ variable, où N est le nombre de particules de notre gaz. Cela est tout simplement hors de portée de la technologie humaine et le restera sans doute encore un certain temps. La seconde raison concerne l'état initial : il est impossible de décrire l'état de 10^{23} particules à un instant donné⁵. On pourrait cependant penser qu'il est possible d'en décrire une approximation et, si l'évolution temporelle était assez stable, c'est-à-dire si elle gommait de petites erreurs faites sur l'état initial au fur et à mesure que le temps passe, cela pourrait être suffisant. C'est malheureusement exactement le contraire qui se passe.

Pour illustrer le second problème (il n'y a guère besoin d'illustrer le premier !), nous allons étudier un exemple très simple, en dimension 2 : le mouvement de réflexion d'une particule piégée à l'intérieur d'un disque circulaire et se réfléchissant sur le bord de D^2 selon la loi (2.3). Notons $D^2 := \{x + iy \in \mathbb{C} / x^2 + y^2 < 1\}$. Alors l'ensemble des trajectoires est

$$\mathcal{E} := \{t \mapsto e^{i\beta} \gamma_\theta(ct) / \theta \in]0, \pi], \beta \in [0, 2\pi[, c \in \mathbb{R}\},$$

où

$$\forall n \in \mathbb{Z}, \forall t \in \left[2n \sin \frac{\theta}{2}, 2(n+1) \sin \frac{\theta}{2} \right], \quad \gamma_\theta(t) = e^{in\theta} \left(1 + i \left(t - 2n \sin \frac{\theta}{2} \right) e^{i\theta/2} \right).$$

(*Exercice.*) La solution γ_θ est celle qui passe en 1 en $t = 0$ et décrit le segment de droite joignant 1 à $e^{i\theta}$ avec une vitesse égale à 1. Le paramètre c est la

⁵Si on est capable de stocker un giga-octet dans un gramme de matière (barrette de mémoire vive par exemple), il faut de l'ordre de un milliard de tonnes de matière pour noter un seul état du système.

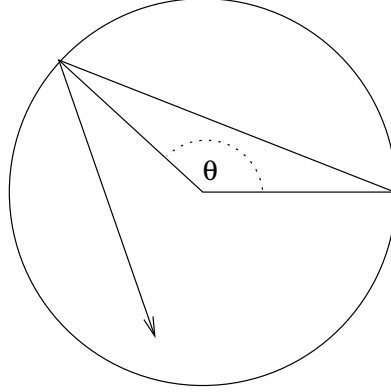


FIG. 2.1 – Une boule de billard dans un disque

célérité et β est une sorte de phase. Enfin la vitesse de γ_θ est donnée par

$$\forall n \in \mathbb{Z}, \forall t \in \left[2n \sin \frac{\theta}{2}, 2(n+1) \sin \frac{\theta}{2} \right], \quad \dot{\gamma}_\theta(t) = ie^{i(n+\frac{1}{2})\theta}.$$

Sur cet exemple, il est facile de se rendre compte que :

- la moindre variation de θ est amplifiée au bout d'un temps suffisamment long. Par exemple, pour deux valeurs $\theta, \theta' \in]0, 2\pi]$ différentes mais *très proches*, les trajectoires dont les positions seraient confondues au point 1 et qui auraient comme vitesses respectives $ie^{i\theta/2}$ et $ie^{i\theta'/2}$ à l'instant 0 finiraient par s'écarter et par avoir des vitesses très différentes au bout d'un temps suffisamment long.
- Si $\theta \in 2\pi\mathbb{Q}$, c'est à dire $\theta = 2\pi p/q$, où $p \in \mathbb{Z}$ et $q \in \mathbb{N}^*$ sont premiers entre eux, les valeurs prises par $\dot{\gamma}_\theta$ sont $\{ie^{i(2k+p)\pi/q} / 0 \leq k < q\}$. Si $\theta \in]0, \pi] \setminus 2\pi\mathbb{Q}$ les valeurs prises par $\dot{\gamma}_\theta$ sont denses dans le cercle S^1 . Notons que le premier cas a une probabilité égale à 0, tandis que le deuxième cas a une probabilité égale à 1. (Cela est physiquement significatif et entraîne ne que dans une expérience pratique, il serait *impossible* de fixer une valeur rationnelle pour θ .)

Pour interpréter cet exemple « physiquement », il est important de distinguer les cas où l'on observe le mouvement sur un temps très court ou sur un temps très long. Dans le premier cas, le système se manifestera de façon déterministe. Dans le deuxième cas, comme il est impossible de prescrire ou de mesurer θ avec une précision infinie, on finira toujours par observer

un comportement chaotique au bout d'un temps suffisamment long. A ce moment, on peut dire que le modèle déterministe n'est plus très utile. En compensation nous pouvons observer le phénomène suivant : si $\theta \in]0, \pi[\setminus 2\pi\mathbb{Q}$, pour tout intervalle ouvert $I \subset S^1$,

$$\lim_{T \rightarrow \infty} \frac{|\{t \in [-T, T] / \dot{\gamma}_\theta(t) \in I\}|}{2T} = \frac{|I|}{2\pi}, \quad (2.4)$$

où ici on désigne la mesure de Lebesgue d'un ensemble mesurable A par $|A|$. (*Exercice.*) Cela signifie que, à défaut de faire des prédictions exactes, nous pouvons faire des prédictions probabilistes. Ici par exemple on pourra observer n'importe quelle vitesse dans S^1 avec une densité de probabilité uniforme.

Des propriétés comme (2.4) sont caractéristiques de l'*ergodicité* de l'évolution temporelle de notre système. Cette ergodicité, lorsqu'elle est présente, est ce qu'il reste de régulier dans une évolution chaotique. Cela nous suggère fortement qu'une description statistique de notre système sera bien plus efficace qu'une description déterministe.

Exercice 3 On considère, pour toute fonction $f \in \mathcal{C}^0([0, 1])$, la limite

$$\mu(f) := \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T f(|\dot{\gamma}_\theta(t)|) dt.$$

Montrer que $\mu(f) = \int_{\sin \frac{\theta}{2}}^1 f(r) \frac{2r dr}{\sqrt{r^2 - \sin^2 \frac{\theta}{2}}}$. Interpréter : quelles prédictions peut-on faire lorsque l'on mesure la distance de $\gamma_\theta(t)$ à 0 connaissant θ ?

Nous voyons sur cet exemple simple que sur des temps très longs, il est impossible de prédire quoi que ce soit de précis sur la direction du vecteur vitesse, et très peu de choses sur la position de la particule⁶. Or il se trouve que dès que l'on remplace le disque par pratiquement n'importe quelle courbe fermée, même très régulière, les phénomènes de chaos seront beaucoup plus importants et concerneront autant les positions que les vitesses : la mesure de probabilité de la position d'une particule sera la mesure uniforme dans le

⁶On pourrait varier ce modèle, par exemple remplacer le disque par l'intérieur d'une ellipse, et on observerait alors des phénomènes similaires. Mais les cercles et les ellipses sont des courbes *très particulières* qui n'existent qu'en approximation dans la nature.

domaine et il en sera de même pour la direction du vecteur vitesse⁷. Ces propriétés d'ergodicité sont difficiles à démontrer mathématiquement en général, mais on peut les vérifier de manière très convaincante par des observations expérimentales (boules de *Loto*) ou par des simulations sur ordinateur (cf. le dernier chapitre du livre de Harthong).

Revenons maintenant au modèle de billard pour N particules : il n'y aura pas de changement qualitatif dans l'émergence du chaos, mais une amplification. Une première raison pour cela est le fait que N est très grand (de l'ordre du nombre d'Avogadro) et que le mouvement des molécules est très rapide (de l'ordre de $1840 \text{ m} \cdot \text{s}^{-1}$ pour l'hydrogène à 273ř K). Par conséquent, même si l'on part d'une configuration « ordonnée », le temps au bout duquel les phénomènes deviendront chaotiques sera *extrêmement court* et de plus, si nous voulons estimer une limite analogue à (2.4) il suffira de prendre un temps d'observation T petit à l'échelle des temps macroscopiques. Une deuxième raison qui est moins immédiate est liée aux propriétés mathématiques (chaotiques) du billard : les chocs entre particules engendrent encore d'avantage d'instabilités dynamiques que les chocs entre les parois. C'est à dire que deux trajectoires différentes qui ont des conditions initiales très proches dans $T^*\Omega_r^N$ s'écartent très rapidement l'une de l'autre (résultats de Y. SINAI). En outre il est impossible de mesurer et de préparer un système physique correspondant à un point particulier (ou même un voisinage extrêmement petit d'un point) dans l'espace de phase $T^*\Omega_r^N$. Pour toutes ces raisons, nous ne pouvons pas décrire de façon déterministe un tel gaz. En revanche une description statistique donnera d'excellents résultats.

En conclusion nous pouvons tirer deux enseignements des considérations précédentes. Le premier est, qu'en dehors des cas exceptionnels (et irréalisables en pratique) où l'enceinte Ω qui enferme le gaz a une géométrie très particulière, il est impossible d'avoir la moindre information précise sur l'état microscopique d'un gaz à l'équilibre, toute information partielle à un instant donné étant immédiatement rendue obsolète par le caractère chaotique de l'évolution de cet état. La seule approche possible du problème est donc une

⁷Seule ici la norme du vecteur vitesse est prédictible, car c'est une constante du mouvement. Noter que si on remplace le domaine bidimensionnel par Ω_r^N comme dans la section précédente et si on interprète le mouvement comme celui d'un système de N particules, la conservation de la norme du vecteur vitesse dans Ω_r^N ne signifie rien d'autre que la conservation de l'énergie cinétique totale du système de particules.

approche statistique.

Le deuxième enseignement vient tempérer le premier : c'est le caractère ergodique de l'évolution microscopique (dans les bons cas), qui se traduit par un phénomène de « balayage uniforme au cours du temps » de l'espace des phases par une trajectoire, illustré par (2.4). Nous aurons besoin au prochain chapitre de supposer que cela est vrai en général (*principe ergodique*).

2.3 Premier usage des probabilités : les gaz dilués

Nous allons commencer par appliquer le « premier enseignement » de la section précédente sur un exemple simple : celui des gaz raréfiés ou dilués. Nous allons ainsi retrouver certaines lois thermodynamiques des gaz parfaits.

Dans ce qui suit, nous supposons pour simplifier que les particules sont ponctuelles (et donc $r = 0$) et nous poserons donc $\Omega' = \Omega$. Imaginons que nous souhaitions calculer une quantité « macroscopique » de la forme

$$A(x_*(t), p_*(t)) := \sum_{i=1}^N a(x_i(t), p_i(t)),$$

où $(x_*, p_*) : \mathbb{R} \longrightarrow T^*\Omega_*^N$ est une solution des équations de Hamilton, $t \in \mathbb{R}$ et a est une fonction sur $T^*\Omega$ (dont les valeurs sont « microscopiques »). Par exemple nous verrons dans la section suivante que si $a(x, p) = \frac{|p|^2}{2m}$, c'est à dire si a est l'énergie cinétique d'une particule, alors $A(x_*(t), p_*(t))$ est la somme de toutes les énergies cinétiques des particules donc (pour un gaz dilué) l'énergie totale du système en t (comme on néglige les chocs, il n'y a pas d'énergie potentielle; noter que, comme l'énergie est conservée, cette quantité ne dépend pas de t). Nous verrons aussi dans la section suivante qu'un autre choix de a permet d'évaluer la force exercée par le gaz sur une portion de la paroi de Ω et donc la pression du gaz.

Comme il est impossible de calculer réellement $A(x_*(t), p_*(t))$, nous allons plutôt en faire une estimation probabiliste, par un calcul de moyenne statistique. Il s'agit donc de construire un espace probabilisé adapté à notre problème. Imaginons que l'on dispose d'un nombre très grand voire infini de

systèmes contenant la même quantité de gaz, « tirés au hasard », que tous les systèmes sont à l'équilibre et identiques du point de vue des grandeurs thermodynamiques. Pour tout ouvert régulier (ou plus généralement tout sous-ensemble mesurable) C de $T^*\Omega_*^N$, on dénombre la proportion de systèmes tels que $(x_*(t), p_*(t)) \in C$ et on note $\mu_t(C)$ cette quantité. En d'autres termes μ_t est une mesure de probabilité sur $T^*\Omega_*^N$ et $\mu_t(C)$ est la probabilité pour que $(x_*(t), p_*(t)) \in C$. Comme le gaz est à l'équilibre il est naturel de supposer que μ_t ne dépend pas de t et donc $\mu_t = \mu$. Nous disposons donc désormais de l'espace probabilisé $(T^*\Omega_*^N, \mu)$, sur lequel la fonction $A(x_*(t), p_*(t))$ est maintenant une variable aléatoire. Nous voulons être capables de calculer des quantités comme sa moyenne, sa variance. Celles-ci dépendent bien entendu de la mesure μ qu'il nous faut essayer de préciser. Dans la suite nous noterons $(x_*, p_*) \simeq (x_*(t), p_*(t))$ et il sera sous-entendu que (x_*, p_*) représente l'état du système à un instant t choisi une fois pour toute.

Soit ρ la densité de μ , de sorte que $d\mu = \rho(x_*, p_*) \prod_{i=1}^N dx_i^1 dx_i^2 dx_i^3 dp_{i,1} dp_{i,2} dp_{i,3}$. L'invariance du problème par rapport au temps⁸ implique aussi que ρ est invariant par le champ de vecteur hamiltonien $\xi_H : d\rho_{(x_*, p_*)}(\xi_H(x_*, p_*)) = \{H, \rho\}(x_*, p_*) = 0$.

Nous faisons maintenant l'hypothèse que le gaz contenu dans une enceinte Ω , isolé et à l'équilibre est *dilué*, ce qui entraîne que les chocs entre particules sont très rares (c'est bien pourquoi par exemple l'air est un excellent isolant). Alors nous pouvons négliger l'influence des chocs et la dynamique du gaz est identique à la superposition des mouvements de N particules qui rebondissent sur les parois de Ω et qui s'ignorent les unes les autres. Une première conséquence de cela est l'approximation $H(x_*, p_*) \simeq \sum_{i=1}^N \frac{|p_i|^2}{2m}$ et donc $\xi_H \simeq \sum_{i=1}^N \sum_{\alpha=1}^3 \frac{p_{i,\alpha}}{m} \frac{\partial}{\partial x_i^\alpha}$ et nous en déduisons que

$$d\rho_{(x_*, p_*)}(\xi_H) = \sum_{i=1}^N \sum_{\alpha=1}^3 \frac{p_{i,\alpha}}{m} \frac{\partial \rho}{\partial x_i^\alpha}(x_*, p_*) = 0. \quad (2.5)$$

Une deuxième conséquence est que, puisque les différentes particules « s'ignorent » entre elles, nous allons faire l'hypothèse (beaucoup moins innocente, comme nous le verrons en 2.6.1) que, pour tout $1 \leq i, j \leq N$ tels que $i \neq j$, les variables aléatoires (x_i, p_i) et (x_j, p_j) sont indépendantes et de même loi. Donc

⁸nous reviendrons sur cette condition de façon plus détaillée au chapitre 3

il existe une densité de probabilité $\phi : T^*\Omega \longrightarrow [0, \infty[$ telle que

$$\rho(x_*, p_*) = \prod_{i=1}^N \phi(x_i, p_i). \quad (2.6)$$

En substituant cette expression dans (2.5) et en divisant par $\rho(x_*, p_*)$, nous obtenons :

$$\sum_{i=1}^N \sum_{\alpha=1}^3 \frac{p_{i,\alpha}}{m\phi(x_i, p_i)} \frac{\partial \phi}{\partial x^\alpha}(x_i, p_i) = 0, \quad \forall (x_*, p_*) \in T^*\Omega_*^N.$$

Cela n'est possible que si $\frac{\partial \phi}{\partial x^\alpha} = 0$. Nous obtenons ainsi que $\phi(x, p) = \psi(p)/|\Omega|$, où $\psi : \mathbb{R}^3 \longrightarrow [0, \infty[$ est une densité de probabilité. Cela signifie que les particules se répartissent uniformément en espace.. Enfin, sur la base des considérations de la section précédente et, compte tenu du fait que les équations du mouvement sont invariantes par rotation dans l'espace \mathbb{R}^3 , nous pouvons aussi faire l'hypothèse que ψ ne dépend que de la norme de p et poser $\psi(p) = f(|p|^2)$, où f est une densité de probabilité sur $[0, \infty[$. Ainsi

$$d\mu = \prod_{i=1}^N \frac{f(|p_i|^2)}{|\Omega|} dx_i^1 dx_i^2 dx_i^3 dp_{i,1} dp_{i,2} dp_{i,3}.$$

L'hypothèse d'indépendance et d'identique distribution entre les particules que nous avons faite ci-dessus est extrêmement importante. Considérons par exemple à nouveau la variable $A = A(x_*, p_*)$. Elle est par définition la *somme* des variables aléatoires *microscopiques* $a_i, i = 1 \dots N$, où a_i est définie par $a_i(x_*, p_*) := a(x_i, p_i)$, ou encore la *moyenne* des variables aléatoires *macroscopiques* $A_i(x_*, p_*) := Na_i(x_*, p_*) = Na(x_i, p_i)$. Cela entraî ne par exemple que

$$\begin{aligned} \langle A(x_*, p_*) \rangle_\mu &= \int_{T^*\Omega_*^N} A(x_*, p_*) d\mu \\ &= \sum_{i=1}^N \int_{T^*\Omega_*^N} a_i(x_*, p_*) d\mu \\ &= \sum_{i=1}^N \langle a_i(x_*, p_*) \rangle_\mu = \frac{1}{N} \sum_{i=1}^N \langle A_i(x_*, p_*) \rangle_\mu. \end{aligned} \quad (2.7)$$

Nous avons pour l'instant seulement écrit le fait que l'espérance d'une somme de variables aléatoires est la somme de leurs espérances. Mais l'hypothèse d'équidistribution assure que la moyenne $\langle a_i(x_*, p_*) \rangle_\mu$ ne dépend pas de i : c'est en fait, pour chaque i , la moyenne de la fonction a sur le i -ième espace $T^*\Omega$, sous la mesure image de μ par l'application i -ième coordonnée $T^*\Omega_*^N \longrightarrow T^*\Omega$. Cette mesure s'écrit plus explicitement

$$\langle a_i(x_*, p_*) \rangle_\mu = \int_{T^*\Omega} a(x, p) \frac{f(|p|^2)}{|\Omega|} dx^1 dx^2 dx^3 dp_1 dp_2 dp_3.$$

L'hypothèse d'indépendance, qu'il faut distinguer de celle d'équidistribution⁹, nous permet d'estimer la variance de $A = A(x_*, p_*)$. Rappelons qu'elle est définie par $\text{Var}(A) = \langle (A - \langle A \rangle)^2 \rangle$. On vérifie aisément que les variances de variables aléatoires indépendantes s'ajoutent et il est par ailleurs immédiat que la variance est une fonction homogène de degré 2. Ainsi,

$$\text{Var}(A) = \text{Var}\left(\sum_{i=1}^N a_i\right) = \sum_{i=1}^N \text{Var}(a_i) = \frac{1}{N^2} \text{Var}(Na).$$

Cela nous montre, puisque N est gigantesque, que A fluctue très peu autour de sa valeur moyenne. Par exemple, l'inégalité de Chebyshev assure que

$$P(|A - \langle A \rangle| > \alpha) \leq \frac{1}{\alpha^2} \text{Var}(A) = \frac{1}{N^2 \alpha^2} \text{Var}(Na). \quad (2.8)$$

La valeur observée de A a donc de très grandes chances d'être proche de $\langle A \rangle$. Nous verrons cependant un peu plus tard des résultats beaucoup plus fins dans cette direction, et nous établirons que A a une probabilité exponentiellement faible en N de s'écarter sensiblement de $\langle A \rangle$.

⁹Pour bien les distinguer, il est peut-être utile de s'apesantir un peu sur la structure probabiliste de l'espace des phases. L'espace $T^*\Omega_*^N$ est muni de N projections qui, physiquement, extraient de l'état global du système l'état d'une particule donnée : pour tout k entre 1 et N , nous avons donc $\pi_k : T^*\Omega_*^N \longrightarrow T^*\Omega$. La donnée de la mesure μ sur $T^*\Omega_*^N$ détermine les N mesures images $(\pi_k)_*\mu$ sur $T^*\Omega$ qui correspondent à chacune des N particules. L'hypothèse d'équirépartition revient à dire que toutes les mesures $(\pi_k)_*\mu$ sont égales. Si les particules sont indiscernables, cela semble être une hypothèse inévitable. L'hypothèse d'indépendance revient, elle, à dire que μ est égale au produit $\otimes_k (\pi_k)_*\mu$. Nous réutiliserons ce formalisme — qui ici peut paraître un peu lourd — lors de l'étude de la distribution des vitesses de Maxwell.

2.4 Théorie cinétique du gaz parfait

Nous allons maintenant appliquer les considérations précédentes au calcul de l'énergie et de la pression d'un gaz dilué à l'équilibre thermodynamique et retrouver ainsi la loi des gaz parfaits.

L'énergie interne est la somme des énergies cinétiques de toutes les molécules :

$$U(x_*, p_*) = \sum_{i=1}^N \frac{|p_i|^2}{2m},$$

et correspond donc à une fonction macroscopique du type précédent avec $a(x, p) := |p|^2/2m$. La relation (2.7) donne

$$\begin{aligned} \frac{1}{N}U(x_*, p_*) &\simeq \frac{1}{N}\langle U(x_*, p_*) \rangle_\mu \\ &= \int_{T^*\Omega} \frac{|p|^2}{2m} \frac{f(|p|^2)}{|\Omega|} dx^1 dx^2 dx^3 dp_1 dp_2 dp_3 \\ &= \int_{\mathbb{R}^3} \frac{|p|^2}{2m} f(|p|^2) dp_1 dp_2 dp_3. \end{aligned} \quad (2.9)$$

Pour le calcul de la pression, considérons une portion $\widehat{\Delta s}$ du bord $\partial\Omega$ et calculons la valeur moyenne $\overline{F}_{\Delta t}$ au cours d'un intervalle de temps Δt de la force résultant des chocs des particules sur $\widehat{\Delta s}$. Nous noterons $\Delta s := |\widehat{\Delta s}|$ l'aire de $\widehat{\Delta s}$. Nous sommes obligés ici de considérer une moyenne sur un intervalle de temps car les chocs ont lieu à des temps discrets. Nous choisissons $\widehat{\Delta s}$ petit par rapport aux dimensions macroscopiques de l'enceinte contenant le gaz, de sorte que nous pouvons assimiler avec une bonne approximation $\widehat{\Delta s}$ à un morceau de surface plane : nous notons alors ν le vecteur normal (extérieur) à $\widehat{\Delta s}$ (nous supposerons aussi dans la suite que ν est $(0, 0, 1)$). De même nous choisissons Δt relativement bref par rapport aux temps macroscopiques, ce qui autorisera à appliquer le résultat de notre calcul pour la définition d'une pression « instantanée » pour des processus d'évolution réversible¹⁰.

¹⁰Noter que, dans la plupart des situations réalistes, le nombre gigantesque de particules (de l'ordre du nombre d'Avogadro) et leur très grandes vitesses permet de supposer en même temps que $\widehat{\Delta s}$ est grand par rapport aux échelles moléculaires et Δt est grand par rapport aux temps de parcours envisagés, de sorte que le nombre de particules qui vont frapper $\widehat{\Delta s}$ pendant l'intervalle de temps Δt est très grand. Nous aurons besoin de cette

Nous commençons par une question simple : supposons qu'une particule se trouve en $x \in \Omega$ avec une impulsion p à un instant t_0 . Quelle force exerce-t-elle sur $\widehat{\Delta s}$ entre t_0 et $t_0 + \Delta t$? Cette force n'est non nulle que si la particule frappe $\widehat{\Delta s}$ pendant l'intervalle $[t_0, t_0 + \Delta t]$, c'est à dire si

$$\begin{aligned} (x, p) \in C_{\Delta t} &:= \{(x, p) \in T^*\Omega / \exists t_c \in [0, \Delta t], x + \frac{p}{m}t_c \in \widehat{\Delta s}\} \\ &= \{(y - \frac{p}{m}t_c, p) \in T^*\Omega / p_3 > 0, y \in \widehat{\Delta s}, t_c \in [0, \Delta t]\}. \end{aligned}$$

Et alors, en supposant le choc élastique et suivant (2.3), la force est $F(t) = 2p_3\delta(t - t_c)$ (il s'agit ici de l'intensité de la force, sa direction est toujours ν). La moyenne de cette force sur $[t_0, t_0 + \Delta t]$ est donc $2\frac{p_3}{\Delta t}$. Ainsi, la moyenne sur

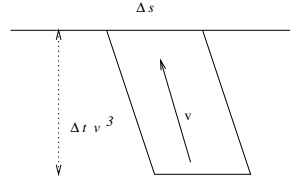


FIG. 2.2 – Le calcul de la pression pour un gaz parfait

$[t_0, t_0 + \Delta t]$ de la force exercée par une particule de condition initiale (x, p) sur $\widehat{\Delta s}$ est $a_{\Delta t}(x, p)$, où :

- si $(x, p) \in C_{\Delta t}$, $a_{\Delta t}(x, p) = 2\frac{p_3}{\Delta t}$
- si $(x, p) \in T^*\Omega \setminus C_{\Delta t}$, $a_{\Delta t}(x, p) = 0$.

Donc pour toute solution $(x_*, p_*) : \mathbb{R} \longrightarrow T^*\Omega_*^N$ des équations de Hamilton représentant l'évolution microscopique d'un gaz dilué, la moyenne sur $[t_0, t_0 + \Delta t]$ de la force exercée par le gaz est

$$\overline{F}_{\Delta t} = \sum_{i=1}^N a_{\Delta t}(x_i, p_i),$$

où $a_{\Delta t}$ a été défini précédemment. Appliquant le principe obtenu¹¹ dans la

hypothèse un peu plus loin (voir la note de bas de page 11). Par ailleurs, moyennant une hypothèse « raisonnable » d'*ergodicité* dont nous reparlerons plus loin, la valeur du nombre de particules divisée par $\Delta t \times \Delta s$ sera un nombre quasi indépendant du temps, de la position et de Δt et Δs (un peu dans le même sens que dans (2.4)).

¹¹pour que l'application de ce principe soit valable, il est nécessaire que le nombre des particules concernées — c'est à dire des particules i telles que $(x_i, p_i) \in C_{\Delta t}$ — soit très grand. Cela implique que Δt doit être choisi pas trop petit, de façon à ce que $N\Delta t$ soit un grand nombre.

section précédente, nous obtenons

$$\begin{aligned}
\overline{F}_{\Delta t} &\simeq \langle \overline{F}_{\Delta t} \rangle_{\mu} \\
&= N \int_{T^*\Omega} a_{\Delta t}(x, p) \frac{f(|p|^2)}{|\Omega|} dx^1 dx^2 dx^3 dp_1 dp_2 dp_3 \\
&= N \int_{C_{\Delta t}} 2 \frac{p_3}{\Delta t} \frac{f(|p|^2)}{|\Omega|} dx^1 dx^2 dx^3 dp_1 dp_2 dp_3.
\end{aligned}$$

Par le changement de variable $(x, p) = (y - \frac{p}{m}t, p)$ cela donne, notant $\mathbb{R}_+^3 := \{(p_1, p_2, p_3) \in \mathbb{R}^3 / p_3 > 0\}$,

$$\begin{aligned}
\overline{F}_{\Delta t} &= N \int_{\mathbb{R}_+^3} \frac{f(|p|^2)}{|\Omega|} \frac{2p_3}{\Delta t} dp_1 dp_2 dp_3 \int_{\widehat{\Delta s}} dy^1 dy^2 \int_0^{\Delta t} \frac{p_3}{m} dt \\
&= 2N \frac{\Delta s}{|\Omega|} \int_{\mathbb{R}_+^3} f(|p|^2) \frac{(p_3)^2}{m} dp_1 dp_2 dp_3.
\end{aligned}$$

Exploitant d'abord le fait que la quantité intégrée est invariante par changement de signe de p_3 , puis que $\int_{\mathbb{R}^3} f(|p|^2) (p_1)^2 = \int_{\mathbb{R}^3} f(|p|^2) (p_2)^2 = \int_{\mathbb{R}^3} f(|p|^2) (p_3)^2 = \frac{1}{3} \int_{\mathbb{R}^3} f(|p|^2) |p|^2$, nous en déduisons que

$$\overline{F}_{\Delta t} = N \frac{\Delta s}{|\Omega|} \int_{\mathbb{R}^3} f(|p|^2) \frac{(p_3)^2}{m} dp_1 dp_2 dp_3 = N \frac{\Delta s}{3|\Omega|} \int_{\mathbb{R}^3} f(|p|^2) \frac{|p|^2}{m} dp_1 dp_2 dp_3.$$

Notons que cette quantité est indépendante de Δt . Ainsi la pression $p := \overline{F}_{\Delta t} / \Delta s$ est

$$p = \frac{2N}{3V} \int_{\mathbb{R}^3} f(|p|^2) \frac{|p|^2}{2m} dp_1 dp_2 dp_3 = \frac{2N}{3V} \int_{\mathbb{R}^3} f(|mv|^2) \frac{m}{2} |v|^2 m^3 dv^1 dv^2 dv^3, \tag{2.10}$$

où $V = |\Omega|$. D'une part, nous constatons que $pV = N \frac{2}{3} \int_{\mathbb{R}^3} f(|p|^2) \frac{|p|^2}{2m} dp_1 dp_2 dp_3$, ce qui, en identifiant $\frac{2}{3} \int_{\mathbb{R}^3} f(|p|^2) \frac{|p|^2}{2m} dp_1 dp_2 dp_3$ avec kT , redonne la relation d'état des gaz parfaits. D'autre part, en comparant avec (2.9), nous obtenons

$$p = \frac{2}{3V} U.$$

Et donc en utilisant la relation $pV = NkT$, nous retrouvons le fait que

$$U = \frac{3}{2} NkT.$$

2.5 La distribution de Maxwell

Nous allons à présent essayer de deviner, par des arguments heuristiques, une expression plus précise de la densité f . Nous faisons l'hypothèse supplémentaire suivante : les variables aléatoires p_1 , p_2 et p_3 sont *indépendantes* deux à deux et de même loi. Soit $f_\alpha(p_\alpha) dp_\alpha$ la mesure de probabilité de la variable p_α . C'est la projection de $f(|p|^2) dp_1 dp_2 dp_3$ sur l'axe des p_α :

$$\forall \alpha = 1, 2, 3, \forall \phi \in L^1(\mathbb{R}), \quad \int_{\mathbb{R}} \phi(p_\alpha) f_\alpha(p_\alpha) dp_\alpha = \int_{\mathbb{R}^3} \phi(p_\alpha) f(|p|^2) dp_1 dp_2 dp_3,$$

on en déduit que f_α est une fonction régulière. Le fait que p_1 , p_2 et p_3 aient la même loi entraîne ne que $f_1 = f_2 = f_3$. Comme on a aussi manifestement $f_\alpha(-p_\alpha) = f_\alpha(p_\alpha)$, nous pouvons donc poser $f_\alpha(p_\alpha) = g((p_\alpha)^2)$. L'hypothèse d'indépendance des lois des composantes de l'impulsion signifie alors que

$$g((p_1)^2)g((p_2)^2)g((p_3)^2) = f((p_1)^2 + (p_2)^2 + (p_3)^2), \quad \forall p \in \mathbb{R}^3.$$

Exercice 4 *Démontrer qu'il existe des constantes $C > 0$ et $a \in \mathbb{R}$ telles que*

$$f(|p|^2) = Ce^{-a|p|^2}. \quad (2.11)$$

Comme nous savons que f est une densité de probabilité, et donc en particulier que $\int_{\mathbb{R}^3} f(|p|^2) dp_1 dp_2 dp_3 = 1$, il est nécessaire que $a > 0$. La mesure $f(|p|^2) dp_1 dp_2 dp_3$ est donc *gaussienne*. De plus la normalisation précédente entraîne ne que $C = \sqrt{\frac{a^3}{\pi}}$ (*Exercice*). Nous pouvons de plus calculer U en fonction de f , pour en déduire une relation supplémentaire sur a . Utilisant (2.9),

$$\begin{aligned} \frac{U}{3} &= \frac{N}{2} \int_{\mathbb{R}^3} f(|p|^2) \frac{|p|^2}{3m} dp_1 dp_2 dp_3 \\ &= \frac{N}{2m} \sqrt{\frac{a}{\pi}}^3 \int_{\mathbb{R}^3} e^{-a|p|^2} (p_1)^2 dp_1 dp_2 dp_3 \\ &= \frac{N}{2m} \sqrt{\frac{a}{\pi}}^3 \left(\int_{\mathbb{R}} e^{-aw^2} dw \right)^2 \int_{\mathbb{R}} e^{-aw^2} w^2 dw \\ &= \frac{N}{2m} \sqrt{\frac{a}{\pi}} \int_{\mathbb{R}} e^{-aw^2} w^2 dw = \frac{N}{m} \sqrt{\frac{a}{\pi}} \int_0^\infty e^{-aw^2} w^2 dw. \end{aligned}$$

Posant $s = aw^2$, la dernière intégrale peut se calculer ainsi¹² :

$$\int_0^\infty e^{-aw^2} w^2 dw = \frac{1}{2a\sqrt{a}} \int_0^\infty e^{-s} \sqrt{s} ds = \frac{1}{2a\sqrt{a}} \Gamma\left(\frac{3}{2}\right) = \frac{1}{2a\sqrt{a}} \frac{\sqrt{\pi}}{2}.$$

on obtient donc

$$\frac{U}{3} = \frac{N}{4ma}.$$

Comme on sait par ailleurs que $U = \frac{3NkT}{2}$, on en déduit que :

$$\frac{N}{4ma} = \frac{NkT}{2} \iff a = \frac{1}{2mkT}.$$

Nous avons obtenu la distribution de Maxwell pour une particule, dont la densité est :

$$\phi(x, v) = \frac{1}{|\Omega| \sqrt{2m\pi kT}^3} e^{-|p|^2/2mkT}.$$

Ce résultat fournit une première interprétation statistique de la température. Il est remarquable que l'exposant de la fonction gaussienne soit

$$-\frac{E_c}{kT} \quad \text{où } E_c = \frac{|p|^2}{2m} \text{ est l'énergie cinétique.}$$

Nous en déduisons, en utilisant (2.6) et $H(x_*, p_*) \simeq \sum_{i=1}^N |p_i|^2/2m$, que la mesure μ sur $T^*\Omega_*^N$ a pour densité

$$\rho(x_*, p_*) = \frac{1}{|\Omega|^N \sqrt{2m\pi kT}^{3N}} e^{-H(x_*, p_*)/kT}. \quad (2.12)$$

2.6 Interprétation statistique de l'irréversibilité

Nous avons vu dans ce qui précède qu'une description probabiliste d'un gaz permettait de faire des prédictions efficaces sur son état macroscopique. On pourrait néanmoins rester un peu frustré de cette approche et considérer qu'elle n'est qu'un pis-aller destiné à pallier temporairement nos trop faibles compétences calculatoires. Ce qui suit va montrer qu'au contraire l'approche probabiliste permet de rendre compte de certains phénomènes macroscopiques qui semblent étrangers à l'approche déterministe hamiltonienne.

¹²On utilise la fonction $\Gamma(z) := \int_0^\infty e^{-x} x^{z-1} dx$ et la relation $\Gamma(z) = (z-1)\Gamma(z-1)$ si $z > 1$.

2.6.1 Le paradoxe de Loschmidt

Considérons une certaine quantité de gaz enfermée dans la moitié gauche d'une enceinte divisée en deux parties.

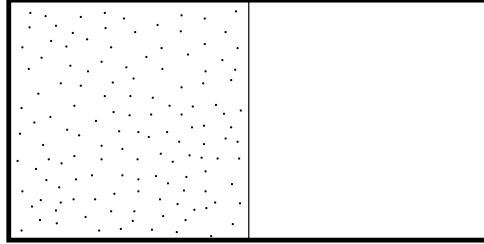


FIG. 2.3 – Le gaz est confiné dans une moitié de l'enceinte.

Si l'on retire brusquement la séparation, le gaz va immédiatement se répandre dans toute l'enceinte. L'expérience quotidienne nous assure que jamais il ne reviendra spontanément se loger dans la demi-enceinte où il était initialement confiné. Si donc nous regardons un film de l'expansion du gaz après que la séparation ait été retirée, nous pourrions à coup sûr dire si ce film est passé à l'endroit ou à l'envers. En revanche, si nous pouvions voir un film du mouvement d'un petit nombre (trois ou quatre) des molécules du gaz prises séparément, rien ne nous permettrait de décider si nous voyons le film à l'envers ou à l'endroit. Autrement dit, bien que les lois du mouvement des particules qui constituent le gaz soient rigoureusement invariantes par inversion du sens d'écoulement du temps, nous observons à l'échelle macroscopique un phénomène irréversible. Cette contradiction apparente est connue sous le nom de *paradoxe de Loschmidt*¹³.

La réponse à cette objection a été fournie par Boltzmann lui-même, quelques années plus tard. Il est indéniable que si, après avoir ôté la cloison et laissé au gaz le temps d'atteindre son état d'équilibre — où il occupe toute l'enceinte — nous transformions instantanément, par un procédé inconnu, l'ensemble de toutes les vitesses des particules en leur opposé, alors, au bout d'un laps de temps égal à celui écoulé depuis le retrait de la cloison, le gaz irait se

¹³Joseph Loschmidt publia en 1876, conformément à la tradition de l'époque, une série de critiques sur l'interprétation moléculaire de l'entropie que L. Boltzmann avait exposée en 1872.

concentrer entièrement dans la moitié gauche de l'enceinte. Cependant, un tel procédé n'existe pas et même s'il existait, la moindre erreur dont serait entachée l'inversion des vitesses se propagerait, par les collisions entre particules, extrêmement vite à l'ensemble du gaz et conduirait irrémédiablement à un échec de l'expérience¹⁴.

La seule expérience qu'il soit possible de faire est la suivante : préparer le gaz dans son état d'équilibre dans l'enceinte complète, et attendre qu'il veuille bien aller se loger dans la moitié gauche. La seule chose que nous puissions contrôler est l'état macroscopique du gaz, mais nous ne pouvons certainement pas prescrire son état microscopique, c'est-à-dire la position et la vitesse de chaque particule. Et sans doute, parmi tous les états microscopiques qui correspondent à l'état macroscopique où le gaz occupe toute l'enceinte, il y en a un certain nombre qui, si on les laisse évoluer, vont conduire à un confinement du gaz dans une moitié de l'espace disponible. Mais combien y en a-t-il ? Lorsque nous préparons l'état macroscopique initial, quelle probabilité y a-t-il que nous préparions du même coup un de ces états microscopiques ? Le raisonnement grossier suivant nous donne un ordre d'idée de la réponse.

Cherchons dans l'ensemble \mathcal{E} des états microscopiques correspondant à une occupation de tout le volume le nombre de ceux qui évoluent vers un confinement du gaz en un temps T fixé. L'ensemble \mathcal{C} de ces états est en bijection avec l'ensemble $\mathcal{E}_{1/2}$ des états microscopiques qui correspondent à une occupation de la moitié gauche de l'enceinte. En effet, l'écoulement du temps définit un flot sur l'ensemble des états microscopiques et l'opérateur qui correspond à l'écoulement d'un temps T , composé avec l'opération consistant à retourner toutes les vitesses, envoie injectivement $\mathcal{E}_{1/2}$ sur \mathcal{C} . Le nombre d'états microscopiques conduisant à un confinement est donc du même ordre de grandeur que le nombre d'états microscopiques correspondant à l'occupation d'une moitié de l'espace¹⁵. Pour chaque particule, la contrainte de se

¹⁴Pour des explications plus détaillées à ce sujet, voir le chapitre 14 du livre de J. Harthong.

¹⁵noter à ce propos que si l'on faisait une expérience statistique analogue à celle décrite en 2.3, mais en ne puisant que des échantillons dans $\mathcal{E}_{1/2}$ (au lieu de \mathcal{E}), on obtiendrait une mesure de probabilité $\mu_{1/2}$ sur $T^*\Omega_*^N$ différente de la mesure μ de Maxwell (une loi de probabilité conditionnelle). Il en résulte que la mesure μ_C est aussi différente de μ . Et en particulier si T est suffisamment grand, μ_C sera un exemple de mesure pour laquelle les lois de distribution de chaque particule sont identiques, mais ces lois ne sont pas indépendantes

trouver dans une moitié de l'enceinte divise par 2 le nombre d'états possibles. Pour un ensemble de N particules, on a donc

$$\frac{|\mathcal{C}|}{|\mathcal{E}|} \approx 2^{-N}.$$

Cela répond à l'objection de Loschmidt : on ne peut absolument pas contrôler l'état microscopique que l'on prépare et on ne peut guère s'en remettre au hasard pour le faire à notre place, à moins bien sûr d'avoir la possibilité de faire un nombre d'essais de l'ordre de $2^{10^{23}}$. Cela nous éclaire aussi sur le sens de l'expression « processus irréversible » : il s'agit d'un processus dont le déroulement à rebours est non pas *impossible* en ce sens qu'il contredirait une quelconque loi physique, mais tellement peu probable qu'il est absolument *in-vraisemblable*. Pour bien comprendre ce point, on peut se demander : à partir de combien de particules vues sur un film se rassembler dans une moitié de l'enceinte estimerait-on que le film a été projeté à l'envers ?

2.6.2 Fluctuations de la densité d'un gaz

Nous allons maintenant nous intéresser de façon plus générale aux fluctuations du nombre de particules qui occupent une région donnée à l'intérieur d'une enceinte remplie de gaz.

Considérons donc une enceinte contenant N molécules d'un gaz parfait. D'après la forme de la distribution de Maxwell, chacune de ces N particules est répartie uniformément dans le volume de l'enceinte, indépendamment de toutes les autres¹⁶. Divisons par la pensée notre enceinte, dont nous noterons V le volume, en 2 sous-enceintes de volumes respectifs V_1 et V_2 . Notons N_1 et N_2 les nombres respectifs de particules dans chacune des sous-enceintes. La

(et donc (2.6) est faux).

¹⁶Bien entendu, dans la mesure où nous raisonnons dans le cadre de la mécanique newtonienne, chaque particule a une position bien définie. Cependant, il est impossible de connaître l'état microscopique du gaz à un instant donné. Le sens de la description probabiliste de la position des particules est le suivant. Considérons un très grand nombre (même un nombre infini) d'enceintes identiques contenant chacune N molécules de notre gaz, et supposons que dans chaque enceinte ces particules soient numérotées de 1 à N . Pour tout k compris entre 1 et N , et pour tout entier m , notons γ_k^m la position de la particule k dans l'enceinte m . (On choisit bien entendu pour chaque enceinte un repère adapté, de telle sorte qu'elle corresponde à une partie fixée Ω de \mathbb{R}^3 , indépendante de l'enceinte.) Alors l'affirmation que chaque particule est répartie uniformément dans la boîte signifie que,

somme de ces nombres est fixe, égale à N , mais chaque N_i est une variable aléatoire de loi binomiale $\mathcal{B}(N, V_i/V)$. En effet, chacune des N particules se trouve dans la i -ième sous-enceinte avec une probabilité V_i/V , que nous noterons désormais p_i , et ceci indépendamment de toutes les autres particules.

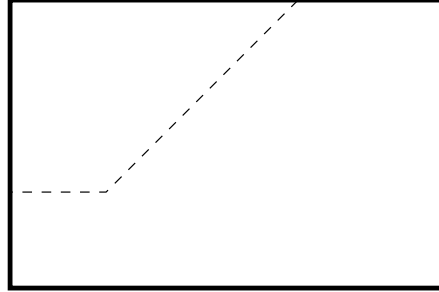


FIG. 2.4 – L'enceinte est divisée en 2 sous-enceintes.

Ainsi, pour tout entier r_1 compris entre 0 et N , on a

$$\mathbb{P}(N_1 = r_1) = C_N^{r_1} p_1^{r_1} (1 - p_1)^{N-r_1}.$$

L'espérance de N_i est égale à $p_i N$. Nous allons étudier de quelle façon N_i fluctue autour de cette valeur moyenne. Ainsi, un regroupement de toutes les particules dans une des sous-enceintes serait un exemple extrême de fluctuation des N_i .

Avant de nous lancer dans une étude précise, commençons par voir ce que peut nous dire le théorème fondamental des probabilités en ce qui concerne les fluctuations d'une variable aléatoire : le théorème de la limite centrale

pour tout k entre 1 et N , on a la convergence faible de mesures

$$\frac{1}{M} \sum_{m=1}^M \delta_{\gamma_k^m} \implies \mathbf{1}_\Omega \frac{dx dy dz}{|\Omega|}$$

lorsque M tend vers l'infini, c'est-à-dire que pour toute fonction continue bornée f sur Ω ,

$$\frac{1}{M} \sum_{m=1}^M f(\gamma_k^m) \xrightarrow{M \rightarrow \infty} \frac{1}{|\Omega|} \int_{\Omega} f dx dy dz.$$

L'indépendance se traduit de façon similaire.

(TCL)¹⁷. En effet, N_1 par exemple est la somme de N variables indépendantes de loi $p_1\delta_1 + (1 - p_1)\delta_0$. Le TCL affirme que si N est assez grand (ou plutôt : lorsque N tend vers l'infini), la variable aléatoire $(N_1 - p_1N)/\sqrt{p_1(1 - p_1)N}$ a presque une loi gaussienne standard. Autrement dit, pour tout réel $a > 0$, on a¹⁸

$$\mathbb{P}(|N_1 - p_1N| \geq a\sqrt{N}) \simeq \frac{2}{\sqrt{2\pi}} \int_a^\infty e^{-t^2/2} dt.$$

Ce résultat dit que des fluctuations de N_1 autour de sa valeur moyenne de l'ordre de $a\sqrt{N}$ sont très improbables si a est grand mais pas si a est petit. En résumé, *les fluctuations de N_1 autour de sa valeur moyenne sont de l'ordre de \sqrt{N} .*

Si nous nous intéressons aux fluctuations relatives, c'est-à-dire aux fluctuations de N_1/N autour de p_1 , elles sont de l'ordre de $1/\sqrt{N}$. Lorsque N est de l'ordre du nombre d'Avogadro, les fluctuations relatives sont de l'ordre de 10^{-12} : pour détecter les fluctuations de pression dues à l'agitation moléculaire dans un bocal d'air au repos, il faudrait une sensibilité de l'ordre du milliardième de millibar.

Nous allons maintenant aborder le problème sous un autre angle et calculer la probabilité d'observer des fluctuations bien plus importantes, de l'ordre non plus de \sqrt{N} , mais de N . Ces probabilités sont bien entendu extrêmement faibles¹⁹ et nous allons montrer qu'elles décroissent exponentiellement vite en fonction de N .

Pour commencer, choisissons un réel q_1 tel que $p_1 < q_1 \leq 1$ et estimons $\mathbb{P}(N_1/N > q_1)$. Cela correspond à une fluctuation de N_1 au-delà de sa moyenne supérieure à $(q_1 - p_1)N$. Nous avons

$$\mathbb{P}\left(\frac{N_1}{N} \geq q_1\right) = \sum_{r \geq q_1 N} C_N^r p_1^r (1 - p_1)^{N-r},$$

d'où on tire, en posant $R = [q_1 N]$, la valeur minimale que doit avoir N_1 pour

¹⁷Ce paragraphe un peu informel ne détermine pas ce qui suit. On n'utilisera pas dans la suite le résultat qu'on y énonce.

¹⁸Nous avons ici délibérément omis un facteur $\sqrt{p_1(1 - p_1)}$ qui ne joue aucun rôle important pour ce qui nous intéresse.

¹⁹L'étude de ces fluctuations très rares s'appelle la théorie des *grandes déviations*.

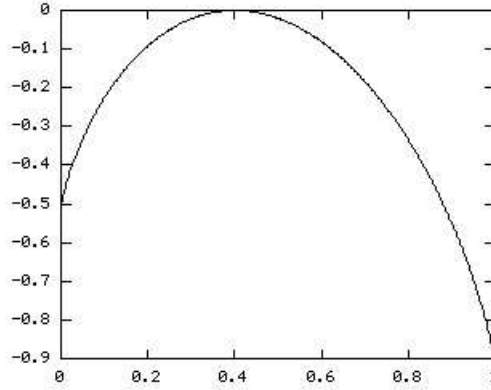


FIG. 2.5 – Le graphe de $q \mapsto -q \log \frac{q}{2/5} - (1 - q) \log \frac{1-q}{3/5}$.

réaliser la fluctuation à laquelle on s'intéresse,

$$C_N^R p_1^R (1 - p_1)^{N-R} \leq \mathbb{P} \left(\frac{N_1}{N} \geq q_1 \right) \leq (N + 1) C_N^R p_1^R (1 - p_1)^{N-R}.$$

La différence entre le minorant et le majorant disparaît à l'échelle logarithmique :

$$\frac{1}{N} \log \mathbb{P} \left(\frac{N_1}{N} \geq q_1 \right) \underset{N \rightarrow \infty}{\sim} \frac{1}{N} \log [C_N^R p_1^R (1 - p_1)^{N-R}].$$

Il est maintenant facile d'établir que le second membre a une limite finie lorsque N tend vers l'infini et d'obtenir le résultat suivant²⁰ :

$$\frac{1}{N} \log \mathbb{P} \left(\frac{N_1}{N} \geq q_1 \right) \underset{N \rightarrow \infty}{\longrightarrow} -q_1 \log \frac{q_1}{p_1} - (1 - q_1) \log \frac{1 - q_1}{1 - p_1}.$$

On démontrerait exactement de la même façon que si $0 \leq q_1 < p_1$, alors

$$\frac{1}{N} \log \mathbb{P} \left(\frac{N_1}{N} \leq q_1 \right) \underset{N \rightarrow \infty}{\longrightarrow} -q_1 \log \frac{q_1}{p_1} - (1 - q_1) \log \frac{1 - q_1}{1 - p_1}.$$

Par exemple, en prenant $q_1 = 0$, on trouve une estimation de la probabilité que la première sous-enceinte soit vide : $\mathbb{P}(N_1 = 0)$ est de l'ordre de $(1 - p_1)^N$. Si $p_1 = 1/2$ comme dans notre problème initial, nous avons une probabilité de l'ordre de $2^{-10^{23}} \approx 10^{-3.10^{22}}$.

²⁰Nous adoptons ici la convention $0 \log 0 = 0$.

2.6.3 Une première approche statistique de l'entropie

Les résultats de décroissance exponentielle obtenus au paragraphe ci-dessus ont des conséquences très importantes à l'échelle macroscopique. En effet, supposons que nous caractérisions l'état macroscopique de notre gaz par la proportion de particules se situant dans chaque sous-enceinte. Alors, comme N est très grand, parmi l'ensemble des états microscopiques possibles (on s'intéresse ici surtout aux positions des particules), une **écrasante majorité** correspond à des valeurs de N_1/N extrêmement voisines de p_1 .

C'est en fait un phénomène très général : lorsqu'on s'intéresse à des grandeurs macroscopiques qui, comme $A(x_*, p_*) = \sum_i a_i(x_*, p_*)$, sont des sommes de grandeurs microscopiques (position, énergie cinétique), il existe *une* valeur pour chaque grandeur macroscopique qui, à une erreur ε près, est réalisée pour *quasiment tous* les états microscopiques, sauf une fraction de l'ordre de $e^{-\alpha\varepsilon^2 N}$, où α est un nombre positif de l'ordre de l'unité²¹. Autrement dit : la valeur moyenne de la grandeur microscopique est aussi celle qui est presque toujours réalisée. On peut donc énoncer le principe suivant :

Un système thermodynamique adopte à l'équilibre l'état macroscopique compatible avec le plus grand « nombre d'états microscopiques »,

sous-entendu, parmi tous les états macroscopiques compatibles avec les contraintes qui lui sont imposées. Nous voulons rapprocher cet énoncé de celui du second principe de la thermodynamique dont une forme est :

Un système thermodynamique adopte à l'équilibre l'état macroscopique qui rend l'entropie maximale.

Cela suggère que l'entropie associée à un état macroscopique est une fonction croissante du nombre d'états microscopiques qui conduisent à cet état. Afin de préciser cela, nous devons en préliminaire :

- définir la liste des grandeurs d'état que nous utilisons pour caractériser un état macroscopique.
- définir clairement ce que nous entendons par « nombres d'états microscopiques ». Il y a là une difficulté, puisqu'en principe l'ensemble

²¹En général, α est la dérivée seconde en un point critique d'une fonction « ordinaire », comme celle dont le graphe est représenté à la figure 2.5.

des états microscopiques qui sont compatibles avec un même état thermodynamique macroscopique est un continuum (et donc a un cardinal infini, de la puissance du continu).

La première tâche dépend du type d'expérience que nous souhaitons réaliser. Ici, dans un souci de simplification, nous supposons que nous avons divisé le domaine Ω contenant un gaz dilué de N particules en deux sous-enceintes Ω_1 et Ω_2 , de volumes respectifs $V_1 := |\Omega_1|$ et $V_2 := |\Omega_2|$ et nous nous intéresserons uniquement aux nombres de particules N_1 et N_2 qui se trouvent respectivement dans Ω_1 et Ω_2 à un certain instant. Donc nous ne tenons pas compte de la pression et de la température et nous pourrions travailler dans la suite avec l'espace des phases microscopique simplifié Ω_*^N au lieu de $T^*\Omega_*^N$, ignorant ainsi les degrés de liberté des impulsions.

Pour résoudre la deuxième difficulté, nous suivons la stratégie suivante, imaginée par les physiciens : nous divisons le domaine Ω représentant l'enceinte en cellules très petites (par exemple des cubes de taille ε). Le nombre M de telles cellules sera de l'ordre de V/ε^3 , où V est le volume de Ω . Pour qu'une telle discrétisation n'induisse pas trop de déformations par rapport au modèle continu, il convient de choisir ε plus petit que la distance moyenne entre les particules, et donc, M plus grand que N . Numérotions les cellules par $\alpha = 1, \dots, M$ et désignons les par q_α . L'idée est de regrouper tous les états microscopiques par paquets : deux états microscopiques x_* et x'_* dans Ω_*^N sont dans le « même paquet » si et seulement si ils correspondent au même nombre de particules dans chaque cellule : $\forall \alpha \in \llbracket 1, M \rrbracket, |\{i \in \llbracket 1, N \rrbracket / x_i \in q_\alpha\}| = |\{i \in \llbracket 1, N \rrbracket / x'_i \in q_\alpha\}|$. Ensuite, étant donné un « état macroscopique », par exemple, spécifier des valeurs (r_1, r_2) pour les variables (N_1, N_2) , on compte le nombre de *paquets* d'états microscopiques qui sont compatibles avec l'état macroscopique correspondant.

Pour formaliser un peu cela, nous introduisons un « espace des phases » discrétisé intermédiaire, l'ensemble des applications de $\llbracket 1, N \rrbracket$ vers $\llbracket 1, M \rrbracket$:

$$\Gamma_\varepsilon := \{\gamma : \llbracket 1, N \rrbracket \longrightarrow \llbracket 1, M \rrbracket\}.$$

En effet à presque tout état microscopique $x_* \in \Omega_*^N$ nous pouvons associer l'application $\gamma \in \Gamma_\varepsilon$ définie par la condition

$$\forall i \in \llbracket 1, N \rrbracket, \quad \gamma(i) = \alpha \quad \text{où } \alpha \in \llbracket 1, M \rrbracket \text{ est l'unique indice tel que } x_i \in q_\alpha.$$

Autrement dit, tout état microscopique x_* peut être identifié avec une application $\llbracket 1, N \rrbracket \rightarrow \Omega$, et γ s'obtient à partir de x_* par la discrétisation de Ω en M cellules. Le nombre total d'états est alors $\sharp_\varepsilon^\Omega(N) := \text{Card}\Gamma_\varepsilon = M^N$.

Ensuite, pour un « état macroscopique » caractérisé par $(N_1, N_2) = (r_1, r_2)$, nous comptons combien d'états microscopiques dans Γ_ε sont compatibles avec cet état macroscopique : si Ω_1 et Ω_2 sont réguliers, les nombres de cellules q_α dans respectivement Ω_1 et Ω_2 sont respectivement $M_1 \simeq p_1 N$ et $M_2 \simeq p_2 N$ (avec une approximation d'autant plus bonne que ε est petit). Comme il y a $M_1^{r_1}$ façons de répartir r_1 particules dans M_1 cellules et $M_2^{r_2}$ façons de répartir r_2 particules dans M_2 cellules, le nombre d'états microscopiques est

$$\sharp_\varepsilon^{\Omega_1, \Omega_2}(r_1, r_2) = \frac{N!}{r_1! r_2!} M_1^{r_1} M_2^{r_2}. \quad (2.13)$$

A ce propos, nous pouvons retrouver, à l'aide de la formule de Stirling, une confirmation du fait déjà illustré dans la section précédente, que parmi les M^N configurations dans Γ_ε , une écrasante majorité de configurations correspondent à $(N_1, N_2) \simeq (p_1 N, p_2 N)$. Il s'agit pour cela d'estimer (2.14) lorsque $(r_1, r_2) = (p_1 N, p_2 N)$. Pour ce faire, dans un premier temps, nous avons, en utilisant $p_1 + p_2 = 1$,

$$\begin{aligned} \frac{N!}{(p_1 N)!(p_2 N)!} &\simeq \frac{N^N e^{-N} \sqrt{2\pi N}}{p_1^{p_1 N} N^{p_1 N} e^{-p_1 N} p_2^{p_2 N} N^{p_2 N} e^{-p_2 N} \sqrt{2\pi p_1 N} \sqrt{2\pi p_2 N}} \\ &= \frac{1}{(p_1^{p_1} p_2^{p_2})^N \sqrt{2\pi p_1 p_2 N}}. \end{aligned}$$

Et comme $M_1^{p_1 N} M_2^{p_2 N} = (p_1 M)^{p_1 N} (p_2 M)^{p_2 N} = (p_1^{p_1} p_2^{p_2})^N M^N$, nous en déduisons que

$$\sharp_\varepsilon^{\Omega_1, \Omega_2}(p_1 N, p_2 N) \simeq \frac{M^N}{\sqrt{2\pi p_1 p_2 N}}. \quad (2.14)$$

Cette valeur diffère de $\sharp_\varepsilon^\Omega(N) = M^N$, mais nous prétendons qu'elle est pratiquement la même ! Par là nous entendons que le facteur $(2\pi p_1 p_2 N)^{-1/2}$ ne change pas l'ordre de grandeur. En fait cette observation est significative parce que la quasi totalité des états microscopiques correspondent à des configurations où la différence entre (N_1, N_2) et $(p_1 N, p_2 N)$ est d'ordre inférieur à \sqrt{N} . Autrement dit et en répétant ce qui a été dit dans la section précédente à propos du paradoxe de Loschmidt, dès que r_1/N diffère de façon significative de p_1 , le rapport $\sharp_\varepsilon^{\Omega_1, \Omega_2}(r_1, r_2) / \sharp_\varepsilon^{\Omega_1, \Omega_2}(p_1 N, p_2 N)$ est extrêmement petit

(plutôt de l'ordre de 10^{-N} que de l'ordre $N^{-1/2}$). Il est donc plus commode de comparer les logarithmes de ces nombres que ces nombres eux-mêmes. Et comme les logarithmes de ces nombres sont encore gigantesques, il est aussi commode de les normaliser en les multipliant par une constante très petite. Nous sommes ainsi conduits à la

Définition 1 *On appelle provisoirement entropie (cinétique) d'un système thermodynamique à l'équilibre la quantité*

$$\tilde{S}_\varepsilon^{\Omega_1, \Omega_2}(N_1, N_2) := k \log \left(\#_\varepsilon^{\Omega_1, \Omega_2}(N_1, N_2) \right), \quad (2.15)$$

où (N_1, N_2) représente l'état macroscopique du système, et où $k \simeq 1,3 \cdot 10^{-23}$ est la constante de Boltzmann.

Ainsi, pour $(N_1, N_2) = (p_1 N, p_2 N)$, on obtient en utilisant (2.14)

$$\tilde{S}_\varepsilon^{\Omega_1, \Omega_2}(p_1 N, p_2 N) \simeq k \log \left(\frac{M^N}{\sqrt{2\pi p_1 p_2 N}} \right) = kN \log M - \frac{k}{2} (\log(2\pi p_1 p_2) + \log N).$$

En ayant en tête que d'une part, $N \simeq 10^{23}$ et donc $\log N \simeq 23 \log 10$ et, d'autre part, k est très petit, nous voyons que $\frac{k}{2} (\log(2\pi p_1 p_2) + \log N)$ est totalement négligeable par rapport à $kN \log M$. Ainsi nous trouvons

$$\tilde{S}_\varepsilon^{\Omega_1, \Omega_2}(p_1 N, p_2 N) \simeq kN \log M = k \log \#_\varepsilon^\Omega(N) = \tilde{S}_\varepsilon^\Omega(N) \quad (2.16)$$

La présence du logarithme dans cette définition entraîne aussi l'extensivité de l'entropie, à cause du facteur N devant le logarithme. De plus une autre vertu du logarithme est la suivante. En effet le nombre $\#_\varepsilon^{\Omega_1, \Omega_2}(N_1, N_2)$ dépend de ε . Par exemple si on divise ε par deux, ce nombre sera multiplié par 2^{3N} et, plus généralement, pour $\tau \in]0, 1[$, $\#_{\tau\varepsilon}^{\Omega_1, \Omega_2}(N_1, N_2) \simeq \tau^{-3N} \#_\varepsilon^{\Omega_1, \Omega_2}(N_1, N_2)$. Ainsi notre définition de l'entropie présente le désagrément de dépendre du choix de l'échelle. Mais nous remarquons que, grâce au logarithme, il en résulte une indétermination de S à une constante additive $-3Nk \log \tau$ près. Cela est relativement satisfaisant, dans la mesure où, dans notre première définition de l'entropie, nous admettions que la valeur de celle-ci soit indéterminée à une constante additive près.

2.6.4 Le paradoxe de Gibbs

La définition précédente de l'entropie souffre d'une incohérence, vis à vis des propriétés d'additivité que l'entropie doit satisfaire. Appliquons en effet la définition (2.15) de l'entropie, aux deux sous-systèmes Ω_1 et Ω_2 séparément. Nous obtenons, en supposant que $(N_1, N_2) = (p_1N, p_2N)$,

$$\tilde{S}_\varepsilon^{\Omega_1}(p_1N) = k \log M_1^{N_1} = kp_1N (\log M + \log p_1)$$

et

$$\tilde{S}_\varepsilon^{\Omega_2}(p_2N) = k \log M_2^{N_2} = kp_2N (\log M + \log p_2).$$

Si le gaz dans Ω_1 est à l'équilibre avec celui dans Ω_2 , c'est à dire si $N_1 = p_1N$ et $N_2 = p_2N$, l'entropie du système total $\tilde{S}_\varepsilon^{\Omega_1, \Omega_2}(p_1N, p_2N)$ doit être la somme des entropies des deux sous-enceintes $\tilde{S}_\varepsilon^{\Omega_1}(p_1N) + \tilde{S}_\varepsilon^{\Omega_2}(p_2N)$. Or nous avons estimé l'entropie du système total à l'équilibre en (2.16) et nous avons trouvé que

$$\begin{aligned} \tilde{S}_\varepsilon^{\Omega_1, \Omega_2}(p_1N, p_2N) &\simeq kN \log M \neq kN (\log M + p_1 \log p_1 + p_2 \log p_2) \\ &= \tilde{S}_\varepsilon^{\Omega_1}(p_1N) + \tilde{S}_\varepsilon^{\Omega_2}(p_2N). \end{aligned}$$

Cette contradiction est appelée *paradoxe de Gibbs*. Pour mieux le comprendre, envisageons une situation un peu plus générale et estimons ce que donne notre définition de l'entropie pour des valeurs arbitraires de $(N_1, N_2) = (r_1, r_2)$:

$$\tilde{S}_\varepsilon^{\Omega_1, \Omega_2}(r_1, r_2) = k \log \left(\frac{N!}{r_1! r_2!} M_1^{r_1} M_2^{r_2} \right),$$

$$\tilde{S}_\varepsilon^{\Omega_1}(r_1) = k \log M_1^{r_1} \quad \text{et} \quad \tilde{S}_\varepsilon^{\Omega_2}(r_2) = k \log M_2^{r_2}.$$

Une façon de corriger ce défaut consiste à faire les substitutions $\#_\varepsilon^{\Omega_1}(r_1) \mapsto \#_\varepsilon^{\Omega_1}(r_1)/r_1!$, $\#_\varepsilon^{\Omega_2}(r_2) \mapsto \#_\varepsilon^{\Omega_2}(r_2)/r_2!$ et $\#_\varepsilon^{\Omega_1, \Omega_2}(r_1, r_2) \mapsto \#_\varepsilon^{\Omega_1, \Omega_2}(r_1, r_2)/N!$. Ainsi nous adoptons une nouvelle définition S^Ω de l'entropie pour laquelle

$$S_\varepsilon^{\Omega_1}(r_1) = k \log \left(\frac{M_1^{r_1}}{r_1!} \right), \quad S_\varepsilon^{\Omega_2}(r_2) = k \log \left(\frac{M_2^{r_2}}{r_2!} \right)$$

et

$$S_\varepsilon^{\Omega_1, \Omega_2}(r_1, r_2) = k \log \left(\frac{M_1^{r_1} M_2^{r_2}}{r_1! r_2!} \right).$$

Et l'on obtient bien alors $S_\varepsilon^{\Omega_1, \Omega_2}(r_1, r_2) = S_\varepsilon^{\Omega_1}(r_1) + S_\varepsilon^{\Omega_2}(r_2)$ pour toute valeur (r_1, r_2) des variables (N_1, N_2) ! Observons que la substitution qui a été opérée

consiste à considérer non pas le nombre de façon de répartir N particules numérotées de 1 à N dans M boîtes, mais le nombre de façon de répartir N particules *indiscernables* dans M boîtes. Cela signifie que les particules sont indiscernables. Autrement dit ce raisonnement semble indiquer qu'il n'est pas possible de numéroter les particules ou de leur attribuer une individualité. Ce comportement un peu étrange des particules sera confirmé par la mécanique quantique : il s'avère que l'indiscernabilité des particules est une propriété profonde de toutes les particules. En conclusion, nous sommes amenés à modifier notre définition de l'entropie comme suit :

Définition 2 On appelle définitivement entropie (cinétique) d'un système thermodynamique à l'équilibre la quantité

$$S_\varepsilon^{\Omega_1, \Omega_2}(N_1, N_2) := k \log \left(\frac{\#_\varepsilon^{\Omega_1, \Omega_2}(N_1, N_2)}{N!} \right), \quad (2.17)$$

où (N_1, N_2) représente l'état macroscopique du système et $N = N_1 + N_2$

Exercice 5 1) Montrer que $\forall (r_1, r_2) \in \mathbb{N}^2$ tel que $r_1 + r_2 = N$, si $r_1 = p_1 N$ et $r_2 = p_2 N$, alors $S_\varepsilon^{\Omega_1, \Omega_2}(r_1, r_2) \simeq \phi(p_1, p_2)$, où

$$\phi(p_1, p_2) := kN (p_1 \log M_1 + p_2 \log M_2 - p_1 \log p_1 - p_2 \log p_2 - \log N + 1).$$

2) Calculer la différentielle de ϕ . Déterminer le maximum de cette fonction lorsque l'on impose la contrainte $p_1 + p_2 = 1$.

3) Retrouver le fait que $\sup_{r_1+r_2=N} S_\varepsilon^{\Omega_1, \Omega_2}(r_1, r_2) \simeq S_\varepsilon^\Omega(N) = k \log (M^N / N!)$.

Remarque — Nous avons vu à propos du paradoxe de Loschmidt que l'état microscopique d'un gaz isolé à l'équilibre pouvait spontanément évoluer vers un état de déséquilibre, mais que la probabilité pour que cela ait lieu est extrêmement petite. Cela a pour conséquence que l'entropie cinétique peut très bien diminuer. Cela semble être en contradiction avec le second principe de la thermodynamique. En fait l'entropie cinétique que nous venons de définir diffère de l'entropie de Clausius que nous avons défini au premier chapitre : les deux quantités ne coïncident qu'à la limite thermodynamique c'est à dire lorsque l'on considère que le nombre de particules N tend vers l'infini.

2.7 Conclusion

Nous avons donc retrouvé un certain nombre de lois qui avaient été établies expérimentalement pour les gaz parfaits, par des estimations probabilistes sur la dynamique d'un gaz de particules élastiques dilué. L'approximation faite n'est bien évidemment pas tout à fait satisfaisante, puisque nous avons négligé les chocs entre particules, alors que nous pressentons intuitivement que les chocs entre particules devraient jouer un rôle capital. En effet, c'est par ces chocs que les particules échangent de l'énergie et « se répartissent » entre elles toute l'énergie cinétique d'une façon à peu près uniforme : une particule beaucoup plus rapide que les autres sera freinée par ses chocs avec ses camarades particules, tandis qu'une particule qui a tendance à rester sur place a de fortes chances d'être réveillée brutalement par une collision. C'est donc la compréhension du mécanisme de « marche vers l'équilibre » du gaz qui nous échappe pour l'instant.

Par ailleurs le modèle probabiliste que nous avons utilisé mène dans certains cas à quelques incohérences apparentes : dans la section 2.3 nous avons fait l'hypothèse que nous observions un très grand nombre d'échantillons identiques de gaz qui ont les mêmes propriétés macroscopiques. Donc en particulier, chacun de ces échantillons a la même température T . Or nous avons aussi établi que l'énergie totale U valait $\frac{3}{2}NkT$. Nous en déduisons que nos échantillons ont tous la même énergie et donc qu'ils correspondent tous à des solutions des équations de Hamilton avec la même valeur de l'hamiltonien $H(x_*, p_*) \simeq \sum_{i=1}^N \frac{p_i^2}{2m}$. Autrement dit tous les échantillons ont des états microscopiques situés dans la même hypersurface $H(x_*, p_*) = \frac{3}{2}NkT$ de $T^*\Omega_*^N$ (qui ici est le produit de Ω_*^N par une sphère dans \mathbb{R}^{3N}). Donc la mesure de probabilité μ doit être concentrée sur cette hypersurface et en particulier s'annuler en dehors de l'hypersurface. Cela n'est évidemment pas compatible avec le résultat (2.12) que nous avons obtenu sur la distribution de Maxwell. Quelle est l'explication de ce paradoxe ? Il y a une erreur de raisonnement ici car en fait, c'est la *moyenne* $\langle U \rangle_\mu$ qui vaut exactement $\frac{3}{2}NkT$ et rien n'oblige à ce que chaque échantillon ait exactement une énergie égale à $\langle U \rangle_\mu$. Mais nous avons néanmoins montré qu'il était extrêmement probable pour U soit très proche de $\langle U \rangle_\mu$. Cela révèle l'existence de deux points de vue possibles : si le gaz dilué *n'est pas isolé*, mais en contact avec un thermostat qui maintient sa température constante, alors l'énergie totale n'est pas nécessairement constante (à cause des éventuels échanges d'énergie avec le thermostat) et la

distribution de Maxwell²² décrit correctement l'état statistique du gaz. En revanche si le système est parfaitement isolé, la mesure μ est bien concentrée²³ sur l'hypersurface $H(x_*, p_*) = \frac{3}{2}NkT$ et la distribution de Maxwell ne convient pas.

Dans ce dernier cas, nous devons utiliser un autre modèle et donc partir d'une hypothèse dans laquelle on ne suppose pas forcément que chaque particule a une loi de probabilité indépendante des autres. Donc, pour un gaz isolé, nous aurons besoin d'employer une mesure de probabilité μ sur $T^*\Omega_*^N$ concentrée sur une hypersurface $H(x_*, p_*) = E$ (où E est une constante). Nous ferons alors l'hypothèse que :

La mesure μ a pour support l'hypersurface $H(x_, p_*) = E$ et a une densité constante par rapport à la mesure naturelle sur cette hypersurface.*

La « mesure naturelle » sera précisée dans le prochain chapitre et est construite à partir de la forme symplectique ω sur $T^*\Omega_*^N$. Cette hypothèse conduit à ce que l'on appelle le formalisme *micro-canonique*. Il nous faudra alors trouver dans un deuxième temps un moyen pour justifier que des prédictions faites grâce à cette hypothèse, sur la base de raisonnements probabilistes (et donc en principe ne s'appliquant qu'à une moyenne sur une collection très grande d'échantillons tirés au hasard) entraînent des prédictions valables sur un seul échantillon ! Dans le cas du gaz dilué, nous avons réussi dans la section 2.3 à le faire grâce à l'hypothèse d'indépendance des lois de chaque particule et en utilisant le fait que le nombre N de particules est très élevé, grâce à la loi des grands nombres. En l'absence d'hypothèse d'indépendance des lois de probabilité de chaque particule, nous aurons besoin de l'*hypothèse ergodique*.

Enfin nous avons vu une première interprétation statistique de l'entropie, dans le cadre de la distribution de Maxwell. Nous aurons besoin dans la suite de généraliser ces idées dans le cadre du formalisme micro-canonique (ou pour d'autres cadres plus généraux). En fait, en apportant une réponse à cette question, Ludwig BOLTZMANN a fait de l'entropie la clef de voute de la physique statistique.

²²ou plus généralement une théorie *canonique*, voir plus loin

²³ce qui correspondra à une théorie *micro-canonique*, voir plus loin

Références : W. Greiner, L. Neise, H. Stöcker, *Thermodynamique et mécanique statistique*, Springer.

J. Harthong, *Probabilités et statistiques*, Diderot Multimédia 1996 ou :
<http://moire4.u-strasbg.fr/bouquins/proba/tabmat1.htm>

Chapitre 3

Mécanique statistique avec interactions entre les particules

3.1 L'ensemble micro-canonique

Notre but est ici de bâtir une théorie statistique applicable à un système **isolé** et susceptible de tenir compte des interactions entre particules. La première étape pour cela est de construire un espace des phases du type $T^*\Omega_\star^N$ avec une famille dépendant du temps de mesures de probabilité $(\nu_t)_{t \in \mathbb{R}}$ qui décrit statistiquement les états microscopiques d'une collection infinie d'échantillons de gaz, choisis « au hasard », au cours du temps t . Ensuite nous verrons comment l'« hypothèse ergodique » justifie le fait que l'on modélise un seul système par une telle mesure. Mais pour qu'une telle description probabiliste soit significative, il faut avant tout que tous les échantillons possèdent les mêmes caractéristiques thermodynamiques.

3.1.1 Le choix de l'espace des phases

Ici il y a lieu de s'interroger : les grandeurs d'états (par exemple pour un gaz dans une enceinte V, p, T, N) de chaque échantillon de la collection devraient heuristiquement être toutes identiques. Cela ne pose pas de problème en principe à réaliser en ce qui concerne le volume V et le nombre de particules N . Et donc, si le système est à l'équilibre thermodynamique, il suffirait alors de fixer T ou p (compte tenu de la relation d'état entre V, p, T et N). Mais le problème est que la pression et la température ne sont définies qu'à l'équilibre. Or nous avons vu, à propos du paradoxe de Loschmidt, qu'un gaz isolé

— même s’il est à l’équilibre thermodynamique — peut en principe évoluer hors de l’équilibre thermodynamique (par exemple si toutes les particules viennent se loger dans une fraction du volume), même si la probabilité pour que cela ait lieu est extrêmement faible. Cela signifie qu’en théorie l’équilibre thermodynamique n’est pas stable ! Cela n’a aucune incidence en pratique. Mais en théorie, et dans la mesure où nous envisageons des expériences par la pensée faisant intervenir une infinité d’échantillons, nous ne pouvons pas exclure les déviations de l’équilibre.

Ici nous sommes sauvés par l’énergie U . En effet à l’équilibre, grâce à la relation d’état entre V, p, T et N et à la relation donnant U en fonction de V, p, T et N , nous pouvons aussi caractériser l’état thermodynamique du gaz en spécifiant V, N et U . L’avantage est que l’énergie U du gaz est définie même en dehors de l’équilibre (puisqu’en particulier U égal la valeur de l’hamiltonien microscopique). Ainsi, suivant L. BOLTZMANN, nous sommes conduits à définir pour tout ouvert $\Omega \subset \mathbb{R}^3$ (enceinte contenant le gaz) et pour toutes valeurs de $N \in \mathbb{N}$ et $E \in \mathbb{R}$ **l’ensemble micro-canonique** : c’est l’ensemble de toutes les configurations microscopiques que peut adopter un gaz de N particules identiques contenu dans Ω et d’énergie totale E .

En termes plus mathématiques, après avoir spécifié l’hamiltonien $H : T^*\Omega_*^N \longrightarrow \mathbb{R}$, nous définissons l’ensemble micro-canonique comme étant l’hypersurface

$$\mathcal{H}_E := \{(x_*, p_*) \in T^*\Omega_*^N / H(x_*, p_*) = E\}.$$

Une variante qui sera très utile est de considérer l’ensemble des états microscopiques sur lequel l’hamiltonien prend des valeurs comprises entre E et $E + \delta E$

$$\mathcal{T}_E^{E+\delta E} := \{(x_*, p_*) \in T^*\Omega_*^N / E < H(x_*, p_*) < E + \delta E\},$$

où $\delta E > 0$ est très petit¹.

Remarquons enfin que la traduction géométrique du fait que l’énergie soit conservée au cours du temps est que le champ de vecteur ξ_H est tangent aux hypersurfaces \mathcal{H}_E , c’est à dire $dH(\xi_H) = 0$. Cela entraîne que toute courbe intégrale de ξ_H est tangente à une hypersurface \mathcal{H}_E .

¹nous verrons en section 3.4.3 comment choisir δE en fonction de E

Enfin, dans un certain nombre d'exemples, comme un gaz monoatomique, nous avons vu à propos du paradoxe de Gibbs que nous devons remplacer l'espace des phases $T^*\Omega_*^N$ par la variété quotient $T^*S_*^N\Omega$, où $S_*^N\Omega$ est le produit symétrisé de Ω , N fois par lui-même, c'est à dire :

$$S_*^N\Omega := \{x_* = (x_1, \dots, x_N) \in \Omega_*^N\} / \sigma_N,$$

où σ_N est le groupe des permutations de N éléments. Par conséquent, dans le formalisme microcanonique, nous remplacerons \mathcal{H}_E par \mathcal{H}_E / σ_N et $\mathcal{T}_E^{E+\delta E}$ par $\mathcal{T}_E^{E+\delta E} / \sigma_N$. Toutefois, dans la suite et lorsqu'il ne sera pas nécessaire de quotienter par σ_N , nous travaillerons dans $T^*\Omega_*^N$, afin de simplifier les notations.

3.1.2 Représentation d'un état par une mesure

L'idée est ici de représenter un état *macroscopique* en évolution² au cours du temps (et donc qui n'est pas nécessairement à l'équilibre) par une famille de mesures $(\nu_t)_{t \in \mathbb{R}}$ sur \mathcal{H}_E ou $\mathcal{T}_E^{E+\delta E}$, selon le modèle envisagé. Pour cela nous imaginons une famille de \mathcal{N} systèmes physiques en évolution, qui sont tous, à chaque instant t , dans le même état macroscopique et « choisis au hasard ». Numérotons chacun des échantillons par $I = 1, \dots, \mathcal{N}$ et notons $(x_*^I, p_*^I) : \mathbb{R} \rightarrow T^*\Omega_*^N$ le chemin décrivant l'évolution de l'échantillon I au cours du temps. Bien entendu chaque (x_*^I, p_*^I) est une solution des équations de Hamilton. Pour toute fonction « observable » $f \in \mathcal{C}^0(T^*\Omega_*^N, \mathbb{R})$, nous supposons que $\forall t \in \mathbb{R}$, la limite

$$\langle f \rangle_{\nu_t} := \lim_{\mathcal{N} \rightarrow \infty} \sum_{I=1}^{\mathcal{N}} \frac{f(x_*^I(t), p_*^I(t))}{\mathcal{N}} \quad (3.1)$$

existe. Cela définit une mesure de probabilité ν_t et nous noterons

$$\int_{T^*\Omega_*^N} f(x_*, p_*) d\nu_t := \langle f \rangle_{\nu_t}.$$

Il y a, dans cette définition, une part d'empirisme, puisque c'est la fortune qui va décider des états de chaque système « choisis au hasard ». Nous verrons un peu plus loin quelques raisonnements théoriques, s'appuyant sur l'hypothèse

²évolution que nous aurons besoin de supposer quasi-statique un peu plus tard

ergodique, permettant dans certains cas de caractériser cette mesure.

Remarquons que la famille $(\nu_t)_{t \in \mathbb{R}}$ ne peut pas être arbitraire, puisque chaque échantillon évolue de façon hamiltonienne. Notons $(e^{t\xi_H})_{t \in \mathbb{R}}$ le flot du champ de vecteur ξ_H , c'est à dire la famille de difféomorphismes $e^{t\xi_H} : T^*\Omega_\star^N \longrightarrow T^*\Omega_\star^N$ telle que $e^{0\xi_H}$ soit l'identité et $\frac{\partial}{\partial t}(e^{t\xi_H}) = \xi_H(e^{t\xi_H})$. Alors pour tout sous-ensemble mesurable $C \subset T^*\Omega_\star^N$, et pour tout instant $t \in \mathbb{R}$, tous les échantillons qui se trouvent dans C à l'instant 0, et seulement ces échantillons, sont dans $e^{t\xi_H}(C)$ à l'instant t . Comme la proportion des échantillons qui se trouvent dans C à l'instant 0 est $\nu_0(C)$ et que la proportion des échantillons qui se trouvent dans $e^{t\xi_H}(C)$ à l'instant t est $\nu_t(e^{t\xi_H}(C))$, on en déduit que

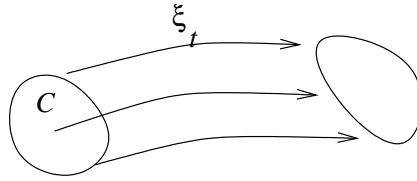


FIG. 3.1 – Convection d'un ouvert C par le flot de ξ_H

$$\forall t \in \mathbb{R}, \quad \nu_0(C) = \nu_t(e^{t\xi_H}(C)). \quad (3.2)$$

Si nous notons $(e^{t\xi_H})^* \nu_t$ la mesure définie pour tout sous-ensemble mesurable $C \subset T^*\Omega_\star^N$ par $(e^{t\xi_H})^* \nu_t(C) := \nu_t(e^{t\xi_H}(C))$, alors la relation (3.1.2) signifie que $\nu_0 = (e^{t\xi_H})^* \nu_t, \forall t \in \mathbb{R}$. Enfin la condition (3.2) sur $(\nu_t)_{t \in \mathbb{R}}$ est aussi équivalente à

$$\forall f \in \mathcal{C}^0(T^*\Omega_\star^N, \mathbb{R}), \quad \int_{T^*\Omega_\star^N} f \circ e^{t\xi_H} d\nu_0 = \int_{T^*\Omega_\star^N} f d\nu_t. \quad (3.3)$$

Cette équation, conséquence de la conservation de la quantité de matière, traduit la conservation de la densité de probabilité.

3.1.3 La construction d'une mesure décrivant l'état d'équilibre

Nous nous intéressons ici à un système macroscopique à l'équilibre thermodynamique et nous cherchons à faire des prédictions théoriques sur toute famille

de mesures $(\nu_t)_{t \in \mathbb{R}}$ modélisant l'évolution d'un tel système. Nous commençons par analyser les propriétés que doit satisfaire une telle mesure.

- le support de ν_t est contenu dans l'ensemble des états microscopiques : \mathcal{H}_E ou $\mathcal{T}_E^{E+\delta E}$, selon le modèle envisagé.
- puisque le gaz est à l'équilibre, la mesure de probabilité ν_t ne dépend pas du temps. Donc, d'après ce qui précède, il existe une mesure de probabilité μ sur $T^*\Omega_*^N$ telle que $\forall t \in \mathbb{R}, \nu_t = \mu$. En vertu de (3.2), cela n'est possible que si $(e^{t\xi_H})^* \mu$ ne dépend pas du temps t . D'où la condition nécessaire sur μ :

$$\frac{d(e^{t\xi_H})^*(\mu)}{dt} = 0 \iff Lie_{\xi_H} \mu = 0. \quad (3.4)$$

Or sur toute variété symplectique $(T^*\mathcal{N}, \omega)$ de dimension $2n$ (où ω est la forme symplectique) il existe une mesure qui satisfait la relation (3.4) pour n'importe quel champ de vecteur hamiltonien ξ_H . C'est la mesure

$$\mu_{Liou} := \frac{(-1)^n}{n!} \omega \wedge \cdots \wedge \omega = (-1)^n \frac{\omega^n}{n!}.$$

En effet nous avons vu à la section 2.1.3 que $Lie_{\xi_H} \omega = 0$. Et comme (*exercice*)

$$Lie_{\xi_H} \omega^n = n (Lie_{\xi_H} \omega) \wedge \omega^{n-1},$$

il est immédiat que $Lie_{\xi_H} \mu_{Liou} = 0$. Ce résultat est dû à Liouville et μ_{Liou} s'appelle la *mesure de Liouville*. Remarquons aussi que la mesure de Liouville est celle dont l'expression est la plus simple dans les coordonnées (x, p) :

$$\mu_{Liou} = (-1)^n \frac{\omega^n}{n!} = dp_1 \wedge \cdots \wedge dp_n \wedge dx^1 \wedge \cdots \wedge dx^n.$$

Revenant à $T^*\Omega_*^N$, nous avons donc ici un candidat naturel pour notre mesure de probabilité sur $\mathcal{T}_E^{E+\delta E}$: la mesure

$$\mu := \frac{1}{Z} \mu_{Liou} = \frac{1}{Z} \frac{\omega^{3N}}{(3N)!}, \quad \text{où } Z = \left(\int_{\mathcal{T}_E^{E+\delta E}} \frac{\omega^{3N}}{(3N)!} \right)^{-1}.$$

(Z est la constante de normalisation.) Dans le cas où nous travaillons sur \mathcal{H}_E , nous obtenons une mesure concentrée sur \mathcal{H}_E et également invariante

par le flot de ξ_H en prenant la limite de la mesure précédente lorsque δE tend vers 0. Cette limite est égale à

$$\mu := \frac{1}{Z} X \lrcorner \mu_{Liou} = \frac{1}{Z} X \lrcorner \frac{\omega^{3N}}{(3N)!},$$

où $X : \mathcal{H}_E \longrightarrow T^*\mathbb{R}^{3N}$ est un champ de vecteur régulier défini le long de \mathcal{H}_E tel que $dH(X) = 1$ partout et $Z := \int_{\mathcal{H}_E} X \lrcorner \omega^{3N} / (3N)!$ est la constante de normalisation (voir exercice ci-dessous).

Exercice 6 Soit C un ouvert dans \mathcal{H}_E . Soit $X : \mathcal{H}_E \longrightarrow T^*\mathbb{R}^{3N}$ un champ de vecteur régulier défini le long de \mathcal{H}_E . On suppose que $dH(X) = 1$ partout (en particulier X est donc transverse à \mathcal{H}_E). Soit, pour $\delta E > 0$ suffisamment petit $C^{\delta E} := \{m + sX(m) / m \in C, s \in \mathbb{R}, E < H(m + sX(m)) < E + \delta E\}$.

1) Montrer que $\lim_{\delta E \rightarrow 0} \mu_{Liou}(C^{\delta E}) / \delta E = X \lrcorner \mu_{Liou}(C^{\delta E})$.

2) Montrer que si $X' : \mathcal{H}_E \longrightarrow T^*\mathbb{R}^{3N}$ est un autre champ de vecteur tel que $dH(X') = dH(X) = 1$, alors $X' \lrcorner \mu_{Liou|_{\mathcal{H}_E}} = X \lrcorner \mu_{Liou|_{\mathcal{H}_E}}$.

3.1.4 Existe-il d'autres mesures décrivant l'équilibre ?

Nous avons construit une mesure de probabilité μ sur \mathcal{H}_E ou $\mathcal{T}_E^{E+\delta E}$ qui est invariante par le flot de ξ_H . A présent reste la question : cette mesure est-elle la bonne ? représente-elle la répartition statistique que « choisirait » la nature pour un gaz à l'équilibre ?

Afin de répondre à cette question nous devons nous demander s'il existe une autre mesure $\tilde{\mu}$ qui soit une mesure de probabilité dont le support est \mathcal{H}_E ou $\mathcal{T}_E^{E+\delta E}$ et qui soit invariant par ξ_H . Pour simplifier, nous allons supposer dans un premier temps que $\tilde{\mu} = \rho\mu$, où $\rho \in \mathcal{C}^1(T^*\Omega_*^N, [0, \infty[)$ est une densité et, pour fixer les idées, nous allons travailler dans $\mathcal{T}_E^{E+\delta E}$. La condition que $\tilde{\mu}$ est une mesure de probabilité signifie alors simplement que

$$\int_{\mathcal{T}_E^{E+\delta E}} \rho\mu = 1$$

et la condition $Lie_{\xi_H} \tilde{\mu} = 0$ donne, en utilisant $Lie_{\xi_H} \mu = 0$ (voir aussi la section 3.3.1) :

$$0 = Lie_{\xi_H}(\rho\mu) = d\rho(\xi_H)\mu \iff d\rho(\xi_H) = \{\rho, H\} = 0.$$

A première vue on aurait envie de dire qu'il y a beaucoup de solutions : il « suffirait de choisir » une hypersurface \mathcal{S} transverse au champ de vecteur ξ_H , de choisir arbitrairement une fonction ρ_0 sur \mathcal{S} et de prolonger ρ_0 par la fonction ρ en dehors de \mathcal{S} de façon à ce que la valeur de ρ sur une trajectoire hamiltonienne soit constante et égale à la valeur prise par ρ_0 au point d'intersection de la trajectoire avec \mathcal{S} . Mais cette construction ne marche pas en général. Pourquoi ? Nous n'avons pris en compte que le comportement local des trajectoires hamiltoniennes, mais rien ne nous dit que la trajectoire ne va pas revenir au bout d'un certain temps et couper \mathcal{S} à un autre endroit. Bien au contraire, le théorème de récurrence de Poincaré semble suggérer que cela risque de se produire un grand nombre de fois. Et cela impose que ρ_0 doit prendre les mêmes valeurs sur tous les points d'intersection d'une trajectoire avec \mathcal{S} .

Exercice 7 On considère le tore $T^2 := \mathbb{R}^2/\mathbb{Z}^2$. Pour tout $(x, y) \in \mathbb{R}^2$ on note $[x, y]$ sa classe d'équivalence dans T^2 . Soit $\tau \in \mathbb{R} \setminus \mathbb{Q}$. On considère le champ de vecteur ξ sur T^2 défini par $\xi([x, y]) := (\tau, 1)$, $\forall [x, y] \in T^2$ et la mesure $\mu = dx dy$. Soit $\rho : T^2 \rightarrow [0, \infty[$ une densité continue, soit $\tilde{\mu} := \rho \mu$ et soit ρ_0 la restriction de ρ à $\mathbb{R} \times \mathbb{Z}/\mathbb{Z}^2$.

- 1) A quelle condition sur ρ a-t-on $\text{Lie}_\xi \tilde{\mu} = 0$ sur $\{[x, y]/x \in \mathbb{R}, y \in [0, 1]\}$?
- 2) A quelle condition sur ρ a-t-on $\text{Lie}_\xi \tilde{\mu} = 0$ sur T^2 ?

Conclure.

Dans l'exercice précédent, la propriété fondamentale du flot de ξ est son *ergodicité*. C'est elle qui entraîne le résultat de la question 2), car elle implique une propriété de récurrence des trajectoires bien plus forte que celle énoncée par le théorème de Poincaré.

3.2 L'ergodicité

3.2.1 Flots ergodiques

Soit (X, \mathfrak{b}, μ) un espace topologique séparable muni d'une mesure borélienne. Soit $(\Phi_t)_{t \in \mathbb{R}}$ une famille d'applications continues $X \rightarrow X$ dépendant continuellement de t et définissant un flot Φ_t qui préserve μ , c'est-à-dire que pour tout $B \in \mathfrak{b}$, $\mu(\Phi_t^{-1}(B)) = \mu(B)$. On suppose que, pour tous $t, s \in \mathbb{R}$, $\Phi_{t+s} = \Phi_t \circ \Phi_s$. L'exemple que nous avons en tête est bien entendu un espace

des phases muni de la mesure de Liouville et d'un flot hamiltonien.

Pour un tel flot, le théorème de Poincaré s'énonce comme suit.

Théorème 1 (Poincaré, 1889) *On suppose que $\mu(X) < \infty$. Soit B une partie borélienne de \mathfrak{b} . Alors, pour presque tout x appartenant à B , la borne supérieure (resp. inférieure) de l'ensemble $\{t \in \mathbb{R} \mid \Phi_t(x) \in B\}$ est $+\infty$ (resp. $-\infty$).*

L'hypothèse $\mu(X) < \infty$ est essentielle : pour s'en convaincre, on peut essayer d'appliquer le théorème à $X = \mathbb{R}$ muni de la mesure de Lebesgue et $\Phi_t(x) = x + t$. Dans le cas qui nous intéresse, la mesure de Liouville de l'espace des phases est finie, parce que l'énergie du système est fixée (le vérifier!).

On suppose désormais que $\mu(X)$ est fini et, quitte à normaliser μ , on suppose que $\mu(X) = 1$. On dit alors que le flot Φ est *ergodique* si toute partie borélienne B de X telle que $\phi_t^{-1}(B) = B$ pour tout $t \in \mathbb{R}$ vérifie $\mu(B) = 0$ ou $\mu(B) = 1$. Autrement dit, un flot ergodique ne laisse aucune partie non-triviale (au sens de la mesure) invariante³. La propriété fondamentale des flots ergodiques est la suivante.

Théorème 2 (Birkhoff, 1931) *Soit Φ un flot ergodique sur (X, \mathfrak{b}, μ) . Soit f une fonction de $L^1(X, \mu)$. Pour tout $x \in X$ et tout $T \in \mathbb{R}_+$, on pose*

$$f_T^*(x) = \frac{1}{T} \int_0^T f(\Phi_t(x)) dt.$$

Alors, pour tout $T \in \mathbb{R}_+$, f_T^ est une fonction de $L^1(X, \mu)$ et la convergence*

$$f_T^* \xrightarrow{T \rightarrow \infty} \int_X f d\mu$$

a lieu μ -presque partout d'une part et dans $L^1(X, \mu)$ d'autre part.

³On montre facilement que cette propriété est équivalente à la suivante, apparemment un peu plus faible : toute partie borélienne B de X telle que $\mu(\Phi_t^{-1}(B) \Delta B) = 0$ pour tout $t \in \mathbb{R}$ vérifie $\mu(B) = 0$ ou $\mu(B) = 1$, où on a utilisé le symbole Δ pour la différence symétrique.

Birkhoff est en fait parti de la conclusion de ce théorème, qui était admise plus ou moins implicitement dans beaucoup de calculs de physique statistique. Ce qu'il a donc découvert, c'est sous quelle hypothèse – remarquablement simple – sur le flot Φ elle était vérifiée. Cependant, il ne faut pas s'y tromper : il est souvent difficile de prouver qu'un flot est ou n'est pas ergodique. L'hypothèse *ergodique* de la mécanique statistique peut s'exprimer comme suit :

Pour tout $E > 0$, le flot hamiltonien est ergodique sur l'hypersurface \mathcal{H}_E ou sur la tranche $\mathcal{T}_E^{E+\delta E}$.

Pour discuter cette hypothèse, remarquons tout d'abord que si Φ est un flot ergodique et si $f \in L^1(X, \mu)$ est invariante, c'est-à-dire si $f \circ \Phi_t = f$ pour tout $t \in \mathbb{R}$, alors f est constante presque partout. (Le vérifier!). Dans le cas où f est continue et où $\text{supp } \mu = X$, cela peut se voir comme une conséquence de la propriété de récurrence suivante, qui renforce le théorème de Poincaré.

Proposition – *Soit Φ un flot ergodique sur (X, \mathfrak{b}, μ) . Pour presque tout $x \in X$, la trajectoire de x repasse dans tout ouvert de X de mesure strictement positive à des temps arbitrairement grands, positifs et négatifs. En particulier, l'adhérence de $\{\Phi_t(x), t \in \mathbb{R}\}$ est égale au support de μ .*

Pour le flot hamiltonien sur \mathcal{H}_E , qui représente l'évolution d'un système mécanique, cela signifie qu'aucune quantité autre que l'énergie totale du système n'est conservée au cours du mouvement : il n'y a aucune autre *intégrale première* du mouvement que l'énergie. Cela est par exemple incompatible avec un modèle microscopique de particules qui s'ignorent et ne se rencontrent pas. En effet, dans un tel système, les énergies cinétiques de chacune des particules seraient inchangées au cours du temps, donc seraient des intégrales premières. Nous pouvons en conclure que les interactions entre les particules sont essentielles lorsqu'on s'intéresse à l'ergodicité du flot hamiltonien. Y. SINAI a démontré que le flot correspondant au mouvement de sphères dures qui se choquent et rebondissent élastiquement dans une enceinte est ergodique. Cependant, ce modèle est très simplifié et on ne sait démontrer pour aucun hamiltonien plus réaliste que le flot correspondant est ergodique.

Revenons à notre problème initial, qui était de déterminer si la mesure de Liouville était la bonne mesure invariante pour le flot $e^{t\xi_H}$ sur une hypersurface d'énergie. L'hypothèse que $e^{t\xi_H}$ est ergodique pour μ nous permet

d'apporter un élément de réponse. En effet, soit ν une autre mesure sur l'hypersurface invariante par le flot hamiltonien. Alors ν se décompose par rapport à μ en une partie absolument continue et une partie singulière. On vérifie que ces deux parties sont invariantes par le flot (le faire!) et que la partie absolument continue est proportionnelle à μ . En effet, si $\rho\mu$ est invariante, alors ρ est une fonction L^1 invariante, donc constante presque partout. Si nous excluons des mesures singulières par rapport à la mesure de Liouville, la seule mesure invariante, donc correspondant à l'équilibre, est donc μ elle-même.

3.2.2 Autres applications de l'hypothèse ergodique

Une reformulation de l'hypothèse ergodique pour un gaz est la suivante. Pour tout ouvert $C \subset T^*\Omega_*^N$ et pour toute condition initiale (x_*^0, p_*^0) pour la configuration microscopique du gaz, soit $(x_*, p_*) : \mathbb{R} \rightarrow T^*\Omega_*^N$ la solution des équations de Hamilton avec $(x_*(0), p_*(0)) = (x_*^0, p_*^0)$. De façon similaire à (2.4) nous considérons

$$\ell(C) := \lim_{\Delta t \rightarrow \infty} \frac{|\{t \in [0, \Delta t] / (x_*(t), p_*(t)) \in C\}|}{\Delta t}. \quad (3.5)$$

Alors l'hypothèse ergodique signifie que $\ell(C) = \mu(C)$, $\forall C$.

En plus de garantir que la mesure de probabilité construite à partir de la mesure de Liouville est la seule qui puisse décrire statistiquement l'équilibre, l'hypothèse ergodique sert aussi à justifier et à élargir le champ d'application de la modélisation statistique au cas d'un seul échantillon. L'idée repose sur le fait que la convergence de la limite définie en (3.5) est suffisamment rapide. Pour cela, nous posons, pour tout $\Delta t > 0$

$$\ell^{\Delta t}(C) := \frac{|\{t \in [0, \Delta t] / (x_*(t), p_*(t)) \in C\}|}{\Delta t}.$$

Et nous faisons l'hypothèse ergodique « forte » :

Pour tout ouvert $C \subset T^\Omega_*^N$, l'approximation $\mu(C) = \ell(C) \simeq \ell^{\Delta t}(C)$ est valable même si l'intervalle de temps Δt est petit.*

Par « petit » on entend ici de l'ordre de grandeur du temps nécessaire pour réaliser une mesure (de l'ordre de la seconde). Cette hypothèse est-elle valable? Cherchons à analyser le pour et le contre :

- d'un côté la très grande rapidité des mouvements des particules d'un gaz plaide en faveur de cette hypothèse
- d'un autre côté l'espace $T^*\Omega_*^N$ est lui-même gigantesque et il nous semble⁴ qu'à cause de la taille et de la dimension gigantesque de cet espace, le temps Δt à partir duquel l'approximation de $\ell(C)$ par une moyenne sur une durée Δt est bonne est très grand. Une estimation grossière de ce temps peut être donnée par le théorème de récurrence de Poincaré et donne en effet une valeur de Δt elle aussi gigantesque.
- cette mauvaise nouvelle doit être tempérée par ce que le paradoxe de Gibbs nous a appris, à savoir que les particules sont indiscernables. Ainsi, si l'hamiltonien H sur $T^*\Omega_*^N$ est invariant par les substitutions des particules entre elles, nous devons remplacer l'espace des phases $T^*\Omega_*^N$ par la variété quotient $T^*S_*^N\Omega$, où $S_*^N\Omega$ est le produit symétrisé de Ω , N fois par lui-même :

$$S_*^N\Omega := \{x_* = (x_1, \dots, x_N) \in \Omega_*^N\} / \sigma_N,$$

où σ_N est le groupe des permutations de N éléments. Cela a pour effet de diviser le volume de l'espace des phases par $N!$, un facteur qui est loin d'être petit !

- enfin il faut probablement essayer de tenir compte aussi d'une persistance du phénomène de mélange mis en évidence dans le cas d'un gaz dilué, même en présence d'interactions.

En réalité nous (les auteurs de ce texte) devons avouer ici notre ignorance d'une justification réellement satisfaisante. L'hypothèse ergodique forte telle que nous l'avons énoncée n'est peut-être pas nécessaire (et peut-être même fausse) et pourrait être remplacée par une version un peu moins forte.

En effet, nous ne sommes pratiquement jamais capables de mesurer l'état microscopique d'une gaz : la plupart des observations (sinon toutes) se font par le biais de grandeurs d'état $A \in \mathcal{C}^0(T^*\Omega_*^N, \mathbb{R})$ que nous cherchons à évaluer. Et ces fonctions sont de la forme $A(x_*, p_*) = \sum_{i=1}^N a(x_i, p_i)$. Donc il suffirait en fait que *l'hypothèse ergodique forte ne soit vérifiée que pour des ouverts $C \subset T^*\Omega_*^N$ « cylindriques », c'est à dire de la forme*

$$T^*\Omega_*^N \cap \left(\cup_{j=1}^N (T^*\Omega)^j \times q \times (T^*\Omega)^{N-j-1} \right),$$

⁴aux auteurs de ce texte

où q est un ouvert de $T^*\Omega_*^N$ (le cas échéant il faut remplacer un tel ouvert par son image dans le quotient par σ_N).

En admettant que l'hypothèse ergodique forte est vraie pour des ouverts cylindriques, ce que nous ferons dans la suite, nous pouvons en déduire les conséquences suivantes.

Observables — Considérons une grandeur d'état $A \in \mathcal{C}^0(T^*\Omega_*^N, \mathbb{R})$ et imaginons que nous fassions *une* mesure de A sur *un seul* échantillon d'un gaz, dans un état microscopique $(x_*, p_*) \in T^*\Omega_*^N$, pendant un certain intervalle de temps Δt , relativement bref. Cela donne le résultat $\frac{1}{\Delta t} \int_t^{t+\Delta t} A(x_*(t), p_*(t)) dt$. Avec quel degré de certitude peut-on dire que $\frac{1}{\Delta t} \int_t^{t+\Delta t} A(x_*(t), p_*(t)) dt = \langle A \rangle_\mu$? Nous avons vu au chapitre précédent que, dans le cas d'un gaz dilué, l'égalité

$$\frac{1}{\Delta t} \int_t^{t+\Delta t} A(x_*(t), p_*(t)) dt \simeq \langle A \rangle_\mu$$

est extrêmement probable. Dans un cas plus général, la justification qu'une telle égalité a lieu presque sûrement est plus délicate à obtenir. Si nous supposons que A est de la forme $A(x_*, p_*) = \sum_{i=1}^N a(x_i, p_i)$, alors nous déduisons de l'hypothèse ergodique forte pour les ouverts cylindriques que, si $t \mapsto (x_*(t), p_*(t))$ est une solution des équations de Hamilton,

$$\frac{1}{\Delta t} \int_t^{t+\Delta t} A(x_*(t), p_*(t)) dt \simeq \langle A \rangle_\mu,$$

avec une très forte probabilité.

Gaz en dehors de l'équilibre — Nous pouvons étendre l'hypothèse ergodique forte, pour les ouverts cylindriques, à des systèmes en évolution quasi-statique (c'est à dire en fait, en très léger déséquilibre). Pour cela nous définissons, pour tout temps t et pour tout ouvert $C \subset T^*\Omega_*^N$,

$$\ell_t^{\Delta t}(C) := \frac{|\{s \in [t, t + \Delta t] / (x_*(s), p_*(s)) \in C\}|}{\Delta t},$$

où Δt est un intervalle de temps « petit » à l'échelle macroscopique et « grand » pour les échelles microscopiques. L'hypothèse est alors que, pour tout temps t et pour tout ouvert $C \subset T^*\Omega_*^N$ *cylindrique*,

$$\ell_t^{\Delta t}(C) \simeq \nu_t(C),$$

où ν_t est la mesure par (3.1).

3.3 Applications du formalisme micro-canonique

3.3.1 Dynamique d'une évolution quasi-statique

Adoptant l'hypothèse ergodique forte pour les ouverts cylindriques, nous sommes capables de modéliser l'évolution quasi-statique d'un gaz par une famille $(\nu_t)_{t \in \mathbb{R}}$ de mesures. Nous avons vu que, à cause de la conservation de la matière, cette famille doit satisfaire la condition (3.2) ou (3.3). Nous allons revisiter cette condition en faisant l'hypothèse supplémentaire que, pour tout t , la mesure ν_t est régulière par rapport à la mesure de l'équilibre μ (construite à partir de la mesure de Liouville). Plus précisément nous supposons qu'il existe une famille de densités $(\rho_t)_{t \in \mathbb{R}}$, telle que

$$\forall t \in \mathbb{R}, \quad \nu_t = \rho_t \mu$$

et $(t, x_*, p_*) \mapsto \rho_t(x_*, p_*)$ est une fonction de classe \mathcal{C}^1 . Nous allons en déduire une équation d'évolution sur la densité ρ_t . Pour simplifier les calculs, nous allons travailler sur $\mathcal{T}_E^{E+\delta E}$ (donc dans ce qui suit, μ s'annule en dehors de $\mathcal{T}_E^{E+\delta E}$).

Pour toute fonction $f \in \mathcal{C}_c^1(\mathcal{T}_E^{E+\delta E}, \mathbb{R})$ (à support compact), nous avons d'une part :

$$\frac{d}{dt} \langle f \rangle_{\nu_t} = \frac{d}{dt} \left(\int_{T^*\Omega_*^N} f \rho_t d\mu \right) = \int_{T^*\Omega_*^N} f \frac{d\rho_t}{dt} d\mu. \quad (3.6)$$

Et d'autre part,

$$\begin{aligned}
\frac{d}{dt}\langle f \rangle_{\nu_t} &= \frac{d}{dt} \left(\lim_{\mathcal{N} \rightarrow \infty} \sum_{I=1}^{\mathcal{N}} f(x_*^I(t), p_*^I(t)) \right) \\
&= \lim_{\mathcal{N} \rightarrow \infty} \sum_{I=1}^{\mathcal{N}} df_{(x_*^I(t), p_*^I(t))} (\xi_H(x_*^I(t), p_*^I(t))) \\
&= \lim_{\mathcal{N} \rightarrow \infty} \sum_{I=1}^{\mathcal{N}} \{f, H\}(x_*^I(t), p_*^I(t)) \\
&= \langle \{f, H\} \rangle_{\nu_t} \\
&= \int_{T^*\Omega_*^N} \{f, H\} \rho_t d\mu \\
&= \int_{T^*\Omega_*^N} \{f \rho_t, H\} d\mu - \int_{T^*\Omega_*^N} \{\rho_t, H\} f d\mu.
\end{aligned} \tag{3.7}$$

A présent nous prétendons que

$$\int_{T^*\Omega_*^N} \{f \rho_t, H\} d\mu = 0. \tag{3.8}$$

En effet,

$$\begin{aligned}
\{f \rho_t, H\} &= \sum_{i=1}^N \sum_{\alpha=1}^3 \frac{\partial(f \rho_t)}{\partial p_{i,\alpha}} \frac{\partial H}{\partial x_i^\alpha} - \frac{\partial(f \rho_t)}{\partial x_i^\alpha} \frac{\partial H}{\partial p_{i,\alpha}} \\
&= \sum_{i=1}^N \sum_{\alpha=1}^3 \frac{\partial}{\partial p_{i,\alpha}} \left(f \rho_t \frac{\partial H}{\partial x_i^\alpha} \right) - \frac{\partial}{\partial x_i^\alpha} \left(f \rho_t \frac{\partial H}{\partial p_{i,\alpha}} \right).
\end{aligned}$$

Et, en intégrant cette quantité sur $\mathcal{T}_E^{E+\delta E}$ et en utilisant le fait que le support de f est compact, on obtient (3.8) par la formule de Stokes. Ainsi (3.7) entraîne que $\frac{d}{dt}\langle f \rangle_{\nu_t} = - \int_{T^*\Omega_*^N} \{\rho_t, H\} f d\mu$. En comparant avec (3.6), on en déduit que

$$\forall f \in \mathcal{C}_c^1(\mathcal{T}_E^{E+\delta E}, \mathbb{R}), \quad \int_{T^*\Omega_*^N} f \frac{d\rho_t}{dt} d\mu = - \int_{T^*\Omega_*^N} \{\rho_t, H\} f d\mu.$$

Cela n'est possible que si

$$\frac{d\rho_t}{dt} = \{H, \rho_t\}. \tag{3.9}$$

Cette équation décrit l'évolution au cours du temps de la densité ρ_t .

3.3.2 Le gaz dilué à l'équilibre

Pour un gaz dilué à l'équilibre nous faisons, comme au chapitre précédent, l'hypothèse que l'on peut négliger les chocs. Cela revient donc à adopter sur l'espace des phases $T^*\Omega_*^N$ l'hamiltonien $H(x_*, p_*) = \sum_{i=1}^N |p_i|^2/2m$. Alors l'ensemble micro-canonique est

$$\mathcal{H}_E := \{(x_*, p_*) \in T^*\Omega_*^N / \sum_{i=1}^N |p_i|^2/2m = E\}$$

c'est à dire le produit de Ω_*^N par la sphère de rayon $\sqrt{2mE}$ dans \mathbb{R}^{3N} . Avec la distribution de Maxwell nous avons calculé que, pour un gaz dilué à l'équilibre à la température T , la valeur moyenne de l'énergie $\langle U \rangle_\mu$ était $\frac{3}{2}NkT$. Nous allons donc supposer ici que $E = \frac{3}{2}NkT$.

La mesure de probabilité sur \mathcal{H}_E est alors

$$\mu = \frac{1}{Z} \left(X \lrcorner \frac{\omega^n}{n!} \right)_{|\mathcal{H}_E}, \quad \text{où } X := \sum_{i=1}^N \sum_{\alpha=1}^3 p_{i,\alpha} \frac{\partial}{\partial p_{i,\alpha}}.$$

Notant $R := \sqrt{2mE} = \sqrt{3Nm kT}$ et renumérotant pour simplifier $(x_*, p_*) = (x^A, p_A)_{A \in [1, 3N]}$, nous avons

$$\mu = \left(\frac{1}{R^{3N-1} |S^{3N-1}|} \sum_{A=1}^{3N} \frac{p_A}{R} \frac{\partial}{\partial p_A} \lrcorner dp_1 \wedge \cdots \wedge dp_{3N} \right) \wedge \left(\frac{1}{|\Omega|^{3N}} dx^1 \wedge \cdots \wedge dx^{3N} \right)_{|\mathcal{H}_E},$$

où, pour tout $n \in \mathbb{N}$, nous notons $S^{n-1} := \{p \in \mathbb{R}^n / |p| = 1\}$ et $|S^{n-1}|$ est le volume de S^{n-1} .

Un calcul de mesure image — Nous allons à présent nous intéresser à N_1 particules dans le gaz et chercher la mesure de probabilité sur les variables de positions et d'impulsion de ces N_1 particules. Nous noterons $N_2 := N - N_1$. Ici nous oublions momentanément notre résolution prise après avoir vu le paradoxe de Gibbs et nous supposons les particules discernables. Nous noterons les coordonnées de positions et d'impulsion de ces N_1 particules sous la forme $(x^A, p_A)_{A \in [1, 3N_1]}$. Il s'agit donc de calculer la mesure μ_{N_1} sur les variables de positions et d'impulsions $(x^A, p_A)_{A \in [1, 3N_1]}$, qui est l'image de

μ par la projection de $T^*\Omega^N$ sur $T^*\Omega^{N_1}$. Il est clair que

$$\mu_{N_1} = \lambda_{N_1} \wedge \left(\frac{1}{|\Omega|^{N_1}} dx^1 \wedge \cdots \wedge dx^{3N_1} \right),$$

où λ_{N_1} est la mesure sur \mathbb{R}^{3N_1} qui est l'image de $\frac{1}{R^{3N_1-1}|S^{3N_1-1}|} \sum_{A=1}^{3N} \frac{p_A}{R} \frac{\partial}{\partial p_A} \lrcorner dp_1 \wedge \cdots \wedge dp_{3N}$. En particulier le support de λ_{N_1} est contenu dans l'adhérence de la boule $B_R^{3N_1} := \{p \in \mathbb{R}^{3N_1} / |p| \leq R\}$.

Pour calculer λ_{N_1} , nous considérons la paramétrisation suivante de S_R^{3N-1} , la sphère de rayon R dans $\mathbb{R}^{3N} = \{(p, q) \in \mathbb{R}^{3N_1} \times \mathbb{R}^{3N_2}\}$:

$$\begin{aligned} \phi : B_R^{3N_1} \times S^{3N_2-1} &\longrightarrow \mathbb{R}^{3N} \\ (p, u) &\longmapsto (p, \sqrt{R^2 - |p|^2}u). \end{aligned}$$

La mesure λ_{N_1} est caractérisée par : $\forall f \in \mathcal{C}^0(B_R^{3N_1})$,

$$\begin{aligned} \int_{B_R^{3N_1}} f(p) d\lambda_{N_1}(p) &= \int_{S_R^{3N-1}} f(p) d\mu(p, q) \\ &= \int_{B_R^{3N_1} \times S^{3N_2-1}} f(p) (\phi^* d\mu)(p, u) \\ &= \int_{B_R^{3N_1} \times S^{3N_2-1}} f(p) |Jac(d\phi)_{(p,u)}| (dp)^{3N_1} d\sigma(u), \end{aligned}$$

où σ est la mesure euclidienne sur S^{3N_2-1} . En un point $(p, u) \in B_R^{3N_1} \times S^{3N_2-1}$, choisissons une base orthonormée $(\epsilon_1, \dots, \epsilon_{3N_1}, e_1, \dots, e_{3N_2-1})$ de $T_{(p,u)}(B_R^{3N_1} \times S^{3N_2-1})$ telle que $(\epsilon_1, \dots, \epsilon_{3N_1})$ soit une base orthonormée de \mathbb{R}^{3N_1} et (e_1, \dots, e_{3N_2-1}) soit une base orthonormée de $T_u S^{3N_2-1}$. Alors $(\epsilon_1, \dots, \epsilon_{3N_1}, e_1, \dots, e_{3N_2-1}, u)$ est une base orthonormée de \mathbb{R}^{3N} et la matrice jacobienne de $d\phi_{(p,u)}$ (application de $T_{(p,u)}(B_R^{3N_1} \times S^{3N_2-1})$ vers \mathbb{R}^{3N}) s'écrit en utilisant ces deux bases :

$$Jac(d\phi)_{(p,u)} = \begin{pmatrix} 1_{3N_1} & 0 \\ 0 & \sqrt{R^2 - |p|^2} 1_{3N_2-1} \\ -\frac{p}{\sqrt{R^2 - |p|^2}} & 0 \end{pmatrix}.$$

Ici $|Jac(d\phi)_{(p,u)}|$ est la norme euclidienne du vecteur dont les composantes sont tous les mineurs $(N-1) \times (N-1)$ extraits de $Jac(d\phi)_{(p,u)}$. Cette quantité est aussi égale à la valeur absolue du déterminant de la matrice $N \times N$ obtenue

en ajoutant à $Jac(d\phi)_{(p,u)}$ le vecteur colonne

$$\frac{1}{R} \begin{pmatrix} p \\ 0 \\ \sqrt{R^2 - |p|^2} \end{pmatrix}$$

(à cause du fait que ce vecteur est de norme 1 et orthogonal à $T_{\phi(p,u)}S_R^{3N-1}$). On trouve ainsi (le calcul peut aussi être simplifier en choisissant la base $(\epsilon_1, \dots, \epsilon_{3N_1})$ de façon à ce que $p = |p|\epsilon_1$) que

$$|Jac(d\phi)_{(p,u)}| = R\sqrt{R^2 - |p|^2}^{3N_2-2}.$$

Exercice — Retrouver cette expression par un raisonnement plus géométrique. ($|Jac(d\phi)_{(p,u)}| = \sqrt{1 - \frac{|p|^2}{R^2}} \sqrt{R^2 - |p|^2}^{3N_2-1}$.)

$$\begin{aligned} \text{Ainsi } & \int_{B_R^{3N_1}} f(p) d\lambda_{N_1}(p) \\ &= \frac{1}{R^{3N-1} |S^{3N-1}|} \int_{B_R^{3N_1}} \int_{S^{3N_2-1}} f(p) R \sqrt{R^2 - |p|^2}^{3N_2-2} (dp)^{3N_1} d\sigma(u) \\ &= \frac{R^{3N_2-1} |S^{3N_2-1}|}{R^{3N-1} |S^{3N-1}|} \int_{B_R^{3N_1}} f(p) \left(1 - \frac{|p|^2}{R^2}\right)^{\frac{3}{2}N_2-1} (dp)^{3N_1} \\ &= \frac{|S^{3N_2-1}|}{R^{3N_1} |S^{3N-1}|} \int_{B_R^{3N_1}} f(p) \left(1 - \frac{|p|^2}{R^2}\right)^{\frac{3}{2}N_2-1} (dp)^{3N_1}. \end{aligned}$$

A présent nous exploitons le fait que N est très grand et supposons aussi que $N_1 \ll N$. Alors, comme

$$\forall n \in \mathbb{N}, \quad |S^{n-1}| = 2 \frac{\pi^{\frac{n}{2}}}{\Gamma(\frac{n}{2})},$$

on obtient en utilisant le fait que $\Gamma(n)/\Gamma(n - \frac{1}{2}) \simeq \sqrt{n}$ lorsque $n \mapsto \infty$ que

$$\frac{|S^{3N_2-1}|}{R^{3N_1} |S^{3N-1}|} = \frac{1}{R^{3N_1}} \frac{\pi^{\frac{3}{2}N_2} \Gamma(\frac{3}{2}N)}{\pi^{\frac{3}{2}N} \Gamma(\frac{3}{2}N_2)} \simeq \frac{\sqrt{\frac{3}{2}N}^{3N_1}}{\sqrt{\pi}^{3N_1} R^{3N_1}}.$$

Par ailleurs, nous pouvons aussi faire l'approximation, pour $|p| \ll R$,

$$\left(1 - \frac{|p|^2}{R^2}\right)^{\frac{3}{2}N_2-1} \simeq e^{-\frac{3}{2}N_2-1 \frac{|p|^2}{R^2}}.$$

Nous concluons que, pour $N_1 \ll N$, $|p| \ll R$, en remplaçant R par sa valeur $\sqrt{3NmkT}$,

$$d\lambda_{N_1}(p) \simeq \frac{1}{\sqrt{2m\pi kT}^{3N_1}} e^{-\frac{|p|^2}{2mkT}} (dp)^{3N_1}.$$

Et donc nous observons que le support de μ_{N_1} est $\Omega^{N_1} \times B_R^{3N_1}$, qui tend vers $\Omega^{N_1} \times \mathbb{R}^{3N_1}$ lorsque N tend vers ∞ , et qu'alors μ_{N_1} tend vers la distribution de Maxwell.

Les conséquences physiques de ce calcul — Premièrement nous pouvons appliquer ce calcul au cas $N_1 = 1$. Alors, nous obtenons que la mesure de probabilité sur la position et l'impulsion d'une particule est $\frac{1}{|\Omega|\sqrt{2m\pi kT}^3} e^{-\frac{|p|^2}{2mkT}} (dp)^3(dx)^3$, c'est à dire la même densité que ce que donne la distribution de Maxwell. Deuxièmement nous voyons que ce résultat s'applique aussi au cas d'un sous-ensemble plus gros de particules. On peut, en extrapolant, imaginer que, parmi un système *isolé* de N particules identiques, les N_1 particules que l'on a ainsi sélectionné forment un système que l'on étudie et que les $N - N_1$ particules restantes forment un thermostat. On déduit alors de ce calcul que, en adoptant une description micro-canonique pour le système total (ce qui est cohérent, puisque ce système est isolé), le sous-système obéit à la statistique de Maxwell ou, en anticipant sur la suite, à une statistique canonique.

3.4 Entropie microcanonique

3.4.1 Entropie statistique sur des ensembles finis

Dans ce paragraphe, nous allons introduire de façon abstraite la notion d'entropie d'une mesure de probabilités. Nous verrons par la suite comment elle s'applique à la physique statistique. Pour commencer, soit $X = \{x_1, \dots, x_r\}$ un ensemble fini. Une mesure de probabilités sur X est un n -uplet $\mu = (\mu_1, \dots, \mu_r)$ de nombres positifs tels que $\sum \mu_i = 1$. Dans tout ce qui suit, on adopte la convention $0 \log 0 = 0$.

Définition 3 Soient μ et ν deux mesures de probabilités sur X .

– On appelle entropie de μ et on note $H(\mu)$ la quantité

$$H(\mu) = - \sum_{i=1}^r \mu_i \log \mu_i.$$

- On appelle entropie relative de ν par rapport à μ et on note $H(\nu|\mu)$ la quantité

$$H(\nu|\mu) = \begin{cases} \sum_{i=1}^r \frac{\nu_i}{\mu_i} \log \frac{\nu_i}{\mu_i} & \text{si } \nu \ll \mu \\ +\infty & \text{sinon.} \end{cases}$$

On montre facilement que $H(\mu)$ est positive, nulle si et seulement si $\mu = \delta_{x_i}$ pour un certain x_i , maximale pour μ uniforme sur X , alors égale à $\log n$. On dit parfois que $H(\mu)$ mesure le *manque d'information* sur l'état réel d'un système lorsqu'on sait seulement que son état est distribué selon la mesure μ . On montre aussi grâce à l'inégalité de Jensen que $H(\nu|\mu)$ est positive, nulle si et seulement si $\nu = \mu$. On peut retenir que, bien que l'entropie relative ne soit pas une distance (elle n'est même pas symétrique), elle mesure l'écart entre deux mesures. Un résultat frappant à cet égard est le théorème de Sanov, qui est le cadre dans lequel s'expriment naturellement les résultats de convergence exponentielle que nous avons rencontré au chapitre précédent.

Soit μ une mesure de probabilités sur X . Nous supposons, quitte à restreindre X , qu'elle charge tout X , c'est-à-dire que $\mu_i > 0$ pour tout $i = 1 \dots r$. Soit X_1, X_2, \dots une suite de v.a. indépendantes, à valeurs dans X , de loi μ . La loi des grands nombres assure que dans un grand échantillon X_1, \dots, X_N , la proportion de x_1 est (presque sûrement) proche de μ_1 , celle de x_2 proche de μ_2 , etc... Le théorème de Sanov estime la probabilité que ces proportions ne soient pas celles auxquelles on s'attend. Plus précisément, on définit la mesure empirique L_N par

$$L_N = \sum_{n=1}^N \delta_{X_n}.$$

C'est une mesure de probabilités aléatoire sur X : si Ω est l'espace de probabilités sur lequel sont définis les X_i et si on note $M_1^+(X)$ l'ensemble des mesures de probabilités sur X , L_N est une fonction $L_N : \Omega \longrightarrow M_1^+(X)$. La loi des grands nombres entraîne la convergence presque-sûre de L_N vers μ lorsque N tend vers l'infini⁵.

⁵On devrait dire : L_N converge presque-sûrement faiblement vers μ , mais c'est une précision inutile puisque toutes les convergences usuelles de mesure sont équivalentes sur un ensemble fini.

Théorème 3 (Théorème de Sanov) Soit $\Gamma \subset M_1^+(X)$ un ensemble quelconque de mesures sur X . Alors

$$\limsup_{N \rightarrow \infty} \frac{1}{N} \log \mathbb{P}(L_N \in \Gamma) \leq - \inf_{\nu \in \Gamma} H(\nu | \mu).$$

Si Γ est un ensemble de mesures assez “éloignées” de μ , c’est-à-dire ayant une grande entropie relative – disons supérieure à h_{min} , la probabilité d’obtenir un échantillon de taille N dont la répartition correspond à une mesure de Γ est plus petite⁶ que $e^{-Nh_{min}}$.

Exercice 8 Supposons que X n’a que deux éléments. Montrer que la fonction dont le graphe est représenté à la figure 2.5 n’est autre que l’entropie relative de la mesure $q\delta_{x_1} + (1 - q)\delta_{x_2}$ par rapport à la mesure $p\delta_{x_1} + (1 - p)\delta_{x_2}$.

3.4.2 Entropie statistique d’un état macroscopique

Nous cherchons maintenant à généraliser et transposer au cadre micro-canonique la définition de l’entropie (cinétique) que nous avons proposée à la fin du chapitre précédent. Nous gardons à l’esprit l’idée originelle que premièrement l’entropie est, à un facteur près, le logarithme du nombre d’états microscopiques compatibles avec un état macroscopique donné et deuxièmement, si on a divisé l’espace en petites cellules de mêmes tailles et si l’on est à l’équilibre, chacune d’elle a la même probabilité d’être occupée.

Mais cette fois le paradigme est un peu différent, car au lieu de découper Ω (voire même $T^*\Omega$) en petites « cases » et de dénombrer la proportion de N particules dans chaque case, c’est une partie de l’espace des phases $T^*\Omega_*^N$ (ou $T^*S\Omega_*^N$) que nous allons découper en petites cellules $(q_I)_{I \in \llbracket 1, M \rrbracket}$ et, comme l’état microscopique du gaz est alors représenté par un point dans l’espace des phases, nous devons faire appel à un artifice supplémentaire : imaginer une expérience avec $\mathcal{N} \rightarrow \infty$ échantillons macroscopiquement identiques, « tirés au hasard ».

En premier nous devons « calibrer » les cellules q_I de façon à ce qu’à l’équilibre, elles soient toutes équiprobables. Grâce à l’hypothèse ergodique qui,

⁶Le théorème de Sanov complet fournit en fait aussi un minorant, qui est du même ordre que le majorant, et même égal si Γ est ouvert.

dans le cas du formalisme micro-canonique, privilégie la mesure de Liouville μ_{Liou} , l'équiprobabilité s'exprime simplement par la condition $\mu_{Liou}(q_I) = \mu_{Liou}(q_{I'})$, $\forall I, I'$. Mais en plus l'entropie est une quantité extensive. Donc ici nous devons choisir une fois pour toute une valeur universelle⁷ pour la quantité $\mu_{Liou}(q_I)$ (qui s'exprime en $(kg \cdot m^2 \cdot s^{-1})^{3N}$). Appelons là h^{3N} , où h est une grandeur⁸ qui s'exprime en $kg \cdot m^2 \cdot s^{-1}$. Alors il n'est pas difficile de définir la valeur de l'entropie à l'équilibre : nous divisons l'espace des phases micro-canonique $\mathcal{T}_E^{E+\delta E}$ ou $\mathcal{T}_E^{E+\delta E}/\sigma_N$ en cellules q_I de tailles $\mu_{Liou}(q_I) = h^{3N}$. Pour cela nous en consommons un nombre $W_h^d := \mu_{Liou}(\mathcal{T}_E^{E+\delta E})/h^{3N}$, si nous travaillons sur $\mathcal{T}_E^{E+\delta E}$ (particules *discernables*) et $W_h^i := \mu_{Liou}(\mathcal{T}_E^{E+\delta E}/\sigma_N)/h^{3N}$ si nous travaillons sur $\mathcal{T}_E^{E+\delta E}/\sigma_N$ (particules *indiscernables*). Alors, pour des particules discernables, l'entropie est :

$$S_h^d := k \log W_h^d = k \log \frac{\mu_{Liou}(\mathcal{T}_E^{E+\delta E})}{h^{3N}} \quad (3.10)$$

et, pour des particules **indiscernables**, l'entropie est :

$$S_h^i := k \log W_h^i = k \log \frac{\mu_{Liou}(\mathcal{T}_E^{E+\delta E}/\sigma_N)}{h^{3N}} = k \log \frac{\mu_{Liou}(\mathcal{T}_E^{E+\delta E})}{h^{3N} N!}. \quad (3.11)$$

L'étape suivante consiste à comprendre comment définir l'entropie d'un système hors de l'équilibre⁹. Un tel système est défini par une mesure ν sur $\mathcal{T}_E^{E+\delta E}$ ou $\mathcal{T}_E^{E+\delta E}/\sigma_N$, qui n'est pas uniforme en général. Ici nous pouvons nous inspirer du paragraphe précédent, où nous avons défini l'entropie relative de deux mesures de probabilité (en adaptant, puisque l'entropie thermodynamique est extensive alors que l'entropie relative de 2 mesures de probabilité ne mesure que le degré d'uniformité de la répartition d'une distribution).

⁷exactement comme pour toutes les quantités extensives (les longueurs, les masses, etc.), pour lesquelles nous devons introduire une unité de mesure (m , kg , etc.).

⁸Par exemple nous choisissons une échelle d'espace minimale δx (exprimées en m) et une échelle d'impulsion minimale δp (exprimée en $kg \cdot m \cdot s^{-1}$) (avec $\delta x \ll \text{diamètre}(\Omega)$ et $\delta p \ll \sqrt{2m\delta E} \ll \sqrt{2mE}$) et nous utilisons des cubes $(q_I)_{I \in [1, M]}$ de la forme $\prod_{i=1}^N \prod_{\alpha=1}^3 [x_i^\alpha, x_i^\alpha + \delta x] \times [p_{i,\alpha}, p_{i,\alpha} + \delta p]$, où $(x_*, p_*) \in T^*\Omega_*^N$. Alors $h = (\delta p)(\delta x)$.

⁹en fait il est nécessaire de supposer que nous sommes *presque* à l'équilibre, par exemple que le système est en évolution quasi-statique, pour pouvoir définir les grandeurs thermodynamiques qui caractérisent nos échantillons ; pour fixer les idées, on peut imaginer que le système est divisé en deux sous-systèmes qui sont pratiquement chacun à l'équilibre et qui évoluent de façon quasi-statique en échangeant du travail et de l'énergie

Notons $\mathcal{P} = \mathcal{T}_E^{E+\delta E}$ ou $\mathcal{T}_E^{E+\delta E}/\sigma_N$ l'espace des phases. Soit $\mu = \mu_{Liou}/\mu_{Liou}(\mathcal{P})$ la mesure de probabilité représentant l'état d'équilibre et soit ν une autre mesure de probabilité sur \mathcal{P} , représentant un autre état. Alors, inspiré par (3.10) ou (3.11) et la section précédente, nous définissons l'entropie correspondant à l'état ν par

$$S_h(\nu) := -k \int_{\mathcal{P}} \frac{d\nu}{d\mu} \log \left(\frac{h^{3N}}{\mu_{Liou}(\mathcal{P})} \frac{d\nu}{d\mu} \right) d\mu.$$

Remarquons que cette quantité peut s'interpréter comme la somme

$$S_h(\nu) = k \log \frac{\mu_{Liou}(\mathcal{P})}{h^{3N}} - k \int_{\mathcal{P}} \frac{d\nu}{d\mu} \log \left(\frac{d\nu}{d\mu} \right) d\mu,$$

c'est à dire la somme de $S(\mu)$, l'entropie à l'équilibre, et de $-k$ fois l'entropie relative $H(\nu|\mu)$. Enfin une troisième expression de l'entropie est :

$$S_h(\nu) := -k \int_{\mathcal{P}} \frac{d\nu}{d\mu_{Liou}} \log \left(h^{3N} \frac{d\nu}{d\mu_{Liou}} \right) d\mu_{Liou}$$

et apparaît comme ($-k$ fois) une généralisation de l'entropie relative de deux mesures de probabilité à des mesures positives qui ne sont pas nécessairement de probabilité. Ici le facteur h^{3N} est essentiel pour que l'argument du logarithme soit sans dimension. Nous remarquons bien sûr que l'entropie dépend du choix de l'échelle h . Ainsi pour tout $\varepsilon > 0$, il est immédiat que $S_{\varepsilon h}(\nu) = S_h(\nu) + 3Nk \log \varepsilon$. Cette indétermination n'est pas trop gênante, puisque, dans la première définition de l'entropie, proposée par Clausius sur la base de l'équation $dS = q/T$ (q : gain infinitésimal de chaleur), l'entropie n'était défini qu'à une constante additive près.

Définition 4 Soit ν la mesure de probabilité sur l'espace des phases $T^*\Omega_*^N$ (discernabilité) ou $T^*\Omega_*^N/\sigma_N$ (indiscernabilité) représentant un état macroscopique donné, et donc caractérisé par des grandeurs thermodynamiques E, V, \dots . Soit $\mathcal{P} \subset T^*\Omega_*^N$ (ou $T^*\Omega_*^N/\sigma_N$) le support de ν . Supposons que, sur \mathcal{P} , ν a une densité régulière ρ par rapport à la mesure de Liouville, c'est à dire : $d\nu = \rho d\mu_{Liou}$. Alors on appelle entropie statistique de cet état macroscopique la quantité

$$S_h(E, V, \dots) = S_h(\nu) = -k \int_{\mathcal{P}} \rho \log (h^{3N} \rho) d\mu_{Liou},$$

où h est une grandeur choisie une fois pour toute qui s'exprime en $\text{kg}\cdot\text{m}^2\cdot\text{s}^{-1}$ et k est la constante de Boltzmann.

Dans la démarche précédente, le choix de l'échelle h apparaît tout à fait arbitraire et est rendu nécessaire pour des raisons d'homogénéité et parce que nous avons besoin de fixer une certaine échelle pour « compter » le nombre d'états. Il se trouve que la nature privilégie en fait une certaine valeur, très petite, qui sera sensible en physique quantique : la constante de Planck.

3.4.3 Entropie microcanonique

Dans le formalisme microcanonique, l'ensemble des états accessibles pour le système qu'on considère est une tranche $\mathcal{T}_E^{E+\delta E}$ de l'espace des phases ou bien son quotient $\mathcal{T}_E^{E+\delta E}/\sigma_N$. La mesure ν associée à un état macroscopique

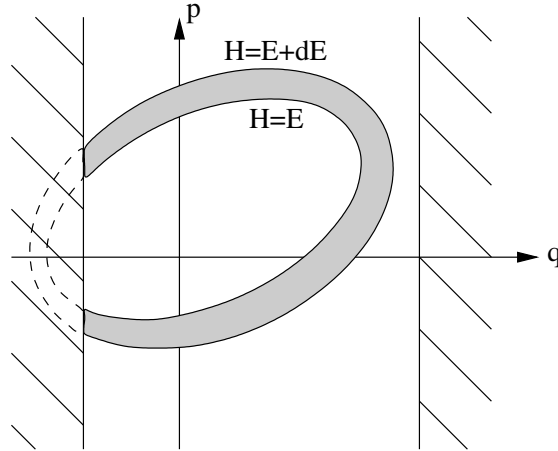


FIG. 3.2 – Représentation schématique des états microscopiques accessibles par un système. Ici, le volume du système est borné, ce qui empêche q de sortir d'un intervalle dont l'extérieur est hachuré. Lors d'une détente de Joule par exemple, cette contrainte est supprimée sans que soit rompu l'isolement du système.

de notre système a donc une densité ρ nulle hors de $\mathcal{T}_E^{E+\delta E}$.

Entropie micro-canonique du gaz dilué

Nous allons maintenant calculer l'entropie pour un gaz dilué à l'équilibre. Notons \mathcal{P} l'espace des états ($\mathcal{T}_E^{E+\delta E}$ ou $\mathcal{T}_E^{E+\delta E}/\sigma_N$). A l'équilibre la mesure de probabilité μ est distribuée uniformément par rapport à la mesure de Liouville μ_{Liou} et donc on a $\mu = \rho\mu_{Liou}$ sur \mathcal{P} , où ρ est constant et est donné par $\int_{\mathcal{P}} \rho d\mu_{Liou} = 1$, c'est à dire $\rho = (\mu_{Liou}(\mathcal{P}))^{-1} = (\int_{\mathcal{P}} d\mu_{Liou})^{-1} = (h^{3N}W_h)^{-1}$. Alors

$$S_h = -k \int_{\mathcal{P}} \rho \log(h^{3N}\rho) d\mu_{Liou} = -k \log \frac{1}{W_h} \int_{\mathcal{P}} \rho d\mu_{Liou} = k \log W_h.$$

Il nous faut donc calculer W_h .

Particules discernables — Nous calculons tout d'abord $W_h^d := h^{-3N} \int_{\mathcal{T}_E^{E+\delta E}} d\mu_{Liou}$, ce qui signifie que nous supposons les particules discernables. Il s'agit de calculer le volume d'une tranche comprise entre deux sphères de rayons respectifs R et $R + \delta R$ (fois le volume de $|\Omega|^N$), où $R = \sqrt{2mE}$ et $R + \delta R = \sqrt{2m(E + \delta E)}$. Cela entraîne $\delta R \simeq \sqrt{2mE}\delta E/2\sqrt{E}$. Nous poserons $\delta E = \alpha E$, où $\alpha \in]0, 1[$, de sorte que $\delta R \simeq \alpha\sqrt{2mE}/2$. Alors, en supposant que $\alpha \ll 1$,

$$\begin{aligned} h^{3N}W_h^d &\simeq (\delta R)R^{3N-1}|S^{3N-1}||\Omega|^N \\ &= \frac{\alpha\sqrt{2mE}}{2}\sqrt{2mE}^{3N-1} \frac{2\pi^{\frac{3N}{2}}}{\Gamma(\frac{3N}{2})}|\Omega|^N \\ &= \alpha\sqrt{2mE}^{3N} \frac{\pi^{\frac{3N}{2}}}{\Gamma(\frac{3N}{2})}V^N. \end{aligned}$$

Donc

$$\begin{aligned} S_h^d &= k \log \left(\alpha h^{-3N} \sqrt{2mE}^{3N} \frac{\pi^{\frac{3N}{2}}}{\Gamma(\frac{3N}{2})} V^N \right) \\ &\simeq k \log(\alpha) + k \log \left(\frac{h^{-3N} (2mE)^{3N/2} \pi^{3N/2} V^N}{(3N/2 - 1)^{3N/2-1} e^{-3N/2+1} \sqrt{\pi(3N-2)}} \right) \\ &\simeq k \log(\alpha) + k \log \left(\frac{h^{-3N} (2\pi mE)^{3N/2} V^N}{(3N/2)^{3N/2} e^{-3N/2}} \right) \\ &= k \log(\alpha) + kN \left(\log \left(\frac{V}{h^3} \left(\frac{4\pi mE}{3N} \right)^{3/2} \right) + \frac{3}{2} \right). \end{aligned}$$

C'est le moment de fixer α : nous supposons que, en plus de $\alpha \ll 1$, $|\log \alpha| \ll N$, ce qui est parfaitement raisonnable. Alors le premier terme dans l'expression qui précède est négligeable. Enfin en utilisant $E = \frac{3}{2}NkT$, on obtient

$$S_h^d \simeq kN \left(\log \left(\frac{V}{h^3} (2\pi mkT)^{3/2} \right) + \frac{3}{2} \right).$$

Cette quantité présente l'inconvénient de ne pas être extensive, notamment à cause de l'argument du logarithme qui croît linéairement avec le volume V . Ce défaut va être corrigé en supposant l'indiscernabilité des particules.

Particules indiscernables — Alors l'espace des phases est $\mathcal{P} = \mathcal{T}_E^{E+\delta E} / \sigma_N$ et nous devons calculer $W_h^i := h^{-3N} \int_{\mathcal{T}_E^{E+\delta E} / \sigma_N} d\mu_{Liou}$. Le calcul est très simple et il suffit de diviser par $N!$: $W_h^i = W_h^d / N!$. Cela a pour effet que $S_h^i = k \log W_h^i$ est obtenue en retranchant la quantité $k \log N!$ à S_h^d . Comme

$$-k \log N! \simeq -kN(\log N - 1),$$

on en déduit que

$$S_h^i \simeq kN \left(\log \left(\frac{V}{Nh^3} (2\pi mkT)^{3/2} \right) + \frac{5}{2} \right). \quad (3.12)$$

Et on peut constater cette fois-ci que S_h^i est une quantité extensive.

Retrouver la température

Envisageons un gaz quelconque, dont la valeur de l'énergie est fixée à E et qui, par conséquent, peut être modélisé par le formalisme microcanonique. Supposons que, par une évolution quasi-statique, le système absorbe une quantité d'énergie infinitésimale et que cette évolution se fait sans changement du nombre de particules N , ni changement du volume V . Cela signifie, sur le plan de la description microscopique, que l'espace des phases $T^*S\Omega_*^N$ et l'hamiltonien $H \in \mathcal{C}^1(T^*S\Omega_*^N)$ ne changent pas. Ce qui va changer, c'est \mathcal{H}_E , c'est à dire la répartition statistique et donc la mesure μ . Sur le plan de la description macroscopique, cet apport d'énergie se fait sous forme de chaleur : $dE = dU = q$ (avec les notations du chapitre 1).

Il découle des considérations thermodynamiques que nous avons faites lorsque nous avons défini l'entropie de Clausius, que $dS^i = q/T = dE/T$. Nous en déduisons une définition de la température à partir de l'entropie statistique micro-canonique

$$\frac{1}{T} := \left(\frac{\partial S_h^i}{\partial E} \right)_{|N,V}. \quad (3.13)$$

À titre d'application du paragraphe précédent, en utilisant l'expression (3.12) pour calculer la variation infinitésimale de l'entropie S_h^i , nous obtenons

$$dS_h^i = kN d \left(\log \left(\frac{V}{Nh^3} (2\pi mk)^{3/2} \right) + \frac{5}{2} \right)_{|N,V} + k \frac{3N}{2} d \log T = k \frac{3N}{2} \frac{dT}{T}.$$

Et, en utilisant $E = \frac{3}{2}NkT$, nous obtenons $dS_h^i = dU/T$, ce qui confirme que la relation (3.13) est en adéquation avec la définition de Clausius de l'entropie.

3.5 Le formalisme canonique

3.5.1 L'ensemble canonique

Le formalisme microcanonique décrit correctement des systèmes **isolés** mais il est en fait bien plus fréquent qu'un système \mathcal{S} soit en mesure d'échanger de la chaleur avec un système beaucoup plus gros que lui \mathcal{T} et dont la température n'est pas sensiblement affectée par ces échanges. Par exemple, c'est le cas d'une enceinte en verre contenant du gaz et placée dans une pièce remplie d'air à une certaine température. Dans ce cas, on appelle \mathcal{T} un *thermostat*. Le système \mathcal{S} peut alors se trouver dans des états microscopiques d'énergie variable et, contrairement à ce qui se passe pour un système isolé, ces états ne sont pas équiprobables. Il faut une nouvelle loi de probabilités pour décrire correctement l'état macroscopique de \mathcal{S} , qui ne sera pas portée par une hypersurface ou une tranche de l'espace des phases, mais par l'espace tout entier.

La présence du thermostat entraîne une autre différence importante entre les formalismes microcanonique et canonique. Le premier s'applique à de grands systèmes, contenant au moins des milliards de particules. Le second peut se comprendre comme l'application du premier au système $\mathcal{S} \cup \mathcal{T}$, qu'on peut

souvent considéré comme isolé. Or, pour que \mathcal{T} puisse jouer son rôle de thermostat, il doit nécessairement être très gros, si bien que \mathcal{S} peut, lui, être aussi petit qu'on veut, et même être constitué d'une seule particule. Ainsi, le formalisme canonique peut s'appliquer à des systèmes arbitrairement petits.

Nous l'avons d'ailleurs déjà fait sans le dire dans la section précédente, où, en projetant la mesure décrivant l'état d'un gaz dilué sur les composantes correspondant à un petit nombre de particules, nous avons (re)trouvé la distribution de Maxwell-Boltzmann. Nous verrons qu'elle est un cas particulier de l'ensemble canonique.

Considérons donc un système de hamiltonien H défini sur l'espace des phases $T^*\Omega_\star^N$ en contact avec un thermostat à la température T . Pour déterminer la mesure $\mu = \rho \mu_{Liou}$ correspondant à l'équilibre de ce système, nous allons considérer que l'action du thermostat revient à fixer la valeur moyenne de l'énergie, à une certaine valeur E . Autrement dit, ρ doit satisfaire la contrainte

$$\int_{T^*\Omega_\star^N} H \rho \, d\mu_{Liou} = E. \quad (3.14)$$

Par ailleurs, ρ doit également vérifier

$$\int_{T^*\Omega_\star^N} \rho \, d\mu_{Liou} = 1, \quad (3.15)$$

pour que μ soit une mesure de probabilités. Pour déterminer ρ parmi toutes les fonctions qui satisfont ces deux contraintes, nous allons également demander qu'elle maximise l'entropie statistique

$$S(\rho) = -k \int_{T^*\Omega_\star^N} \rho \log(h^{3N} \rho) \, d\mu_{Liou}. \quad (3.16)$$

Pour résoudre, un peu formellement, ce problème, on prend ρ qui satisfait les deux conditions (3.15), (3.14) et maximise l'entropie. Alors, pour toute petite variation $\delta\rho$ de ρ telle que $\rho + \delta\rho$ satisfasse encore (3.15) et (3.14), on doit avoir $S(\rho + \delta\rho) = S(\rho)$ à des termes du second ordre près. Or la variation de S s'écrit :

$$S(\rho + \delta\rho) - S(\rho) = \int_{T^*\Omega_\star^N} \delta\rho (\log(h^{3N} \rho) + 1) \, d\mu_{Liou} + O(\delta\rho^2).$$

D'autre part, le fait que $\rho + \delta\rho$ satisfait encore (3.15) et (3.14) s'écrit

$$\int_{T^*\Omega_*^N} \delta\rho d\mu_{Liou} = \int_{T^*\Omega_*^N} H\delta\rho d\mu_{Liou} = 0.$$

Définissons $\psi(\delta\rho) = \int \delta\rho(\log(h^{3N}\rho) + 1) d\mu_{Liou}$, $\phi_1(\delta\rho) = \int \delta\rho d\mu_{Liou}$ et $\phi_2(\delta\rho) = \int H\delta\rho d\mu_{Liou} = 0$. On a trois formes linéaires ψ, ϕ_1 et ϕ_2 sur un espace de fonctions, la première dépendant de ρ qui est inconnue. On demande que le noyau de ψ contienne l'intersection de ceux de ϕ_1 et ϕ_2 . Cela est équivalent à demander que ψ soit une combinaison linéaire de ϕ_1 et ϕ_2 , ce qui signifie qu'il existe deux constantes α et β telles que $\log(h^{3N}\rho) + 1 = \alpha - \beta H$, autrement dit

$$\rho = h^{-3N} C e^{-\beta H}.$$

Ces deux constantes doivent être déterminées, selon l'expression de H , par les conditions (3.15) et (3.14). On aura toujours en particulier

$$\frac{1}{C} = \frac{1}{h^{3N}} \int_{T^*\Omega_*^N} e^{-\beta H} d\mu_{Liou},$$

et on notera Z ce nombre, qui s'appelle aussi la *fonction de partition*.

3.5.2 Identification de β

Dans le cas d'un gaz dilué, où $H(x_*, p_*) = \frac{|p_*|^2}{2m}$, nous pouvons facilement calculer β en fonction de E . En effet,

$$\begin{aligned} \frac{1}{Z} \int_{T^*\Omega_*^N} H e^{-\beta H} d\mu_{Liou} &= \frac{1}{2m} \frac{1}{\int_{\mathbb{R}^{3N}} e^{-\frac{\beta}{2m} \sum |p_i|^2} (dp)^{3N}} \int_{\mathbb{R}^{3N}} \sum_{i=1}^{3N} |p_i|^2 e^{-\frac{\beta}{2m} \sum |p_i|^2} (dp)^{3N} \\ &= \frac{3N}{2m} \frac{1}{\int_{-\infty}^{\infty} e^{-\frac{\beta}{2m} p_1^2} dp_1} \int_{-\infty}^{\infty} p_1^2 e^{-\frac{\beta}{2m} p_1^2} dp_1 \\ &= \frac{3N}{2m} \frac{m}{\beta} = \frac{3N}{2\beta}. \end{aligned}$$

Pour un gaz dilué isolé, on sait que E est reliée à la température par $E = \frac{3}{2} N k T$. En utilisant cette égalité dans le formalisme canonique, on obtient la relation fondamentale

$$\beta = \frac{1}{kT}.$$

On définit donc l'ensemble canonique pour un système de hamiltonien H mis en contact avec un thermostat à la température T comme étant la mesure

$$\mu_{can} = \frac{1}{h^{3N} Z} e^{-\frac{H}{kT}} \mu_{Liou}.$$

Exemple 1 Dans une enceinte contenant un gaz dilué à la température T , choisissons une molécule par la pensée. Elle ne constitue pas un système isolé : elle échange constamment de l'énergie avec le reste du gaz au cours des chocs qu'elle subit avec les autres molécules. Cependant, la température du système formé par ces autres molécules n'est pas sensiblement affectée par ces chocs : ils mettent en jeu une énergie de l'ordre de kT , alors que l'énergie du reste des molécules est de l'ordre de NkT . La molécule que nous avons choisie est donc en situation canonique, à la température T . Son hamiltonien s'écrit $H(x, p) = |p|^2/2m$ donc son état peut être représenté par la mesure

$$\frac{1}{hZ} e^{-\frac{|p|^2}{2mkT}} dx dp,$$

qui est la distribution de Maxwell-Boltzmann.

3.5.3 Répartition de l'énergie

Le facteur exponentiel dans la mesure canonique indique qu'entre deux états microscopiques, le plus énergétique est aussi le moins probable. Cependant, entre deux énergies données, celle à laquelle on est le plus susceptible de trouver le système n'est pas forcément la plus basse. En effet, les états d'énergie plus élevée sont moins probables mais aussi beaucoup plus nombreux. Nous allons vérifier dans le cas du gaz dilué que ces deux phénomènes contraires se compensent autour de la valeur moyenne imposée à l'énergie. La distribution de l'énergie du système présente un pic autour de cette valeur moyenne dont nous allons aussi évaluer la largeur.

Commençons par exprimer la loi de l'énergie du système sous la mesure μ_{can} . Pour toute fonction f , on a

$$\langle f(H) \rangle_{can} = \frac{1}{Z} \int f\left(\frac{r^2}{2m}\right) r^{3N-1} e^{-\frac{r^2}{2mkT}} dr |S^{3N-1}| d\sigma,$$

en se plaçant en coordonnées sphériques et en notant $d\sigma$ la mesure uniforme de masse 1 sur la sphère unité. Comme l'intégrande est invariant par rotation, la partie sphérique disparaît et il ne reste que

$$\begin{aligned}\langle f(H) \rangle_{can} &= \frac{1}{Z} \int f\left(\frac{r^2}{2m}\right) r^{3N-1} e^{-\frac{r^2}{2mkT}} dr \\ &= \frac{1}{\int s^{\frac{3N-1}{2}} e^{-\frac{s}{kT}} ds} \int_0^\infty f(s) s^{\frac{3N-1}{2}} e^{-\frac{s}{kT}} ds.\end{aligned}$$

Posons $d(s) = s^{\frac{3N-1}{2}} e^{-\frac{s}{kT}}$. On vérifie facilement que d' s'annule en $s_{max} = \frac{3N-1}{2}kT \simeq \frac{3}{2}NkT = E$. Ainsi, la densité est bien maximale au voisinage de l'énergie moyenne qui est imposée au système.

On peut aussi calculer d'' et vérifier que

$$d''(s_{max}) = -(kT)^{\frac{3N-5}{2}} e^{-\frac{3N-1}{2}} \left(\frac{3N-1}{2}\right)^{\frac{3N-3}{2}}.$$

En négligeant les termes d'ordre supérieur à deux et donc en identifiant localement le graphe de la densité à une parabole, on peut estimer la largeur du pic de la distribution par la formule

$$\delta E \simeq \sqrt{\left| \frac{d(s_{max})}{d''(s_{max})} \right|}.$$

Le rapport sous la racine carrée est égal à $\frac{3N-1}{2}N(kT)^2$, c'est-à-dire qu'il est de l'ordre de grandeur de $N(kT)^2 \simeq \frac{E^2}{N}$. Ainsi, δE est de l'ordre de $\frac{E}{\sqrt{N}}$ et on obtient

$$\frac{\delta E}{E} \simeq \frac{1}{\sqrt{N}}.$$

Ce calcul indique, lorsque le système est macroscopique (ce qui, encore une fois, n'est pas toujours le cas dans le formalisme canonique), que l'énergie fluctue très peu autour de sa valeur moyenne. A la limite où N tend vers l'infini, qui s'appelle aussi *limite thermodynamique*, l'énergie est toujours égale à sa valeur moyenne : la distribution canonique se concentre sur une hypersurface d'énergie et devient finalement identique à la distribution microcanonique¹⁰. Cette propriété entre dans le cadre de l'équivalence des différents ensembles (microcanonique, canonique, ...) à la limite thermodynamique.

¹⁰La signification mathématique de cette remarque demanderait à être précisée, car lorsque N varie, on se trouve dans des espaces des phases de plus en plus gros.

Appendice

La fonction Γ — Elle est définie par $\Gamma(z) := \int_0^\infty e^{-x} x^{z-1} dx$. Elle satisfait la relation $\Gamma(z) = (z-1)\Gamma(z-1)$ si $z > 1$. On en déduit $\Gamma(n) = (n-1)!, \forall n \in \mathbb{N}^*$.

La formule de Stirling

$$\Gamma(z+1) \simeq z^z e^{-z} \sqrt{2\pi z},$$

au sens où $\lim_{z \rightarrow \infty} \frac{z^z e^{-z} \sqrt{2\pi z}}{\Gamma(z+1)} = 1$.

Le volume de la sphère $S^{n-1} := \{x \in \mathbb{R}^n / |x| = 1\}$ — Il vaut

$$|S^{n-1}| = 2 \frac{\pi^{\frac{n}{2}}}{\Gamma(\frac{n}{2})}.$$

Pour le démontrer on calcule de 2 façons différentes $\int_{\mathbb{R}^n} e^{-|x|^2} dx^1 \cdots dx^n$: par le théorème de Fubini et en utilisant $e^{-|x|^2} = e^{-(x^1)^2} \cdots e^{-(x^n)^2}$,

$$\int_{\mathbb{R}^n} e^{-|x|^2} dx^1 \cdots dx^n = \left(\int_{\mathbb{R}} e^{-t^2} dt \right)^n = \sqrt{\pi}^n.$$

En intégrant en coordonnées sphériques :

$$\int_{\mathbb{R}^n} e^{-|x|^2} dx^1 \cdots dx^n = |S^{n-1}| \int_0^\infty e^{-r^2} r^{n-1} dr.$$

Alors on fait le changement de variable $r^2 = t$, $dr = dt/2\sqrt{t}$:

$$\int_{\mathbb{R}^n} e^{-|x|^2} dx^1 \cdots dx^n = |S^{n-1}| \int_0^\infty e^{-t} t^{\frac{n}{2}-1} \frac{dt}{2} = \frac{|S^{n-1}|}{2} \Gamma\left(\frac{n}{2}\right).$$

On en déduit ainsi le résultat.

Approximation de $(1 + \frac{x}{n})^n$, lorsque $n \rightarrow \infty$ —

$$\left(1 + \frac{x}{n}\right)^n \simeq e^x.$$

Pour le démontrer on écrit que $\log\left(\left(1 + \frac{x}{n}\right)^n e^{-x}\right) = n \log\left(1 + \frac{x}{n}\right) - x$, quantité qui tend vers 0 lorsque $n \rightarrow \infty$.

Equivalent de $\frac{\Gamma(n)}{\Gamma(n-1/2)}$ lorsque $n \rightarrow \infty$ —

$$\frac{\Gamma(n)}{\Gamma(n-1/2)} \simeq \sqrt{n}.$$

Pour le démontrer

$$\begin{aligned} \frac{\Gamma(n)}{\Gamma(n-1/2)} &\simeq \frac{(n-1)^{(n-1)} e^{-n+1} \sqrt{2\pi(n-1)}}{(n-3/2)^{(n-3/2)} e^{-n+3/2} \sqrt{2\pi(n-3/2)}} \\ &\simeq \frac{n^{n-1} \left(1 - \frac{1}{n}\right)^{n\left(1-\frac{1}{n}\right)}}{n^{n-3/2} \left(1 - \frac{3}{2n}\right)^{n\left(1-\frac{3}{2n}\right)} e^{1/2}} \\ &\simeq \sqrt{n} \frac{e^{-\left(1-\frac{1}{n}\right)}}{e^{-\frac{3}{2}\left(1-\frac{3}{2n}\right)} e^{1/2}} \simeq \sqrt{n}. \end{aligned}$$

Chapitre 4

Interactions thermodynamiques entre la lumière et la matière

Lorsque Max Planck, à la fin du XIX^{ème} siècle, aborda l'étude des échanges thermodynamiques entre la matière et la lumière, il était alors un spécialiste de la thermodynamique « classique » et, comme la plupart de ses contemporains, relativement peu convaincu par la physique statistique de Boltzmann. A tel point qu'un de ses buts était de mettre en évidence des contradictions dans la théorie de Boltzmann. Il ne se doutait pas que, grâce notamment à sa grande honnêteté intellectuelle, ses travaux déboucheraient sur une confirmation éclatante des idées de Boltzmann et surtout sur une découverte qui allait bouleverser la physique.

4.1 Préliminaire : l'équipartition de l'énergie

Commençons par examiner une conséquence de la physique statistique : nous envisageons un système avec M degrés de liberté indépendants possédant chacun une énergie quadratique. Cela peut être par exemple :

Un gaz dilué de N particules — L'ensemble des états d'un système (sans tenir compte de l'indiscernabilité) est $T^*\Omega_*^N$, où Ω est un ouvert borné de \mathbb{R}^3 et l'hamiltonien est :

$$H(x_*, p_*) := \sum_{i=1}^N \frac{|p_i|^2}{2m},$$

et alors $M = 3N$. Ou bien :

Un système de N oscillateurs harmoniques — Cela peut modéliser par exemple les petits mouvement d'atomes répartis en un réseau cristallin autour de leurs positions d'équilibre. Alors N est égal à 3 fois (degrés de liberté) le nombre d'atomes. L'ensemble des états du système est $T^*\mathbb{R}^N$, avec

$$H(x_*, p_*) := \sum_{i=1}^N \frac{(p_i)^2}{2m} + m\omega^2 \frac{(x^i)^2}{2}.$$

Alors $M = 6N$.

Dans les deux cas, la situation est modélisable par un système de $M + L$ degrés de liberté $(z_A)_{A \in \llbracket 1, M+L \rrbracket}$ (qui représentent les coordonnées $p_{i,\alpha}$, x_α^i dans le premier exemple) et l'hamiltonien

$$H(z) := \sum_{A=1}^M c_A \frac{z_A^2}{2},$$

où chaque coefficient c_A est dans $]0, \infty[$. Nous supposons ici que, pour tout $A \in \llbracket 1, M \rrbracket$, z_A prend ses valeurs dans \mathbb{R} . Notons que les variables z_A , pour $A \in \llbracket M+1, M+L \rrbracket$, n'apportent aucune contribution dans l'énergie (comme c'est le cas pour les variables x_α^i dans le premier exemple). On doit donc supposer que $(z_A)_{A \in \llbracket M+1, M+L \rrbracket}$ prend ses valeurs dans un ensemble borné K , pour que la mesure canonique que nous allons considérer plus loin soit de masse totale finie.

Nous considérons le formalisme canonique et introduisons pour cela la fonction de partition

$$\begin{aligned} Z(\beta) &= \frac{1}{\sigma} \int_{\mathbb{R}^M \times K} e^{-\beta H(z)} \prod_{A=1}^{M+L} dz_A = \frac{|K|}{\sigma} \int_{\mathbb{R}^M} e^{-\beta H(z)} \prod_{A=1}^M dz_A \\ &= \frac{|K|}{\sigma} \prod_{A=1}^M \int_{\mathbb{R}} e^{-\beta c_A z_A^2 / 2} dz_A, \end{aligned}$$

où $\sigma = h^{M+L}$ est l'« unité de mesure » sur $\mathbb{R}^M \times K$. La mesure de probabilité canonique sur $\mathbb{R}^M \times K$ est alors

$$d\mu := \frac{1}{\sigma Z(\beta)} e^{-\beta H(z)} \prod_{A=1}^{M+L} dz_A.$$

Lemme 1 Pour tout $A \in \llbracket 1, M \rrbracket$, la valeur moyenne de l'énergie correspondant au degré de liberté z_A vaut :

$$\left\langle c_A \frac{z_A^2}{2} \right\rangle = \frac{kT}{2},$$

où $\beta kT = 1$.

Preuve — Il s'agit d'un calcul :

$$\begin{aligned} \left\langle c_A \frac{z_A^2}{2} \right\rangle &= \frac{1}{\sigma Z(\beta)} \int_{\mathbb{R}^{M \times K}} c_A \frac{z_A^2}{2} e^{-\beta \sum_{B=1}^M c_B z_B^2 / 2} \prod_{B=1}^{M+L} dz_B \\ &= \frac{|K|}{\sigma Z(\beta)} \prod_{B \neq A; B=1}^M \left(\int_{\mathbb{R}} e^{-\beta c_B z_B^2 / 2} dz_B \right) \int_{\mathbb{R}} c_A \frac{z_A^2}{2} e^{-\beta c_A z_A^2 / 2} dz_A \\ &= \left(\int_{\mathbb{R}} e^{-\beta c_A z_A^2 / 2} dz_A \right)^{-1} \int_{\mathbb{R}} c_A \frac{z_A^2}{2} e^{-\beta c_A z_A^2 / 2} dz_A. \end{aligned}$$

On pose alors $t = \sqrt{\beta c_A / 2} z_A$, $dt = \sqrt{\beta c_A / 2} dz_A$, et

$$\begin{aligned} \left\langle c_A \frac{z_A^2}{2} \right\rangle &= \left(\int_{\mathbb{R}} e^{-t^2} \sqrt{\frac{2}{\beta c_A}} dt \right)^{-1} \int_{\mathbb{R}} \frac{t^2}{\beta} e^{-t^2} \sqrt{\frac{2}{\beta c_A}} dt \\ &= \frac{1}{\beta} \left(\int_{\mathbb{R}} e^{-t^2} dt \right)^{-1} \int_{\mathbb{R}} t^2 e^{-t^2} dt = \frac{1}{2\beta}. \end{aligned}$$

d'où le résultat. ■

4.2 Le problème du rayonnement du corps noir

A la fin du XIX^{ème} siècle, les interactions entre la lumière et la matière sont la source d'énigmes face auxquelles on disposait d'assez peu de moyens théoriques. Par exemple il n'était pas évident de savoir comment la thermodynamique s'applique à la lumière. Et pourtant on peut expérimenter des relations évidentes entre la température d'un corps et l'intensité et la distribution en longueur d'onde du rayonnement lumineux qu'il émet. De même il est évident que le simple fait de rester exposé au soleil conduit à une augmentation de la température. A l'époque les travaux de Maxwell avaient permis de montrer de façon convaincante que la lumière était une onde du

champ électromagnétique (voir l'appendice de ce chapitre). M. Planck, qui ne croyait pas à la théorie de L. Boltzmann et en particulier à son interprétation statistique de l'entropie, pensait, parce que les équations de Maxwell sont linéaires et donc ses solutions sont superposables, que l'étude de la lumière lui permettrait de mettre en défaut cette interprétation statistique.

4.2.1 La pensée thermodynamique de Planck

M. Planck envisage une expérience relativement facile à mettre en oeuvre, dans laquelle on considère une boîte totalement fermée constituée par un matériau opaque (que nous supposons être un conducteur parfait). L'intérieur de la boîte est vide et le tout est laissé au repos, à une certaine température T pendant un temps suffisamment long pour qu'un équilibre thermodynamique ait lieu, les ondes électromagnétiques à l'intérieur pouvant échanger de l'énergie avec les molécules qui constituent la boîte. En pratiquant une petite ouverture dans la boîte, on peut alors, par des mesures sur l'intensité du rayonnement qui en sort en déduire la densité volumique d'énergie électromagnétique. Rappelons que toute onde lumineuse peut être décomposée en une somme infinie (transformée de Fourier) d'ondes monochromatiques pour lesquelles les champs électriques et magnétiques sont de la forme :

$$\vec{E}(t, x) = \text{Re} \left(\mathbf{E} e^{i(\omega t + \vec{k} \cdot \vec{x})} \right) \quad \vec{B}(t, x) = \text{Re} \left(\mathbf{B} e^{i(\omega t + \vec{k} \cdot \vec{x})} \right),$$

où ω est la pulsation ($\omega = 2\pi\nu$, où ν est la fréquence), $\vec{k} \in (\mathbb{R}^3)^*$ est le vecteur d'onde et satisfait $c|\vec{k}| = \omega$. De plus, si on se donne une base orthonormée directe $(\epsilon_1, \epsilon_2, \epsilon_3)$ de \mathbb{R}^3 telle que $c\vec{k} = \omega\epsilon_3$, alors on a nécessairement $\mathbf{E} = e^{i\varphi}(-b^2\epsilon_1 + b^1\epsilon_2)$ et $\mathbf{B} = e^{i\varphi}(b^1\epsilon_1 + b^2\epsilon_2)$, où φ , b^1 et b^2 sont des constantes réelles (voir l'appendice de ce chapitre). Les deux degrés de liberté dans le plan engendré par (ϵ_1, ϵ_2) correspondent à l'existence de deux polarisations possibles. Il est à remarquer qu'il est possible expérimentalement de sélectionner des valeurs du nombre d'onde¹ ω (ou de façon équivalente de la longueur d'onde $\lambda = 2\pi c/\omega$). Nous pouvons ainsi définir la densité d'énergie électromagnétique par unité de volume d'espace et par unité de nombre d'onde : c'est une fonction $\rho(x, \omega, T)$, où T est la température, telle que la quantité d'énergie

¹et même des valeurs du vecteur d'onde \vec{k} (dont la direction est celle de la propagation du rayon lumineux et la longueur dépend de la longueur d'onde) et de la polarisation (grâce à des filtres)

électromagnétique contenue dans un voisinage d'un point $x \in \mathbb{R}^3$ de volume $(\delta x)^3$ et dont le nombre d'onde est contenu dans l'intervalle $[\omega, \omega + \delta\omega]$ (ou de façon équivalente la longueur d'onde est contenue dans $[\lambda, \lambda + \delta\lambda]$, avec $\lambda\omega = (\lambda + \delta\lambda)(\omega + \delta\omega) = 2\pi c$) est égale à :

$$\rho(x, \omega, T)(\delta x)^3 \delta\omega.$$

Dans la suite nous ne considérerons que des états d'équilibre, pour lesquels la valeur de ρ ne dépend pas de la position, et nous supposons donc que $\rho(x, \omega, T) = \rho(\omega, T)$. Comme il est possible² de mesurer expérimentalement la densité $\rho(\omega, T)$, nous pouvons considérer le sous-système thermodynamique obtenu en isolant par la pensée les ondes lumineuses dont le nombre d'onde est dans un intervalle $[\omega, \omega + \delta\omega]$. Pendant un temps relativement court (de façon à ce que les échanges entre la lumière et la boîte aient des effets négligeables), ce « sous-système » peut légitimement être considéré comme isolé parce que précisément les équations de Maxwell sont linéaires et donc il n'y a pas d'échange directs entre différents modes. Ainsi, il est possible d'estimer l'entropie pour ce sous-système, en adaptant la relation de Clausius $dS = q/T$: nous définissons la densité d'entropie par unité de volume d'espace et unité de nombre d'onde comme étant la quantité $s(\omega, T)$ caractérisée par la relation

$$ds(\omega, T) = \frac{d\rho(\omega, T)}{T},$$

où ici la différentielle d agit sur les variables d'états thermodynamiques u et T du système (et non sur \vec{k} qui joue ici le rôle d'un paramètre).

L'originalité du raisonnement³ de M. Planck a été de s'appuyer sur la relation entre la densité d'énergie ρ et la densité d'entropie s , alors que la plupart

²en pratique on mesure la quantité d'énergie transmise à travers un élément de surface dS pendant une durée dt , dans des directions comprises dans un angle solide $d\Omega$, dans un intervalle de nombre d'onde $d\omega$ et selon une certaine polarisation, quantité notée $J(\omega, T)dtd\omega dSd\Omega$, où $J(\omega, T)$ est la *radiance*. Cette densité est reliée à la densité d'énergie électromagnétique par unité de volume d'espace et par unité de nombre d'onde $\rho(\omega, T)$ par la relation $J(\omega, T) = c\rho(\omega, T)/8\pi$.

³en réalité M. Planck a travaillé sur un autre type de densité d'énergie et d'entropie. En effet il a suivi une approche un peu plus compliquée : il n'a pas raisonné sur la lumière directement mais sur l'état thermodynamique de la matière lorsqu'elle est à l'équilibre avec la lumière et il a modélisé les fluctuations de la position des particules chargées de la matière par de petits oscillateurs harmoniques échangeant de l'énergie avec le champ électromagnétique. Cela se comprend dans la mesure où il ignorait si les lois de la thermo-

de ses contemporains cherchaient en vain à deviner ce qui se passait en travaillant directement sur le lien entre la température et la fréquence $\omega/2\pi$. Il est parti du principe que l'entropie doit être une fonction concave de l'énergie et donc en particulier que $\frac{d^2s}{(d\rho)^2} < 0$ et cela pour deux raisons : parce qu'il avait foi en la thermodynamique et parce que le fait que $\frac{d^2s}{(d\rho)^2} < 0$ était une des rares informations que l'on avait pu vérifier expérimentalement.

4.2.2 Position du problème

Nous supposons pour simplifier que la boîte est un cube de côtés de longueur ℓ . Soit $\mathbb{N}_2^3 := (\mathbb{N}^*)^3$ et $\mathbb{N}_1^3 := \{\vec{n} \in \mathbb{N}^3 / \exists! \alpha \in \llbracket 1, 3 \rrbracket, n_\alpha = 0\}$. Nous considérons l'ensemble dénombrable d'indices :

$$\Gamma := \{(\vec{n}, 1) / \vec{n} \in \mathbb{N}_1^3\} \cup \{(\vec{n}, A) / \vec{n} \in \mathbb{N}_2^3, A \in \{1, 2\}\}.$$

Une décomposition de Fourier des champs électriques et magnétiques à l'intérieur de la boîte, tenant compte des conditions au bord, et une formulation variationnelle des équations de Maxwell conduisent à identifier la dynamique des équations de Maxwell dans la boîte à celles d'une assemblée d'oscillateurs harmoniques numérotés par les indices $\gamma = (\vec{n}, A)$ parcourant Γ . Physiquement à chaque $\gamma \in \Gamma$ correspond un mode stationnaire de vecteur d'onde $\vec{k} = \pm\pi\vec{n}/\ell$ (l'ambiguïté du signe est due au fait que, comme il s'agit d'ondes stationnaires, un mode \vec{k} est superposé avec le mode $-\vec{k}$) et où les deux modes de polarisation possibles sont repérés par l'indice A . Nous renvoyons le lecteur à l'appendice de ce chapitre pour tous les détails. L'espace des phases est alors l'espace vectoriel de dimension infinie (ici un espace de Hilbert)

$$E := \{(b^\gamma, p_\gamma)_{\gamma \in \Gamma} / \forall \gamma \in \Gamma, b^\gamma \in \mathbb{R}, p_\gamma \in \mathbb{R}, H(b^\gamma, p_\gamma) < \infty\}.$$

Sur cet espace, nous définissons l'hamiltonien :

$$H(b^\gamma, p_\gamma) := \frac{1}{2} \sum_{\gamma \in \Gamma} \frac{c^2 \pi^2 |\vec{n}|^2}{\ell^2} (p_\gamma)^2 + (b^\gamma)^2$$

dynamique s'appliquaient à la lumière (extrapoler la thermodynamique aux ondes électromagnétiques peut nous sembler naturel une fois que nous avons assimilé l'interprétation statistique des phénomènes thermodynamiques, mais il faut se souvenir qu'à l'époque M. Planck n'avait pas confiance en la théorie de Boltzmann). Ainsi le raisonnement que nous présentons diffère sensiblement de celui de Planck, mais reproduit néanmoins une partie de son cheminement.

et la forme symplectique

$$\Omega := \sum_{\gamma \in \Gamma} dp_\gamma \wedge db^\gamma.$$

Alors les équations d'évolution du champ électromagnétique à l'intérieur de la boîte sont équivalentes aux équations de Hamilton pour une assemblée dénombrable d'oscillateurs harmoniques dont la pulsation est $\omega = c\pi|\vec{n}|/\ell$.

Si nous essayons d'appliquer la théorie de Boltzmann et d'utiliser le formalisme canonique à ce système, nous butons sur une difficulté très sérieuse : nous avons vu au début de ce chapitre (lemme 1) que chaque oscillateur harmonique avait en moyenne une énergie égale à kT . Comme nous sommes en présence d'une infinité dénombrable d'oscillateurs harmoniques, l'énergie moyenne totale devrait être :

$$\sum_{\gamma \in \Gamma} kT,$$

une quantité infinie ! Par chance les mesures expérimentales permettaient de connaître la densité d'énergie par unité de nombre d'onde et donc d'avoir des informations assez précises sur la répartition statistique des niveaux d'énergie selon la longueur d'onde.

Ignorons donc momentanément le problème de l'énergie totale moyenne infinie et explorons la répartition de l'énergie dans une certaine gamme de fréquence. Nous nous intéressons aux « oscillateurs harmoniques » dont le nombre d'onde est compris entre les valeurs ω et $\omega + \delta\omega$. Nous ferons l'hypothèse que tous les oscillateurs harmoniques γ dans la tranche $[\omega, \omega + \delta\omega]$ ont approximativement la même énergie moyenne, égale à

$$\langle \epsilon_\omega \rangle := \langle \omega^2 (p_\gamma)^2 + (b^\gamma)^2 \rangle, \quad \text{où } \gamma = (\vec{k}, A) \text{ et } \omega = c|\vec{k}| = \frac{c\pi|\vec{n}|}{\ell}.$$

Nous devons d'abord dénombrer le nombre d'indices $\gamma = (\vec{n}, A)$ tels que $\omega < c|\vec{k}| < \omega + \delta\omega$, où $\vec{k} = \pm\pi\vec{n}/\ell$, c'est à dire tels que

$$\frac{\ell\omega}{c\pi} < |\vec{n}| < \frac{\ell\omega}{c\pi} + \frac{\ell\delta\omega}{c\pi}.$$

En tenant compte de la polarisation et en négligeant les restrictions lorsque $\vec{n} \in \mathbb{N}_2^3$ (où il n'y a qu'une seule polarisation), on trouve que le nombre de

modes γ dans la tranche $[\omega, \omega + \delta\omega]$ est approximativement

$$N \simeq 2 \cdot \frac{4\pi}{8} |\vec{n}|^2 \delta|\vec{n}| = \pi \frac{\ell^3 \omega^2}{c^3 \pi^3} \delta\omega.$$

Alors la valeur moyenne de l'énergie totale de tous les oscillateurs dans cette tranche est approximativement $\frac{\ell^3 \omega^2}{c^3 \pi^2} \langle \epsilon_\omega \rangle \delta\omega$ (ℓ^3 représente ici le volume du cube).

A présent supposons que le formalisme canonique s'applique : alors le lemme 1 entraîne que $\langle \epsilon_\omega \rangle = kT$. Donc la valeur moyenne de l'énergie totale de tous les oscillateurs dans la tranche $[\omega, \omega + \delta\omega]$ est approximativement $\ell^3 \rho_{RJ}(\omega, T) \delta\omega = \frac{\ell^3 \omega^2}{c^3 \pi^2} kT \delta\omega$. Cette relation conduit à définir la densité de *Rayleigh–Jeans*

$$\rho_{RJ}(\omega, T) := \frac{\omega^2}{c^3 \pi^2} kT.$$

Bien évidemment nous savions à l'avance que cette densité doit conduire à une énergie totale moyenne infinie, ce que confirme le fait que $\int_{[0, \ell]^3} \int_0^\infty \rho_{RJ}(\omega, T) d^3x d\omega = \infty$. Il est néanmoins intéressant de confronter cette prédiction aux observations expérimentales : celles-ci nous enseignent que la densité ρ_{RJ} est très proche de la valeur expérimentale de la densité d'énergie par unité de volume et unité de nombre d'onde $\rho(\omega, T)$ pour les *petites valeurs de ω* (plus précisément si $h\omega \ll kT$). Nous en concluons que le raisonnement suivi a une certaine validité pour ω petit.

Lorsque les valeurs de ω sont grandes, les observations expérimentales ont permis d'établir la loi (empirique) de *Wien*, à savoir que la densité d'énergie coïncide avec une bonne approximation avec :

$$\rho_W(\omega, T) := A\omega^3 e^{-\frac{h\omega}{kT}},$$

où A et h sont des constantes empiriques positives⁴.

4.2.3 Interpolation entre les lois de Rayleigh–Jeans et de Wien

La première démarche de M. Planck consiste à essayer de trouver une loi qui interpole entre les deux lois présentées précédemment. Pour cela il s'appuie

⁴Le fait de désigner par h la constante devant ω/kT est une anticipation de ce qui va suivre

sur un raisonnement thermodynamique classique dans l'esprit de ce que nous avons décrit à la section 1.2. L'idée est d'éliminer la température et d'étudier la relation entre la densité d'énergie et la densité d'entropie définie à la section 1.2. Dans le cas d'application de la loi de Rayleigh–Jeans, nous avons

$$T = \frac{\rho_{RJ} c^3 \pi^2}{k \omega^2},$$

et dans celui de la loi de Wien,

$$T = -\frac{\hbar \omega}{k} \frac{1}{\log\left(\frac{\rho_W}{A \omega^3}\right)}.$$

Partant de $ds = d\rho/T$ (ici la chaleur reçue est égale à la variation de l'énergie), on obtient deux densités d'entropie possibles, selon les lois adoptées :

$$ds_{RJ} = \frac{d\rho_{RJ}}{T} = \frac{k \omega^2}{\rho_{RJ} c^3 \pi^2} d\rho_{RJ} \quad \text{et} \quad ds_W = \frac{d\rho_W}{T} = -\frac{k \log\left(\frac{\rho_W}{A \omega^3}\right)}{\hbar \omega} d\rho_W.$$

On en déduit que

$$\frac{\partial^2 s_{RJ}}{(\partial \rho_{RJ})^2} = -\frac{k \omega^2}{c^3 \pi^2 \rho_{RJ}^2} \quad \text{et} \quad \frac{\partial^2 s_W}{(\partial \rho_W)^2} = -\frac{k}{\hbar \omega \rho_W}.$$

Dans les deux cas M. Planck observe que $\frac{\partial^2 s}{(\partial \rho)^2} < 0$, ce qui est en accord avec la concavité de l'entropie. L'idée est alors de chercher à interpoler entre ces deux lois. Là M. Planck a probablement dû faire plusieurs tentatives avant de trouver la bonne façon d'interpoler : en posant $\frac{\partial^2 s}{(\partial \rho)^2} = \left(\left(\frac{\partial^2 s}{(\partial \rho)^2} \right)_{RJ}^{-1} + \left(\frac{\partial^2 s}{(\partial \rho)^2} \right)_W^{-1} \right)^{-1}$. Cela nous conduit à supposer que :

$$\frac{\partial^2 s}{(\partial \rho)^2} = \frac{k}{\hbar \omega} \left(\frac{1}{\rho + \frac{\hbar \omega^3}{c^3 \pi^2}} - \frac{1}{\rho} \right).$$

Après, nous revenons en arrière, c'est à dire nous intégrons cette relation pour obtenir

$$\frac{1}{T} = \frac{\partial s}{\partial \rho} = \frac{k}{\hbar \omega} \left(\log \left(1 + \frac{\hbar \omega^3}{c^3 \pi^2 \rho} \right) + a \right),$$

où a est une constante d'intégration. On déduit de cette relation :

$$\rho(\omega, T) = \frac{\hbar \omega^3}{c^3 \pi^2 \left(e^{-a} e^{\frac{\hbar \omega}{kT}} - 1 \right)}.$$

Il reste à comparer cette nouvelle loi avec les deux lois que l'on connaît. Sachant que la loi de Rayleigh–Jeans est en parfait accord avec les mesures expérimentales pour les petites valeurs de ω , il est primordial de vérifier si la valeur de ρ obtenue est proche de ρ_{RJ} lorsque $\omega \rightarrow 0$: cela est bien le cas à condition que $a = 0$. Nous en déduisons que ρ devrait être égal à

$$\rho_P(\omega, T) = \frac{\hbar\omega^3}{c^3\pi^2 \left(e^{\frac{\hbar\omega}{kT}} - 1 \right)}. \quad (4.1)$$

C'est la loi de Planck. Ici $h = 2\pi\hbar$ est appelée la *constante de Planck*. Sur le plan théorique le but est atteint : on vérifie immédiatement que $\rho_P \simeq \rho_{RJ}$ pour les petites valeurs de ω et $\rho_P \simeq \rho_W$ pour les grandes valeurs de ω .

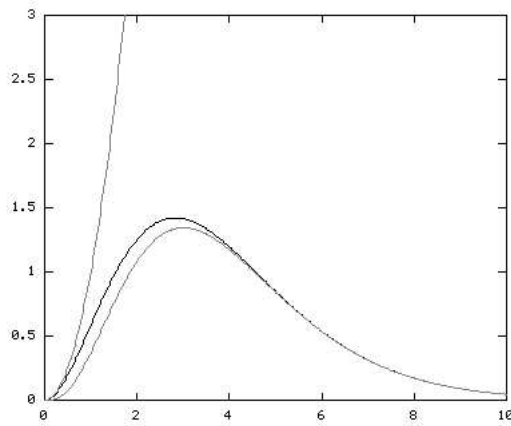


FIG. 4.1 – La loi de Planck (en noir) et les lois de Rayleigh-Jeans et Wien (en gris).

Mais le moment de vérité, c'est quand on compare avec l'expérience, et là ça marche ! M. Planck a trouvé une loi de densité pour le rayonnement qui concorde avec les mesures expérimentales et, de plus, l'énergie totale moyenne est bien bornée, puisque $\int_0^\infty \rho_P(\omega, T) d\omega < \infty$.

4.2.4 Comprendre la loi de Planck

L'histoire n'est pas finie : comme Planck, nous aimerions comprendre l'origine de cette formule. Pour cela essayons de remonter en arrière dans le calcul et

estimer quelle valeur de $\langle \epsilon_\omega \rangle$ conduit à la densité d'énergie de Planck. Pour cela il suffit de résoudre l'équation

$$\frac{\ell^3 \omega^2}{c^3 \pi^2} \langle \epsilon_\omega \rangle \delta\omega = \rho_P(\omega, T) \ell^3 \delta\omega = \frac{\ell^3 \hbar \omega^3}{c^3 \pi^2 \left(e^{\frac{\hbar \omega}{kT}} - 1 \right)} \delta\omega,$$

ce qui donne

$$\langle \epsilon_\omega \rangle = \frac{\hbar \omega}{e^{\frac{\hbar \omega}{kT}} - 1}.$$

Un mathématicien exercé reconnaîtrait dans cette expression que

$$\frac{1}{e^{\frac{\hbar \omega}{kT}} - 1} = \frac{e^{-\frac{\hbar \omega}{kT}}}{1 - e^{-\frac{\hbar \omega}{kT}}} = e^{-\frac{\hbar \omega}{kT}} \sum_{n=0}^{\infty} e^{-\frac{n\hbar \omega}{kT}} = \sum_{n=1}^{\infty} e^{-\frac{n\hbar \omega}{kT}}.$$

Est-ce cela qui a guidé Planck ? Nous n'en savons rien. Et est-ce que Planck s'est souvenu de ce que lui avait dit Boltzmann en 1891 :

« Je en vois aucune raison pour laquelle l'énergie ne pourrait pas être regardée comme divisée atomiquement » ?

En tout cas Planck fait alors l'hypothèse que l'énergie prise par un oscillateur est quantifiée et montre dans ce cas que l'énergie moyenne peut être obtenue par un procédé tout à fait analogue au formalisme canonique de Boltzmann : supposons que l'énergie d'un oscillateur harmonique de nombre d'onde ω ne soit autorisé à prendre **que** les valeurs $n\hbar\omega$, avec $n \in \mathbb{N}^*$ et que la valeur $n\hbar\omega$ soit atteinte avec la probabilité

$$\frac{e^{-\frac{n\hbar \omega}{kT}}}{\sum_{n=1}^{\infty} e^{-\frac{n\hbar \omega}{kT}}}.$$

Alors la valeur moyenne de l'énergie est bien $\frac{\hbar \omega}{e^{\frac{\hbar \omega}{kT}} - 1}$. Pour le vérifier, le plus simple est de poser $\beta = 1/kT$ et

$$Z(\beta) := \sum_{n=1}^{\infty} e^{-\beta n \hbar \omega}.$$

Calculons de deux façons différentes $-\frac{d}{d\beta} \log Z(\beta)$: d'une part

$$-\frac{d}{d\beta} \log Z(\beta) = \frac{\sum_{n=1}^{\infty} n \hbar \omega e^{-\beta n \hbar \omega}}{\sum_{n=1}^{\infty} e^{-\beta n \hbar \omega}}.$$

D'autre part $Z(\beta) = (1 - e^{-\beta\hbar\omega})^{-1}$ implique que

$$-\frac{d}{d\beta} \log Z(\beta) = \frac{\hbar\omega e^{-\beta\hbar\omega}}{1 - e^{-\beta\hbar\omega}} = \frac{\hbar\omega}{e^{\beta\hbar\omega} - 1},$$

qui est précisément la valeur que nous cherchons à comprendre. Ce calcul va amener Planck à remettre en question sa méfiance vis à vis de la théorie de Boltzmann.

Remarque — Il est intéressant de noter que, premièrement, lorsque $\omega \rightarrow 0$, $\frac{\hbar\omega}{e^{\beta\hbar\omega} - 1} \rightarrow \frac{1}{\beta} = kT$ et, deuxièmement, si l'on supposait que, au lieu d'être discrétisée, l'énergie pouvait prendre toutes les valeurs ϵ dans $[0, \infty[$ avec la densité de probabilité $e^{-\beta\epsilon} / \int_0^\infty e^{-\beta\epsilon} d\epsilon$, alors la valeur moyenne de l'énergie serait

$$\beta \int_0^\infty \epsilon e^{-\beta\epsilon} d\epsilon = \frac{1}{\beta} = kT.$$

Il faut souligner le fait qu'à l'époque, Planck est resté pendant longtemps très prudent. Sa première conclusion a été que les *échanges d'énergie entre la lumière et la matière étaient quantifiés*. Le pas supplémentaire a été franchi par Einstein en 1905 : c'est lui qui a supposé le premier que le champ électromagnétique était quantifié et il a considérablement clarifié les calculs de Planck. Ainsi la présentation qui a été faite plus haut est largement inspirée du travail d'Einstein, ne serait-ce que parce que, dès le départ, nous avons appliqué les idées de physique statistique à la lumière et non à la matière.

4.3 L'effet photoélectrique

Lorsqu'Einstein revisite le calcul de Planck en 1905 et qu'il fait l'hypothèse que le champ électromagnétique est quantifié, il montre que cette même hypothèse permet de résoudre une énigme insoluble depuis plusieurs années : l'effet photoélectrique.

Ce phénomène physique a été observé pour la première fois en 1887, par Hertz. L'expérience consiste en deux électrodes à une certaine distance l'une de l'autre, l'une chargée négativement (anode) et l'autre positivement (cathode). La différence de potentiel est non nulle, mais relativement faible, de sorte qu'aucun arc électrique n'apparaît entre les deux électrodes. En revanche, lorsque l'anode est éclairée par de la lumière, on peut observer un

arc électrique. L'effet photoélectrique a été étudié ensuite par Millikan. Le plus étrange est la façon dont le courant dépend de l'intensité lumineuse et de sa longueur d'onde.

Si le nombre d'onde ω de la lumière est inférieur à une certaine valeur ω_0 , il

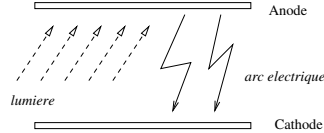


FIG. 4.2 – L'effet photoélectrique

ne se passe rien, même si l'intensité lumineuse est très grande. En revanche, si ω dépasse ω_0 (et donc la longueur d'onde est inférieure à $c/2\pi\omega_0$), un arc électrique est observé. L'intensité du courant électrique augmente avec l'intensité lumineuse.

Einstein comprend que ce mécanisme est profondément relié au calcul de Planck : il faut apporter une certaine énergie E_0 pour « arracher » l'électron à l'anode et qu'ensuite il puisse se laisser entraîner par la force électrique qui s'exerce sur lui entre les deux électrodes. Cette énergie est apportée dans cette expérience par la lumière. Mais la lumière ne peut transmettre son énergie que sous forme de « *quantas* », c'est à dire par paquets de niveau $\hbar\omega$. C'est pour cela que si $\omega < E_0/\hbar$, la lumière n'arrive pas à transmettre de paquet d'énergie suffisante pour arracher un électron à l'anode. En revanche, si $\omega > E_0/\hbar$, des électrons sont arrachés, leur nombre dépend du nombre de quanta de lumière incidents, donc de l'intensité lumineuse. En 1926, Lewis proposera d'appeler « *photons* » ces quanta.

4.4 Loi de déplacement de Wien et loi de Stefan-Boltzmann

Si l'on compare les distributions $\rho_P(\omega, T)$ pour différentes valeurs de T , on constate que la pulsation pour laquelle la densité d'énergie est maximale croît avec T . Ceci s'appelle la loi de déplacement de Wien, qui correspond au fait intuitif qu'un corps plus chaud rayonne principalement des fréquences plus élevées (ne dit-on pas : chauffer au rouge, chauffer à blanc ?).

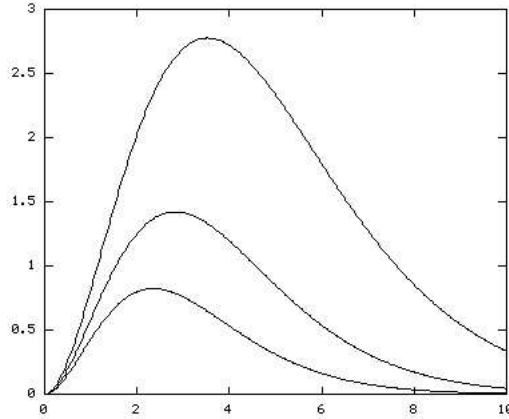


FIG. 4.3 – La distribution de Planck pour trois valeurs différentes de T .

Cette loi se démontre très facilement, en dérivant ρ_P . En posant $x = \frac{h\omega}{kT}$, on trouve que $\frac{d}{d\omega}\rho_P = 0$ équivaut à $e^x(x - 3) + 3 = 0$, ce qui a lieu pour une seule valeur non nulle de x , que nous notons a et qui vaut environ 2,82. On en déduit immédiatement que

$$\omega_{max} = a \frac{kT}{\hbar}.$$

On écrit en général cette loi sous la forme

$$\lambda_{max}T = \text{cste} \simeq 0,29 \text{ cm K}.$$

Par exemple, pour une température de 6000 Kelvin (température de surface du Soleil), la longueur d'onde dominante est d'environ 480 nm, ce qui correspond à la lumière jaune. Cependant, cette application numérique n'a qu'une valeur très limitée, car le Soleil ne peut être assimilé à un corps noir : le rayonnement qui en sort n'est pas nécessairement à l'équilibre thermodynamique.

La loi de Stefan-Boltzmann concerne l'énergie (volumique) totale du rayonnement du corps noir. Nous avons vu que, contrairement à la loi de Rayleigh-Jeans, celle de Planck prévoit une énergie finie. Calculons-la : il suffit d'intégrer ρ_P par rapport à ω . En utilisant à nouveau le changement de variable

$x = \hbar\omega/kT$, on trouve

$$\begin{aligned} \int_0^\infty \rho_P(\omega, T) d\omega &= \frac{k^4}{c^3 \hbar^3 \pi^2} T^4 \int_0^\infty \frac{x^3}{e^x - 1} dx \\ &= \frac{6\zeta(4)k^4}{c^3 \hbar^3 \pi^2} T^4 \\ &= \frac{\pi^2 k^4}{15c^3 \hbar^3} T^4. \end{aligned}$$

La constante $\pi^2 k^4 / 15c^3 \hbar^3$ vaut environ $7,56 \text{ J} \cdot \text{m}^{-3} \cdot \text{K}^{-4}$. On retiendra que l'énergie totale est proportionnelle à la puissance quatrième de la température.

Voici une application un peu fantaisiste de cette loi, attribuée à Richard Feynman. Pour faire refroidir une tasse de thé brûlante, vaut-il mieux ajouter du lait et attendre un moment, ou attendre un moment puis ajouter du lait ? La loi de Stefan-Boltzmann indique qu'il vaut mieux attendre puis ajouter le lait au moment de boire, car le thé, resté plus chaud plus longtemps, aura rayonné plus d'énergie et sera donc finalement moins chaud que si l'on avait ajouté le lait tout de suite. Avis aux amateurs...

4.5 Une approche combinatoire de la loi de Planck

Il est possible de retrouver la forme générale de la loi de Planck par des arguments combinatoires simples. Considérons une population de particules *indiscernables* (nous pensons bien sûr aux photons), en contact avec un thermostat et à l'équilibre thermodynamique. Cela signifie que la valeur de l'énergie totale de ce système ne fluctue pas : elle est fixée.

Appelons Γ l'ensemble des états accessibles pour chacune de ces particules. Par exemple, pour des photons dans une boîte cubique, on peut identifier \mathcal{E} avec l'ensemble discret $\{(\vec{n}, 1)/\vec{n} \in \mathbb{N}_1^3\} \cup \{(\vec{n}, A)/\vec{n} \in \mathbb{N}_2^3, A \in \{1, 2\}\}$, c'est-à-dire l'ensemble des modes et des polarisations compatibles avec les données au bord. Sur Γ est défini un hamiltonien $H : \Gamma \rightarrow \mathbb{R}_+$ et nous choisissons un pas $\delta > 0$ de discrétisation de l'énergie. Il ne s'agit pas du tout d'un pas de quantification de l'énergie, mais juste d'une longueur de découpage : posons $\varepsilon_i = i\delta$ pour $i = 0, 1, \dots$. Appelons \mathcal{T}_i l'ensemble des états $\gamma \in \Gamma$ tels que

$H(\gamma)$ appartient à $[\varepsilon_i, \varepsilon_{i+1}]^5$. Notons m_i le cardinal de \mathcal{T}_i . Nous supposons que δ est choisi de telle façon que tous les m_i soient à la fois grands devant 1 et petits devant le cardinal de Γ .

L'état macroscopique du système est complètement déterminé par la suite (n_0, n_1, \dots) des nombres d'occupations des paquets d'états \mathcal{T}_i : il y a n_0 particules dans des états appartenant au paquet \mathcal{T}_0 , etc... L'état microscopique du système, lui, est la donnée du nombre $n(\gamma)$ de particules dans chacun des états de Γ (rappelons que les particules sont indiscernables). La valeur de l'énergie $\sum_{\gamma \in \Gamma} n(\gamma)H(\gamma)$ du système étant fixée, nous pouvons faire un raisonnement micro-canonique et affirmer que tous les états microscopiques sont équiprobables. Si nous supposons que le phénomène désormais familier de concentration exponentielle de la loi de distribution des quantités macroscopiques a lieu, nous pouvons affirmer que l'état macroscopique du système est presque certainement celui qui est compatible avec le plus grand nombre d'états microscopiques. La première question est donc : combien y a-t-il d'états microscopiques compatibles avec l'état macroscopique (n_0, n_1, \dots) ? Ce nombre N peut s'écrire $N = \prod_i B(n_i \rightarrow m_i)$, où $B(n \rightarrow m)$ est le nombre de façons de placer n boules indiscernables dans m boîtes. Il faut calculer $B(n \rightarrow m)$.

Pour cela, on peut remarquer qu'il est équivalent de mettre n boules dans m boîtes ou d'aligner dans un ordre arbitraire n petits ronds (représentant les boules) et $m - 1$ bâtons (les séparations entre les boîtes). Par exemple, la ligne

$$| \bullet \bullet || \bullet | \bullet \bullet \bullet | \bullet$$

correspond à la disposition suivante de sept boules dans six boîtes : première boîte vide, seconde contient deux boules, troisième boîte vide, la quatrième contient une boule, la cinquième trois, la sixième une. Le nombre de telles lignes de ronds et de bâtons est égal à C_{n+m-1}^n , puisqu'il suffit de choisir la place des ronds parmi les $n + m - 1$ caractères. Nous savons donc maintenant que

$$N = \prod_i \frac{(n_i + m_i - 1)!}{n_i!(m_i - 1)!}.$$

Pour trouver quelle suite (n_0, n_1, \dots) maximise N , nous allons imiter le calcul

⁵dans le cas des photons dans la boîte cubique $H(\gamma) = \hbar\omega$, où $\omega = c|\vec{k}|$ et $\gamma = (\vec{k}, A)$.

différentiel et faire subir aux n_i une petite variation : nous allons remplacer chaque n_i par $n_i + j_i$. Les j_i ne sont cependant pas tout-à-fait arbitraires, car l'énergie totale du système est fixée⁶. La contrainte sur les j_i s'écrit donc $\sum_i j_i \varepsilon_i = 0$. Appelons N' le nombre d'états microscopiques compatibles avec l'état macroscopique modifié $(n_0 + j_0, n_1 + j_1, \dots)$. Pour calculer N'/N , nous étudions le rapport $B(n_i + j_i \rightarrow m_i)/B(n_i \rightarrow m_i)$.

$$\begin{aligned} \frac{B(n_i + j_i \rightarrow m_i)}{B(n_i \rightarrow m_i)} &= \left(\frac{n_i + m_i}{n_i} \right)^{j_i} \frac{\left(1 + \frac{1}{n_i + m_i}\right) \cdots \left(1 + \frac{j_i - 1}{n_i + m_i}\right)}{\left(1 + \frac{1}{n_i}\right) \cdots \left(1 + \frac{j_i}{n_i}\right)} \\ &\simeq \left(\frac{n_i + m_i}{n_i} \right)^{j_i} \exp\left(\frac{1 + \cdots + (j_i - 1)}{n_i + m_i} - \frac{1 + \cdots + j_i}{n_i}\right) \\ &\simeq \left(\frac{n_i + m_i}{n_i} \right)^{j_i} \exp\left(\frac{j_i^2}{2(n_i + m_i)} - \frac{j_i^2}{2n_i}\right) \\ &\simeq \left(\frac{n_i + m_i}{n_i} \right)^{j_i} \exp\left(-\frac{m_i j_i^2}{2n_i(n_i + m_i)}\right). \end{aligned}$$

Ainsi,

$$\frac{N'}{N} = \exp\left(-\sum_i \frac{m_i j_i^2}{2n_i(n_i + m_i)}\right) \prod_i \left(1 + \frac{m_i}{n_i}\right)^{j_i}.$$

Pour que (n_0, n_1, \dots) soit la distribution qui maximise N , il faut que N'/N soit égal à 1 « au premier ordre en les j_i ». Comme le terme exponentiel est maximal pour $j_i = 0$, il faut donc que le terme produit soit stationnaire. Sans contrainte sur les j_i , cela ne pourrait être le cas pour aucune valeur des n_i . Mais avec la contrainte exprimée plus haut, nous pouvons choisir un indice entier a tel que $\varepsilon_a \neq 0$ (en fait, tout indice non nul convient) et exprimer j_a en fonction des autres j_i :

$$j_a = -\frac{1}{\varepsilon_a} \sum_{i \neq a} \varepsilon_i j_i.$$

Nous pouvons maintenant réécrire le produit sous la forme

$$\prod_i \left(1 + \frac{m_i}{n_i}\right)^{j_i} = \left(1 + \frac{m_a}{n_a}\right)^{-\frac{1}{\varepsilon_a} \sum_{i \neq a} \varepsilon_i j_i} \prod_{i \neq a} \left(1 + \frac{m_i}{n_i}\right)^{j_i}$$

⁶Le nombre total de particules en revanche n'est pas fixé : il peut y avoir émission ou absorption de photons.

$$= \prod_{i \neq a} \left[\frac{1 + \frac{m_i}{n_i}}{\left(1 + \frac{m_a}{n_a}\right)^{\frac{\varepsilon_i}{\varepsilon_a}}} \right]^{j_i}.$$

Comme tous les j_i , pour $i \neq a$, peuvent varier sans contrainte, ce dernier produit n'est stationnaire que si chaque terme entre crochets est égal à 1, c'est-à-dire que pour chaque i ,

$$\left(1 + \frac{m_i}{n_i}\right)^{\frac{1}{\varepsilon_i}} = \text{cste} = \left(1 + \frac{m_a}{n_a}\right)^{\frac{1}{\varepsilon_a}}.$$

En appelant K cette constante, on en déduit la forme de n_i :

$$n_i = \frac{m_i}{K^{\varepsilon_i} - 1}.$$

En remplaçant m_i par le nombre de modes dans une tranche donnée de pulsations $[\omega, \omega + \delta\omega[$ et ε_i par l'énergie correspondante $\hbar\omega$, on retrouve une expression qui est presque celle de Planck :

$$\rho_{stat}(\omega) = \frac{\hbar\omega^3}{c^3\pi^2(K^{\hbar\omega} - 1)},$$

où il ne reste « plus qu'à » poser $K = e^{\frac{1}{kT}}$.

4.6 Prologue

Le formalisme canonique avait pour conséquence le fait que la valeur moyenne de l'énergie d'un oscillateur harmonique dans un système ergodique à l'équilibre est égale à kT , et ce, quelles que soient les constantes caractérisant l'oscillateur. L'étude de la lumière a conduit à faire une autre hypothèse, à savoir que la valeur prise par l'énergie de l'oscillateur harmonique est nécessairement $n\hbar\omega$, pour $n \in \mathbb{N}^*$, et que la loi de probabilité de ces niveaux d'énergie est un analogue discret des mesures de probabilités canoniques. Il existe toutefois un principe qui rétablit une autre forme de démocratie entre oscillateurs harmoniques dont les constantes sont différentes, même lorsque l'énergie est quantifiée. Pour cela revenons à la formulation hamiltonienne

d'un oscillateur harmonique : sur l'espace des phases \mathbb{R}^2 , avec les variables (p, q) et la forme symplectique $\Omega := dp \wedge dq$, on considère l'hamiltonien

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

Exercice — 1) Démontrer que les solutions des équations de Hamilton sont de la forme

$$q(t) = A \cos \omega t + B \sin \omega t \quad \text{et} \quad p(t) = m\omega(-A \sin \omega t + B \cos \omega t).$$

Il est immédiat que toutes les solutions des équations de Hamilton sont périodiques, de période $2\pi/\omega$ (et elles décrivent des ellipses dans le plan). Nous nous intéressons à la quantité

$$\mathcal{A} := \int_0^{2\pi/\omega} p(t) dq(t),$$

où $t \mapsto (q(t), p(t))$ est une solution des équations de Hamilton.

2) On suppose que $t \mapsto (q(t), p(t))$ est une solution des équations de Hamilton telle que $H(q(t), p(t)) = n\hbar\omega$, où $n \in \mathbb{N}^*$. Démontrer que $\mathcal{A} = nh$. ■

Le résultat de cette exercice montre que, pour un oscillateur harmonique, la condition de quantification est équivalente à demander que l'aire enfermée à l'intérieur d'une trajectoire hamiltonienne périodique soit un multiple entier de h .

Appendice : les équations de Maxwell

4.6.1 Généralités

Ces équations gouvernent le comportement des champs électrique et magnétique, en présence de charges électriques distribuées dans l'espace, en fonction du temps. Soit \vec{E} et \vec{B} sont respectivement le champ électrique et le champ magnétique. On peut les considérer comme des champs de vecteurs (à valeurs dans \mathbb{R}^3) définis sur l'espace-temps $\mathbb{R} \times \mathbb{R}^3$. Les équations de Maxwell sont :

$$\left\{ \begin{array}{l} \text{rot } \vec{E} + \frac{1}{c} \frac{\partial \vec{B}}{\partial t} = 0 \\ \text{div } \vec{B} = 0 \end{array} \right. , \quad \left\{ \begin{array}{l} \text{rot } \vec{B} - \frac{1}{c} \frac{\partial \vec{E}}{\partial t} = \frac{4\pi}{c} \vec{j} \\ \text{div } \vec{E} = 4\pi\rho, \end{array} \right.$$

où $\rho : \mathbb{R} \times \mathbb{R}^3 \longrightarrow \mathbb{R}$ est la densité de charge électrique par unité de volume, $\vec{j} : \mathbb{R} \times \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ est le vecteur densité de courant électrique et c est la vitesse de la lumière.

Rappelons que l'on peut considérer \vec{E} et \vec{B} comme les composantes d'une paire de 2-formes sur l'espace-temps $\mathbb{R} \times \mathbb{R}^3$, et \vec{j} et ρ comme les composantes d'une 3-forme sur $\mathbb{R} \times \mathbb{R}^3$. Les 2-formes sont :

$$F := (E_1 dx^1 + E_2 dx^2 + E_3 dx^3) \wedge c dt + (B_1 dx^2 \wedge dx^3 + B_2 dx^3 \wedge dx^1 + B_3 dx^1 \wedge dx^2),$$

et

$$\star F := -(B_1 dx^1 + B_2 dx^2 + B_3 dx^3) \wedge c dt + (E_1 dx^2 \wedge dx^3 + E_2 dx^3 \wedge dx^1 + E_3 dx^1 \wedge dx^2).$$

La 3-forme est :

$$J := \rho dx^1 \wedge dx^2 \wedge dx^3 - (j_1 dx^2 \wedge dx^3 + j_2 dx^3 \wedge dx^1 + j_3 dx^1 \wedge dx^2) \wedge dt.$$

Alors les deux sous-systèmes d'équations de Maxwell sont équivalents respectivement aux deux relations $dF = 0$ et $d(\star F) = 4\pi J$. De plus la relation $dF = 0$ entraîne, grâce au lemme de Poincaré, qu'il existe une 1-forme Γ sur $\mathbb{R} \times \mathbb{R}^3$ telle que $d\Gamma = F$. En écrivant

$$\Gamma = V c dt + \sum_{\alpha=1}^3 A_\alpha dx^\alpha,$$

on obtient ainsi que \vec{E} et \vec{B} dérivent du potentiel V et du potentiel-vecteur \vec{A} (de composantes A_α) par les relations

$$\vec{E} = \text{grad} V - \frac{1}{c} \frac{\partial \vec{A}}{\partial t} \quad \text{et} \quad \vec{B} = \text{rot} \vec{A}.$$

4.6.2 La jauge de Coulomb

La forme Γ , que nous appellerons *potentiel de jauge* n'est pas unique, car pour toute fonction $\phi : \mathbb{R} \times \mathbb{R}^3 \longrightarrow \mathbb{R}$, $\Gamma + d\phi$ satisfait également $d(\Gamma + d\phi) = F$ et donc est un autre potentiel de jauge. Cette ambiguïté est source de difficultés, car cela signifie que les composantes de Γ contiennent des degrés de liberté qui ne sont pas physiquement observables. Pour remédier à cela, il est possible d'imposer des conditions supplémentaires sur Γ , que l'on appelle

conditions de jauge. Une possibilité est de supposer que $\frac{1}{c} \frac{\partial V}{\partial t} - \operatorname{div} \vec{A} = 0$: c'est la jauge de Lorentz. Elle peut être imposée dans n'importe quelle situation.

Ici nous sommes intéressés par la lumière dans une cavité, c'est à dire les ondes électromagnétiques, et la jauge la plus adaptée est celle de Coulomb :

$$V = 0 \quad \text{et} \quad \operatorname{div} \vec{A} = 0.$$

Démontrons que, *en l'absence de charges et de courants électriques*, une telle condition peut toujours être imposée. Nous envisageons la situation sur un ouvert $\Omega \subset \mathbb{R}^3$ dont le bord est régulier (ou bien régulier par morceau, comme c'est le cas pour un cube). Nous allons montrer qu'il est possible de construire \vec{A} satisfaisant la condition de jauge de Coulomb et qu'en plus nous pouvons supposer que $\vec{A} \cdot \vec{n} = 0$ sur le bord $\partial\Omega$ de Ω , où \vec{n} est le vecteur normal à $\partial\Omega$.

La première étape consiste à annuler V . Pour cela nous prenons une fonction ϕ sur $\mathbb{R} \times \Omega$ telle que

$$V = \frac{1}{c} \frac{\partial \phi}{\partial t} \quad \Longleftrightarrow \quad d\phi = V c dt + \sum_{\alpha=1}^3 \frac{\partial \phi}{\partial x^\alpha} dx^\alpha.$$

Par exemple $\phi(t, x) = \int_0^t cV(s, x) ds$. Alors $\Gamma' := \Gamma - d\phi$ satisfait à :

$$\Gamma' = \sum_{\alpha=1}^3 A'_\alpha dx^\alpha \quad (\text{c'est à dire } V' = 0) \quad \text{et} \quad d\Gamma' = d\Gamma = F.$$

Utilisons l'hypothèse qu'il n'y a pas de charges électriques. Cela signifie que $\operatorname{div} \vec{E} = 0$. On en déduit que

$$\frac{\partial}{\partial t} \left(\sum_{\alpha=1}^3 \frac{\partial A'_\alpha}{\partial x^\alpha} \right) = \sum_{\alpha=1}^3 \frac{\partial}{\partial x^\alpha} \left(\frac{\partial A'_\alpha}{\partial t} \right) = - \sum_{\alpha=1}^3 \frac{\partial}{\partial x^\alpha} (cE_\alpha) = 0.$$

Donc $\operatorname{div} \vec{A}' := \sum_{\alpha=1}^3 \frac{\partial A'_\alpha}{\partial x^\alpha}$ est indépendant du temps t .

Pour la deuxième étape, nous choisissons $f : \Omega \rightarrow \mathbb{R}$ solution de

$$\begin{cases} \Delta f &= \operatorname{div} \vec{A}' & \text{sur } \Omega \\ df(\vec{n}) &= \Gamma'(\vec{n}) & \text{sur } \partial\Omega. \end{cases}$$

Remarquons que f existe et est unique à une constante additive près. Alors $\Gamma'' := \Gamma' - df$ satisfait :

$$\begin{cases} \operatorname{div} \vec{A}'' &= \operatorname{div} \vec{A}' - \Delta f &= 0 & \text{sur } \Omega \\ \Gamma''(\vec{n}) &= \Gamma'(\vec{n}) - df(\vec{n}) &= 0 & \text{sur } \partial\Omega. \end{cases}$$

Conclusion : Γ'' satisfait la jauge de Coulomb. De plus Γ'' est unique si en plus nous imposons une condition telle que $\int_{\Omega} \Gamma'' = 0$.

4.6.3 Formulation variationnelle des équations de Maxwell dans le vide

Les équations de Maxwell sont les équations d'Euler–Lagrange d'un problème variationnel sur l'espace-temps. Pour tout compact $K \subset \mathbb{R} \times \mathbb{R}^3$, nous considérons l'espace

$$\mathcal{E}_K := \left\{ \vec{E}, \vec{B} \in \mathcal{C}^1(K, \mathbb{R}^3) / \vec{E} + \frac{1}{c} \frac{\partial \vec{B}}{\partial t} = 0, \operatorname{div} \vec{B} = 0 \right\}.$$

(Noter que les contraintes imposées sont équivalentes à la condition $dF = 0$.) Nous définissons sur \mathcal{E}_K la fonctionnelle :

$$\mathcal{A}_K(\vec{E}, \vec{B}) := \frac{1}{2} \int_K (|\vec{E}|^2 - |\vec{B}|^2) dt dx^1 dx^2 dx^3.$$

Proposition 1 *Les champs $(\vec{E}, \vec{B}) \in \mathcal{E}_K$ sont points critiques de \mathcal{A}_K , pour des variations à support compact, si et seulement si*

$$\begin{cases} \operatorname{rot} \vec{B} - \frac{1}{c} \frac{\partial \vec{E}}{\partial t} &= 0 \\ \operatorname{div} \vec{E} &= 0. \end{cases}$$

Preuve — Toute variation infinitésimale de $(\vec{E}, \vec{B}) \in \mathcal{E}_K$ peut s'écrire sous la forme

$$(\vec{E}_\varepsilon, \vec{B}_\varepsilon) := \left(\vec{E} - \varepsilon \frac{1}{c} \frac{\partial \vec{U}}{\partial t} + \varepsilon \operatorname{grad} \varphi, \vec{B} + \varepsilon \operatorname{rot} \vec{U} \right),$$

où $\varphi \in \mathcal{C}_c^2(K, \mathbb{R})$ et $\vec{U} \in \mathcal{C}_c^2(K, \mathbb{R}^3)$. Nous écrivons alors $\mathcal{A}_K(\vec{E}_\varepsilon, \vec{B}_\varepsilon) = \mathcal{A}_K(\vec{E}, \vec{B}) + 2\varepsilon \delta_{(\varphi, \vec{U})} \mathcal{A}_K(\vec{E}, \vec{B}) + o(\varepsilon)$ où :

$$\delta_{(\varphi, \vec{U})} \mathcal{A}_K(\vec{E}, \vec{B}) = \int_K \left(\vec{E} \cdot \left(\vec{\text{grad}}\varphi - \frac{1}{c} \frac{\partial \vec{U}}{\partial t} \right) - \vec{B} \cdot \vec{\text{rot}}\vec{U} \right) dt dx^1 dx^2 dx^3.$$

En utilisant la relation $\vec{B} \cdot \vec{\text{rot}}\vec{U} = \text{div}(\vec{U} \times \vec{B}) + \vec{U} \cdot \vec{\text{rot}}\vec{B}$, nous obtenons

$$\begin{aligned} \delta_{(\varphi, \vec{U})} \mathcal{A}_K(\vec{E}, \vec{B}) &= \int_K \left(\text{div}(\varphi \vec{E} - \vec{U} \times \vec{B}) - \frac{1}{c} \frac{\partial}{\partial t} (\vec{E} \cdot \vec{U}) \right) dt dx^1 dx^2 dx^3 \\ &\quad + \int_K \left(\vec{U} \cdot \left(\frac{1}{c} \frac{\partial \vec{E}}{\partial t} - \vec{\text{rot}}\vec{B} \right) - \varphi \text{div}\vec{E} \right) dt dx^1 dx^2 dx^3. \end{aligned}$$

En vertu du théorème de Stokes, le premier terme dans le membre de droite est égal à une intégrale sur le bord de K . Mais comme φ et \vec{U} sont ici à support compact, cette intégrale s'annule. A présent nous concluons que (\vec{E}, \vec{B}) est point critique de \mathcal{A}_K si et seulement si la dernière intégrale est nulle pour n'importe quel choix de (φ, \vec{U}) , ce qui entraîne le résultat.

4.6.4 La lumière comme oscillation du champ électromagnétique

Toujours en l'absence de charges électriques, les champs électrique et magnétique satisfont l'équation des ondes. Pour cela il suffit par exemple d'appliquer l'opérateur rotationnel à $\vec{\text{rot}}\vec{E} + \frac{1}{c} \frac{\partial \vec{B}}{\partial t} = 0$, ce qui donne

$$\vec{\text{grad}}(\text{div}\vec{E}) - \Delta\vec{E} + \frac{1}{c} \frac{\partial \vec{\text{rot}}\vec{B}}{\partial t} = 0.$$

En utilisant l'équation $\vec{\text{rot}}\vec{B} - \frac{1}{c} \frac{\partial \vec{E}}{\partial t} = 0$ pour éliminer $\vec{\text{rot}}\vec{B}$ et en utilisant aussi $\text{div}\vec{E} = 0$, on en déduit que

$$\frac{1}{c^2} \frac{\partial^2 \vec{E}}{(\partial t)^2} - \Delta\vec{E} = 0..$$

Une manipulation similaire permet d'obtenir l'équation $\frac{1}{c^2} \frac{\partial^2 \vec{B}}{(\partial t)^2} - \Delta\vec{B} = 0$. Cela signifie que les deux champs, en oscillant, propagent des ondes se dé-

plaçant à la vitesse⁷ c . Enfin on peut aussi montrer que si on choisit la jauge de Coulomb, alors le potentiel vecteur \vec{A} satisfait également l'équation $\frac{1}{c^2} \frac{\partial^2 \vec{A}}{(\partial t)^2} - \Delta \vec{A} = 0$.

Des solutions très simples sont les *ondes planes*. Elles sont obtenues en recherchant des solutions particulières des équations de Maxwell dans le vide, pour lesquelles nous supposons que \vec{E} et \vec{B} sont de la forme⁸

$$\vec{E}(t, x) = \mathbf{E} e^{i(\omega t + \vec{k} \cdot \vec{x})} \quad \vec{B}(t, x) = \mathbf{B} e^{i(\omega t + \vec{k} \cdot \vec{x})},$$

où \mathbf{E} et \mathbf{B} sont des vecteurs constants dans \mathbb{C}^3 . Ici ω est une constante réelle (la *pulsation*, égale à $2\pi\nu$, où ν est la fréquence) et \vec{k} est un vecteur constant (le *vecteur d'onde*, dont la direction est celle de la propagation des ondes et la longueur est $|\vec{k}| = 2\pi/\lambda$, où λ est la longueur d'onde). Pour que ces quantités soient solutions de l'équation des ondes, il faut et il suffit que la relation

$$\omega = c|\vec{k}|$$

soit satisfaite. Alors les équations de Maxwell deviennent

$$\left\{ \begin{array}{l} \vec{k} \times \mathbf{E} + \frac{\omega}{c} \mathbf{B} = 0 \\ \vec{k} \cdot \mathbf{B} = 0 \end{array} \right. , \quad \left\{ \begin{array}{l} \vec{k} \times \mathbf{B} - \frac{\omega}{c} \mathbf{E} = 0 \\ \vec{k} \cdot \mathbf{E} = 0. \end{array} \right.$$

En particulier \mathbf{E} et \mathbf{B} sont orthogonaux entre eux et avec \vec{k} . Soit ϵ_1 et ϵ_2 deux vecteurs de \mathbb{R}^3 tels que $(\epsilon_1, \epsilon_2, \vec{k}/|\vec{k}|)$ soit une base orthonormée directe de \mathbb{R}^3 . Alors les solutions du système précédent sont :

$$\mathbf{B} = b(\cos \theta \epsilon_1 + \sin \theta \epsilon_2) \quad \mathbf{E} = b(-\sin \theta \epsilon_1 + \cos \theta \epsilon_2)$$

Ici b est un nombre complexe (dont le module représente l'intensité et l'argument une phase) tandis que $\theta \in \mathbb{R}/2\pi\mathbb{Z}$ paramétrise un degré de liberté supplémentaire qui est la *polarisation* de la lumière. Notons que, dans la jauge de Coulomb, le potentiel vecteur s'écrit $\vec{A} = \mathbf{A} e^{i(\omega t + \vec{k} \cdot \vec{x})}$, où $\mathbf{A} \in \mathbb{C}^3$ satisfait les conditions $\vec{k} \cdot \mathbf{A} = 0$ et $\mathbf{B} = i\vec{k} \times \mathbf{A}$. Cela entraîne ne que $\mathbf{A} = i\vec{k}/|\vec{k}|^2 \times \mathbf{B} = ib \frac{c^2}{\omega^2} (-\sin \theta \epsilon_1 + \cos \theta \epsilon_2)$.

⁷si on avait écrit les équations de Maxwell en faisant intervenir les constantes ϵ_0 (permittivité électrique dans le vide) et μ_0 (perméabilité magnétique dans le vide), on aurait obtenu que c est reliée à ces constantes par la relation $\epsilon_0 \mu_0 c^2 = 1$

⁸en fait \vec{E} et \vec{B} sont les parties réelles de ces vecteurs complexes

4.6.5 Les ondes électromagnétiques à l'intérieur d'une cavité cubique

Nous imaginons ici une boîte vide dont les parois sont parfaitement conductrice. Nous supposons qu'elle a une forme de cube et nous la modélisons par l'ouvert $Q :=]0, \ell[^3$, où ℓ est la longueur des arêtes du cube. Les parois sont à une certaine température T et nous supposons que le champ électromagnétique à l'intérieur de la boîte a interagit suffisamment longtemps avec les particules chargées qui constituent la paroi, pour qu'il soit à l'équilibre thermodynamique avec la paroi. Nous disons alors que le champ électromagnétique à l'intérieur est à la température T . Les conditions satisfaites par le champ électromagnétique sur le bord du bord sont que \vec{E} est parallèle au vecteur normal à la surface $\vec{\nu}$ et \vec{B} est orthogonal à $\vec{\nu}$.

Préliminaire : décomposition de Fourier sur un intervalle

Nous considérons ici des fonctions $f \in \mathcal{C}^0([0, \ell]) \cap L^2([0, \ell])$ et nous rappelons comment il est possible de les décomposer en séries de Fourier. Nous distinguons deux cas.

Dans le cas général, on peut étendre toute fonction $f \in \mathcal{C}^0([0, \ell]) \cap L^2([0, \ell])$ en une fonction 2ℓ -périodique et *paire* sur \mathbb{R} , c'est à dire :

$$\forall x \in \mathbb{R}, \quad f(-x) = f(x) \quad \text{et} \quad f(x + 2\ell) = f(x),$$

tout en satisfaisant la condition $f \in \mathcal{C}^0(\mathbb{R}) \cap L^2_{loc}(\mathbb{R})$. Alors nous pouvons écrire

$$f(x) = \frac{a_0}{\sqrt{\ell}} + \sum_{n=1}^{\infty} a_n \sqrt{\frac{2}{\ell}} \cos\left(\frac{n\pi x}{\ell}\right).$$

Dans le cas particulier où $f(0) = f(\ell) = 0$, alors il est possible d'étendre f en une fonction dans $\mathcal{C}^0(\mathbb{R}) \cap L^2_{loc}(\mathbb{R})$, 2ℓ -périodique et *impaire*, c'est à dire :

$$\forall x \in \mathbb{R}, \quad f(-x) = -f(x) \quad \text{et} \quad f(x + 2\ell) = f(x).$$

Alors nous avons une autre décomposition de Fourier :

$$f(x) = \sum_{n=1}^{\infty} b_n \sqrt{\frac{2}{\ell}} \sin\left(\frac{n\pi x}{\ell}\right).$$

Dans chacun des cas, la décomposition se fait suivant une base hilbertienne orthonormée, qui est soit $(c_n)_{n \in \mathbb{N}}$, soit $(s_n)_{n \in \mathbb{N}^*}$, où les c_n et s_n sont des fonctions définies sur $[0, \ell]$ par

$$\forall x \in [0, \ell], \quad c_0(x) := \frac{1}{\sqrt{\ell}}, \quad c_n(x) := \sqrt{\frac{2}{\ell}} \cos\left(\frac{n\pi x}{\ell}\right), \quad s_n(x) := \sqrt{\frac{2}{\ell}} \sin\left(\frac{n\pi x}{\ell}\right).$$

En particulier $\|\sum_{n=0}^{\infty} a_n c_n\|_{L^2}^2 = \sum_{n=0}^{\infty} |a_n|^2$ pour les fonctions paires et $\|\sum_{n=1}^{\infty} b_n s_n\|_{L^2}^2 = \sum_{n=1}^{\infty} |b_n|^2$ pour les fonctions impaires.

Décomposition de Fourier dans un cube

Revenons au champ électromagnétique dans le cube $Q :=]0, \ell[^3$. Nous supposons ici qu'à tout instant t , les champs électrique et magnétique sont continus sur \overline{Q} et sont dans $L^2(Q, \mathbb{R}^3)$, c'est à dire $\int_Q |\vec{E}(t, \vec{x})|^2 dx^1 dx^2 dx^3$ et $\int_Q |\vec{B}(t, \vec{x})|^2 dx^1 dx^2 dx^3$ sont bornés $\forall t$. Cette dernière hypothèse est très naturelle en raison du fait que l'énergie électromagnétique à l'instant t est la quantité

$$\frac{1}{2} \int_Q \left(|\vec{E}(t, \vec{x})|^2 + |\vec{B}(t, \vec{x})|^2 \right) dx^1 dx^2 dx^3. \quad (4.2)$$

Rappelons que, pour toutes fonctions $f_1, f_2, f_3 \in L^2([0, \ell])$, on note $f_1 \otimes f_2 \otimes f_3$ le produit tensoriel de ces trois fonctions, c'est à dire la fonction définie sur $[0, \ell]^3$ par $f_1 \otimes f_2 \otimes f_3(x^1, x^2, x^3) := f_1(x^1) f_2(x^2) f_3(x^3)$. On note $L^2([0, \ell]) \otimes L^2([0, \ell]) \otimes L^2([0, \ell])$ l'espace vectoriel des combinaisons linéaires finies de fonctions de la forme $f_1 \otimes f_2 \otimes f_3$, où chaque fonction f_α est dans $L^2([0, \ell])$. Alors $L^2([0, \ell]) \otimes L^2([0, \ell]) \otimes L^2([0, \ell])$ est dense dans $L^2(Q)$. Une conséquence de cela est que, puisque $(c_n)_{n \in \mathbb{N}}$ est une base hermitienne orthogonale de $L^2([0, \ell])$, alors $(c_{n_1} \otimes c_{n_2} \otimes c_{n_3})_{(n_1, n_2, n_3) \in \mathbb{N}^3}$ est une base hermitienne orthogonale de $L^2(Q)$. Et si par exemple on s'intéresse aux fonctions $g \in L^2(Q) \cap \mathcal{C}^0(\overline{Q})$ qui s'annulent lorsque $x^1 = 0$ et $x^1 = \ell$, alors il peut être plus judicieux de décomposer g dans la base $(s_{n_1} \otimes c_{n_2} \otimes c_{n_3})_{n_1 \in \mathbb{N}^*, (n_2, n_3) \in \mathbb{N}^2}$.

Nous appliquons cela aux composantes des champs électrique et magnétique :

$$\vec{E}(t, \vec{x}) = \sum_{\vec{n} \in \mathbb{N}^3} \vec{E}^{\vec{n}}(t, \vec{x}), \quad \text{où} \quad \vec{E}^{\vec{n}}(t, \vec{x}) := \begin{pmatrix} E_1^{\vec{n}}(t) c_{n_1}(x^1) s_{n_2}(x^2) s_{n_3}(x^3) \\ E_2^{\vec{n}}(t) s_{n_1}(x^1) c_{n_2}(x^2) s_{n_3}(x^3) \\ E_3^{\vec{n}}(t) s_{n_1}(x^1) s_{n_2}(x^2) c_{n_3}(x^3) \end{pmatrix},$$

et

$$\vec{B}(t, \vec{x}) = \sum_{\vec{n} \in \mathbb{N}^3} \vec{B}^{\vec{n}}(t, \vec{x}), \quad \text{où} \quad \vec{B}^{\vec{n}}(t, \vec{x}) := \begin{pmatrix} B_1^{\vec{n}}(t) s_{n_1}(x^1) c_{n_2}(x^2) c_{n_3}(x^3) \\ B_2^{\vec{n}}(t) c_{n_1}(x^1) s_{n_2}(x^2) c_{n_3}(x^3) \\ B_3^{\vec{n}}(t) c_{n_1}(x^1) c_{n_2}(x^2) s_{n_3}(x^3) \end{pmatrix}.$$

Ici on convient que $s_0 = 0$ et que tous les coefficients $E_\alpha^{\vec{n}}(t)$ ou $B_\alpha^{\vec{n}}(t)$ sont nuls dès qu'ils apparaissent en facteur de s_0 . C'est à dire : $\forall \alpha, \beta \in \llbracket 1, 3 \rrbracket$, si $\alpha \neq \beta$ et $n_\beta = 0$, alors $E_\alpha^{\vec{n}}(t) = 0$ et $\forall \alpha \in \llbracket 1, 3 \rrbracket$, si $n_\alpha = 0$, alors $B_\alpha^{\vec{n}}(t) = 0$. L'intérêt d'une telle décomposition est que chaque terme $\vec{E}^{\vec{n}}(t, \vec{x})$ ou $\vec{B}^{\vec{n}}(t, \vec{x})$ dans la décomposition de $\vec{E}(t, \vec{x})$ ou $\vec{B}(t, \vec{x})$ satisfait les bonnes conditions sur le bord de Q , à savoir $\vec{E}^{\vec{n}}(t, \vec{x}) // \vec{\nu}$ et $\vec{B}^{\vec{n}}(t, \vec{x}) \perp \vec{\nu}$ sur chaque face de ∂Q .

Dynamique du champ électromagnétique dans un cube

Nous considérons l'espace

$$\mathcal{E}_{\mathbb{R} \times Q}^* := \left\{ (\vec{E}, \vec{B}) \in \mathcal{C}^1(\mathbb{R} \times Q, (\mathbb{R}^3)^2) \cap \mathcal{C}^1(\mathbb{R}, \mathcal{C}^0(\overline{Q}, (\mathbb{R}^3)^2)) / \right. \\ \left. \text{rot} \vec{E} + \frac{1}{c} \frac{\partial \vec{B}}{\partial t} = 0 \text{ et } \text{div} \vec{B} = 0, \vec{E} // \vec{\nu}, \vec{B} \perp \vec{\nu} \text{ sur } \mathbb{R} \times \partial Q \right\}$$

et, pour tout compact $K \subset \mathbb{R}$, la fonctionnelle sur $\mathcal{E}_{\mathbb{R} \times Q}^*$:

$$\mathcal{A}_{K \times Q}(\vec{E}, \vec{B}) := \frac{1}{2} \int_{K \times Q} (|\vec{E}|^2 - |\vec{B}|^2) dt dx^1 dx^2 dx^3.$$

Nous nous intéressons donc aux champs dans $\mathcal{E}_{\mathbb{R} \times Q}^*$ qui sont points critiques de $\mathcal{A}_{K \times Q}$ pour tout compact $K \subset \mathbb{R}$. Les développements en séries de Fourier nous permettent de paramétriser les champs (\vec{E}, \vec{B}) dans $\mathcal{E}_{\mathbb{R} \times Q}^*$ par une famille dénombrable de fonctions $E_\alpha^{\vec{n}}$ et $B_\alpha^{\vec{n}}$ dépendant du temps $t \in \mathbb{R}$. Nous devons d'abord trouver les conditions sur ces fonctions pour que les contraintes $\text{rot} \vec{E} + \frac{1}{c} \frac{\partial \vec{B}}{\partial t} = 0$ et $\text{div} \vec{B} = 0$ soient satisfaites. Pour cela nous calculons que

$$\text{rot} \vec{E}(t, \vec{x}) + \frac{1}{c} \frac{\partial \vec{B}}{\partial t}(t, \vec{x}) = \frac{\pi}{\ell} \sum_{\vec{n} \in \mathbb{N}^3} \begin{pmatrix} (n_2 E_3^{\vec{n}}(t) - n_3 E_2^{\vec{n}}(t) + \frac{\ell}{c\pi} \frac{d}{dt} B_1^{\vec{n}}(t)) s_{n_1}(x^1) c_{n_2}(x^2) c_{n_3}(x^3) \\ (n_3 E_1^{\vec{n}}(t) - n_1 E_3^{\vec{n}}(t) + \frac{\ell}{c\pi} \frac{d}{dt} B_2^{\vec{n}}(t)) c_{n_1}(x^1) s_{n_2}(x^2) c_{n_3}(x^3) \\ (n_1 E_2^{\vec{n}}(t) - n_2 E_1^{\vec{n}}(t) + \frac{\ell}{c\pi} \frac{d}{dt} B_3^{\vec{n}}(t)) c_{n_1}(x^1) c_{n_2}(x^2) s_{n_3}(x^3) \end{pmatrix}$$

et

$$\text{div} \vec{B}(t, \vec{x}) = \frac{\pi}{\ell} \sum_{\vec{n} \in \mathbb{N}^3} (n_1 B_1^{\vec{n}}(t) + n_2 B_2^{\vec{n}}(t) + n_3 B_3^{\vec{n}}(t)) c_{n_1}(x^1) c_{n_2}(x^2) c_{n_3}(x^3).$$

Nous en déduisons que

$$\vec{E}^{\vec{n}}(t) := \begin{pmatrix} E_1^{\vec{n}}(t) \\ E_2^{\vec{n}}(t) \\ E_3^{\vec{n}}(t) \end{pmatrix} \quad \text{et} \quad \vec{B}^{\vec{n}}(t) := \begin{pmatrix} B_1^{\vec{n}}(t) \\ B_2^{\vec{n}}(t) \\ B_3^{\vec{n}}(t) \end{pmatrix}.$$

doivent satisfaire les conditions :

$$\forall \vec{n}, \quad \vec{n} \times \vec{E}^{\vec{n}}(t) + \frac{\ell}{c\pi} \frac{d\vec{B}^{\vec{n}}(t)}{dt} = 0 \quad \text{et} \quad \vec{n} \cdot \vec{B}^{\vec{n}}(t) = 0. \quad (4.3)$$

Pour tout $\vec{n} \in \mathbb{N}^3$, soit $(\epsilon_1, \epsilon_2, \epsilon_3)$ une base orthonormée de \mathbb{R}^3 telle que $\epsilon_3 = \vec{n}/|\vec{n}|$. Les solutions de (4.3) sont alors

$$\vec{B}^{\vec{n}}(t) = b^{\vec{n},1}(t)\epsilon_1 + b^{\vec{n},2}(t)\epsilon_2, \quad \vec{E}^{\vec{n}}(t) = \frac{\ell}{c\pi|\vec{n}|} \left(-\frac{db^{\vec{n},2}(t)}{dt}\epsilon_1 + \frac{db^{\vec{n},1}(t)}{dt}\epsilon_2 \right) + e^{\vec{n},3}(t)\epsilon_3,$$

où $b^{\vec{n},1}$, $b^{\vec{n},2}$ et $e^{\vec{n},3}$ sont des fonctions \mathcal{C}^1 arbitraires. Les deux degrés de liberté correspondant à $b^{\vec{n},1}$ et $b^{\vec{n},2}$ correspondent à la polarisation de la lumière.

Revenons aux conditions aux bords. Elles avaient pour conséquence sur les coefficients $B_\alpha^{\vec{n}}$ que, si $\alpha \in \llbracket 1, 3 \rrbracket$ et $\vec{n} \in \mathbb{N}^3$ sont tels que $n_\alpha = 0$ alors $B_\alpha^{\vec{n}}(t) = 0$. Cela peut être réexprimé de la façon suivante : soit $(\vec{\nu}_1, \vec{\nu}_2, \vec{\nu}_3)$ une base orthonormée de \mathbb{R}^3 associée au cube Q , c'est à dire constituée par des vecteurs orthogonaux aux faces du cube, alors $\forall \alpha \in \llbracket 1, 3 \rrbracket$, $\vec{\nu}_\alpha \cdot \vec{n} = 0$ entraîne ne que $\vec{\nu}_\alpha \cdot \vec{B}^{\vec{n}} = 0$. Donc, en plus d'être orthogonal à \vec{n} , $\vec{B}^{\vec{n}}$ doit parfois être orthogonal à d'autres direction, selon les valeurs de \vec{n} . Cela a les conséquences suivantes.

Soit $\mathbb{N}_2^3 := (\mathbb{N}^*)^3$, $\mathbb{N}_1^3 := \{\vec{n} \in \mathbb{N}^3 / \exists! \alpha \in \llbracket 1, 3 \rrbracket, n_\alpha = 0\}$ et $\mathbb{N}_0^3 := \{\vec{n} \in \mathbb{N}^3 / \exists \alpha, \beta \in \llbracket 1, 3 \rrbracket \text{ avec } \alpha \neq \beta \text{ et } n_\alpha = n_\beta = 0\}$, de sorte que \mathbb{N}^3 est la réunion disjointe de \mathbb{N}_0^3 , \mathbb{N}_1^3 et \mathbb{N}_2^3 . Alors si $\vec{n} \in \mathbb{N}_0^3$, \vec{n} est orthogonal à au moins deux vecteurs $\vec{\nu}_\alpha$ et $\vec{\nu}_\beta$, avec $\alpha \neq \beta$; on peut alors choisir sans perte de généralité $\epsilon_1 = \vec{\nu}_\alpha$ et $\epsilon_2 = \vec{\nu}_\beta$, mais alors comme $\vec{\nu}_\alpha \cdot \vec{B}^{\vec{n}} = \vec{\nu}_\beta \cdot \vec{B}^{\vec{n}} = 0$, on a nécessairement $b^{\vec{n},1} = b^{\vec{n},2} = 0$. Si $\vec{n} \in \mathbb{N}_1^3$, \vec{n} est orthogonal à un unique vecteur $\vec{\nu}_\alpha$; on peut alors choisir sans perte de généralité $\epsilon_2 = \vec{\nu}_\alpha$ et alors comme $\vec{\nu}_\alpha \cdot \vec{B}^{\vec{n}} = 0$, on a nécessairement $b^{\vec{n},2} = 0$. Enfin si $\vec{n} \in \mathbb{N}_2^3$, il n'y a aucune contrainte supplémentaire sur $b^{\vec{n},1}$ et $b^{\vec{n},2}$. Conclusion⁹ : pour $\vec{n} \in \mathbb{N}_0^3$,

⁹de même on obtient que $\forall \vec{n} \in \mathbb{N}_0^3 \cup \mathbb{N}_1^3$, $e^{\vec{n},3} = 0$.

il n'y a pas de mode, pour $\vec{n} \in \mathbb{N}_1^3$, il n'y a qu'une seule polarisation possible et pour $\vec{n} \in \mathbb{N}_2^3$ il y a deux polarisations.

Nous pouvons maintenant écrire l'action $\mathcal{A}_{K \times Q}(\vec{E}, \vec{B})$ en fonction des variables $b^{\vec{n},1}$, $b^{\vec{n},2}$ et $e^{\vec{n},3}$: nous notons $\Gamma := \{(\vec{n}, 1)/\vec{n} \in \mathbb{N}_1^3\} \cup \{(\vec{n}, A)/\vec{n} \in \mathbb{N}_2^3, A \in \{1, 2\}\}$ l'ensemble des indices γ pour lesquels b^γ est non nul. Alors

$$\mathcal{A}_{K \times Q}(\vec{E}, \vec{B}) = \frac{1}{2} \int_K \left(\sum_{\gamma \in \Gamma} \frac{\ell^2}{c^2 \pi^2 |\vec{n}|^2} \left(\frac{db^\gamma(t)}{dt} \right)^2 - (b^\gamma(t))^2 + \sum_{\vec{n} \in \mathbb{N}_2^3} (e^{\vec{n},3}(t))^2 \right) dt.$$

Les équations d'Euler-Lagrange de cette action sont :

$$\forall \gamma \in \Gamma, \quad \frac{d^2 b^\gamma(t)}{(dt)^2} + \frac{c^2 \pi^2 |\vec{n}|^2}{\ell^2} b^\gamma(t) = 0,$$

$$\forall \vec{n} \in \mathbb{N}_2^3, \quad e^{\vec{n},3}(t) = 0.$$

La dernière équation exprime le fait que $\text{div} \vec{E} = 0$ (loi de Gauss) et nous retrouvons ainsi une des équations de Maxwell en l'absence de charge. Les autres équations sur les variables b^γ montrent que notre problème a une formulation semblable à celle d'une assemblée dénombrable d'oscillateurs harmoniques.

Formulation hamiltonienne

Il est maintenant aisé de formuler les équations de Maxwell dans notre cube sous une forme hamiltonienne. Nous laissons de côté les variables $e^{\vec{n},3}$, puisqu'elles sont nulles¹⁰. La transformation de Legendre est très simple ici puisqu'il suffit de changer le signe de l'« énergie potentielle » et de faire la substitution $p_\gamma = \frac{\ell^2}{c^2 \pi^2 |\vec{n}|^2} \frac{d}{dt} b^\gamma$. Nous obtenons ainsi l'hamiltonien :

$$H(b^\gamma, p_\gamma) := \frac{1}{2} \sum_{\gamma \in \Gamma} \frac{c^2 \pi^2 |\vec{n}|^2}{\ell^2} (p_\gamma)^2 + (b^\gamma)^2.$$

¹⁰le fait que l'on retrouve ces équations où la dérivée par rapport au temps n'intervient pas (i.e. $e^{\vec{n},3}(t) = 0$) parmi les équations de la dynamique est caractéristique des théories de jauge et est relié au fait qu'il existe une infinité de potentiel de jauge pour le champ électromagnétique. Ces relations sont un obstacle à la formulation hamiltonienne (car la transformation de jauge est alors dégénérée) et c'est pourquoi elles font l'objet d'un traitement spécial : nous annulons ici simplement toutes les quantités $e^{\vec{n},3}$ et nous ignorons donc ces composantes dans la suite.

Il est remarquable¹¹ que cette quantité coïncide avec l'énergie totale définie en (4.2).

Références :

J. Harthong, *Probabilités et statistiques*, Diderot Multimédia 1996 ou :

<http://moire4.u-strasbg.fr/bouquins/proba/tabmat1.htm>

J. Kouneiher, Cours de Mécanique quantique, Université de Marne-la-Vallée.

¹¹en fait ça n'est pas un hasard et il est possible d'arriver au même résultat directement sans passer par des séries de Fourier

Chapitre 5

Vers la quantification des états d'énergie de la matière

Le problème de l'équilibre thermodynamique de la lumière et de la matière, résolu par Planck en 1900 et Einstein en 1905, n'était qu'une partie des difficultés théoriques rencontrées pour concilier la mécanique de Galilée et Newton et l'électromagnétisme de Maxwell. Deux autres problèmes se posaient :

- la mécanique newtonienne et l'électromagnétisme se comportent très différemment par changement de référentiel. Suite aux travaux de G.F. Fitzgerald et H.A. Lorentz et aux expériences de A.A. Michelson et E.W. Morley, la réponse à ces difficultés fut apportée par la théorie de la relativité restreinte, construite par A. Einstein et H. Poincaré en 1905.
- les raies spectrales : lorsqu'ils sont excités, des atomes d'hydrogène n'émettent de la lumière que dans certaines fréquences, qui forment un sous-ensemble discret de toutes les fréquences envisageables, le spectre de l'atome d'hydrogène. Ce phénomène avait fait l'objet de nombreuses expériences assez précises. Toutefois on ne connaissait aucune explication théorique à ces raies spectrales.

Nous allons voir plus en détail ces deux énigmes et examiner comment elles ont été résolues. Au passage nous découvrirons comment les réponses apportées ont contribué à l'édifice de la mécanique quantique.

5.1 Quelques éléments de relativité

L'histoire remonte en fait à Galilée (1564–1642) qui observe et formule son *principe d'inertie* :

Le mouvement d'un corps massif qui n'est soumis à aucune force est un mouvement de translation rectiligne uniforme.

Tel qu'il est énoncé, ce principe est incomplet, car nous n'avons pas précisé dans quelles conditions et plus précisément dans quel repère il s'applique. Il est clair que l'on peut le vérifier avec une bonne approximation sur la terre ferme (à condition d'avoir pu se « débarasser » de la force de gravité et des forces de frottement), par exemple en faisant rouler des billes sur une surface parfaitement plane et horizontale. Mais si la surface et l'observateur sont embarqués sur un bateau soumis à divers mouvements, c'est bien évidemment faux. Galilée prend ainsi conscience de la nécessité de préciser les conditions d'application du principe d'inertie. En termes mathématiques (avec les acquis des coordonnées de Descartes et beaucoup plus tard de la géométrie différentielle) nous pouvons modéliser ces conditions par un repère, c'est à dire un système de coordonnées sur l'espace-temps. Notons E l'espace-temps, ou l'ensemble des « événements », c'est à dire de la donnée d'une position dans l'espace et d'un instant au cours du temps. Un système de coordonnées est une bijection

$$\begin{aligned}\varphi : E &\longrightarrow \mathbb{R} \times \mathbb{R}^3 \\ e &\longmapsto (t, x^1, x^2, x^3).\end{aligned}$$

Jusqu'à la fin du XIX^{ème} siècle, comme on supposait implicitement que E a une structure d'espace affine, on n'envisageait que des systèmes de coordonnées affines, donc complètement caractérisés par la donnée d'une origine 0 (un point dans l'espace à une certaine date), d'une échelle de temps (donnée par une horloge qui se trouve en 0 à l'instant 0) et de trois axes de coordonnées, c'est à dire un *repère d'espace-temps*.

Galilée avait donc supposé, pour compléter son principe d'inertie, qu'il existait des repères que l'on dit être *inertiels* dans lesquels son principe d'inertie s'appliquait. Noter qu'à partir du moment où l'on sait qu'il existe au moins un système de coordonnées inertielles, il est facile d'en trouver une infinité : soit $\varphi' : E \longrightarrow \mathbb{R} \times \mathbb{R}^3$ un autre système de coordonnées affines et supposons que $\varphi' = T \circ \varphi$, où $T : \mathbb{R} \times \mathbb{R}^3 \longrightarrow \mathbb{R} \times \mathbb{R}^3$ est une bijection affine. Alors si

le repère R associé à φ est inertiel, le repère R' associé à φ' est aussi inertiel et vice-versa. A l'époque il était naturel de se restreindre à des bijections affines T qui ne modifient pas l'échelle de temps (sauf à changer l'origine du calendrier) et qui respectent les longueurs (sinon cela aurait pour effet de changer la définition de la masse d'un référentiel à un autre). L'ensemble des transformations T qui transforment un système de coordonnées inertielles en un autre (et qui préservent l'orientation) s'identifie au groupe :

$$\mathcal{G} := \left\{ T : \begin{pmatrix} t \\ \vec{x} \end{pmatrix} \mapsto \begin{pmatrix} 1 & 0 \\ -\vec{v} & R \end{pmatrix} \begin{pmatrix} t \\ \vec{x} \end{pmatrix} + \begin{pmatrix} t_0 \\ \vec{x}_0 \end{pmatrix} / t_0 \in \mathbb{R}, \vec{x}_0, \vec{v} \in \mathbb{R}^3, R \in SO(3) \right\},$$

que l'on appelle *groupe de Galilée*.

Une fois cette observation faite se pose la question suivante, de nature plus philosophique : dans chaque repère inertiel les lois de la physique sont les mêmes, donc apparemment aucun ne semble être privilégié par les lois de la physique. Est-ce que cela signifie qu'il n'existe aucun repère inertiel privilégié ou absolu, comme celui (approximativement inertiel) associé à la terre (voire au soleil) ? Allant contre le sens commun et les intérêts de l'Eglise, Galilée suppose que non : il n'y a pas de repère privilégié. C'est le principe de *relativité galiléenne*.

Ce principe fut admis et il était en particulier compatible avec les lois de la mécanique newtonienne. Rien ne semblât le remettre en question pendant longtemps. Mais au XIXème siècle les progrès accomplis dans la compréhension de la lumière soulevèrent de nouvelles difficultés. En effet les expériences de Fresnel et de Young montraient de façon irréfutable que la lumière est un phénomène ondulatoire, c'est à dire analogue à la propagation des ondes à la surface d'un étang ou aux ondes sonores. Dans les deux comparaisons, le phénomène ondulatoire a lieu au sein d'un milieu matériel (qui est l'eau pour les ondes à la surface de l'étang et l'air pour les ondes acoustiques) et la propagation d'une onde est la transmission d'une énergie (potentielle et cinétique) à travers ce milieu. L'idée la plus naturelle qui se présenta fut d'imaginer que la lumière avait une nature similaire. Cela conduisit à l'hypothèse qu'il existait un milieu matériel fluide emplissant l'univers, que l'on appela *éther*, dont les vibrations pouvaient s'interpréter comme les ondes lumineuses. Bien évidemment on ne disposait à l'époque d'aucunes données expérimentales sur cet éther. La propagation de la lumière pouvait alors être modélisée par

l'équation des ondes :

$$\frac{1}{c^2} \frac{\partial^2 f}{(\partial t)^2} - \Delta f = 0,$$

dans laquelle on pourrait concevoir la fonction f comme une sorte de pression locale de l'éther.

La même époque voit l'établissement des lois de l'électromagnétisme par M. Faraday et J.C. Maxwell et l'explication de la lumière comme un phénomène électromagnétique. Ainsi les propriétés de l'éther devaient aussi expliquer l'électromagnétisme. Ce n'est que plus tard que le concept, relativement plus abstrait, de champ électromagnétique fut introduit et remplaça la notion d'éther.

En fait la théorie de l'éther était en contradiction avec la relativité galiléenne. En effet l'équation des ondes n'est pas invariante par le groupe de Galilée, comme l'illustre l'exercice qui suit.

Exercice 9 On considère la transformation $T \in \mathcal{G}$ définie par

$$T : \begin{pmatrix} t \\ x^1 \\ x^2 \\ x^3 \end{pmatrix} \mapsto \begin{pmatrix} 1 & 0 & 0 & 0 \\ -v & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} t \\ x^1 \\ x^2 \\ x^3 \end{pmatrix}.$$

Soit $F : E \rightarrow \mathbb{R}$ une composante du champ électromagnétique (dont l'évolution décrit une onde lumineuse) et R et \tilde{R} deux référentiels galiléens. On note φ (respectivement $\tilde{\varphi}$) la carte locale et (t, x^1, x^2, x^3) (respectivement $(\tilde{t}, \tilde{x}^1, \tilde{x}^2, \tilde{x}^3)$) le système de coordonnées associé à R (respectivement \tilde{R}) et f (respectivement \tilde{f}) la représentation de F dans R (respectivement \tilde{R}), de sorte que

$$F = f \circ \varphi = \tilde{f} \circ \tilde{\varphi}.$$

Démontrer que, si $\tilde{\varphi} = T \circ \varphi$,

$$\left(\frac{1}{c^2} \frac{\partial^2 f}{(\partial t)^2} - \Delta f \right) (t, \vec{x}) = \left(\frac{1}{c^2} \frac{\partial^2 \tilde{f}}{(\partial \tilde{t})^2} - \tilde{\Delta} \tilde{f} - \frac{2v}{c} \frac{\partial^2 \tilde{f}}{\partial \tilde{t} \partial \tilde{x}^1} + \frac{v^2}{c^2} \frac{\partial^2 \tilde{f}}{(\partial \tilde{x}^1)^2} \right) \circ T(t, \vec{x}),$$

$$\text{où } \tilde{\Delta} = \frac{\partial^2}{(\partial \tilde{x}^1)^2} + \frac{\partial^2}{(\partial \tilde{x}^2)^2} + \frac{\partial^2}{(\partial \tilde{x}^3)^2}.$$

Cet exercice montre que, si jamais f est solution de l'équation des ondes, alors \tilde{f} n'est pas une solution de l'équation des ondes en générale. Cela montre que l'équation des ondes lumineuse est altérée lorsque l'on passe d'un référentiel galiléen à un autre. C'est a fortiori aussi le cas pour les équations de Maxwell.

Les physiciens de l'époque se trouvèrent face à un choix : il fallait soit abandonner le principe de relativité de Galilée, soit renoncer à l'interprétation ondulatoire de la lumière, qui reposait à l'époque sur la théorie de l'éther. Comme la nature ondulatoire de la lumière est un phénomène expérimental beaucoup plus évident que l'existence ou l'absence d'un repère absolu, il semblait préférable de simplement supposer que le principe de relativité de Galilée n'était pas vrai. Et alors il devait exister un repère privilégié, du moins du point de vue des phénomènes électromagnétiques, dans lequel l'éther était au repos.

Une troisième possibilité, permettant de réconcilier le principe de relativité de Galilée et la nature ondulatoire de la lumière, fut mise au point à la fin du XIXème siècle et au début du XXème siècle et est l'objet de la théorie de la relativité. La première étape vers cette théorie est la découverte d'un groupe de changements de coordonnées d'espace-temps compatible avec les équations de l'électromagnétisme et pouvant se substituer au groupe de Galilée. De telles transformations furent proposées d'abord par G.F. Fitzgerald et H.A. Lorentz comme lois phénoménologiques décrivant comment la matière est "déformée" lorsque l'on passe d'un repère à un autre, qui est en mouvement de translation à vitesse constante par rapport au premier. Cette théorie n'avait pas l'ambition de remettre en question la théorie de l'éther mais reposait sur l'hypothèse physique que les particules de matières étaient des phénomènes électromagnétiques, donc supportés par l'éther. Un peu plus tard H. Poincaré remarqua que ces règles forment un groupe de transformations, tout comme les changements de repère galiléens. En même temps ces transformations laissent invariantes les équations de Maxwell, ce qui n'est pas surprenant puisque précisément la démarche de Fitzgerald et Lorentz s'appuyait sur une vision "électromagnétique" de la matière. L'ensemble de ces transformations est

$$SO^\uparrow(1, 3) := \left\{ T : \begin{pmatrix} t \\ \vec{x} \end{pmatrix} \mapsto \begin{pmatrix} \text{ch}\theta & -\frac{1}{c}\text{sh}\theta\vec{n} \\ -c\text{sh}\theta\vec{n} & R\text{ch}\theta \end{pmatrix} \begin{pmatrix} t \\ \vec{x} \end{pmatrix} / \theta \in \mathbb{R}, \vec{n} \in S^2, R \in SO(3) \right\}.$$

H. Poincaré appela cet ensemble *groupe de Lorentz* et montra que l'on peut

étendre ce groupe en en faisant le produit semi-direct par les translations dans l'espace-temps (ce groupe étendu est appelé *groupe de Poincaré*). L'exercice suivant permet de vérifier que l'équation des ondes est invariante par le groupe de Lorentz.

Exercice 10 *Démontrer que, pour toute transformation $R \in SO^\uparrow(1,3)$, si $f = \tilde{f} \circ T$, alors*

$$\frac{1}{c^2} \frac{\partial^2 f}{(\partial t)^2} - \Delta f = \left(\frac{1}{c^2} \frac{\partial^2 \tilde{f}}{(\partial \tilde{t})^2} - \tilde{\Delta} \tilde{f} \right) \circ T.$$

Une des particularités de ces transformations (et qui sera à l'origine des résistances psychologiques opposées à la théorie de la relativité ultérieurement) est que l'échelle de temps peut se retrouver « contractée » ou « dilatée » d'un système de coordonnées à un autre, semblant passer plus ou moins vite dans l'un des systèmes. Et il en est de même pour certaines coordonnées d'espace. Mais pour Lorentz ou Fitzgérald il ne s'agissait que d'effets phénoménologiques, explicables par les forces de freinage que devait exercer l'éther sur tout corps qui se déplaçait par rapport à lui.

L'hypothèse de l'éther suggérait une expérience dont le but serait de déterminer le référentiel absolu, autrement dit de mesurer le déplacement du référentiel lié à un laboratoire construit sur terre par rapport à l'éther, supposé immobile. Une telle expérience fut réalisée par A.A. Michelson et E.W. Morley en 1887. Le résultat fut un choc : tout se passait comme si la terre était immobile ! Cela semble être une coïncidence hautement improbable, sachant que la terre tourne autour d'elle-même, autour du soleil et que le soleil n'a aucune raison d'être lui-même immobile, etc. Sans rentrer dans une description détaillée de l'expérience, il faut mentionner qu'elle reposait sur une mesure très fine (grâce à des mesures d'interférences lumineuses) des variations de la vitesse de la lumière lorsqu'elle se propage dans différentes directions. A vrai dire, ce qu'indiquait clairement le résultat de l'expérience, c'est que la vitesse de la lumière est constante dans toute les directions dans le référentiel lié à la terre.

Il était difficile de supposer que la terre soit « immobile » : c'eût été revenir à la thèse officielle de l'église à l'époque de G. Bruno et G. Galilée. Une solution radicale fut proposée par Einstein en 1905. Elle nécessite tout d'abord

d'abandonner l'idée de l'éther, mais en contrepartie elle rétablit le principe de relativité de Galilée qui avait été momentanément écarté. Mais il s'agit bien sûr d'adapter le principe de relativité, afin qu'il soit compatible avec les lois de l'électromagnétisme et les résultats de l'expérience de Michelson et Morley. Cela fut accompli en partant du principe que *la vitesse de la lumière est la même dans tous les référentiels inertiels*. Mathématiquement ce postulat de base amène à remplacer le groupe de Galilée par le groupe de Lorentz ou le groupe de Poincaré. Alors les transformations entre deux référentiels inertiels laissent invariantes les équations de Maxwell. Cela est à la base de la théorie de la *relativité*. Une version un peu différente de cette théorie fut proposée indépendamment et à la même époque par H. Poincaré.

D'un point de vue mathématique, il est utile d'observer, comme le firent H. Poincaré et H. Minkowski, que le groupe de Poincaré est le groupe des isométries linéaires qui laissent invariantes la forme quadratique¹ (« *produit scalaire de Minkowski* »)

$$\langle (t_1, x_1^1, x_1^2, x_1^3), (t_2, x_2^1, x_2^2, x_2^3) \rangle := c^2 t_1 t_2 - x_1^1 x_2^1 - x_1^2 x_2^2 - x_1^3 x_2^3.$$

L'espace-temps M^4 muni d'une structure affine et de la distance obtenue à partir de ce produit scalaire est appelé *espace de Minkowski*. Cette propriété est très utile du point de vue mathématique, puisqu'il est possible de déduire l'invariance des équations de Maxwell dans le vide par $SO^\uparrow(1, 3)$ à partir du fait que les transformations dans $SO^\uparrow(1, 3)$ préservent le produit scalaire de Minkowski².

Cela étant acquis, il restait à reconstruire ce qui avait été démolé et trouver une théorie de la dynamique pouvant se substituer à la théorie de Newton sui-

¹plus exactement $SO^\uparrow(1, 3)$ est la composante connexe de l'identité du groupe $O(1, 3)$ des transformations linéaires qui sont des « isométries » de \mathbb{R}^4 muni du produit scalaire de Minkowski. Ce groupe a 4 composantes connexes qui se distinguent entre elles selon que les transformations préservent ou non l'orientation de l'espace-temps et le sens d'écoulement du temps.

²un procédé pour démontrer cela consiste à écrire les équations de Maxwell dans le vide sous la forme $dF = d(\star F) = 0$, où F est la 2-forme sur l'espace-temps de Minkowski dont les composantes représentent le champ électromagnétique et à s'apercevoir que $\star F$ est l'image de F par l'opérateur de Hodge $\star : \Lambda^2 \mathbb{R}^4 \longrightarrow \Lambda^2 \mathbb{R}^4$ qui à toute 2-forme α associe d'unique 2-forme $\star \alpha$ telle que, pour toute 2-forme β , $\beta \wedge \star \alpha = \langle \beta, \alpha \rangle c dt \wedge dx^1 \wedge dx^2 \wedge dx^3$. Ici $\langle \beta, \alpha \rangle$ est le « produit scalaire » sur les 2-formes construit à partir du produit scalaire de Minkowski.

vant ces nouveaux principes. Nous allons la retrouver en nous aidant d'une formulation variationnelle. Nous nous contentons ici d'étudier le mouvement d'une particule dans l'espace-temps de Minkowski qui n'est soumise à aucune force. Fidèles aux principes énoncés par Galilée, nous postulons qu'une telle particule doit avoir un mouvement de translation rectiligne uniforme, dont la trajectoire dans l'espace-temps de Minkowski est une droite. Parmi les fonctionnelles définies sur l'espace des paramétrisations de trajectoires $\mathcal{C}^2(\mathbb{R}, M^4)$ qui sont invariantes par le groupe de Poincaré, celles qui sont les plus simples géométriquement et dont les points critiques sont des droites sont : la longueur $\int_{\mathbb{R}} \sqrt{\langle \dot{\gamma}(\tau), \dot{\gamma}(\tau) \rangle} d\tau$ et l'énergie $\int_{\mathbb{R}} \frac{1}{2} \langle \dot{\gamma}(\tau), \dot{\gamma}(\tau) \rangle d\tau$, où $\dot{\gamma} := \frac{d}{d\tau} \gamma$. Nous remarquons ici que la quantité $\langle \dot{\gamma}(\tau), \dot{\gamma}(\tau) \rangle$ n'est pas toujours positive : cette « norme au carré » est nulle si $\dot{\gamma}(\tau)$ prend ses valeurs dans un cône, appelé *cône de lumière*, positive à l'« intérieur » de cône et négative à l'« extérieur » de ce cône. L'inégalité $\langle \dot{\gamma}(\tau), \dot{\gamma}(\tau) \rangle > 0$ signifie que la vitesse de la particule qui suit la trajectoire est strictement inférieure à celle de la lumière. C'est bien la situation la plus réaliste pour une particule matérielle, c'est pourquoi nous considérons les fonctionnelles sur l'espace :

$$\Gamma := \{ \gamma \in \mathcal{C}^2(\mathbb{R}, M^4) / \forall \tau \in \mathbb{R}, \langle \dot{\gamma}(\tau), \dot{\gamma}(\tau) \rangle > 0 \}$$

des trajectoires « de type temps ». Cela permet d'éviter de considérer $\sqrt{\langle \dot{\gamma}(\tau), \dot{\gamma}(\tau) \rangle}$ lorsque $\langle \dot{\gamma}(\tau), \dot{\gamma}(\tau) \rangle$ est négatif. Les deux types d'action (longueur ou énergie) donnent essentiellement les mêmes résultats. Prenons par exemple le cas de la longueur :

$$\mathcal{L}[\gamma] = A \int_{\mathbb{R}} \sqrt{\langle \dot{\gamma}(\tau), \dot{\gamma}(\tau) \rangle} d\tau.$$

Cette fonctionnelle est invariante par changement de paramétrisation de γ : pour tout difféomorphisme $\varphi : \mathbb{R} \rightarrow \mathbb{R}$, $\mathcal{L}[\gamma \circ \varphi] = \mathcal{L}[\gamma]$ (*Exercice : vérifier*). Ici la constante A peut être ajustée de façon à essayer de retrouver les principes de la physique newtonienne lorsque la particule a une vitesse très petite par rapport à la vitesse de la lumière. Envisageons une trajectoire γ qui soit un point critique de \mathcal{L} , c'est à dire dont l'image dans M^4 est une droite. En vertu de l'invariance par changement de paramétrisation, nous sommes libres de choisir une paramétrisation γ telle que $\gamma(\tau) = {}^t(\tau, v^1\tau, v^2\tau, v^3\tau)$. Alors

$$\begin{aligned} \mathcal{L}[\gamma] &= A \int_{\mathbb{R}} \sqrt{c^2 - |\vec{v}|^2} d\tau = A \int_{\mathbb{R}} c \left(1 - \frac{|\vec{v}|^2}{c^2} \right)^{\frac{1}{2}} d\tau \\ &= A \int_{\mathbb{R}} c \left(1 - \frac{|\vec{v}|^2}{2c^2} + O(c^{-4}) \right) d\tau = \int_{\mathbb{R}} \left(Ac - A \frac{|\vec{v}|^2}{2c} + O(c^{-3}) \right) d\tau. \end{aligned}$$

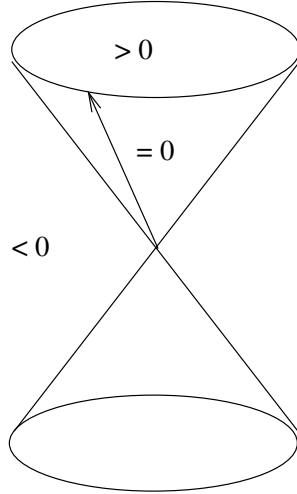


FIG. 5.1 – Le cône de lumière et le signe de $\langle v, v \rangle$ selon que v soit à l'intérieure, à l'extérieure ou sur le cône de lumière.

Nous voyons que le premier terme n'apporte aucune information tandis que le deuxième coïncide avec l'action de Maupertuis à condition que $A = -mc$. Nous choisissons donc cette valeur : $\mathcal{L}[\gamma] := -mc \int_{\mathbb{R}} \sqrt{c^2 - |\vec{v}|^2} d\tau$.

Il est intéressant de calculer les valeurs des impulsions données par le lagrangien $L(\dot{\gamma}(\tau)) = -mc\sqrt{\langle \dot{\gamma}(\tau), \dot{\gamma}(\tau) \rangle}$, toujours pour une trajectoire $\gamma(\tau) = {}^t(\tau, v^1\tau, v^2\tau, v^3\tau)$:

$$\frac{\partial L}{\partial \dot{t}} = -\frac{mc^3}{\sqrt{c^2 - |\vec{v}|^2}} = -\frac{mc^2}{\sqrt{1 - |\vec{v}|^2/c^2}} = -mc^2 - m\frac{|\vec{v}|^2}{2} + O(c^{-2}),$$

$$\frac{\partial L}{\partial \dot{x}^\alpha} = \frac{mcv^\alpha}{\sqrt{c^2 - |\vec{v}|^2}} = \frac{mv^\alpha}{\sqrt{1 - |\vec{v}|^2/c^2}}.$$

Cela nous donne les valeurs du moment cinétique $p_\alpha = \frac{mv^\alpha}{\sqrt{1 - |\vec{v}|^2/c^2}}$ et de l'énergie $E = -\frac{\partial L}{\partial \dot{t}} = \frac{mc^2}{\sqrt{1 - |\vec{v}|^2/c^2}} = mc^2 + m\frac{|\vec{v}|^2}{2} + O(c^{-2})$ pour une particule en mécanique relativiste. A des vitesses petites par rapport à c , il apparaît que l'énergie est approximativement la somme de deux contributions : le fameux terme mc^2 qui s'interprète comme une énergie au repos et l'énergie cinétique classique. Dans le point de vue géométrique de Minkowski, il est

naturel d'interpréter ces quantités comme les composantes $(p_0, p_1, p_2, p_3) := \left(-\frac{mc^2}{\sqrt{1-|\vec{v}|^2/c^2}}, \frac{mv^\alpha}{\sqrt{1-|\vec{v}|^2/c^2}} \right)$ d'un covecteur « quadri-impulsion » dans T^*M^4 , que l'on a coutume d'appeler « quadrivecteur »³. Ce covecteur satisfait une relation importante, $\frac{1}{c^2}(p_0)^2 - (p_1)^2 - (p_2)^2 - (p_3)^2 = m^2c^2$, que les physiciens ont l'habitude d'écrire sous la forme :

$$E^2 - c^2|\vec{p}|^2 = (mc^2)^2. \quad (5.1)$$

5.2 Retour aux quanta de lumières

La relation (5.1) a joué un rôle important dans le développement de la mécanique quantique. En la comparant avec la relation de dispersion pour la lumière

$$\hbar^2\omega^2 - c^2\hbar^2|\vec{k}|^2 = 0, \quad (5.2)$$

qui relie la pulsation ω au vecteur d'onde \vec{k} , une analogie saute aux yeux. Nous avons ici volontairement multiplié le terme $\omega^2 - c^2|\vec{k}|^2$ par \hbar^2 . Sachant que $\hbar\omega$ s'interprète comme l'énergie d'un quantum de lumière (photon), nous sommes tentés de faire le rapprochement et d'imaginer que :

Le photon est une particule dont la masse au repos est nulle, l'énergie est $E = \hbar\omega$ et dont l'impulsion est $\vec{p} = \hbar\vec{k}$.

Cette hypothèse fut proposée par Einstein en 1905. Elle attribue ainsi au quantum de lumière le statut de particule. Elle sera confirmée en 1923 par la découverte expérimentale de l'effet Compton, observé lors de la diffraction de rayons X (photons très énergétiques) sur une mince feuille d'aluminium : les électrons sont expulsés et le rayonnement est diffusé *avec une longueur d'onde différente de la longueur d'onde incidente*. Ce phénomène ne peut pas s'expliquer avec un point de vue ondulatoire sur la lumière, mais corpusculaire. Les observations expérimentales confirment que l'on doit appliquer les mêmes règles de conservation d'énergie et de moment cinétique que pour les collisions de particules relativistes massives, avec la valeur $\vec{p} = \hbar\vec{k}$ pour le photon.

Plus tard nous reviendrons sur la comparaison entre (5.1) et (5.2).

³un géomètre différentiel dirait plutôt « quadricovecteur ».

5.3 Le spectre des atomes

Un autre problème se posait à la physique, celui des raies spectrales des atomes. Soulignons qu'à la différence des questions à l'origine de la théorie de la relativité restreinte, rien ne laissait présager au départ que la compréhension des raies spectrales nécessitait de bouleverser les règles de la physique. De plus, jusqu'au début du XX^{ème} siècle, on ne disposait pas de modèle décrivant l'atome de façon suffisamment complète pour pouvoir être testé⁴.

Le phénomène est le suivant : les atomes n'émettent de la lumière (ou plus généralement des rayonnements électromagnétiques) qu'à certaines longueurs d'onde. Par exemple l'atome d'hydrogène ne peut émettre qu'à des fréquences $\nu = c/\lambda$ particulières. En 1885 Balmer (il n'était pas physicien, mais un professeur de lycée, passionné de numérogie) découvre que la longueur d'onde λ de la lumière émise par l'hydrogène dans le spectre visible satisfait nécessairement la relation :

$$\frac{1}{\lambda} = \frac{1}{L} \left(1 - \frac{4}{n^2} \right), \quad \text{où } n \in \mathbb{N} \text{ et } n \geq 3.$$

Ultérieurement d'autres valeurs de λ ont pu être mesurées (en dehors du spectre visible) : elles obéissent à la règle de Rydberg, généralisant celle de Balmer :

$$\frac{1}{\lambda} = R \left(\frac{1}{m^2} - \frac{1}{n^2} \right), \quad \text{où } n, m \in \mathbb{N}^* \text{ et } m < n,$$

où R est la *constante de Rydberg* $R = 4/L$ (L étant la « longueur de Balmer »). L'étude du spectre des atomes était une discipline scientifique à part entière, mais relativement ignorée des physiciens, à tel point que Bohr, dont nous parlerons plus loin n'eut connaissance de la règle de Balmer que quelques semaines avant de faire sa découverte.

En 1911, grâce des expériences qui consistent à bombarder des plaques métalliques par des rayonnements α ⁵, E. Rutherford découvre que la structure

⁴toutefois un modèle avait été bâti par Lorentz, en partant d'une idée proposée par Thomson. Dans ce modèle les électrons étaient agités d'un mouvement de vibration à l'intérieur de l'atome. C'est notamment sur ce modèle que Planck s'appuya dans son travail sur l'interaction de la matière avec la lumière.

⁵il s'agit de noyaux d'Hélium, c'est à dire constitués de 2 protons et 2 neutrons

de l'atome est différente de ce que l'on pensait jusqu'alors : les charges positives sont en fait concentrées dans un volume très petit par rapport au reste de l'atome. Rutherford conçoit alors son *modèle planétaire* de l'atome, calqué sur le système solaire : des électrons très légers (que l'on imagine comme des particules quasiment ponctuelles) gravitent autour d'un noyau chargé positivement et plusieurs milliers de fois plus massif que l'électron. La taille du noyau est de l'ordre de 10^4 fois plus petite que celle de l'atome. Cela autorise une description plus précise de l'atome d'hydrogène : soit m_p la masse du noyau (proton) et m_e celle de l'électron, soit q la charge du noyau et $-q$ celle de l'électron. L'évolution du système atomique est décrite par les points critiques de

$$\int_{\mathbb{R}} \left(m_p \frac{|\dot{\gamma}_p|^2}{2} + m_e \frac{|\dot{\gamma}_e|^2}{2} + \frac{q^2}{4\pi\epsilon_0|\gamma_p - \gamma_e|} \right) dt.$$

Exercice 11 On pose $C(t) := \frac{m_p\gamma_p(t)+m_e\gamma_e(t)}{m_p+m_e}$ (centre de masse), $\gamma := \gamma_e - \gamma_p$ (position relative), $m := \frac{m_p m_e}{m_p + m_e}$ (masse réduite) et $e^2 = \frac{q^2}{4\pi\epsilon_0}$.

1) Démontrer que l'action se réécrit :

$$\int_{\mathbb{R}} (m_p + m_e) \frac{|\dot{C}|^2}{2} dt + \int_{\mathbb{R}} \left(m \frac{|\dot{\gamma}|^2}{2} + \frac{e^2}{|\gamma|} \right) dt.$$

On notera $\mathcal{A}(\gamma)$ la deuxième intégrale.

2) Démontrer que les équations d'Euler-Lagrange de cette action sont :

$$(m_p + m_e)\ddot{C} = 0 \quad \text{et} \quad m\ddot{\gamma} = -\frac{e^2\gamma}{|\gamma|^3}.$$

Expliquer pourquoi on peut se ramener à l'étude de γ (et laisser de côté C).

3) On recherche les solutions circulaires, de la forme $\gamma(t) = (r \cos \omega t, r \sin \omega t, 0)$. Démontrer qu'une telle expression est solution des équations du mouvement si et seulement si :

$$m\omega^2 r = \frac{e^2}{r}. \tag{5.3}$$

4) L'énergie cinétique est $E_c = m \frac{|\dot{\gamma}|^2}{2}$, l'énergie potentielle est $E_p = -\frac{e^2}{|\gamma|}$ et l'énergie totale $E = E_c + E_p$. Calculer la valeur de E_c pour une solution circulaire en fonction de e^2 et r . En déduire que $E = -e^2/2r$.

5) L'intensité du moment angulaire⁶ pour une trajectoire circulaire est $J =$

⁶ $\vec{J} := m\dot{\gamma} \times \gamma$

$m\omega r^2$. Exprimer J en fonction de m , e^2 et r pour une solution circulaire (on pourra utiliser (5.3)). En déduire la relation suivante entre l'énergie totale et J :

$$E = -\frac{me^4}{2J^2}. \quad (5.4)$$

Le modèle de Rutherford permettait de faire un certain nombre de calculs mais, pas d'expliquer les raies spectrales, c'est à dire l'apparition de valeurs discrètes pour les fréquences de la lumière émise par les atomes. De plus ce modèle souffrait d'une difficulté que nous n'avons pas le temps de détailler ici : le mouvement de l'électron est nécessairement accéléré (que la trajectoire de l'électron soit circulaire ou elliptique). Or on peut démontrer (en utilisant les équations de Maxwell avec source) que *toute particule électriquement chargée rayonne de l'énergie sous forme d'onde électromagnétique*. Or la relation $E = -e^2/2r$ (pour une trajectoire circulaire) montre que si l'électron perd de l'énergie sa distance au noyau diminue. En d'autres termes le rayonnement électromagnétique émis par l'électron « freine » celui-ci et l'on peut calculer que l'électron devrait s'écraser sur le noyau en une fraction de seconde. Cela est évidemment en contradiction avec la très grande stabilité de l'atome. Ainsi nous sommes face à une incohérence sérieuse, qui n'est pas sans rappeler les difficultés rencontrées pour comprendre le rayonnement du corps noir.

En 1912 N. Bohr a connaissance de ces problèmes, de la quantification de la lumière par Planck et Einstein et enfin, tout récemment il vient d'apprendre les règles de Balmer. Il conçoit alors en quelques semaines le modèle suivant :

Comme pour la lumière, les niveaux d'énergie de l'électron dans l'atome sont quantifiés. Plus précisément, en dehors des phénomènes transitoires, l'électron est dans un état stationnaire et les états stationnaires ne peuvent prendre que certains niveaux d'énergies, discrets :

$$E_1 < E_2 < \dots < E_n < \dots < 0.$$

Lorsqu'un électron passe d'un état stationnaire avec un niveau d'énergie E_n à un autre état stationnaire avec un niveau d'énergie E_m , avec $m < n$, il émet un photon dont l'énergie est $\hbar\omega = E_n - E_m$.

L'idée est donc en quelque sorte de se débarrasser du problème de la perte d'énergie par rayonnement en postulant que les orbites de l'électron dans

l'atome n'obéissent pas aux lois de la physique classique, mais à d'autres règles qui, comme pour la quantification de la lumière, restent encore mystérieuses. L'intérêt est aussi d'interpréter les raies spectrales : la règle de Rydberg (associée avec la quantification de Planck et Einstein de la lumière) nous enseigne que les photons émis par l'atome d'hydrogène ont une énergie

$$\hbar\omega = \frac{hc}{\lambda} = hcR \left(\frac{1}{m^2} - \frac{1}{n^2} \right), \quad \text{où } n, m \in \mathbb{N}^* \text{ et } m < n.$$

Une telle expression est bien du type $E_n - E_m$, où :

$$E_n = -\frac{hcR}{n^2}, \quad E_m = -\frac{hcR}{m^2} \quad \text{pour } n, m \in \mathbb{N}^* \text{ et } m < n.$$

Il reste à interpréter le sens de cette relation. Nous pouvons comparer la règle de quantification $E_n = -\frac{hcR}{n^2}$ avec le résultat du calcul classique donné en (5.4) : $E = -\frac{me^4}{2J^2}$. Cela donne :

$$-\frac{hcR}{n^2} = -\frac{me^4}{2J^2} \iff J = \sqrt{\frac{me^4}{2hcR}}n.$$

Rappelons que la constante de Rydberg R était phénoménologique. Une observation cruciale est que la valeur du coefficient $\sqrt{\frac{me^4}{2hcR}}$ coïncide avec... $\hbar = h/2\pi$.

Conclusion — N. Bohr adopte comme « règle de quantification » des niveaux d'énergie de l'électron dans l'atome d'hydrogène que le moment cinétique angulaire J est égal à J_n où :

$$J_n := \hbar n \iff E_n = -\frac{me^4}{2\hbar^2} \frac{1}{n^2}, \quad \text{pour } n \in \mathbb{N}^*.$$

Exercice 12 Démontrer que, si $\gamma(t)$ est une solution classique circulaire de $m\ddot{\gamma} = -e^2\gamma/|\gamma|^3$ de période $2\pi/\omega$ et si $J = \hbar n$, alors

$$\int_0^{2\pi/\omega} \sum_{\alpha=1}^3 p_\alpha dq^\alpha = nh,$$

où $q(t) = \gamma(t)$ est la position et $p_\alpha(t) = \frac{\partial L}{\partial v^\alpha}(\gamma(t), \dot{\gamma}(t))$ (avec $L(x, v) := m|v|^2/2 + e^2/|x|$) est l'impulsion.

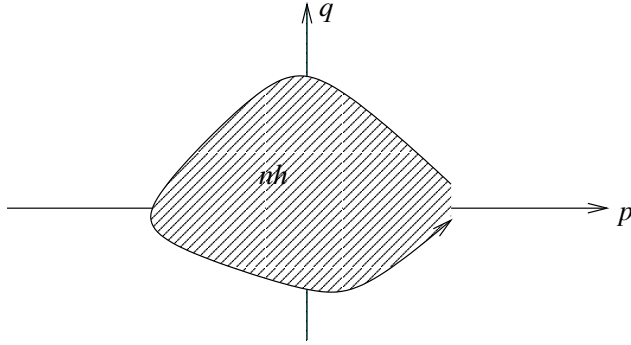


FIG. 5.2 – Une interprétation semi-classique de la « quantification » d'un mouvement périodique : l'aire entourée par la trajectoire dans l'espace symplectique est un multiple entier de h .

Le résultat de cet exercice montre que la règle de quantification de Bohr a une interprétation de même nature que celle de Planck (où l'on avait assimilé un mode électromagnétique à un oscillateur harmonique) : dans les deux cas, l'intégrale $\int pdq$ sur une orbite périodique (donc l'aire symplectique enfermée par la trajectoire hamiltonienne) est un multiple entier de h (il s'agit d'une observation due à Einstein).

5.4 L'équation d'Hamilton–Jacobi

Nous allons remonter un peu en arrière et revenir à une question que s'était posée Hamilton : comment concilier le principe de Huygens et celui de Fermat pour la lumière ? Rappelons que, selon le point de vue de Huygens, la lumière est une onde qui se propage dans l'espace un peu à la manière des ronds que l'on observe à la surface de l'eau d'un étang lorsqu'on jette un caillou. Cette hypothèse fut confirmée expérimentalement par les expériences de diffraction de Young et Fresnel et est totalement prise en compte dans la description de la lumière comme une oscillation du champ électromagnétique donnée par la théorie de Maxwell. Suivant Huygens, nous pouvons donc associer à une onde lumineuse des surfaces (« fronts d'onde », que dans le cadre de la théorie de Maxwell nous pouvons interpréter comme les surfaces de niveau d'une des composantes du champ électromagnétique satisfaisant l'équation des ondes) se propageant dans l'espace. Il argua qu'en connaissant

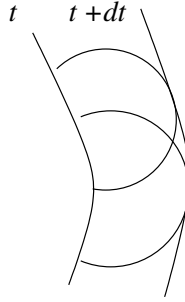


FIG. 5.3 – Le principe de Huygens : le front d’onde à l’instant $t + dt$ est l’enveloppe des sphères de rayon $c dt$ centrées sur tous les points du front d’onde à l’instant t .

la position d’un front d’onde à un certain instant t et la direction dans laquelle la lumière se propage, il est possible de construire géométriquement la position du front d’onde à un instant infinitésimalement ultérieur $t + dt$, comme étant l’enveloppe de sphères infinitésimales de rayon $c dt$ centrées sur tous les points du front d’onde. Cette construction est le *principe de Huygens*.

Ce principe était concurrent d’une autre description de la lumière, envisagée comme la propagation d’un rayon suivant le plus court chemin, comme le prévoit le principe de Fermat. Il semblait toutefois que les partisans d’un tel point de vue aient perdu la bataille au début du XIXème siècle, lorsque les expériences de diffraction avaient imposé le point de vue de Huygens⁷.

Cependant Hamilton cherchait à comprendre comment les deux aspects antagonistes de la lumière pouvaient être compris simultanément : il s’agit de bâtir une description intégrant la propagation de fronts d’onde et de « rayons » ponctuelles. L’idée est que chaque sphère infinitésimale de Huygens de rayon $c dt$ soit simplement le lieu des extrémités des chemins de longueur extrémale (donc des segments de droite) issus d’un même point A sur le front d’onde « initial » (à l’instant t). Le front d’onde « ultérieur » (à l’instant $t + dt$) étant l’enveloppe de toutes les petites sphères, il est nécessairement tangent aux sphères infinitésimales. Donc le segment de droite infinitésimal qui joint

⁷Bien sûr les découvertes de Planck et Einstein viendront restaurer partiellement la pertinence d’une théorie corpusculaire de la lumière

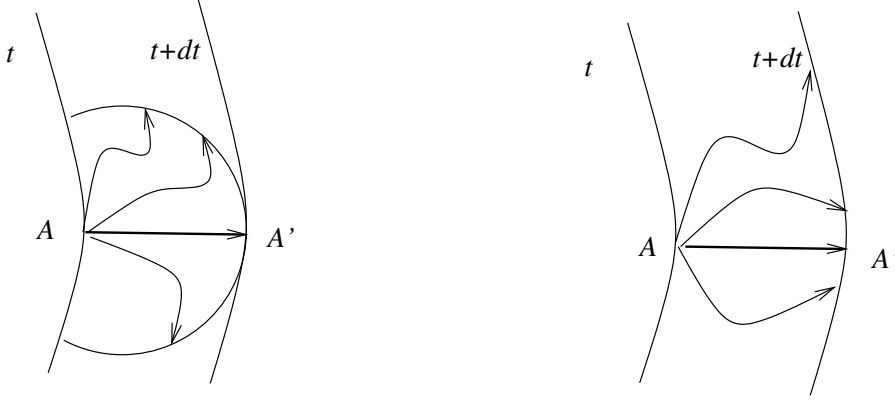


FIG. 5.4 – Le chemin de A à A' est à la fois le plus court chemin entre A et le point de contact de la sphère de Huygens avec le front d'onde ultérieur (en $t + dt$) et le plus court chemin entre A et le front d'onde ultérieur.

le point A sur la surface initiale au point de contact de la sphère centrée en A avec la surface ultérieure est *aussi* le plus court chemin qui joint A à la surface ultérieure (voir figure 5.4).

Nous donnons dans la suite une présentation mathématique de cette idée et de sa généralisation (en partie due à Carathéodory) pour un problème variationnel à une variable (le temps). Dans la suite on pourra imaginer que la densité lagrangienne est $L(\gamma, \dot{\gamma}) = \frac{m}{2} |\dot{\gamma}|^2 - V(\gamma)$, c'est à dire est celle du principe de Maupertuis.

Principe — Nous considérons un ouvert Ω de $\mathbb{R} \times \mathbb{R}^3$ (espace-temps) et une densité lagrangienne $L \in \mathcal{C}^1(\Omega \times \mathbb{R}^3)$ (une fonction des variables $(t, x) \in \Omega$ et $v \in \mathbb{R}^3$). Cela définit une action \mathcal{A} pour toute trajectoire (I, Γ) (où I est un intervalle de \mathbb{R} et $I \ni t \mapsto \Gamma(t) = (t, \gamma(t)) \in \Omega$ paramétrise une courbe de classe \mathcal{C}^1) par :

$$\mathcal{A}[I, \Gamma] := \int_I L(t, \gamma(t), \dot{\gamma}(t)) dt.$$

Nous cherchons :

- une famille d'hypersurfaces $(H_s)_{s \in [s, \bar{s}]}$ qui forme un feuilletage de Ω (les « fronts d'onde » dans l'espace-temps)
- une famille de trajectoires $(I_q, \Gamma_q)_{q \in \omega}$ (où ω est un ouvert de \mathbb{R}^3) qui

forme un feuilletage de Ω , où $I_q \ni t \mapsto \Gamma_q(t) = (t, \gamma_q(t)) \in \Omega$.

Avec les propriétés que

- (i) chaque trajectoire Γ_q coupe chaque hypersurface H_s en un seul point $(t_q^s, x_q^s) := (t_q^s, \gamma_q(t_q^s))$ et
- (ii) si $I_q^s :=]-\infty, t_q^s] \cap I_q$, $(I_q^s, \Gamma_q|_{I_q^s})$ est point critique de \mathcal{A} pour des variations qui autorisent le point extrémité $(t_q^s, \gamma_q(t_q^s))$ à « glisser » le long de H_s .

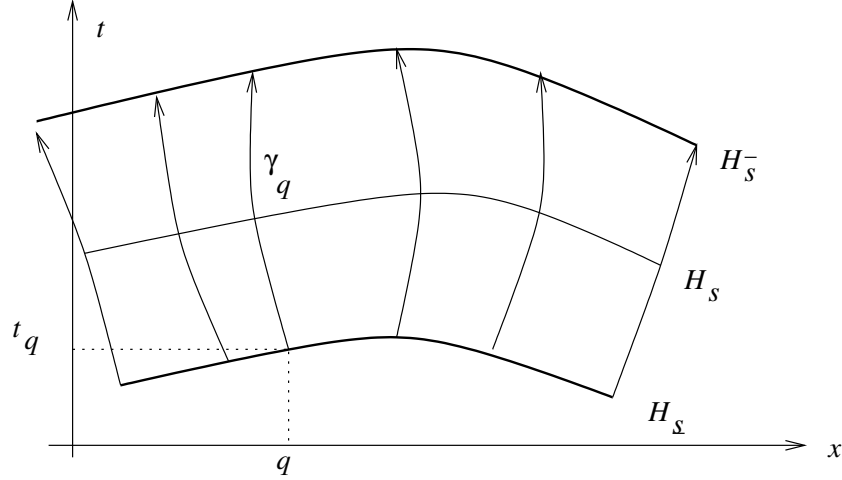


FIG. 5.5 – Le feuilletage de Ω par la famille de fronts d'onde $(H_s)_{s \in [\underline{s}, \bar{s}]}$ et la famille de trajectoires $(I_q, \Gamma_q)_{q \in \omega}$.

La propriété (ii) est une généralisation du principe de Huygens. Précisons un peu les hypothèses. Sans perte de généralité, nous pouvons choisir l'indice q de façon à ce qu'il désigne le point de l'espace dont est issue la trajectoire Γ_q : I_q est un intervalle de la forme $[t_q^s, t_q^{\bar{s}}]$ et $\gamma_q(t_q^s) = q$. Nous supposons de plus que $\{(t_q^s, q)/q \in \omega\}$ coïncide avec une hypersurface $H_{\underline{s}}$ et $\{(t_q^{\bar{s}}, x_q^{\bar{s}})/q \in \omega\}$ avec une autre hypersurface $H_{\bar{s}}$.

Pour tout $s \in]\underline{s}, \bar{s}]$ nous notons

$$\mathcal{E}_q^s := \{([t_q^s, t], \Gamma)/\Gamma \in \mathcal{C}^1([t_q^s, t], \Omega), \forall \tau \in [t_q^s, t], \Gamma(\tau) = (\tau, \gamma(\tau)), \gamma(t_q^s) = q, \Gamma(t) \in H_s\}$$

et, pour toute valeur de (s', s'') telle que $\underline{s} \leq s' < s'' \leq s \leq \bar{s}$, nous définissons

la fonctionnelle sur \mathcal{E}_q^s

$$\mathcal{A}_{s'}^{s''}(I, \Gamma) = \int_{(t, \gamma(t)) \in H_{s'}}^{(t, \gamma(t)) \in H_{s''}} L(t, \gamma(t), \dot{\gamma}(t)) dt.$$

Noter que l'hypothèse (ii) signifie que $(I_q^s, \Gamma_{q|I_q^s})$ est point critique de $\mathcal{A}_{\underline{s}}^s$ sur \mathcal{E}_q^s .

Cela va entraîner la propriété suivante : pour tout $q_1, q_2 \in H_{\underline{s}}$ les trajectoires Γ_{q_1} et Γ_{q_2} issues respectivement de $(t_{q_1}^s, q_1)$ et $(t_{q_2}^s, q_2)$ ont même action entre deux valeurs $s' < s''$. C'est à dire, $\forall s', s''$, telles que $\underline{s} \leq s' < s'' \leq \bar{s}$,

$$\mathcal{A}_{s'}^{s''}(I_{q_1}, \Gamma_{q_1}) = \mathcal{A}_{s'}^{s''}(I_{q_2}, \Gamma_{q_2}). \quad (5.5)$$

Preuve — Supposons dans un premier temps que $\underline{s} < s'$. Soit $]-\varepsilon_0, \varepsilon_0[\ni \varepsilon \mapsto q(\varepsilon)$ un chemin de classe \mathcal{C}^1 dans ω et considérons la famille de trajectoires à un paramètre $(I_{q(\varepsilon)}, \Gamma_{q(\varepsilon)})_\varepsilon$ correspondantes. Nous posons $q := q(0)$. Nous considérons une famille de déformations $(\tilde{I}_\varepsilon, \tilde{\Gamma}_\varepsilon)$ de (I_q, Γ_q) , où $\tilde{I}_\varepsilon = [t_q^s, t_{q(\varepsilon)}^{s''}]$, telle que $\tilde{\Gamma}_\varepsilon$ coïncide avec Γ_q sur un voisinage de t_q^s dans $[t_q^s, t_{q(\varepsilon)}^{s'}[$, et coïncide avec $\Gamma_{q(\varepsilon)}$ sur $[t_{q(\varepsilon)}^{s'}, t_{q(\varepsilon)}^{s''}]$. De

$$\mathcal{A}_{\underline{s}}^{s''}(\tilde{I}_\varepsilon, \tilde{\Gamma}_\varepsilon) = \mathcal{A}_{\underline{s}}^{s''}(I_q, \Gamma_q) + o(\varepsilon)$$

et

$$\mathcal{A}_{\underline{s}}^{s'}(\tilde{I}_\varepsilon, \tilde{\Gamma}_\varepsilon) = \mathcal{A}_{\underline{s}}^{s'}(I_q, \Gamma_q) + o(\varepsilon),$$

nous déduisons par soustraction (puisque $\mathcal{A}_{\underline{s}}^{s''} = \mathcal{A}_{\underline{s}}^{s'} + \mathcal{A}_{s'}^{s''}$) que

$$\mathcal{A}_{s'}^{s''}(I_{q(\varepsilon)}, \Gamma_{q(\varepsilon)}) = \mathcal{A}_{s'}^{s''}(\tilde{I}_\varepsilon, \tilde{\Gamma}_\varepsilon) = \mathcal{A}_{s'}^{s''}(I_q, \Gamma_q) + o(\varepsilon).$$

Cela entraîne donc que $\mathcal{A}_{s'}^{s''}(I_q, \Gamma_q)$ est constant lorsque q parcourt $H_{\underline{s}}$ (car la dérivée de cette expression par rapport à ε est nulle).

L'identité (5.5) pour $\underline{s} < s' < s'' \leq \bar{s}$ s'étend au cas $\underline{s} = s'$ de la façon suivante : puisque γ_q est \mathcal{C}^1 sur $[t_q^s, t_q^{\bar{s}}]$, $\lim_{s' \rightarrow \underline{s}} \mathcal{A}_{s'}^{s''}(\gamma_q) = \mathcal{A}_{\underline{s}}^{s''}(\gamma_q)$. Il suffit donc de passer à la limite dans (5.5) lorsque $s' \rightarrow \underline{s}$. ■

Une conséquence de (5.5) est la suivante : puisque $\mathcal{A}_{s'}^{s''}(I_q, \Gamma_q)$ ne dépend que de s' et s'' , on peut poser $\mathcal{A}_{s'}^{s''}(I_q, \Gamma_q) = F(s', s'')$. Et à cause de la propriété d'additivité $\mathcal{A}_{s'}^{s'''}(I_q, \Gamma_q) = \mathcal{A}_{s'}^{s''}(I_q, \Gamma_q) + \mathcal{A}_{s''}^{s'''}(I_q, \Gamma_q)$, on a $F(s', s''') =$

$$\begin{aligned}
& \mathcal{A} \left(\begin{array}{c} \text{---} \\ s'' \\ \vdots \\ \text{---} \\ s' \\ \vdots \\ \text{---} \\ \text{---} \end{array} \right) - \mathcal{A} \left(\begin{array}{c} \text{---} \\ s'' \\ \vdots \\ \text{---} \\ s' \\ \vdots \\ \text{---} \\ \text{---} \end{array} \right) = \\
& \mathcal{A} \left(\begin{array}{c} \text{---} \\ s'' \\ \vdots \\ \text{---} \\ s' \\ \vdots \\ \text{---} \\ \text{---} \end{array} \right) - \mathcal{A} \left(\begin{array}{c} \text{---} \\ s'' \\ \vdots \\ \text{---} \\ s' \\ \vdots \\ \text{---} \\ \text{---} \end{array} \right) + \mathcal{A} \left(\begin{array}{c} \text{---} \\ s'' \\ \vdots \\ \text{---} \\ s' \\ \vdots \\ \text{---} \\ \text{---} \end{array} \right) - \mathcal{A} \left(\begin{array}{c} \text{---} \\ s'' \\ \vdots \\ \text{---} \\ s' \\ \vdots \\ \text{---} \\ \text{---} \end{array} \right) \\
& = o(\varepsilon)
\end{aligned}$$

FIG. 5.6 – L'idée du calcul de $\mathcal{A}_{s'}^{s''}(I_{q(\varepsilon)}, \Gamma_{q(\varepsilon)}) - \mathcal{A}_{s'}^{s''}(I_q, \Gamma_q)$.

$F(s', s'') + F(s'', s''')$ et donc $F(s', s'') = f(s'') - f(s')$, pour une certaine fonction f . Supposons que f soit une fonction monotone⁸ (et donc un difféomorphisme), alors, quitte à changer s en $f(s)$, nous pouvons supposer que $f(s) = s$. Donc nous nous ramenons ainsi au cas où

$$\mathcal{A}_{\underline{s}}^s(I_q, \Gamma_q) = s - \underline{s}. \quad (5.6)$$

Définissons la fonction $S : \Omega \rightarrow \mathbb{R}$ par $\forall s \in [\underline{s}, \bar{s}]$, $\forall (t, x) \in H_s$, $S(t, x) = s$. Autrement dit les ensembles de niveau de S sont exactement les hypersurfaces H_s . La définition de S et la propriété (5.6) peuvent être résumées par

$$\forall q \in \omega, \forall s \in [\underline{s}, \bar{s}], \quad S(t_q^s, x_q^s) = s = \underline{s} + \int_{t_q^{\underline{s}}}^{t_q^s} L(\tau, \gamma_q(\tau), \dot{\gamma}_q(\tau)) d\tau. \quad (5.7)$$

Cette relation entraîne, en remplaçant (t_q^s, x_q^s) par $(t_q^s + t, \gamma_q(t_q^s + t))$ et en

⁸cela est vrai si par exemple $L(t, \gamma_q, \dot{\gamma}_q) > 0$, $\forall q \in \omega$, $\forall t$, une propriété satisfaite dans la plupart des exemples empruntés à la mécanique.

dérivant par rapport à t , en $t = 0$:

$$\frac{\partial S}{\partial t}(t_q^s, x_q^s) + \sum_{\alpha=1}^3 \frac{\partial S}{\partial x^\alpha}(t_q^s, x_q^s) \dot{\gamma}_q^\alpha(t_q^s) = L(t_q^s, x_q^s, \dot{\gamma}_q^\alpha(t_q^s)). \quad (5.8)$$

Equation d'Euler-Lagrange pour (I_q, Γ_q) — Nous allons à présent chercher la condition sur les trajectoires (I_q, Γ_q) qui caractérise le fait que ces trajectoires sont points critiques de $\mathcal{A}_{\underline{s}}^s$, pour tout $s \in]\underline{s}, \bar{s}]$. Pour cela nous envisageons des variations de (I_q^s, Γ_q) , où $I_q^s = [t_q^s, t_q^s]$ de la forme $([t_q^s, t_q^s + \tau(\varepsilon)], \Gamma_q + \varepsilon\Phi)$, où $\varepsilon \in]-\varepsilon_0, \varepsilon_0[$, $\tau \in \mathcal{C}^1(]-\varepsilon_0, \varepsilon_0[,]-T, T])$ satisfait $\tau(0) = 0$ et $\Phi \in \mathcal{C}_c^1([t_q^s, t_q^s + T], \mathbb{R}^4)$ est de la forme $\Phi(t) = (0, \varphi(t))$. Nous voulons que $([t_q^s, t_q^s + \tau(\varepsilon)], \Gamma_q + \varepsilon\Phi)$ reste dans \mathcal{E}_q^s , et donc il est nécessaire que

$$\Gamma_q(t_q^s + \tau(\varepsilon)) + \varepsilon\Phi(t_q^s + \tau(\varepsilon)) \in H_s.$$

Comme $\Gamma_q(t_q^s + \tau(\varepsilon)) + \varepsilon\Phi(t_q^s + \tau(\varepsilon)) = (t_q^s, x_q^s) + \varepsilon(\tau'(0), \dot{\gamma}_q(t_q^s)\tau'(0) + \varphi(t_q^s)) + o(\varepsilon)$, cette condition entraîne à l'ordre 1 en ε que $(\tau'(0), \dot{\gamma}_q(t_q^s)\tau'(0) + \varphi(t_q^s)) \in T_{(t_q^s, x_q^s)}H_s$ ou encore :

$$\frac{\partial S}{\partial t}(t_q^s, x_q^s)\tau'(0) + \sum_{\alpha=1}^3 \frac{\partial S}{\partial x^\alpha}(t_q^s, x_q^s) (\dot{\gamma}_q^\alpha(t_q^s)\tau'(0) + \varphi^\alpha(t_q^s)) = 0. \quad (5.9)$$

A présent nous calculons la variation première de $\mathcal{A}_{\underline{s}}^s$ en (I_q, Γ_q) :

$$\begin{aligned} & \mathcal{A}_{\underline{s}}^s([t_q^s, t_q^s + \tau(\varepsilon)], \Gamma_q + \varepsilon\Phi) \\ &= \int_{t_q^s}^{t_q^s + \tau(\varepsilon)} L(t, \gamma_q(t) + \varepsilon\varphi(t), \dot{\gamma}_q(t) + \varepsilon\dot{\varphi}(t)) dt \\ &= \int_{t_q^s}^{t_q^s} L(t, \gamma_q(t), \dot{\gamma}_q(t)) dt + \varepsilon\tau'(0)L(t, x_q^s, \dot{\gamma}_q(t_q^s)) \\ & \quad + \varepsilon \int_{t_q^s}^{t_q^s} \left(\frac{\partial L}{\partial x^\alpha}(t, \gamma_q(t), \dot{\gamma}_q(t))\varphi^\alpha(t) + \frac{\partial L}{\partial v^\alpha}(t, \gamma_q(t), \dot{\gamma}_q(t))\dot{\varphi}^\alpha(t) \right) dt + o(\varepsilon) \\ &= \mathcal{A}_{\underline{s}}^s([t_q^s, t_q^s], \Gamma_q) + \varepsilon\tau'(0)L(t, x_q^s, \dot{\gamma}_q(t_q^s)) \\ & \quad + \varepsilon \left[\frac{\partial L}{\partial v^\alpha}(t, \gamma_q(t), \dot{\gamma}_q(t))\varphi^\alpha(t) \right]_{t_q^s}^{t_q^s} \\ & \quad + \varepsilon \int_{t_q^s}^{t_q^s} \left(\frac{\partial L}{\partial x^\alpha}(t, \gamma_q(t), \dot{\gamma}_q(t)) - \frac{d}{dt} \left(\frac{\partial L}{\partial v^\alpha}(t, \gamma_q(t), \dot{\gamma}_q(t)) \right) \right) \varphi^\alpha(t) dt + o(\varepsilon). \end{aligned}$$

Pour exploiter cette relation, nous pouvons dans un premier temps supposer que $\tau = 0$ et φ s'annule sur un voisinage de t_q^s . Alors la condition (5.9) est bien évidemment satisfaite. Tous les termes à l'ordre 1 dans le membre de droite de la dernière expression s'annulent sauf la dernière intégrale. Pour que la variation première soit nulle pour toutes les valeurs de φ , il faut donc que γ_q soit solution de l'équation d'Euler–Lagrange

$$\frac{\partial L}{\partial x^\alpha}(t, \gamma_q(t), \dot{\gamma}_q(t)) - \frac{d}{dt} \left(\frac{\partial L}{\partial v^\alpha}(t, \gamma_q(t), \dot{\gamma}_q(t)) \right) = 0.$$

Dans un deuxième temps, on utilise l'équation que nous venons d'obtenir pour simplifier la variation première de $\mathcal{A}_{\underline{s}}$ qui vaut maintenant

$$\delta \mathcal{A}_{\underline{s}} = L(t, x_q^s, \dot{\gamma}_q(t_q^s))\tau'(0) + \frac{\partial L}{\partial v^\alpha}(t_q^s, x_q^s, \dot{\gamma}_q(t_q^s))\varphi^\alpha(t_q^s).$$

Il faut que cette expression soit nulle pour toutes les valeurs de $(\tau'(0), \varphi^\alpha(t_q^s))$ satisfaisant (5.9). Cela est vrai si et seulement si il existe un réel λ (multiplicateur de Lagrange) tel que $L(t, x_q^s, \dot{\gamma}_q(t_q^s)) = \lambda \left(\frac{\partial S}{\partial t}(t_q^s, x_q^s) + \frac{\partial S}{\partial x^\alpha}(t_q^s, x_q^s)\dot{\gamma}_q^\alpha(t_q^s) \right)$ et $\frac{\partial L}{\partial v^\alpha}(t_q^s, x_q^s, \dot{\gamma}_q(t_q^s)) = \lambda \frac{\partial S}{\partial x^\alpha}(t_q^s, x_q^s)$. Inversant ces relations nous obtenons

$$\begin{cases} \frac{\partial S}{\partial x^\alpha}(t_q^s, x_q^s) &= \frac{1}{\lambda} \frac{\partial L}{\partial v^\alpha}(t_q^s, x_q^s, \dot{\gamma}_q(t_q^s)) \\ \frac{\partial S}{\partial t}(t_q^s, x_q^s) &= \frac{1}{\lambda} \left(L(t, x_q^s, \dot{\gamma}_q(t_q^s)) - \frac{\partial L}{\partial v^\alpha}(t_q^s, x_q^s, \dot{\gamma}_q(t_q^s))\dot{\gamma}_q(t_q^s) \right) \end{cases}$$

Enfin en substituant ces valeurs dans (5.8) on déduit immédiatement que $\lambda = 1$. Nous en concluons que S satisfait :

$$\begin{cases} \frac{\partial S}{\partial x^\alpha}(t_q^s, x_q^s) &= \frac{\partial L}{\partial v^\alpha}(t_q^s, x_q^s, \dot{\gamma}_q(t_q^s)) \\ \frac{\partial S}{\partial t}(t_q^s, x_q^s) &= L(t, x_q^s, \dot{\gamma}_q(t_q^s)) - \frac{\partial L}{\partial v^\alpha}(t_q^s, x_q^s, \dot{\gamma}_q(t_q^s))\dot{\gamma}_q(t_q^s) \end{cases} \quad (5.10)$$

Nous observons alors que $\frac{\partial S}{\partial x^\alpha}(t_q^s, x_q^s)$ coïncide avec la valeur de l'impulsion, image de $(t_q^s, x_q^s, \dot{\gamma}_q(t_q^s))$ par la transformée de Legendre, tandis que $-\frac{\partial S}{\partial t}(t_q^s, x_q^s)$ coïncide avec la valeur de l'hamiltonien en $(t_q^s, x_q^s, \frac{\partial L}{\partial v^\alpha}(t_q^s, x_q^s, \dot{\gamma}_q(t_q^s)))$. En éliminant $\dot{\gamma}_q(t_q^s)$, nous obtenons

$$\frac{\partial S}{\partial t}(t_q^s, x_q^s) + H \left(t_q^s, x_q^s, \frac{\partial S}{\partial x^\alpha}(t_q^s, x_q^s) \right) = 0.$$

Il faut noter que pour tout $(t, x) \in \Omega$ il existe un unique q et un unique s tels que $(t, x) = (t_q^s, x_q^s)$. Donc nous obtenons que S est solution de l'équation d'Hamilton–Jacobi sur Ω :

$$\frac{\partial S}{\partial t}(t, x) + H\left(t, x, \frac{\partial S}{\partial x^\alpha}(t, x)\right) = 0. \quad (5.11)$$

5.5 Les débuts de la « mécanique ondulatoire »

Le travail de Bohr avait permis une compréhension plus synthétique des mécanismes d'absorption et d'émission de la lumière par les atomes. Toutefois il n'expliquait pas comment et pourquoi les électrons dans l'atome devaient rester piégés dans des modes stationnaires dont l'énergie est quantifiée. Il y avait besoin d'une idée nouvelle, qui fut apportée par L. de Broglie en 1923. Commençons avec quelques considérations intuitives qui ont certainement guidé de Broglie.

Une première piste est de penser à une analogie entre l'atome et un tuyau d'orgue ou une corde de violon : les sons émis par une corde vibrante sont une superposition d'harmoniques dont les fréquences sont aussi discrétisées parce que les extrémités de la corde étant fixes, la longueur d'onde λ de chaque mode propre doit satisfaire $n\lambda = \ell$, où ℓ est la longueur de la corde. D'où l'hypothèse : *le « mouvement » de l'électron dans un état stationnaire dans l'atome a une nature ondulatoire*. Pour l'instant cette idée est encore un peu vague. Deux observations vont la rendre plus précise :

- L. de Broglie remarque la profonde analogie entre le principe de Fermat pour l'optique et le principe de Leibniz–Maupertuis pour la mécanique. Sachant que l'optique admettait deux descriptions théoriques complémentaires et toutes deux pertinentes sur le plan physique (l'*optique géométrique*, incarnée par le principe de Maupertuis, et l'*optique ondulatoire*, que l'on peut dériver du principe de Huygens), de Broglie a l'intuition que la matière elle-même est à la fois susceptible d'une description « géométrique », qui est le contenu de la mécanique de Newton et du principe de Maupertuis et une description « ondulatoire », dont la formulation était à trouver.
- La ressemblance entre la relation (5.1) pour une particule relativiste et la relation (5.2) pour une onde lumineuse : de façon réciproque à la proposition d'Einstein d'associer des particules au champ électroma-

gnétique, de Broglie suppose que l'on peut associer à toute particule d'impulsion \vec{p} une onde dont le vecteur d'onde est $\vec{k} = \vec{p}/\hbar$.

Voyons comment en partant de ces deux principes nous pouvons retrouver les conditions de quantification de Bohr. Nous suivons pour cela un raisonnement qui peut apparaître un peu biaisé : nous faisons comme si l'électron, lorsqu'il « tourne » autour du noyau de l'atome ne subissait aucune accélération⁹. Cela semble extravagant, puisque nous avons imaginé au départ que l'électron tournait autour du noyau. Mais après tout pourquoi pas, puisque N. Bohr avait déjà mis en évidence le fait que ce mouvement accéléré « échappe aux lois de la physique classique » qui prévoient qu'une particule chargée dont le mouvement est accéléré rayonne de l'énergie et est freinée. Sur la base de cette constatation, nous allons donc assimiler le mouvement de l'électron dans l'atome à celui d'une translation rectiligne uniforme dans $\mathbb{R}/\ell\mathbb{Z}$, où ℓ est la longueur de la trajectoire. Nécessairement l'onde associée a une longueur d'onde λ telle que $\ell = n\lambda$, où $n \in \mathbb{N}^*$. La règle proposée par de Broglie nous dit alors que le vecteur d'onde $k = 2\pi/\lambda$ de l'onde dépend de l'impulsion de la particule p par la relation $p = \hbar k = h/\lambda$. Nous en déduisons que

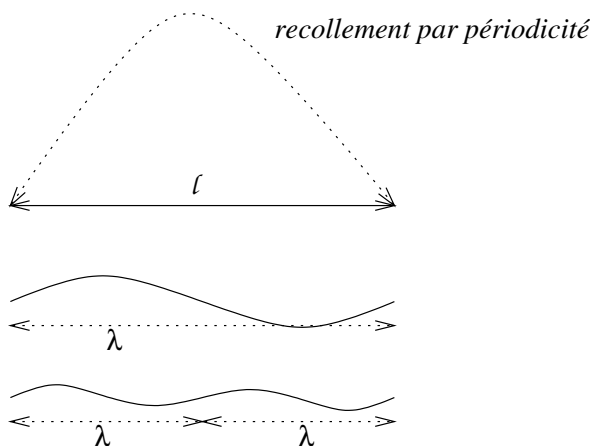


FIG. 5.7 – Représentation schématique de la condition de quantification de de Broglie.

⁹de Broglie n'a pas procédé ainsi et a suivi un chemin beaucoup plus long, reposant sur des considérations thermodynamiques. La présentation faite ici est donc un raccourci du travail de de Broglie.

$$\int_{\mathbb{R}/\ell\mathbb{Z}} pdq = \int_{\mathbb{R}/\ell\mathbb{Z}} \frac{h}{\lambda} dq = \frac{h\ell}{\lambda} = nh.$$

Nous retrouvons ainsi la condition de quantification de Bohr–Sommerfeld.

Conclusion — L. de Broglie a supposé que :

A toute particule d'impulsion \vec{p} et d'énergie E est associée une onde de vecteur d'onde $\vec{k} = \vec{p}/\hbar$ et d'impulsion $\omega = E/\hbar$.

Et cette hypothèse permet d'expliquer et retrouver par un mécanisme ondulatoire les conditions de quantification de Bohr.

Nous allons revisiter cet argument de façon un peu plus précise en suivant une approche inspirée par celle développée ultérieurement par E. Schrödinger, qui utilise l'équation d'Hamilton–Jacobi. Nous adopterons à nouveau un point de vue abusif, dans lequel nous devons à un certain moment « ignorer » l'accélération que subit l'électron. Rappelons que l'équation d'Hamilton–Jacobi nous enseigne comment associer une fonction S sur l'espace-temps (dont les hypersurfaces de niveau décrivent des fronts d'onde) à un faisceau de trajectoires de particules virtuelles (satisfaisant certaines conditions de compatibilité).

Imaginons donc un faisceau de trajectoires périodiques circulaires tournant autour du noyau de l'atome d'hydrogène, associé à une solution S de l'équation d'Hamilton–Jacobi. En vérité une telle solution n'existe pas (à cause des différents rayons de courbure des trajectoires, qui interdisent l'intégrabilité de dS), mais à nouveau nous choisissons de passer outre cette incohérence, arguant que, puisque l'électron ignore l'accélération et donc la courbure de sa trajectoire, sa trajectoire est assimilable à une ligne droite¹⁰. A priori S n'a pas de caractère oscillant (c'est plutôt une fonction monotone du temps et de l'angle dans l'espace), il est donc naturel de supposer que l'amplitude de l'onde recherchée est plutôt une quantité comme $e^{iS/A}$, où A est une constante à déterminer. Remarquons que S a les dimensions d'une action (i.e. s'exprime

¹⁰intuitivement une telle hypothèse a des connexions avec la relativité générale qui explique la gravité universelle par la courbure de l'espace-temps. Dans cette théorie une particule soumise à la force gravitationnelle n'est pas accélérée, mais se déplace selon une géodésique de l'espace-temps, l'analogue d'une ligne droite en géométrie riemannienne.

en $kg \cdot m^2 \cdot s^{-1}$) et donc a la même homogénéité que la constante de Planck h . Nous en déduisons que A pourrait s'exprimer en fonction de h .

A présent, considérons un chemin $c : [0, 1] \longrightarrow \mathbb{R}^3$ dont l'image coïncide avec une trajectoire circulaire (et tel que $c(0) = c(1)$), et une « détermination » de S le long de ce chemin (c'est à dire, un peu comme en analyse complexe, une primitive de dS le long de ce chemin). Il semble nécessaire de supposer qu'à chaque instant t , $e^{iS/A}$ soit continue et qu'en particulier $e^{iS(t,c(0))/A} = e^{iS(t,c(1))/A}$. Cela implique que $S(t, c(1)) - S(t, c(0))$ soit de la forme $2\pi An$, $n \in \mathbb{N}^*$.

Cette condition est intéressante car la relation (5.10) que nous avons obtenue sur S signifie essentiellement que $dS = p_\alpha dq^\alpha - Hdt$. La condition obtenue sur $S(t, c(1)) - S(t, c(0))$ peut donc s'écrire

$$2\pi An = S(t, c(1)) - S(t, c(0)) = \int_{c([0,1])} p_\alpha dq^\alpha - Hdt = \int_{c([0,1])} p_\alpha dq^\alpha, \quad \forall t.$$

Cette relation est identique à la relation $\int p_\alpha dq^\alpha = hn$ qui caractérise les orbites stationnaires de Bohr, à condition de supposer que $A = h/2\pi = \hbar$. Nous avons ainsi obtenu heuristiquement une expression pour l'amplitude de l'onde de de Broglie : $\psi = e^{iS/\hbar}$. A nouveau cette expression est à prendre dans un sens intuitif, puisque la fonction S qui intervient ici n'existe pas en réalité, à cause de la courbure des trajectoires.

Remarques — 1) Il faut faire attention que le principe énoncé par de Broglie n'est pas tout à fait conforme à l'utilisation de l'équation d'Hamilton–Jacobi que nous avons faite : cette équation n'associe pas une onde à une trajectoire, mais à un *faisceau de trajectoires* (plus ou moins virtuelles) *satisfaisant certaines conditions de compatibilité*. L'assimilation d'une particule ou de sa trajectoire à une onde est assez trompeuse et est source de confusion en mécanique quantique. Il est possible que de Broglie lui-même se soit un peu fourvoyé dans son travail ultérieur à cause de cela¹¹.

¹¹L. de Broglie, comme Einstein, était insatisfait de l'interprétation probabiliste qui fut donnée à l'équation de Schrödinger et rechercha toute sa vie une autre interprétation de l'onde associée à la matière, dans laquelle la matière reste toujours sous forme de particule, mais est « guidée » par une onde. Ces tentatives n'ont jamais rien donné. Cependant la question posée par de Broglie de savoir si on peut éviter d'avoir recours aux probabilités reste sans doute d'actualité.

2) Le petit raisonnement heuristique que nous avons fait sur le mouvement de l'électron dans l'atome est bien évidemment assez fragile et nous verrons ultérieurement qu'il faudra l'abandonner au profit d'une description beaucoup plus réaliste. Il est néanmoins « correct » sur deux points : il explique les raies spectrales et est une étape indispensable pour la suite.

5.6 L'équation de Schrödinger

Suite au travail de de Broglie il était urgent d'identifier l'onde et de trouver une équation décrivant sa dynamique, afin de confirmer d'avantage cette hypothèse et de l'appliquer à des questions plus sophistiquées. Ce fut l'oeuvre d'Erwin Schrödinger en 1926. L'idée se situe dans le prolongement des considérations de la section précédente : nous imaginons que l'amplitude de l'onde ressemble sous certaines hypothèses (notamment si la particule est en mouvement de translation rectiligne) avec la fonction

$$\psi(t, x) := e^{iS(t,x)/\hbar},$$

où S est une solution des équations d'Hamilton–Jacobi. Nous pouvons déjà remarquer que, si une telle relation est satisfaite, alors

$$\bar{\psi}(t, x) \left(-i\hbar \frac{\partial \psi}{\partial x^\alpha}(t, x) \right) = \frac{\partial S}{\partial x^\alpha}(t, x) = p_\alpha(t, x), \quad (5.12)$$

la valeur de l'impulsion de la trajectoire *faisant partie du faisceau de trajectoires associé à S* qui passe au point x à l'instant t . De même,

$$\bar{\psi}(t, x) \left(i\hbar \frac{\partial \psi}{\partial t}(t, x) \right) = -\frac{\partial S}{\partial t}(t, x) = H \left(x, \frac{\partial S}{\partial x^\alpha}(t, x) \right) = E, \quad (5.13)$$

la valeur de l'énergie de la trajectoire *faisant partie du faisceau de trajectoires associé à S* qui passe au point x à l'instant t .

Nous allons chercher une loi d'évolution dynamique, c'est à dire une équation différentielle spécifiant comment varie ψ au cours du temps. Avant toute chose il convient aussi de savoir de quelle théorie « classique »¹² nous partons, c'est à dire quel est le lagrangien et donc quel est l'hamiltonien de départ.

¹²au sens « non quantique »

Exercice 13 On considère : la densité lagrangienne de Maupertuis $L_M(x, v) := m\frac{|v|^2}{2} - V(x)$ sur $T\mathbb{R}^3$, et les deux densités lagrangiennes pour le mouvement d'une particule relativiste soumise à aucune force : $L_1(\dot{\gamma}) := -mc\sqrt{\langle \dot{\gamma}, \dot{\gamma} \rangle}$ et $L_2(\dot{\gamma}) := \frac{m}{2}\langle \dot{\gamma}, \dot{\gamma} \rangle$ sur TM^4 (où $\gamma(\tau) = {}^t(t(\tau), x(\tau))$ et $\langle \dot{\gamma}, \dot{\gamma} \rangle = c^2\dot{t}^2 - |\dot{x}|^2$).

1) Démontrer que les hamiltoniens $H_M : T^*\mathbb{R}^3 \rightarrow \mathbb{R}$, $H_1 : T^*M^4 \rightarrow \mathbb{R}$ et $H_2 : T^*M^4 \rightarrow \mathbb{R}$, images respectivement de L_M , L_1 et L_2 par la transformation de Legendre, sont $H_M(q, p) = \frac{|p|^2}{2m} + V(q)$, $H_1 = 0$ et $H_2(p) = \frac{p_0^2}{2mc^2} - \frac{|p|^2}{2m}$.

2) Expliciter les équations de Hamilton–Jacobi : $\frac{\partial S}{\partial t} + H_M(x, \frac{\partial S}{\partial x^\alpha}) = 0$ (pour une fonction $S : \mathbb{R} \times \mathbb{R}^3 \rightarrow \mathbb{R}$) et $H_2(\frac{\partial S}{\partial x^\mu}) = E$ (pour une fonction $S : M^4 \rightarrow \mathbb{R}$).

3) Calculer les solutions pour les deux équations d'Hamilton–Jacobi précédentes qui sont de la forme $S(t, x) = -Et + p_1x^1 + p_2x^2 + p_3x^3$, où E et les p_α sont des constantes.

L'idée la plus naturelle qui vient à l'esprit est d'utiliser l'hamiltonien relativiste H_2 pour définir l'équation d'Hamilton–Jacobi et de chercher ensuite une équation ressemblant à celle que doit satisfaire $e^{iS(t,x)/\hbar}$ lorsque S est solution de cette équation d'Hamilton–Jacobi. Schrödinger obtint ainsi une équation linéaire relativiste sur ψ :

$$\frac{1}{c^2} \frac{\partial^2 \psi}{(\partial t)^2} - \Delta \psi + \mu^2 \psi = 0.$$

Mais, fait relativement intrigant, il ne fut pas capable d'exploiter cette équation pour, par exemple, l'appliquer à l'étude de l'atome d'hydrogène¹³.

A contre-cœur (et on le comprend) il revint alors vers une version non relativiste, en partant de l'hamiltonien

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

hérité de l'action de Maupertuis. L'équation d'Hamilton–Jacobi est alors

$$\frac{\partial S}{\partial t}(t, x) + \frac{|\nabla S|^2}{2m}(t, x) + V(x) = 0.$$

¹³Cette équation fut redécouverte plus tard par Klein et Gordon, dont elle porte le nom. C'est beaucoup plus tard que l'on comprit comment exploiter et interpréter cette équation dans le cadre de la théorie quantique des champs.

Considérons à nouveau $\psi(t, x) = e^{iS(t,x)/\hbar}$. Nous avons alors :

$$\begin{aligned} i\hbar \frac{\partial \psi}{\partial t} &= -\frac{\partial S}{\partial t} e^{iS(t,x)/\hbar} = H(x, \nabla S) \psi \\ &= \frac{|\nabla S|^2}{2m} \psi + V(x) \psi. \end{aligned} \quad (5.14)$$

D'un autre côté,

$$\begin{aligned} -\frac{\hbar^2}{2m} \Delta \psi &= -\frac{\hbar^2}{2m} \operatorname{div} \left(\frac{i}{\hbar} \nabla S e^{iS(t,x)/\hbar} \right) \\ &= -\frac{\hbar^2}{2m} \left(-\frac{|\nabla S|^2}{\hbar^2} e^{iS(t,x)/\hbar} + i \frac{\Delta S}{\hbar} e^{iS(t,x)/\hbar} \right) \\ &= \frac{|\nabla S|^2}{2m} \psi + i\hbar \frac{\Delta S}{2m} \psi. \end{aligned} \quad (5.15)$$

Comparant les deux équations, nous en déduisons que

$$i\hbar \frac{\partial \psi}{\partial t} = \left(-\frac{\hbar^2}{2m} \Delta + V(x) \right) \psi - i \frac{\hbar \Delta S}{2m} \psi. \quad (5.16)$$

L'équation (5.16) est linéaire en ψ , à l'exception du dernier terme (puisque S dépend de ψ). Rappelons que S (et donc $e^{iS/\hbar}$) n'existe pas toujours (notamment pour l'onde associée à l'électron dans l'atome) et que donc il est vain de chercher à imposer que l'équation (5.16) soit toujours satisfaite, sauf peut-être pour les particules en mouvement de translation rectiligne uniforme. Or dans ce dernier cas, il est facile de construire une solution S de l'équation d'Hamilton–Jacobi, associée à un faisceau de trajectoires parallèles, et alors S est une fonction affine de (t, x) . En particulier $\Delta S = 0$ et dans ce cas le dernier terme dans l'équation (5.16) s'annule.

Cela nous amène, suivant Schrödinger, à supposer que la fonction ψ , censée modéliser les véritables phénomènes quantiques, satisfait l'équation de Schrödinger

$$i\hbar \frac{\partial \psi}{\partial t} = \left(-\frac{\hbar^2}{2m} \Delta + V(x) \right) \psi. \quad (5.17)$$

Il n'y a pas de justification précise pour décider de supprimer le dernier terme dans (5.16), à part les considérations précédentes sur les ondes de de Broglie planes et le souci de trouver une équation différentielle relativement simple et, en particulier, linéaire (noter aussi que cela revient à négliger les termes

en $O(\hbar)$ dans les termes de droite de (5.14) et (5.15)).

In fine la justification pour ce choix est donnée a posteriori : en l'appliquant à l'étude de phénomènes physiques, elle donne des prédictions qui s'accordent remarquablement bien avec les résultats expérimentaux... tant que les effets relativistes sont négligeables (et donc que les vitesses envisagées sont petites par rapport à la vitesse de la lumière).

Conclusion — Nous avons obtenu une équation qui s'avérera être parfaitement adaptée pour décrire des phénomènes « quantiques » non relativistes (mais aussi bien évidemment les phénomènes « classique », qui prédominent lorsque la valeur de \hbar est négligeable par rapport aux autres grandeurs en jeu). Avant de l'appliquer, il reste à mettre en place un cadre mathématique pour écrire correctement l'équation et surtout... construire un cadre conceptuel pour l'interpréter physiquement. Cette dernière tâche est la plus complexe : le paradigme qui fut proposé un peu plus tard par l'« école de Copenhague » remplit totalement cette fonction, sans la moindre incohérence logique, ni la moindre contradiction avec l'expérience, mais il n'a convaincu ni Einstein, ni de Broglie, ni Schrödinger.

Références

- J.-L. Basdevant, *Cours de mécanique quantique*, Ecole Polytechnique.
H. Rund, *The Hamilton–Jacobi theory in the calculus of variations, its role in mathematics and physics*, Krieger Pub. Comp., 1973.

Chapitre 6

L'équation de Schrödinger et son interprétation

Il nous reste maintenant à interpréter l'équation de Schrödinger (5.17) et à comprendre comment l'utiliser. Notamment, en formulant les choses d'une façon un peu naïve, on peut se demander comment extraire des informations telles que la position, la vitesse ou l'énergie d'une particule « décrite » par une solution ψ de cette équation.

Une première approche, due à de Broglie, consiste à privilégier des solutions de l'équation de Schrödinger dont le comportement serait analogue à celui d'un corpuscule. L'idée est que, pour ces solutions, l'« essentiel » de la fonction d'onde ψ soit concentré dans une portion relativement petite de l'espace, où serait « localisée » la particule. De telles solutions existent et sont appelées « paquets d'onde », puisque, grâce à la transformée de Fourier, on peut les représenter comme une superposition d'ondes planes monochromatiques.

Puis nous chercherons à analyser l'équation de Schrödinger avec un point de vue plus global. Cela nécessite une petite réflexion préliminaire sur la notion de *quantité observable* (nous anticipons ici sur certaines idées proposées par Heisenberg, dont nous reparlerons plus loin) : tant que le système physique que l'on étudie est un ballon ou une voiture, les choses sont relativement claires pour nous et nous n'avons pas besoin d'un effort conceptuel important pour comprendre ce que sont la position ou la vitesse d'un tel objet. Mais si le système est un électron, nos conceptions habituelles ne suffisent plus : comme première observation, force est d'admettre que toute mesure effectuée

sur un électron est extrêmement indirecte : si l'électron est isolé, il réagit à un choc avec un photon comme une particule relativiste (effet Compton), s'il est lié à un atome (ou dans un matériau conducteur) il réagit comme un système oscillatoire, passant d'un mode stationnaire à un autre. Dans tous les cas, on observe jamais directement l'électron, mais plutôt le photon, ou du moins sa variation d'énergie et d'impulsion, c'est à dire un changement de la fréquence de la lumière diffusée et de sa direction.

L'exemple précédemment évoqué (une mesure sur l'état d'un électron par le biais d'interactions avec un photon) illustre un fait général en physique : notre perception et donc notre compréhension du monde reposent sur la mesure de quantités (par exemple l'énergie et la quantité de mouvement) associées à un système (par exemple un photon) qui sont conservées si le système n'interagit pas (par exemple si le photon « voyage » dans le vide) mais qui, au cours d'une expérience spécialement conçue pour cela (par exemple la rencontre du photon avec l'électron), seront modifiées. Pourquoi nécessairement utiliser des quantités qui ont la propriété d'être conservées lorsqu'il n'y a pas d'interactions ? Précisément parce qu'une expérience doit être reproductible et donc reposer sur la mesure de quantités stables et inaltérables pour un système isolé. Ainsi une grande partie de la physique moderne repose sur l'identification des « quantités conservées », comme l'énergie, le moment cinétique, la quantité de matière ou la charge électrique.

Cela est à rapprocher d'un autre fait : les lois de la nature sont gouvernées par des *principes variationnels*. Or dans ce contexte nous savons, grâce au théorème de Noether, qu'une quantité conservée est associée à une symétrie, c'est à dire une invariance des lois de la physique par l'action d'un groupe (par exemple le groupe de Poincaré). Cela explique le rôle très important que joue la théorie des groupes en physique, incarné par la définition (en apparence abstraite) d'une particule proposée par E. Wigner : *une particule est la représentation d'une symétrie*.

6.1 Du paquet d'onde à l'interprétation probabiliste de l'équation de Schrödinger

6.1.1 Le paquet d'onde

Pour mettre nos nouvelles connaissances à l'épreuve, il nous faudrait quelques solutions explicites de cette équation ! Commençons donc par essayer de la résoudre dans le vide, c'est-à-dire en l'absence de potentiel. Nous allons chercher des solutions particulières, telles que l'état physique du système n'évolue pas au cours du temps. Autrement dit, nous cherchons une fonction $\psi(t, x)$ dont la dépendance en t ne consiste qu'à multiplier une fonction $\chi(x)$ donnée par un scalaire variable $\alpha(t)$. L'équation de Schrödinger peut alors se mettre sous la forme

$$\chi \frac{\partial \alpha}{\partial t} = \frac{i\hbar}{2m} \alpha \Delta \chi.$$

Il doit donc exister une constante ω telle que

$$\begin{cases} \frac{\partial \alpha}{\partial t} = -i\omega \alpha \\ -\frac{\hbar^2}{2m} \Delta \chi = \hbar\omega \chi. \end{cases}$$

Les solutions que nous cherchons, qui s'appellent des solutions *stationnaires*, sont donc périodiques en temps, de la forme $\psi(t, x) = \chi(x)e^{-i\omega t}$. En cherchant χ sous la forme $\chi(x^1, x^2, x^3) = \phi_1(x^1)\phi_2(x^2)\phi_3(x^3)$, un petit calcul montre que ψ se met, à normalisation près, sous la forme

$$\psi(t, x) = e^{-\frac{i}{\hbar}(\frac{p^2}{2m}t - \mathbf{p}\cdot x)},$$

où \mathbf{p} est un vecteur quelconque tel que $\mathbf{p}^2/2m = \hbar\omega$. On reconnaît dans cette expression l'onde de de Broglie associée à une particule libre d'impulsion \mathbf{p} .

Cependant cette expression ne nous satisfait pas complètement : si en effet il est facile d'associer une impulsion à une telle solution (à savoir \mathbf{p}), nous ne voyons pas du tout « où se trouve » la particule. Nous verrons plus loin que le principe d'Heisenberg fixe une limite théorique infranchissable dans la précision avec laquelle nous pouvons déterminer simultanément la position et l'impulsion d'une particule (voir l'exercice à la fin de ce paragraphe). Cela dit nous aimerions trouver des solutions un peu mieux « localisées »

dans l'espace. Par ailleurs nous verrons au prochain paragraphe des raisons théoriques (confirmées par l'interprétation de N. Bohr qui est la seule qui, jusqu'à maintenant, s'accorde avec l'expérience) pour demander à ce que la densité $\rho := |\psi|^2$ soit intégrable. Or nous voyons qu'à aucun instant la fonction d'onde ψ obtenue précédemment n'est de carré intégrable par rapport aux variables d'espace. Ce problème est similaire à celui que l'on rencontre lorsqu'on résout les équations de Maxwell dans le vide et que l'on trouve des ondes électromagnétiques planes dont l'énergie totale est infinie. Et de même qu'en électromagnétisme on est amené à considérer des superpositions d'ondes planes, nous allons exploiter le fait que l'équation de Schrödinger est linéaire et envisager la solution suivante :

$$\psi(t, x) = \int_{\mathbb{R}^3} \phi(\mathbf{p}) e^{-\frac{i}{\hbar}(Et - \mathbf{p} \cdot x)} \frac{d\mathbf{p}}{(2\pi\hbar)^{3/2}},$$

où on a naturellement posé $E = \mathbf{p}^2/2m$ et où ϕ est une fonction à déterminer.

A nouveau pour des raisons que nous verrons plus loin, la contrainte principale sur ϕ est que ψ doit être de carré sommable. Or, à quelques constantes près, et à t fixé, les fonctions ϕ et ψ sont transformées de Fourier l'une de l'autre. Comme la transformation de Fourier réalise une isométrie de $L^2(\mathbb{R}^3)$ sur lui-même, ψ est L^2 si et seulement si ϕ l'est.

Cette symétrie entre ϕ et ψ ainsi que le rapprochement avec l'onde de de Broglie que nous avons fait plus haut conduisent à postuler que la fonction ϕ joue vis-à-vis de l'impulsion de la particule le même rôle que ψ pour la position.

Si ϕ est extrêmement concentrée autour d'une certaine valeur de \mathbf{p} , on s'approche de la situation de l'onde de de Broglie et on parle de paquet d'onde localisé en impulsion. Si en revanche ψ est très concentrée autour d'une certaine valeur de x , c'est ϕ qui s'approche d'une onde plane et on parle de paquet localisé en position.

Une conséquence fondamentale de l'interprétation de ϕ comme amplitude de répartition de l'impulsion est due à une inégalité très générale liant deux fonctions qui sont transformées de Fourier l'une de l'autre. L'exercice suivant illustre ce phénomène en une dimension d'espace.

Exercice 14 Soient f et g deux fonctions de $L^2(\mathbb{R})$ qui sont transformées de Fourier l'une de l'autre. On définit deux nombres $\langle f \rangle$ et Δf (ce Δ n'a rien à voir avec le laplacien) par

$$\langle f \rangle = \int_{\mathbb{R}} x |f(x)|^2 dx \quad , \quad (\Delta f)^2 = \int_{\mathbb{R}} (x - \langle f \rangle)^2 |f(x)|^2 dx.$$

En particulier, on suppose que $\langle f \rangle$ existe. On définit de même $\langle g \rangle$ et Δg . Montrer que

$$\Delta f \Delta g \geq \frac{1}{2}.$$

On pourra commencer par supposer que $\langle f \rangle = \langle g \rangle = 0$ et étudier le signe de l'intégrale $\int_{\mathbb{R}} |xf(x) + \lambda f'(x)|^2 dx$, où λ est un paramètre réel.

Plus généralement en dimension 3, notons par exemple Δx^1 la dispersion de l'abscisse, définie par

$$(\Delta x^1)^2 = \int_{\mathbb{R}^3} (x^1 - \langle x^1 \rangle)^2 |\psi(x)|^2 dx.$$

De même, notons Δp_1 la dispersion de l'impulsion selon $\frac{\partial}{\partial x^1}$. Alors l'inégalité

$$\Delta x^1 \Delta p_1 \geq \hbar/2$$

a lieu. Autrement dit, il y a une limite de principe (et non due à un manque de capacités techniques) à la localisation simultanée en position et en impulsion d'une fonction d'onde : on ne peut préparer une particule dans un état où on connaît simultanément et exactement sa position et son impulsion. Nous reviendrons sur ce point plus tard.

6.1.2 Une loi de conservation

Revenons à l'idée exposée en introduction de privilégier les quantités conservées et recherchons comment associer de telles quantités à une solution de l'équation de Schrödinger

$$i\hbar \frac{\partial \psi}{\partial t} = \widehat{H}\psi = \left(-\frac{\hbar^2}{2m} \Delta + V(x) \right) \psi. \quad (6.1)$$

qui se conservent. Nous supposons dans un premier temps que $\psi \in \mathcal{C}^\infty(\mathbb{R} \times \mathbb{R}^3, \mathbb{C})$. Une idée est de considérer la quantité $\rho(t, x) := \overline{\psi(t, x)}\psi(t, x)$. La

dérivée par rapport au temps de ρ est $\frac{\partial \rho}{\partial t} = \bar{\psi} \frac{\partial \psi}{\partial t} + \frac{\partial \bar{\psi}}{\partial t} \psi$. Elle peut donc s'exprimer en fonction de ψ et de ses dérivées spatiales en utilisant l'équation de Schrödinger et sa conjuguée. Cela donne :

$$\begin{aligned} i\hbar \frac{\partial \rho}{\partial t} &= \bar{\psi} \left[\left(-\frac{\hbar^2}{2m} \Delta + V(x) \right) \psi \right] - \left[\left(-\frac{\hbar^2}{2m} \Delta + V(x) \right) \bar{\psi} \right] \psi \\ &= -\frac{\hbar^2}{2m} (\bar{\psi}(\Delta\psi) - (\Delta\bar{\psi})\psi) \\ &= -\frac{\hbar^2}{2m} \sum_{\alpha=1}^3 \frac{\partial}{\partial x^\alpha} \left(\bar{\psi} \frac{\partial \psi}{\partial x^\alpha} - \frac{\partial \bar{\psi}}{\partial x^\alpha} \psi \right). \end{aligned}$$

Cela entraîne que

$$\frac{\partial \rho}{\partial t} + \sum_{\alpha=1}^3 \frac{\partial j_\alpha}{\partial x^\alpha} = 0, \quad (6.2)$$

où

$$j_\alpha := \frac{\hbar}{2im} \left(\bar{\psi} \frac{\partial \psi}{\partial x^\alpha} - \frac{\partial \bar{\psi}}{\partial x^\alpha} \psi \right).$$

Une équation du type (6.2) porte le nom de *loi de conservation*. Elle signifie que la 3-forme

$$\alpha := \rho dx^1 \wedge dx^2 \wedge dx^3 - dt \wedge (j_1 dx^2 \wedge dx^3 + j_2 dx^3 \wedge dx^1 + j_3 dx^1 \wedge dx^2)$$

est fermée. Par conséquent on a $\int_{\partial U} \alpha = \int_U d\alpha = 0$ sur tout ouvert régulier $U \subset \mathbb{R} \times \mathbb{R}^3$, en vertu du théorème de Stokes.

Dans le cas où l'on suppose en plus que ψ et ses dérivées spatiales décroissent vers 0, lorsque $|x| \rightarrow \infty$, cette relation a une conséquence globale.

Proposition 2 *Soit $\psi \in C^\infty(\mathbb{R} \times \mathbb{R}^3, \mathbb{C})$ une solution de l'équation de Schrödinger (6.1). Supposons que $\forall t \in \mathbb{R}, \int_{\mathbb{R}^3} (|\psi(t, x)|^2 + |\nabla \psi(t, x)|^2) dx \leq C < \infty$. Alors la quantité*

$$P(t) := \int_{\mathbb{R}^3} \rho(t, x) dx = \int_{\mathbb{R}^3} |\psi(t, x)|^2 dx$$

est constante.

Preuve — Nous observons tout d'abord que, en notant $|j| = \sqrt{j_1^2 + j_2^2 + j_3^2}$, l'inégalité de Young implique que $\forall t \in \mathbb{R}$,

$$\begin{aligned} \int_{\mathbb{R}^3} |j(t, x)| dx &\leq \frac{\hbar^2}{2m} \int_{\mathbb{R}^3} (3|\psi(t, x)|^2 + |\nabla\psi(t, x)|^2) dx \\ &\leq C' := \frac{3\hbar^2}{2m} C. \end{aligned}$$

Soit $\chi \in \mathcal{C}^2(\mathbb{R}, \mathbb{R})$ telle que $\chi(s) = 0 \forall s \in \mathbb{R} \setminus [-2, 2]$, $\chi(s) = 1 \forall s \in [-1, 1]$ et $|\chi'(s)| \leq 2$ partout. Posons, pour $R \in]0, \infty[$ $\chi_R(x) := \chi(|x|/R)$ et $B_{2R} := \{x \in \mathbb{R}^3 / |x| < 2R\}$. Nous considérons la 3-forme $\chi_R \alpha$, où α a été défini un peu plus haut. Soit, pour $t_1 < t_2$, l'ouvert $U :=]t_1, t_2[\times B_{2R}$. Comme $\chi_R \alpha$ s'annule sur $]t_1, t_2[\times \partial B_{2R}$, on a

$$\begin{aligned} \int_{\{t_2\} \times B_{2R}} \chi_R \alpha - \int_{\{t_1\} \times B_{2R}} \chi_R \alpha &= \int_{\partial U} \chi_R \alpha \\ &= \int_U d(\chi_R \alpha) = \int_U d\chi_R \wedge \alpha. \end{aligned}$$

Cela entraîne

$$\begin{aligned} \left| \int_{\mathbb{R}^3} \chi_R(x) (\rho(t_2, x) - \rho(t_1, x)) dx \right| &= \left| \int_U d\chi_R \wedge \alpha \right| \\ &\leq \int_{]t_1, t_2[\times \mathbb{R}^3} |\nabla \chi_R(x)| |j(t, x)| dt dx \\ &\leq \frac{2}{R} \int_{]t_1, t_2[\times \mathbb{R}^3} |j(t, x)| dx \leq \frac{2C'(t_2 - t_1)}{R}. \end{aligned}$$

En faisant tendre $R \rightarrow \infty$ on déduit de cette inégalité en utilisant le théorème de convergence dominée de Lebesgue (observer que $\rho \in L^1(\mathbb{R}^3)$, puisque $\int_{\mathbb{R}^3} |\psi(t, x)|^2 dx < \infty$) que

$$\int_{\mathbb{R}^3} (\rho(t_2, x) - \rho(t_1, x)) dx = 0,$$

c'est à dire $P(t_1) = P(t_2)$. ■

Nous avons donc obtenu que, sous certaines hypothèses de « décroissance » à l'infini, la quantité $P(t)$ est constante. Si l'on se réfère aux idées discutées au début du chapitre, $P(t)$ devrait donc correspondre à une quantité

physiquement significative. Comme le résultat ne dépend pas du choix du potentiel de force V dans l'hamiltonien, la première interprétation qui vient à l'esprit est que $P(t)$ représente la « quantité de matière » ou de charge électrique. Et donc ρ serait alors une densité volumique de matière ou de charge électrique. C'est en effet l'hypothèse qui fut proposée par de Broglie et Schrödinger. Mais cette interprétation s'avéra être imparfaite car en fait, en plus de son comportement corpusculaire, la matière doit aussi maintenant manifester une nature ondulatoire.

6.1.3 Le dogme de l'« école de Copenhague »

Nous sommes en face de la question difficile de savoir si la matière est ondulatoire ou corpusculaire. En fait on ne peut pas répondre à cette question, mais simplement admettre que la matière a un comportement tantôt ondulatoire, tantôt corpusculaire. On parle de « *dualité onde-particule* ». L'interprétation proposée par N. Bohr est que $\rho(t, x)$ est une *densité de probabilité* :

Pour tout ouvert $\Omega \subset \mathbb{R}^3$ et pour tout temps $t \in \mathbb{R}$, la probabilité pour que la particule se trouve dans l'ouvert Ω à l'instant t est égale à $\int_{\Omega} |\psi(t, x)|^2 dx^1 dx^2 dx^3$.

Comme la particule doit sûrement se trouver quelque part dans l'espace, il est donc naturel de supposer que ψ est normalisé de façon à ce que $P(t) = 1$, $\forall t$. Autrement dit la norme de ψ dans $L^2(\mathbb{R}^3, \mathbb{C})$ est constante et égale à 1. Ce paradigme s'est imposé car il correspond effectivement à ce que l'on observe et toutes les expériences qui ont été faites jusqu'à présent n'ont fait que confirmer cette thèse !

Une conséquence immédiate de ce nouveau principe est que la fonction ψ , appelée *fonction d'onde* ne donne pas d'information déterministe sur la position de la particule, mais uniquement des informations probabilistes. Il s'agit d'un premier exemple d'une situation tout à fait universelle en physique quantique : *l'état physique d'une « particule », ici décrit par la fonction d'onde ψ , ne donne que des informations partielles et statistiques sur les quantités mesurables (ici la position).*

Une telle conclusion, révolutionnaire et relativement choquante, a amené les physiciens à une réflexion philosophique profonde sur la nature du monde

physique¹. Elle les a notamment forcé à faire la distinction entre d'une part *l'état d'un système physique* et d'autre part les *observations et les mesures* que l'on en fait.

Une interprétation du principe qui précède est la suivante : supposons que nous disposions d'un grand nombre N de systèmes qui sont tous dans le même état quantique $\psi \in L^2(\mathbb{R}^3, \mathbb{C})$, et que nous mesurons la position de la particule dans chaque système, nous obtenons des résultats $x_1, \dots, x_N \in \mathbb{R}^3$. La valeur moyenne est $\frac{1}{N}(x_1 + \dots + x_N)$, sa limite lorsque $N \rightarrow \infty$, $\langle x \rangle$, doit alors satisfaire :

$$\langle x \rangle = \int_{\mathbb{R}^3} x \rho(x) dx = \int_{\mathbb{R}^3} \overline{\psi(x)} x \psi(x) dx.$$

6.1.4 Le principe de superposition

Ce principe est, d'après Dirac, « l'idée fondamentale nouvelle de la mécanique quantique ». Il est mathématiquement équivalent à la linéarité de l'équation de Schrödinger : toute combinaison linéaire à coefficients complexes de fonctions d'ondes est encore une fonction d'onde. Cela signifie, étant donné deux états possibles d'un système, que toute superposition de ces états est également un état possible. Cette propriété est tout à fait nouvelle pour la

¹Einstein en particulier est resté convaincu jusqu'à la fin de sa vie qu'une théorie essentiellement non-déterministe ne pouvait être au mieux qu'une théorie provisoire. Il pensait qu'il devait exister une description plus fine du monde dans laquelle l'aléa dont la mécanique quantique rend compte apparaîtrait comme une dépendance par rapport à des degrés de libertés encore inconnus, les *variables cachées*. Avec B. Podolsky et N. Rosen, il a imaginé en 1935 une expérience, irréalisable à l'époque, mais simple dans son principe, qui lui semblait de nature à prendre en défaut l'interprétation de Copenhague. En 1966, J.S. Bell démontra qu'une certaine quantité statistique, qui pouvait être calculée à partir de mesures dans l'expérience EPR, devait, s'il existait des variables cachées, satisfaire une certaine inégalité, alors que la mécanique quantique « orthodoxe » prévoyait qu'elle serait violée. Il ne restait plus qu'à faire l'expérience... ce qu'a fait Alain Aspect au début des années 1980, à Orsay. Et il a observé que les inégalités de Bell étaient violées. Le débat semble tranché, et pourtant tous ne sont pas convaincus. Les hypothèses plus ou moins implicites sous lesquelles Bell a démontré son théorème, les conditions dans lesquelles A. Aspect a conclu que les inégalités étaient violées, tout cela prête encore à de nombreuses discussions et publications : on trouve environ 250 articles avec le mot Bell dans le titre dans la rubrique quant-ph d'arxiv pour ces cinq dernières années (Le lecteur intéressé trouvera plus de détails et une bibliographie à la page math.ucr.edu/home/baez/physics/Quantum/bells_inequality.html).

mécanique puisque les équations de Newton étaient non linéaires.

Cela pose immédiatement la question de comprendre les implications physiques de ce principe. Une situation simple est lorsque l'on suppose que deux solutions ψ_1 et ψ_2 de l'équation de Schrödinger sont telles qu'à un certain instant t les supports² de ψ_1 et de ψ_2 sont disjoints, i.e. $\text{supp}(\psi_1) \cap \text{supp}(\psi_2) = \emptyset$. Alors, pour tous coefficients $\lambda, \mu \in \mathbb{C}$,

$$\int_{\mathbb{R}^3} |\lambda\psi_1 + \mu\psi_2|^2 dx = |\lambda|^2 \int_{\mathbb{R}^3} |\psi_1|^2 dx + |\mu|^2 \int_{\mathbb{R}^3} |\psi_2|^2 dx = |\lambda|^2 + |\mu|^2.$$

En supposant que $|\lambda|^2 + |\mu|^2 = 1$, il est possible d'appliquer le postulat de Bohr et notamment d'interpréter la densité $|\lambda\psi_1 + \mu\psi_2|^2$ comme une densité de probabilité de présence de l'électron. Dans le cas particulier qui a été considéré ici cela signifie que l'on a une probabilité égale à $|\lambda|^2$ pour que l'électron se trouve dans $\text{supp}(\psi_1)$ et une probabilité égale à $|\mu|^2$ pour que l'électron se trouve dans $\text{supp}(\psi_2)$.

Dirac propose alors d'étendre la portée de ce principe. En préliminaire Dirac donne une définition pragmatique de l'état d'un système : celui-ci est déterminé lorsqu'on connaît la manière dont a été préparé le système et que cette préparation est reproductible. Pour une particule, un tel état est complètement caractérisé par une fonction d'onde. Le sens d'une superposition de deux états est alors le suivant : « un état 0 peut être formé par superposition de deux états 1 et 2 lorsqu'il existe une probabilité non nulle pour qu'un résultat quelconque, obtenu au moyen d'une observation effectuée sur le système dans l'état 0, puisse se retrouver identiquement par la même observation effectuée sur le système dans l'un (au moins) des états 1 et 2 ». Autrement dit : « ... si le système a une probabilité non nulle — disons p — de se comporter comme s'il était dans l'état 1 et une probabilité également non nulle $1 - p$ de se comporter comme s'il était dans l'état 2 ».

Donnons tout de suite la forme quantitative de ce principe. Si ψ_1 et ψ_2 sont des fonctions d'onde, c'est-à-dire des solutions de l'équation de Schrödinger, correspondant aux états 1 et 2 et si α_1 et α_2 sont des nombres complexes, alors

$$\psi_0 = \alpha_1\psi_1 + \alpha_2\psi_2$$

²le support $\text{supp}(\varphi)$ d'une fonction continue φ est le complémentaire du plus grand ouvert Ω tel que φ s'annule sur Ω

est la fonction d'onde qui décrit l'état dans lequel le système se comporte comme s'il était dans l'état 1 avec probabilité

$$p_1 = \frac{|\alpha_1|^2}{|\alpha_1|^2 + |\alpha_2|^2}$$

et comme s'il était dans l'état 2 avec probabilité

$$p_2 = 1 - p_1 = \frac{|\alpha_2|^2}{|\alpha_1|^2 + |\alpha_2|^2}.$$

D'après cette définition de la superposition, la superposition d'un état avec lui-même est encore ce même état. Autrement dit, pour tout α complexe non nul, les fonctions d'onde ψ et $\alpha\psi$ décrivent le *même* état physique.

6.2 Observables

La question suivante est de définir une notion d'impulsion ou d'énergie pour une particule décrite par une fonction d'onde, qui aurait une connexion avec la notion classique. Cela n'est pas non plus immédiat. A ce sujet il peut être utile de revenir sur la forme mathématique que revêt l'idée de déterminisme dans les descriptions classiques et quantiques d'une particule. Cette idée signifie que si l'on connaît exactement l'état d'un système isolé (ou bien même de l'univers) à un certain instant, on devrait être en mesure de déterminer toute son évolution future en utilisant les équations de la dynamique. Pour être formulée mathématiquement de façon cohérente, cette idée nécessite d'avoir identifié les équations de la dynamique (les équations de Newton ou bien de Hamilton en physique classique, l'équation de Schrödinger en mécanique quantique non relativiste), afin d'en déduire une définition de « condition initiale » satisfaisante, c'est à dire telle qu'à une condition initiale corresponde une et une seule évolution³. En physique classique, ces conditions initiales sont simplement la position et la vitesse si l'on travaille avec les équations de Newton, ou bien la position et l'impulsion si l'on travaille avec les équations de Hamilton. L'équation de Schrödinger se comporte différemment au sens où elle est du premier ordre en temps et donc il suffit de connaître la valeur initiale de ψ pour déterminer son évolution ultérieure.

³cette question rejoint par conséquent une préoccupation naturelle des mathématiciens travaillant sur les problèmes d'évolution, à savoir si un problème de Cauchy est bien posé.

Ces considérations nous conduisent à supposer que l'« impulsion quantique » de la particule doit pouvoir être déduite de la connaissance de ψ (et non de $\frac{\partial\psi}{\partial t}$). Comment ? Cela ne semble pas évident a priori, mais nous pouvons nous aider de ce que nous avons vu à propos du lien entre l'équation d'Hamilton–Jacobi et l'équation de Schrödinger : la valeur ponctuelle de l'impulsion d'un faisceau de trajectoires classiques associé à une solution S de l'équation d'Hamilton–Jacobi s'obtient à partir de la fonction $\psi = e^{iS/\hbar}$ par la relation (5.12), i.e. $p_\alpha = \overline{\psi} \left(\frac{\hbar}{i} \frac{\partial\psi}{\partial x^\alpha} \right)$. Nous devons transposer cela dans le contexte de l'interprétation quantique de l'équation de Schrödinger. De façon analogue à ce qui précède nous considérons cette fois-ci la quantité

$$\langle p_\alpha \rangle := \int_{\mathbb{R}^3} \overline{\psi(x)} \left(\frac{\hbar}{i} \frac{\partial\psi}{\partial x^\alpha}(x) \right) dx.$$

Quel sera le sens de $\langle p_\alpha \rangle$? Ce sera la valeur moyenne de la composante α de l'impulsion d'une particule dans un « état quantique » ψ . Cela signifie que si nous disposons d'un grand nombre N de systèmes qui sont tous dans l'état $\psi \in L^2(\mathbb{R}^3, \mathbb{C})$, et que nous mesurons l'impulsion de la particule dans chaque système, nous obtenons des résultats $p_1, \dots, p_N \in (\mathbb{R}^3)^*$. Alors $\lim_{N \rightarrow \infty} \frac{1}{N}(p_1 + \dots + p_N)$ est égal à $\langle p \rangle$, dont les composantes sont $\langle p_\alpha \rangle$. Cette interprétation est à nouveau totalement conforme avec l'expérience.

Ainsi la valeur $\langle p_\alpha \rangle$ doit aussi être comprise comme une valeur moyenne, à savoir l'espérance mathématique de la variable aléatoire p_α . Il faut noter qu'il est totalement impossible de déduire cette valeur de la densité de probabilité ρ comme nous l'avions fait pour la position. En particulier, et en dépit de ce que suggèrent les notations, $\langle p_\alpha \rangle$ et $\langle x^\alpha \rangle$ sont deux espérances mathématiques qui correspondent à *deux lois de probabilité différentes*. Cette propriété curieuse est à nouveau reliée au principe d'incertitude d'Heisenberg.

Un raisonnement similaire peut être fait en ce qui concerne l'énergie, à partir de la relation (5.13) : la quantité

$$\begin{aligned} \langle H \rangle &:= \int_{\mathbb{R}^3} \overline{\psi(x)} \left(\widehat{H}\psi(x) \right) dx \\ &= \int_{\mathbb{R}^3} \overline{\psi(x)} \left(-\frac{\hbar^2}{2m} \Delta\psi(x) + V(x)\psi(x) \right) dx \end{aligned}$$

sera la valeur moyenne de l'énergie.

Les trois exemples de quantités envisagés (position, impulsion, énergie) ont en commun que leur valeur moyenne est de la forme

$$\langle a \rangle := \int_{\mathbb{R}^3} \overline{\psi(x)} \left(\widehat{A}\psi(x) \right) dx,$$

où \widehat{A} est un opérateur qui agit sur les fonctions d'onde ψ :

- l'opérateur position $\widehat{x}^\alpha : \psi \mapsto [x \mapsto x^\alpha \psi(x)]$
- l'opérateur impulsion $\widehat{p}_\alpha : \psi \mapsto [x \mapsto \frac{\hbar}{i} \frac{\partial}{\partial x^\alpha} \psi(x)]$
- l'opérateur énergie ou l'hamiltonien $\widehat{H} : \psi \mapsto [x \mapsto -\frac{\hbar^2}{2m} \Delta \psi(x) + V(x)\psi(x)]$

Ces opérateurs sont, comme nous allons le voir, des opérateurs auto-adjoints hermitiens. Mais en plus il faut signaler la difficulté qu'ils ne sont pas définis sur tout $L^2(\mathbb{R}^3, \mathbb{C})$. Il est toutefois possible de leur donner un sens comme des *opérateurs non bornés* sur $L^2(\mathbb{R}^3, \mathbb{C})$.

6.2.1 Une parenthèse mathématique : opérateurs auto-adjoints hermitiens

Opérateurs auto-adjoints hermitiens bornés

Soit \mathcal{H} un espace vectoriel sur \mathbb{C} . Une forme hermitienne sur \mathcal{H} est une application

$$\langle \cdot, \cdot \rangle : \mathcal{H} \times \mathcal{H} \longrightarrow \mathbb{C},$$

satisfaisant les propriétés

- $\forall u, v, w \in \mathcal{H}, \forall \lambda, \mu \in \mathbb{C}, \langle u, \lambda v + \mu w \rangle = \lambda \langle u, v \rangle + \mu \langle u, w \rangle$
- $\forall u, v \in \mathcal{H}, \langle v, u \rangle = \overline{\langle u, v \rangle}$
- $\forall u \in \mathcal{H}, \langle u, u \rangle \geq 0$ et $\langle u, u \rangle = 0 \implies u = 0$.

Une conséquence de ces relations est que $\langle \lambda u, \mu v \rangle = \overline{\lambda} \mu \langle u, v \rangle$. Nous avons adopté ici la convention des physiciens, malheureusement inverse de celle utilisée par les mathématiciens⁴. La dernière propriété garantit le fait que $\|\cdot\| : \mathcal{H} \longrightarrow [0, \infty[$, défini par $\|u\|^2 = \langle u, u \rangle$, est une norme. Si $(\mathcal{H}, \|\cdot\|)$ est complet, nous dirons que $(\mathcal{H}, \langle \cdot, \cdot \rangle)$ est un *espace de Hilbert hermitien*.

Dans la suite nous ne considérerons que des espaces de Hilbert hermitiens qui sont soit de dimension finie, soit séparables (ce qui entraîne qu'ils admettent

⁴qui écrivent plutôt $\langle \lambda u, \mu v \rangle = \lambda \overline{\mu} \langle u, v \rangle$.

une base hilbertienne dénombrable).

Exemples — L'espace $\mathcal{H} = \mathbb{C}^n$, muni du produit hermitien $\langle u, v \rangle_{\mathbb{C}^n} := \sum_{j=1}^n \overline{u^j} v^j$, où $u = {}^t(u^1, \dots, u^n)$ et $v = {}^t(v^1, \dots, v^n)$.

L'espace $\mathcal{H} := L^2(\mathbb{R}^3, \mathbb{C})$ muni du produit hermitien

$$\langle \varphi, \psi \rangle_{L^2} := \int_{\mathbb{R}^3} \overline{\varphi(x)} \psi(x) dx.$$

Opérateur adjoint — Soit \mathcal{H}_1 et \mathcal{H}_2 deux espaces de Hilbert, \mathcal{H}_1^* et \mathcal{H}_2^* respectivement leurs duals topologiques et $T : \mathcal{H}_1 \rightarrow \mathcal{H}_2$ un opérateur linéaire continu. Son adjoint est l'opérateur $T^* : \mathcal{H}_2^* \rightarrow \mathcal{H}_1^*$ défini par

$$\forall \beta \in \mathcal{H}_2^*, \forall u \in \mathcal{H}_1, \quad (T^* \beta) u = \beta(Tu).$$

En présence du produit hermitien, l'adjoint T^* peut être représenté par un opérateur de \mathcal{H}_2 dans \mathcal{H}_1 . En effet d'après le théorème de Riesz–Fréchet il existe des isomorphismes $R_1 : \mathcal{H}_1 \rightarrow \mathcal{H}_1^*$ et $R_2 : \mathcal{H}_2 \rightarrow \mathcal{H}_2^*$ définis par

$$\forall u \in \mathcal{H}_a, \forall v \in \mathcal{H}_a, \quad (R_a u) v = \langle u, v \rangle, \quad \text{pour } a = 1 \text{ ou } 2.$$

Il découle de cette définition que l'isomorphisme de Riesz–Fréchet R_a est \mathbb{R} -linéaire et \mathbb{C} -antilinéaire, c'est à dire

$$\forall u \in \mathcal{H}_a, \forall \lambda \in \mathbb{C}, \quad R_a(\lambda u) = \overline{\lambda} R_a u.$$

Ces isomorphismes permettent alors de construire l'opérateur $T^\dagger : \mathcal{H}_2 \rightarrow \mathcal{H}_1$ par $T^\dagger := R_1^{-1} \circ T^* \circ R_2$. Par exemple si $\mathcal{H}_1 = \mathcal{H}_2 = \mathbb{C}^n$, on peut identifier T avec une matrice $n \times n$ complexe et T^\dagger s'identifie alors à la transposée de la conjuguée de T : $T^\dagger = {}^t \overline{T}$.

Définition 5 *Soit \mathcal{H} un espace de Hilbert hermitien. Un opérateur $T : \mathcal{H} \rightarrow \mathcal{H}$ est dit hermitien auto-adjoint si et seulement si $T^\dagger = T$. Cette propriété est équivalente à*

$$\forall u, v \in \mathcal{H}, \quad \langle Tu, v \rangle = \langle u, Tv \rangle.$$

Les bra- et les -cket de Dirac

Un système de notation fut introduit par P.A.M. Dirac. Il s'avère être relativement commode et est très largement utilisée par les physiciens. Etant donné un espace de Hilbert hermitien \mathcal{H} (censé représenter un espace des états physiques), si $\psi \in \mathcal{H}$, on note

- $|\psi\rangle := \psi$
- $\langle\psi| := R\psi \in \mathcal{H}^*$, où $R : \mathcal{H} \longrightarrow \mathcal{H}^*$ est l'isomorphisme de Riesz-Fréchet.

Si de plus $T : \mathcal{H} \longrightarrow \mathcal{H}$, on notera

- $T|\psi\rangle := T\psi$ l'image de ψ par T
- $\langle\psi|T := T^*(R\psi) = R(T\psi)$, l'image de $\langle\psi|$ par T^* .

Alors le produit hermitien entre φ et $\psi \in \mathcal{H}$ se note

$$\langle\varphi, \psi\rangle = (R\varphi)\psi = \langle\varphi|(|\psi\rangle) =: \langle\varphi|\psi\rangle.$$

Ne résistant pas au plaisir d'un jeu de mot⁵, Dirac appelle alors un « bra- » tout $\langle\psi| \in \mathcal{H}^*$ et un « -cket » tout $|\psi\rangle$. Un intérêt de cette notation est qu'un opérateur T est auto-adjoint hermitien si et seulement si

$$\forall|\psi\rangle \in \mathcal{H}, \forall\langle\varphi| \in \mathcal{H}^*, \quad (\langle\varphi|T)|\psi\rangle = \langle\varphi|(T|\psi\rangle),$$

qui peut s'interpréter comme une propriété d'associativité. On notera donc, pour tout opérateur T tel que $T = T^\dagger$,

$$\langle\varphi|T|\psi\rangle := (\langle\varphi|T)|\psi\rangle = \langle\varphi|(T|\psi\rangle).$$

Par exemple si $\mathcal{H} = L^2(\mathbb{R}^3, \mathbb{C})$, alors si $\varphi, \psi \in \mathcal{H}$ et si $T : \mathcal{H} \longrightarrow \mathcal{H}$ est autoadjoint, on note

$$\langle\varphi|T|\psi\rangle = \int_{\mathbb{R}^3} \overline{\varphi} T\psi \, dx.$$

En outre cette notation s'avérera être commode lorsque, dans la suite, on désignera un vecteur par des grandeurs qui le caractérisent complètement (modulo la multiplication par une constante complexe de module égal à un, qui n'a pas de sens physique). Par exemple supposons qu'un opérateur auto-adjoint hermitien (l'hamiltonien) $\widehat{H} : \mathcal{H} \longrightarrow \mathcal{H}$ soit « diagonalisable » avec des sous-espaces propres de dimension 1 sur \mathbb{C} , c'est à dire : $\mathcal{H} = \bigoplus_{n=0}^{\infty} V_n$ (somme hilbertienne), chaque V_n est une droite complexe de vecteurs propres de \widehat{H} pour la valeur propre E_n et les valeurs propres E_n sont distinctes deux à deux. Alors on notera $|E_n\rangle$ ou même de façon plus concise $|n\rangle$ un vecteur engendrant V_n , de sorte que $\widehat{H}|n\rangle = E_n|n\rangle$. Si la connaissance de la valeur propre E_n ne suffit pas, on peut ajouter d'autres nombres, correspondant à des valeurs propres d'autres opérateurs commutant avec \widehat{H} . Nous verrons cela plus en détail plus loin (à propos des Ensembles Complets d'Observables qui Commutent).

⁵le mot « crochet » se dit « bracket » en anglais

Opérateurs non bornés

Malheureusement les seuls exemples d'opérateurs auto-adjoint hermitiens que nous avons rencontrés au début de ce chapitre ne sont pas définis sur l'espace de Hilbert hermitien sur lequel il serait naturel de travailler, à savoir $L^2(\mathbb{R}^3, \mathbb{C})$. Par exemple l'opérateur hamiltonien $\widehat{H} = -\frac{\hbar^2}{2m}\Delta + V$ de l'équation de Schrödinger est un opérateur différentiel du second ordre, défini a priori sur le sous-espace $\mathcal{C}^2 \cap L^2(\mathbb{R}^3, \mathbb{C})$. Il est possible de faire un peu mieux et de donner un sens à cet opérateur sur un espace un peu plus large, l'espace de Sobolev

$$H^2(\mathbb{R}^3, \mathbb{C}) := \left\{ \psi \in L^2(\mathbb{R}^3, \mathbb{C}) / \exists \psi_\alpha, \psi_{\alpha\beta} \in L^2(\mathbb{R}^3, \mathbb{C}) (\alpha, \beta = 1, 2, 3), \right. \\ \left. \forall f \in \mathcal{C}_c^2(\mathbb{R}^3), \int_{\mathbb{R}^3} \frac{\partial f}{\partial x^\alpha} \psi + f \psi_\alpha = 0, \int_{\mathbb{R}^3} \frac{\partial f}{\partial x^\beta} \psi_\alpha + f \psi_{\alpha\beta} = 0 \right\}.$$

Les fonctions ψ_α sont appelées les *dérivées faibles* ou *dérivées au sens des distributions* de ψ , les $\psi_{\alpha\beta}$ sont les *dérivées secondes au sens des distributions*. On les note respectivement $\frac{\partial \psi}{\partial x^\alpha} := \psi_\alpha$, $\frac{\partial^2 \psi}{\partial x^\alpha \partial x^\beta} := \psi_{\alpha\beta}$, bien qu'il ne s'agisse pas de dérivées partielles au sens classique.

On peut vérifier que si $V \in \mathcal{C}^\infty(\mathbb{R}^3)$ et si V et ses dérivées premières et secondes sont uniformément bornés sur \mathbb{R}^3 , alors \widehat{H} applique continument $H^2(\mathbb{R}^3, \mathbb{C})$ sur $L^2(\mathbb{R}^3, \mathbb{C})$. Mais il n'est pas possible d'étendre cet opérateur à tout $L^2(\mathbb{R}^3, \mathbb{C})$.

Définition 6 Soit \mathcal{H}_1 et \mathcal{H}_2 deux espaces de Hilbert munis de structures hermitiennes $\langle \cdot, \cdot \rangle_1$ et $\langle \cdot, \cdot \rangle_2$, respectivement. Un opérateur non borné de \mathcal{H}_1 dans \mathcal{H}_2 est la donnée d'un sous-espace vectoriel $D(T)$ de \mathcal{H}_1 (le domaine de l'opérateur) et d'un opérateur $T : D(T) \longrightarrow \mathcal{H}_2$.

Nous ne considérerons que le cas $\mathcal{H}_1 = \mathcal{H}_2 = \mathcal{H}$ et des opérateurs non bornés T tels que :

- $D(T)$ est muni d'une structure hermitienne $\langle \cdot, \cdot \rangle_T$, a priori différente de celle de $(\mathcal{H}, \langle \cdot, \cdot \rangle)$
- $D(T)$ est fermé pour la topologie induite par $\langle \cdot, \cdot \rangle_T$
- $D(T)$ est dense dans $(\mathcal{H}, \langle \cdot, \cdot \rangle)$
- T est un « opérateur non borné borné », c'est à dire : $T : (D(T), \langle \cdot, \cdot \rangle_T) \longrightarrow (\mathcal{H}, \langle \cdot, \cdot \rangle)$ est continu.

Toutes ces conditions sont satisfaites dans $\mathcal{H} = L^2(\mathbb{R}^3, \mathbb{C})$ pour $T = \widehat{H}$ et $D(T) = H^2(\mathbb{R}^3, \mathbb{C})$ muni du produit hermitien $\langle \varphi, \psi \rangle_{H^2} := \int_{\mathbb{R}^3} \overline{\varphi} \psi + \sum_{\alpha} \frac{\partial \overline{\varphi}}{\partial x^{\alpha}} \frac{\partial \psi}{\partial x^{\alpha}} + \sum_{\alpha, \beta} \frac{\partial^2 \overline{\varphi}}{\partial x^{\alpha} \partial x^{\beta}} \frac{\partial^2 \psi}{\partial x^{\alpha} \partial x^{\beta}}$, à condition que V et ses dérivées premières et secondes soient bornées.

Deux autres exemples d'opérateurs bornés de $\mathcal{H} = L^2(\mathbb{R}^3, \mathbb{C})$ dans lui-même sont :

- l'opérateur \widehat{x}^{α} , dont le domaine est $D(\widehat{x}^{\alpha}) := \{\psi \in L^2(\mathbb{R}^3, \mathbb{C}) / \widehat{x}^{\alpha} \psi \in L^2(\mathbb{R}^3, \mathbb{C})\}$ avec le produit hermitien $\langle \varphi, \psi \rangle_{\widehat{x}^{\alpha}} := \int_{\mathbb{R}^3} \overline{\varphi} \psi + (x^{\alpha})^2 \overline{\varphi} \psi$
- l'opérateur \widehat{p}_{α} , dont le domaine est $D(\widehat{p}_{\alpha}) := \{\psi \in L^2(\mathbb{R}^3, \mathbb{C}) / \frac{\partial \psi}{\partial x^{\alpha}} \in L^2(\mathbb{R}^3, \mathbb{C})\}$ avec le produit hermitien $\langle \varphi, \psi \rangle_{\widehat{p}_{\alpha}} := \int_{\mathbb{R}^3} \overline{\varphi} \psi + \frac{\partial \overline{\varphi}}{\partial x^{\alpha}} \frac{\partial \psi}{\partial x^{\alpha}}$

Adjoint d'un opérateur non borné — Soit $T : \mathcal{H}_1 \longrightarrow \mathcal{H}_2$ un opérateur non borné entre deux espaces de Hilbert dont le domaine est $D(T)$. Nous supposons que $D(T)$ est dense dans \mathcal{H}_1 et que $T : D(T) \longrightarrow \mathcal{H}_2$ est continu. On définit :

$$D(T^*) := \{\alpha \in \mathcal{H}_2^* / \exists C > 0 \text{ tel que } |\alpha(T\psi)| \leq C \|\psi\|_1, \forall \psi \in D(T)\}.$$

Pour tout $\alpha \in D(T^*)$ nous définissons $T^*\alpha$, une forme linéaire sur $D(T)$, de la façon suivante :

$$\forall \psi \in D(T), \quad (T^*\alpha)\psi = \alpha(T\psi).$$

Il est clair que, à cause de la définition de $D(T^*)$, $T^*\alpha$ est une forme linéaire sur $D(T)$ qui est continue pour la topologie induite par $\langle \cdot, \cdot \rangle_1$. Comme $D(T)$ est dense dans \mathcal{H}_1 pour cette topologie, il existe une unique forme linéaire sur \mathcal{H}_1 qui prolonge $T^*\alpha$. Nous noterons également $T^*\alpha$ ce prolongement⁶. Alors l'opérateur

$$\begin{array}{ccc} T^* : D(T^*) & \longrightarrow & \mathcal{H}_1^* \\ \alpha & \longmapsto & T^*\alpha \end{array}$$

est l'adjoint de T .

Définition 7 Soit $(\mathcal{H}, \langle \cdot, \cdot \rangle)$ un espace de Hilbert hermitien et T un opérateur non borné de $(\mathcal{H}, \langle \cdot, \cdot \rangle)$ dans lui-même. On suppose que $D(T)$ est dense dans $(\mathcal{H}, \langle \cdot, \cdot \rangle)$ et que $T : (D(T), \langle \cdot, \cdot \rangle_{D(T)}) \longrightarrow (\mathcal{H}, \langle \cdot, \cdot \rangle)$ est continu. Alors

⁶dans le cas où $D(T)$ n'est pas dense dans \mathcal{H}_1 , il est possible de construire un prolongement (non unique) de $T^*\alpha$ en utilisant le théorème de Hahn–Banach.

on dit que T est auto-adjoint hermitien si et seulement si $D(T^*)$ est l'image de $D(T)$ par l'isomorphisme de Riesz-Fréchet $R : \mathcal{H} \longrightarrow \mathcal{H}^*$ et $T^* \circ R = R \circ T$ sur $D(T)$.

Exercice 15 On considère l'opérateur hamiltonien $\widehat{H} = -\frac{\hbar^2}{2m}\Delta + V$, où $V \in C^\infty(\mathbb{R}^3)$ est uniformément borné sur \mathbb{R}^3 , ainsi que ses dérivées premières et secondes. Son domaine est $H^2(\mathbb{R}^3, \mathbb{C}) \subset L^2(\mathbb{R}^3, \mathbb{C})$. On rappelle que $\mathcal{C}_c^\infty(\mathbb{R}^3, \mathbb{C})$, l'espace des fonctions de C^∞ de \mathbb{R}^3 vers \mathbb{C} à support compact, est dense dans $(L^2(\mathbb{R}^3, \mathbb{C}), \langle \cdot, \cdot \rangle_{L^2})$ et dans $(H^2(\mathbb{R}^3, \mathbb{C}), \langle \cdot, \cdot \rangle_{H^2})$.

1) Démontrer que $\forall \varphi \in \mathcal{C}_c^\infty(\mathbb{R}^3, \mathbb{C}), \forall \psi \in H^2(\mathbb{R}^3, \mathbb{C})$,

$$\langle \varphi, \widehat{H}\psi \rangle_{L^2} = \langle \widehat{H}\varphi, \psi \rangle_{L^2}.$$

2) En déduire que $D(\widehat{H}^*) = R_{L^2}(H^2(\mathbb{R}^3, \mathbb{C}))$ et que \widehat{H} est auto-adjoint hermitien.

3) Démontrer de même que \widehat{p}_α et \widehat{x}^α sont des opérateurs auto-adjoints hermitiens non bornés de $L^2(\mathbb{R}^3, \mathbb{C})$ dans lui-même.

6.3 Le théorème d'Ehrenfest et ses conséquences

6.3.1 Le théorème d'Ehrenfest

Nous revenons à l'équation de Schrödinger et allons rechercher à comprendre les « observables » considérées à la section 6.2 de ce chapitre. Nous allons trouver d'autres lois de conservation, ou même des relations d'un type plus général, qui permettent de prédire la variation de certaines quantités.

Nous notons ici $\mathcal{H} := L^2(\mathbb{R}^3, \mathbb{C})$ et nous considérons un opérateur linéaire auto-adjoint hermitien \widehat{A} , non borné en général, agissant sur \mathcal{H} . Nous considérons une solution $\psi \in C^\infty(\mathbb{R} \times \mathbb{R}^3, \mathbb{C})$ de l'équation de Schrödinger et nous faisons aussi l'hypothèse que $\forall t \in \mathbb{R}, \widehat{A}\psi(t, \cdot) \in C^\infty(\mathbb{R}^3, \mathbb{C})$. Nous associons à ψ et à \widehat{A} la densité

$$\rho^{\widehat{A}}(t, x) := \overline{\psi(t, x)} \widehat{A}\psi(t, x).$$

Observer qu'en général, et à la différence du cas où \widehat{A} est l'identité, cette fonction n'est pas réelle positive, mais prend ses valeurs dans \mathbb{C} . Calculons sa dérivée par rapport au temps : à partir de $\frac{\partial \rho^{\widehat{A}}}{\partial t} = \overline{\psi} \widehat{A} \frac{\partial \psi}{\partial t} + \frac{\partial \overline{\psi}}{\partial t} \widehat{A} \psi$ et de

l'équation de Schrödinger, nous obtenons

$$\begin{aligned}
i\hbar \frac{\partial \rho^{\hat{A}}}{\partial t} &= \bar{\psi} \hat{A} \left[\left(-\frac{\hbar^2}{2m} \Delta + V(x) \right) \psi \right] - \left[\left(-\frac{\hbar^2}{2m} \Delta + V(x) \right) \bar{\psi} \right] \hat{A} \psi \\
&= -\frac{\hbar^2}{2m} \left(\bar{\psi} \hat{A} (\Delta \psi) - (\Delta \bar{\psi}) \hat{A} \psi \right) + \bar{\psi} \left(\hat{A} (V \psi) - V (\hat{A} \psi) \right) \\
&= -\frac{\hbar^2}{2m} \left(\bar{\psi} \Delta (\hat{A} \psi) - (\Delta \bar{\psi}) \hat{A} \psi \right) \\
&\quad -\frac{\hbar^2}{2m} \left(\bar{\psi} \hat{A} (\Delta \psi) - \bar{\psi} \Delta (\hat{A} \psi) \right) + \bar{\psi} \left(\hat{A} (V \psi) - V (\hat{A} \psi) \right) \\
&= -\frac{\hbar^2}{2m} \sum_{\alpha=1}^3 \frac{\partial}{\partial x^\alpha} \left(\bar{\psi} \frac{\partial (\hat{A} \psi)}{\partial x^\alpha} - \frac{\partial \bar{\psi}}{\partial x^\alpha} \hat{A} \psi \right) + \bar{\psi} \left(\hat{A} \hat{H} \psi - \hat{H} \hat{A} \psi \right).
\end{aligned}$$

Cela entraîne que

$$i\hbar \left(\frac{\partial \rho^{\hat{A}}}{\partial t} + \sum_{\alpha=1}^3 \frac{\partial j_\alpha^{\hat{A}}}{\partial x^\alpha} \right) = \bar{\psi} \left([\hat{A}, \hat{H}] \psi \right), \quad (6.3)$$

où

$$j_\alpha^{\hat{A}} := \frac{\hbar}{2im} \left(\bar{\psi} \frac{\partial (\hat{A} \psi)}{\partial x^\alpha} - \frac{\partial \bar{\psi}}{\partial x^\alpha} \hat{A} \psi \right)$$

et $[\hat{A}, \hat{H}] := \hat{A} \hat{H} - \hat{H} \hat{A}$ est le commutateur des deux opérateurs.

En procédant de la même façon que pour la proposition 2 on en déduit le résultat suivant.

Proposition 3 (*Théorème d'Ehrenfest*) Soit $\psi \in \mathcal{C}^\infty(\mathbb{R} \times \mathbb{R}^3, \mathbb{C})$ une solution de l'équation de Schrödinger (6.1). Supposons que $\forall t \in \mathbb{R}$, $\psi(t, \cdot)$ et $\hat{A} \psi(t, \cdot) \in H^1(\mathbb{R}^3, \mathbb{C})$. Alors

$$i\hbar \frac{d}{dt} \int_{\mathbb{R}^3} \overline{\psi(t, x)} \hat{A} \psi(t, x) dx = \int_{\mathbb{R}^3} \overline{\psi(t, x)} \left([\hat{A}, \hat{H}] \psi \right) (t, x) dx. \quad (6.4)$$

Remarque — Il est possible d'étendre ce résultat à des solutions ψ de l'équation de Schrödinger qui sont moins régulières, en supposant par exemple qu'à chaque instant t , $\psi(t, \cdot)$ prenne ses valeurs dans $\{\phi \in L^2(\mathbb{R}^3, \mathbb{C}) / \hat{A} \phi, \hat{H} \phi, \hat{A} \hat{H} \phi, \hat{H} \hat{A} \phi \in L^2(\mathbb{R}^3, \mathbb{C})\}$.

Le résultat peut alors se déduire de la proposition 6.4 en utilisant le fait que fonctions qui sont contenues dans ce sous-ensemble et qui sont \mathcal{C}^∞ sont denses dans ce sous-ensemble.

6.3.2 Interprétation et exemples d'applications

Un principe général est qu'à toute quantité observable a est associé un opérateur auto-adjoint hermitien, en général non borné \widehat{A} . Alors si, à un certain instant t , la particule est dans un état physique quantique décrit par la fonction d'onde $\psi \in D(\widehat{A}) \subset L^2(\mathbb{R}^3, \mathbb{C})$, les valeurs prises par la quantité observable a sont aléatoires et la valeur moyenne de a vaut :

$$\langle a \rangle = \int_{\mathbb{R}^3} \overline{\psi} \widehat{A} \psi \, dx.$$

Alors si l'état physique de la particule évolue en étant gouverné par l'hamiltonien H , c'est à dire si ψ est une solution de l'équation de Schrödinger $i\hbar \frac{d\psi}{dt} = \widehat{H}\psi$, l'évolution de $\langle a \rangle$ est donnée en vertu du théorème d'Ehrenfest par

$$i\hbar \frac{d\langle a \rangle}{dt} = \int_{\mathbb{R}^3} \overline{\psi} [\widehat{A}, \widehat{H}] \psi \, dx. \quad (6.5)$$

Il est particulièrement intéressant d'examiner ce que donne ce résultat lorsque \widehat{A} est \widehat{x}^α , \widehat{p}_α ou \widehat{H} .

Exercice 16 Démontrer les relations suivantes : $[\widehat{H}, \widehat{H}] = 0$,

$$[\widehat{x}^\alpha, \widehat{H}] = i\hbar \frac{\widehat{p}_\alpha}{m} \quad \text{et} \quad [\widehat{p}_\alpha, \widehat{H}] = \frac{\hbar}{i} \frac{\partial \widehat{V}}{\partial x^\alpha}.$$

En appliquant les résultats de l'exercice précédent, on obtient que

$$\frac{d\langle H \rangle}{dt} = \frac{d}{dt} \left(\int_{\mathbb{R}^3} \overline{\psi} \widehat{H} \psi \, dx \right) = 0 \quad (\text{conservation de l'énergie}),$$

$$\frac{d\langle x^\alpha \rangle}{dt} = \frac{d}{dt} \left(\int_{\mathbb{R}^3} \overline{\psi} x^\alpha \psi \, dx \right) = \int_{\mathbb{R}^3} \overline{\psi} \frac{\widehat{p}_\alpha}{m} \psi \, dx = \frac{\langle p_\alpha \rangle}{m},$$

et

$$\frac{d\langle p_\alpha \rangle}{dt} = \frac{d}{dt} \left(\int_{\mathbb{R}^3} \overline{\psi} \widehat{p}_\alpha \psi \, dx \right) = - \int_{\mathbb{R}^3} \overline{\psi} \frac{\partial V}{\partial x^\alpha} \psi \, dx = - \left\langle \frac{\partial V}{\partial x^\alpha} \right\rangle.$$

Ces relations permettent de « retrouver » les lois de Newton pour la mécanique. Mais il est encore plus intéressant de les comparer avec la formulation hamiltonienne de la mécanique, à savoir : $\frac{dH}{dt} = \{H, H\} = 0$,

$$\frac{dx^\alpha}{dt} = \frac{p_\alpha}{m} = \{H, x^\alpha\} \quad \text{et} \quad \frac{dp_\alpha}{dt} = - \frac{\partial V}{\partial x^\alpha} = \{H, p^\alpha\}.$$

Plus généralement, pour toute fonction $F \in \mathcal{C}^\infty(T^*\mathcal{N}, \mathbb{R})$ (observable classique) et toute trajectoire hamiltonienne $t \mapsto (q(t), p(t))$,

$$\frac{dF(q(t), p(t))}{dt} = \{F, H\}(q(t), p(t)), \quad (6.6)$$

équation dont l'analogie avec (6.5) est frappante.

6.3.3 Un principe de correspondance

Récapitulons les observations faites jusqu'à présent : les lois de la physique classique et de la mécanique quantique peuvent se présenter toutes les deux comme

- la donnée d'un espace des phases : l'espace \mathcal{E} de tous les états physiques e que peut adopter un système (décrivant une particule par exemple) à un certain instant. Cette description est déterministe car la connaissance de l'état physique à un certain instant détermine en principe les états physiques futurs (et passés) du système.
- une loi d'évolution : celle-ci permet justement de préciser comment déduire les états dans le futur et le passé à partir de la connaissance de l'état du système à un certain instant. Mathématiquement cette loi d'évolution est décrite par un système dynamique, c'est à dire une équation $\frac{de}{dt} = X(e(t))$, où X est un champ de vecteur défini sur l'espace des phases \mathcal{E} .
- des grandeurs observables : il s'agit de quantités a que nous sommes capables de mesurer expérimentalement. Les exemples les plus courants sont l'énergie, la position, l'impulsion d'une particule⁷. Ces grandeurs observables sont modélisées mathématiquement par des fonction ou des opérateurs sur l'espace des phases \mathcal{E} .

⁷Notons à ce sujet que trouver une définition concise et purement mathématique de cette notion est une question difficile. De plus si nous revenons aux considérations présentées à l'introduction de chapitre et si nous réfléchissons bien, nous remarquons qu'il existe des quantités plus directement observables (la longueur d'onde d'un photon incident ou diffusé par exemple) que d'autres, qui nécessitent des calculs de physique plus ou moins sophistiqués afin de pouvoir être déduites des mesures expérimentales.

Concept	Mécanique classique	Mécanique quantique
Espace des phases :	Variété symplectique $T^*\mathbb{R}^3$ ou $T^*\mathcal{N}$	Espace de Hilbert $\mathcal{H} = L^2(\mathbb{R}^3, \mathbb{C})$ ou $L^2(\mathcal{N}, \mathbb{C})$
un état :	(q, p)	$ \psi\rangle$ avec $\langle\psi \psi\rangle = 1$
Loi d'évolution :	$\frac{d}{dt}(q, p) = \xi_H(q, p)$ $H(q, p) = \frac{ p ^2}{2m} + V(q)$	$i\hbar\frac{\partial\psi}{\partial t} = \widehat{H}\psi$ $\widehat{H} = -\frac{\hbar^2}{2m}\Delta + V(x)$
Observables :	$F \in C^\infty(T^*\mathcal{N}, \mathbb{R})$	$\widehat{A} : \mathcal{H} \longrightarrow \mathcal{H},$ $\widehat{A}^\dagger = \widehat{A}$
Lien avec des mesures expérimentales :	$F(q, p) =$ valeur observée	$\langle\psi \widehat{A} \psi\rangle = \langle a \rangle =$ valeur moyenne de a
Evolution de la valeur d'une observable :	$\frac{d}{dt}(F(q, p)) = \{F, H\}(q, p)$	$i\hbar\frac{d}{dt}\langle a \rangle = \langle\psi [\widehat{A}, \widehat{H}] \psi\rangle$

Malgré ces parallèles il existe des différences profondes entre la mécanique classique et la mécanique quantique ici, notamment en ce qui concerne le lien entre état physique et quantité observable : d'une part en physique classique, pour un état physique donné, si on connaît la valeur de toutes les observables en cet état, on peut en déduire l'état physique correspondant. En mécanique quantique, on ne peut pas connaître les valeurs de $\overline{\psi\widehat{A}\psi}$ pour tous les opérateurs observables simultanément, mais seulement pour une famille d'opérateurs $\widehat{A}_1, \dots, \widehat{A}_p$ qui commutent entre eux (cf. Chapitre 7). De plus, on ne connaît au mieux ψ qu'à une phase près, puisque toutes les quantités observables sont des fonctions de ψ invariantes par les transformations $\psi \longmapsto e^{i\theta}\psi$, où θ est une constante réelle⁸. Enfin il faut se rappeler que $|\psi\rangle$ ne donne que des informations probabilistes sur les observations.

⁸On peut en fait remédier à cela en posant que l'ensemble des états physiques est le quotient de $\{|\psi\rangle \in \mathcal{H} / \langle\psi|\psi\rangle = 1\}$ par $\{e^{i\theta}/\theta \in \mathbb{R}\}$, ou encore le quotient de $\mathcal{H} \setminus \{0\}$ par \mathbb{C}^* , c'est à dire l'espace projectif complexe $P\mathcal{H}$.

Chapitre 7

Applications de la mécanique quantique

Après avoir exposé quelques uns des principes fondateurs de la mécanique quantique, nous allons maintenant essayer de « jouer » avec ces nouvelles règles et examiner la description quantique de quelques systèmes physiques simples.

7.1 L'effet tunnel

Dans ce paragraphe, nous allons montrer que l'équation de Schrödinger prévoit qu'une particule dotée d'une certaine énergie peut, avec une probabilité non nulle, franchir une barrière de potentiel plus haute que son énergie. Cet *effet tunnel*, purement quantique, est vérifié et mis en pratique quotidiennement, en particulier par son application aux semi-conducteurs.

Nous allons considérer un problème unidimensionnel, ce qui ne change rien au phénomène physique mais simplifie les calculs. Commençons par nous placer dans un domaine où le potentiel est constant, égal à V . Si nous cherchons à nouveau une solution stationnaire, sous la forme $\psi(t, x) = \psi(x)e^{-i\omega t}$, l'équation de Schrödinger s'écrit

$$-\frac{\hbar^2}{2m}\psi'' + (V - E)\psi = 0,$$

où on a posé $E = \hbar\omega$. Les solutions de cette équation sont de la forme

$$\psi(x) = Ae^{i\frac{px}{\hbar}} + Be^{-i\frac{px}{\hbar}},$$

où $p^2/2m = E - V$. Si $E > V$, c'est-à-dire, d'après l'interprétation inspirée de de Broglie, si l'énergie de la particule est supérieure au potentiel, la fonction d'onde est une superposition de deux ondes planes monochromatiques, l'une se déplaçant vers la droite, l'autre vers la gauche. En revanche, si $E < V$, p est imaginaire. Cette situation, interdite en mécanique classique, est possible en mécanique quantique : la fonction d'onde est alors une somme d'exponentielles réelles.

Lorsque le potentiel est constant sur un intervalle non borné, nous sommes à nouveau confrontés à des problèmes de normalisabilité de notre solution. Dans le cas d'exponentielles réelles, il suffit de ne conserver que celle des deux exponentielles qui reste bornée (le cas où $E < V$ sur \mathbb{R} tout entier est dénué de sens physique). Dans le cas d'exponentielles imaginaires, la meilleure façon de procéder est d'interpréter une onde plane monochromatique $e^{-\frac{i}{\hbar}(Et - px)}$ comme un flux ininterrompu de particules, se déplaçant vers la droite ou vers la gauche suivant le signe de p (resp. positif ou négatif). Ce tour de passe-passe permet de se contenter de la forme simple des solutions que nous avons écrites plus haut et d'éviter d'avoir à utiliser des paquets d'onde explicites.

Supposons maintenant que le potentiel soit défini comme suit : $V(x) = 0$

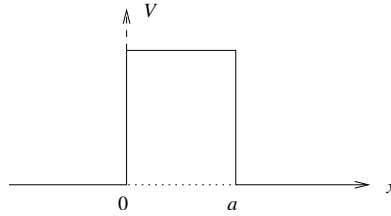


FIG. 7.1 – Un exemple de potentiel pour l'équation de Schrödinger

pour $x < 0$, $V(x) = V_0 > 0$ pour $0 < x < a$ et $V(x) = 0$ pour $x > a$. Imaginons maintenant qu'un flux de particules d'énergie $E = \hbar\omega < V_0$ et venu de $-\infty$ vient frapper cette barrière de potentiel. Classiquement, le flux est complètement réfléchi et aucune particule n'est transmise. Voyons ce que donne l'équation de Schrödinger. Pour cela, posons $k = \sqrt{2mE}/\hbar$ et $k' = \sqrt{2m(V_0 - E)}/\hbar$. La solution s'écrit $\psi(t, x) = e^{-i\omega t}\psi(x)$ avec

$$\psi(x) = \begin{cases} e^{ikx} + \alpha e^{-ikx} & x < 0 \\ \gamma e^{-k'x} + \delta e^{k'x} & 0 < x < a \\ \beta e^{ikx} & a < x. \end{cases}$$

On a normalisé à 1 le coefficient de l'onde incidente et on a éliminé d'emblée dans le domaine $a < x$ la partie de la solution correspondant à un flux issu de $+\infty$. Pour déterminer les constantes, on demande que ψ soit de classe C^1 sur \mathbb{R} . On trouve la valeur suivante pour le coefficient β :

$$\beta = \frac{4ikk'e^{-ika}}{(k + ik')^2 e^{k'a} - (k - ik')^2 e^{-k'a}}.$$

Lorsque $k'a \gg 1$, ce qui est le cas dans la plupart des applications, on a

$$|\beta| \simeq \frac{4kk'}{k^2 + k'^2} e^{-k'a}.$$

En particulier, β n'est pas nul : les particules incidentes ont une probabilité non nulle de franchir la barrière de potentiel. Par exemple, pour un flux d'électrons d'énergie 1eV arrivant sur une barrière de hauteur 2eV et de largeur 1Å, la probabilité de traverser est de 0,8 environ ! (Dans ce cas par exemple, on n'a pas $k'a \gg 1$.)

7.2 L'oscillateur harmonique

7.2.1 Première méthode

Cet exemple fondamental va nous permettre d'illustrer les principes du chapitre précédent. Il s'agit d'une particule de masse m plongée dans un potentiel unidimensionnel quadratique $V(x) = \frac{1}{2}m\omega^2 x^2$, où ω est une constante. Nous cherchons les états stationnaires, c'est-à-dire les fonctions propres du hamiltonien :

$$-\frac{\hbar^2}{2m}\psi''(x) + \frac{1}{2}m\omega^2 x^2 \psi(x) = E\psi(x).$$

Posons $\epsilon = E/(\hbar\omega)$, $\alpha = m\omega/\hbar$ et $\phi(y) = \psi(y/\sqrt{\alpha})$. Avec ces changements de variable, l'équation devient

$$\frac{1}{2} \left(y^2 - \frac{d^2}{dy^2} \right) \phi(y) = \epsilon \phi(y).$$

Exercice 17 En posant $\phi(y) = f(y)e^{-\frac{y^2}{2}}$, montrer que l'équation satisfaite par f s'écrit $f'' - 2yf' + (2\epsilon - 1)f = 0$. On peut chercher des solutions sous forme de séries entières $f(y) = \sum a_n y^n$. Montrer alors que les coefficients

satisfont la relation $a_{n+2} = -\frac{2\epsilon-1-2n}{(n+1)(n+2)}a_n$. Montrer que si f n'est pas un polynôme, ψ n'est pas de carré intégrable en montrant que $\phi(y) \sim e^{\frac{y^2}{2}}$. Montrer que f n'est un polynôme que lorsque $E = (n + 1/2)\hbar\omega$.

Cette équation n'a de solutions L^2 que lorsque $E = (n + 1/2)\hbar\omega$ pour un certain entier n . La fonction d'onde ψ_n correspondante est la fonction

$$\psi_n(x) = C_n e^{-\alpha x^2/2} H_n(x\sqrt{2\alpha}),$$

où H_n est le n -ième polynôme de Hermite. Ainsi, les énergies que peut prendre l'oscillateur forment-elles un ensemble discret, on dit que l'énergie est *quantifiée*, d'où le nom de la théorie. Remarquons aussi que l'énergie minimale de l'oscillateur est strictement positive, ce qui est à rapprocher des relations d'incertitude : une particule ne peut être parfaitement au repos en un point parfaitement déterminé.

7.2.2 Deuxième méthode

Voyons maintenant comment une approche plus algébrique permet d'obtenir les mêmes résultats. Le hamiltonien

$$\hat{H} = \frac{\hat{p}^2}{2m} + \frac{1}{2}m\omega^2\hat{x}^2$$

peut se factoriser de la façon suivante. Posons $\hat{X} = \sqrt{\frac{m\omega}{\hbar}}\hat{x}$, $\hat{P} = \frac{1}{\sqrt{m\hbar\omega}}\hat{p}$ et $\hat{H} = \hbar\omega\hat{\mathcal{H}}$. Alors

$$\hat{\mathcal{H}} = \frac{1}{2}(\hat{X}^2 + \hat{P}^2).$$

Posons $\hat{a} = \frac{1}{\sqrt{2}}(\hat{X} + i\hat{P})$ et $\hat{a}^* = \frac{1}{\sqrt{2}}(\hat{X} - i\hat{P})$. Ces deux opérateurs sont adjoints l'un de l'autre. Essayons de les utiliser pour exprimer $\hat{\mathcal{H}}$:

$$\hat{\mathcal{H}} = \frac{1}{\sqrt{2}}(\hat{X} - i\hat{P})\frac{1}{\sqrt{2}}(\hat{X} + i\hat{P}) - \frac{i}{2}[\hat{X}, \hat{P}] = \hat{a}^*\hat{a} - \frac{i}{2}[\hat{X}, \hat{P}].$$

Il y a un « reste » qu'il nous faut calculer. Tout d'abord, on vérifie que

$$[\hat{x}, \hat{p}_x] = i\hbar,$$

ce qui signifie que, pour toute fonction ψ , $x \left(\frac{\hbar}{i} \frac{\partial}{\partial x} \right) \psi - \frac{\hbar}{i} \frac{\partial}{\partial x} (x\psi) = i\hbar\psi$. On en déduit immédiatement que $[\hat{X}, \hat{P}] = i$. Ainsi, nous obtenons

$$\hat{\mathcal{H}} = \hat{a}^* \hat{a} + \frac{1}{2}.$$

Posons $\hat{\mathcal{N}} = \hat{a}^* \hat{a}$. Les opérateurs \hat{a} , \hat{a}^* et $\hat{\mathcal{N}}$ s'appellent respectivement opérateurs d'annihilation, de création et de nombre, pour une raison qui sera plus claire dans un instant.

Nous cherchons les vecteurs propres de $\hat{\mathcal{H}}$, ou, ce qui revient au même, ceux de $\hat{\mathcal{N}}$. L'opérateur $\hat{\mathcal{N}}$ est positif, car, pour tout ψ , $(\psi | \hat{\mathcal{N}} | \psi) = \| \hat{a}\psi \|^2 \geq 0$. Soit ψ un vecteur propre de $\hat{\mathcal{N}}$ pour une certaine valeur propre n . Les relations fondamentales pour ce qui suit sont les suivantes (Exercice : les vérifier !) :

$$[\hat{\mathcal{N}}, \hat{a}] = -\hat{a} \quad [\hat{\mathcal{N}}, \hat{a}^*] = \hat{a}^*.$$

En effet, ces relations nous permettent de calculer

$$\hat{\mathcal{N}}(\hat{a}\psi) = \hat{a}(\hat{\mathcal{N}}\psi) + [\hat{\mathcal{N}}, \hat{a}]\psi = (n-1)\hat{a}\psi,$$

$$\hat{\mathcal{N}}(\hat{a}^*\psi) = \hat{a}^*(\hat{\mathcal{N}}\psi) + [\hat{\mathcal{N}}, \hat{a}^*]\psi = (n+1)\hat{a}^*\psi.$$

Ces deux calculs simples sont fondamentaux. Ils indiquent que lorsque ψ est un état propre de $\hat{\mathcal{H}}$ pour la valeur propre $n + 1/2$, $\hat{a}\psi$ et $\hat{a}^*\psi$, s'ils ne sont pas nuls, sont encore vecteurs propres de $\hat{\mathcal{H}}$, pour les valeurs propres $n - 1/2$ et $n + 3/2$ respectivement. La résolution du problème aux valeurs propres est maintenant possible. Commençons par le niveau fondamental : $n = 0$. La fonction ϕ_0 doit satisfaire $\hat{a}\phi_0 = 0$, puisque -1 n'est pas une valeur propre de l'opérateur nombre. Ceci s'écrit

$$\frac{m\omega}{\hbar} x \phi_0 + \frac{\partial \phi_0}{\partial x} = 0,$$

et se résout en

$$\phi_0(x) = C_0 e^{-\frac{m\omega}{\hbar} \frac{x^2}{2}}.$$

On obtient des états excités en appliquant un nombre arbitraire de fois l'opérateur \hat{a}^* à ψ_0 . On trouve

$$\phi_n = C_n \frac{1}{\sqrt{2^n}} \left(x \sqrt{\frac{m\omega}{\hbar}} - \sqrt{\frac{\hbar}{m\omega}} \frac{d}{dx} \right)^n \phi_0.$$

Exercice 18 1) Soit $P(x)$ un polynôme. Montrer que

$$\frac{1}{\sqrt{2}} \left(x \sqrt{\frac{m\omega}{\hbar}} - \sqrt{\frac{\hbar}{m\omega}} \frac{d}{dx} \right) \left[P(x) e^{-\frac{m\omega}{\hbar} \frac{x^2}{2}} \right] = \left[\left(x \sqrt{\frac{2m\omega}{\hbar}} - \sqrt{\frac{\hbar}{2m\omega}} \frac{d}{dx} \right) P \right] (x) e^{-\frac{m\omega}{\hbar} \frac{x^2}{2}}.$$

Montrer que si l'on pose $R(x) = P \left(x \sqrt{\frac{\hbar}{2m\omega}} \right)$, alors

$$\left[\left(x \sqrt{\frac{2m\omega}{\hbar}} - \sqrt{\frac{\hbar}{2m\omega}} \frac{d}{dx} \right)^n P \right] (x) = [(x - \partial_x)^n R] \left(x \sqrt{\frac{2m\omega}{\hbar}} \right).$$

En déduire que

$$\psi_n = C_n [(x - \partial_x)^n 1] \left(x \sqrt{\frac{2m\omega}{\hbar}} \right) e^{-\frac{m\omega}{\hbar} \frac{x^2}{2}},$$

où 1 désigne le polynôme constant égal à 1.

2) On définit maintenant le produit scalaire

$$(f, g) = \int_{\mathbb{R}} f(x)g(x) e^{-\frac{x^2}{2}} \frac{dx}{\sqrt{2\pi}}.$$

Autrement dit, on considère l'espace L^2 de la mesure gaussienne standard sur \mathbb{R} . Montrer que l'adjoint formel de l'opérateur ∂_x par rapport à ce produit scalaire est l'opérateur $\partial_x^* = x - \partial_x$. On définit maintenant une suite de polynômes H_n par la relation

$$h_n(x) = (\partial_x^*)^n 1.$$

Montrer que, si $m < n$, alors les fonctions x^m et h_n sont orthogonales. En déduire que les h_n forment une famille orthogonale pour le produit scalaire (\cdot, \cdot) . Montrer que

$$\|h_n\|^2 = \int_{\mathbb{R}} h_n(x)^2 e^{-\frac{x^2}{2}} \frac{dx}{\sqrt{2\pi}} = n!.$$

On pose donc $H_n(x) = \frac{1}{\sqrt{n!}} h_n(x)$. Ces polynômes s'appellent les polynômes de Hermite et ce sont les polynômes orthogonaux de la mesure gaussienne standard.

3) Nous allons maintenant montrer que la famille $H_n(x) e^{-\frac{x^2}{2}}$ est totale dans

l'espace de Hilbert usuel $L^2(\mathbb{R}, dx)$. Tout d'abord, montrer que ceci est une conséquence du fait, à démontrer, que la famille des H_n est totale dans l'espace L^2 de la mesure gaussienne standard. Nous nous plaçons donc désormais dans cet espace $L^2(\mathbb{R}, e^{-\frac{x^2}{2}} dx / \sqrt{2\pi})$. Montrer que, pour tout t réel, la série $\sum (itx)^n / n!$ y converge vers e^{-itx} . En déduire que si une fonction $u(x)$ est orthogonale à tous les polynômes de Hermite, elle satisfait la relation

$$\int_{\mathbb{R}} u(x) e^{itx} d\gamma(x),$$

où on a défini la mesure γ par

$$d\gamma(x) = e^{-\frac{x^2}{2}} dx / \sqrt{2\pi}.$$

Comme $u \in L^2(\mathbb{R}, d\gamma)$ et $\gamma(\mathbb{R}) = 1 < \infty$, on a aussi $u \in L^1(\mathbb{R}, d\gamma)$. Ainsi, la mesure $u\gamma$ existe et, d'après la relation précédente, elle satisfait, pour tout t réel,

$$\widehat{u\gamma}(t) = 0,$$

où le chapeau indique la transformation de Fourier, c'est-à-dire la fonction caractéristique. Comme celle-ci est nulle identiquement, la mesure $u\gamma$ est la mesure nulle, donc $u = 0$ γ -p.p., donc $u = 0$.

Nous avons donc complètement identifié les fonctions propres du hamiltonien :

$$\phi_n(x) = H_n \left(\sqrt{\frac{2m\omega}{\hbar}} x \right) e^{-\frac{x^2}{2}}.$$

Il n'y en a en effet pas d'autres : nous avons démontré que la famille ϕ_0, ϕ_1, \dots était totale dans $L^2(\mathbb{R})$. Il n'y a donc plus de place pour d'autres vecteurs propres !

Ces fonctions satisfont la relation $\widehat{H}\phi_n = (n + 1/2)\hbar\omega\phi_n$. Les états stationnaires de l'oscillateur harmonique à une dimension s'écrivent donc

$$\psi_n(t, x) = e^{-i(n+\frac{1}{2})\omega t} H_n \left(\sqrt{\frac{2m\omega}{\hbar}} x \right) e^{-\frac{x^2}{2}}.$$

La forme exacte du résultat est moins importante ici que la façon dont on l'a obtenu. En particulier, cette seconde méthode illustre l'importance fondamentale des relations algébriques entre les observables et tout particulièrement leurs relations de commutation.

7.3 Ensembles complets d'observables qui commutent

Comment peut-on préparer un système dans un état bien défini ? Lorsqu'on effectue une mesure sur un système, le résultat donne une information qui peut être partielle sur le système, car une valeur propre donnée d'une observable peut être dégénérée, c'est-à-dire de multiplicité plus grande que 1. On peut savoir que la fonction d'onde se trouve dans un espace propre donné, sans toutefois la connaître à une constante près. Une mesure d'une seconde quantité peut alors affiner notre connaissance du système, à condition qu'elle ne détruise pas l'information fournie par la première : il faut que les observables en jeu commutent.

On appelle *ensemble complet d'observables qui commutent* (ECOC en abrégé) une famille $\hat{a}_1, \hat{a}_2, \dots$ (finie ou infinie) d'observables qui commutent deux à deux et telles que, si V_1, V_2, \dots sont respectivement des espaces propres de $\hat{a}_1, \hat{a}_2, \dots$, alors $\bigcap_{n \geq 1} V_n$ est au plus de dimension 1. Notons qu'il existe une analogue classique de la notion de famille d'observables qui commutent : on le retrouve en revenant au principe de correspondance énoncé à la fin du chapitre précédent. Il s'agit en effet d'une famille de fonctions F_1, F_2, \dots appartenant à $\mathcal{C}^\infty(T^*\mathcal{N})$ telles que $\forall i, j, \{F_i, F_j\} = 0$.

En termes physiques, les quantités a_1, a_2, \dots sont mesurables simultanément et leur connaissance détermine complètement l'état du système.

Par exemple, pour un oscillateur harmonique, le hamiltonien \hat{H} forme à lui seul un ECOC : la condition de commutation est vraie par défaut et les valeurs propres de \hat{H} sont simples. Nous allons voir au paragraphe suivant un exemple plus sophistiqué, où le choix d'un ECOC est particulièrement utile.

7.4 Le moment cinétique

Le moment cinétique joue un rôle important dans l'étude de problèmes mécaniques invariants par rotation. Un exemple classique est celui du problème de Képler, où l'on étudie le mouvement de deux masses ponctuelles (figurant des astres) interagissant par la force de gravitation universelle. Après

s'être ramené au mouvement d'un point mobile autour d'un point fixe¹, on remarque que les trois composantes du vecteur moment cinétique par rapport à l'origine (où se trouve le centre de masse) :

$$\vec{L} = \begin{pmatrix} L_x \\ L_y \\ L_z \end{pmatrix} := \begin{pmatrix} x \\ y \\ z \end{pmatrix} \times \begin{pmatrix} p_x \\ p_y \\ p_z \end{pmatrix}$$

sont constantes. Cela est très utile : on en déduit dans un premier temps que la trajectoire est entièrement contenue dans un plan (qui est celui passant par l'origine, orthogonal à \vec{L}). Donc il suffit d'envisager un mouvement plan : sans perte de généralité on peut, grâce à un choix judicieux du système de coordonnées, supposer que ce plan est celui contenant les axes x et y . Alors il nous reste encore une information utile, à savoir que $L_z = xp_y - yp_x$ est constant. Sachant également que l'énergie totale est conservée, il est alors possible de déterminer toutes les trajectoires (les trajectoires elliptiques, paraboliques et hyperboliques de Képler).

Si ce problème a pu être simplifié, c'est parce que, outre la relation $\{H, H\} = 0$ où H est l'hamiltonien, qui entraîne la conservation de l'énergie, on a :

$$\{\vec{L}, H\} = 0.$$

On dit que H est *en involution* avec chacune des composantes de \vec{L} . Cette relation permet en effet de retrouver le fait que les composantes du moment cinétique sont conservées.

Lorsque l'on étudie la version quantique de ce type de problème, comme par exemple l'atome d'hydrogène (mouvement de « rotation » d'un électron autour d'un proton selon le modèle de Rutherford), des relations analogues entre l'hamiltonien \hat{H} et les composantes de l'opérateur moment cinétique

$$\hat{\vec{L}} = \begin{pmatrix} \hat{L}_x \\ \hat{L}_y \\ \hat{L}_z \end{pmatrix} := \begin{pmatrix} \hat{x} \\ \hat{y} \\ \hat{z} \end{pmatrix} \times \begin{pmatrix} \hat{p}_x \\ \hat{p}_y \\ \hat{p}_z \end{pmatrix} = \begin{pmatrix} \hat{y}\hat{p}_z - \hat{z}\hat{p}_y \\ \hat{z}\hat{p}_x - \hat{x}\hat{p}_z \\ \hat{x}\hat{p}_y - \hat{y}\hat{p}_x \end{pmatrix}$$

¹par réduction dans le référentiel lié au centre de masse, voir Exercice 11 au chapitre 5, section 5.3

peuvent être obtenues et sont aussi d'une grande utilité. Il s'agit des relations

$$\left[\widehat{L}, \widehat{H} \right] = \left[\widehat{L}^2, \widehat{H} \right] = \left[\widehat{L}, \widehat{L}^2 \right] = 0,$$

où ici les crochets de Poisson ont été remplacés par des commutateurs et $\widehat{L}^2 := (\widehat{L}_x)^2 + (\widehat{L}_y)^2 + (\widehat{L}_z)^2$. En particulier nous pouvons choisir trois opérateurs qui commutent tous deux à deux, par exemple \widehat{H} , \widehat{L}^2 et $n^1 \widehat{L}_x + n^2 \widehat{L}_y + n^3 \widehat{L}_z$, où (n^1, n^2, n^3) sont les composantes d'un vecteur non nul (sans perte de généralité nous choisirons $(n^1, n^2, n^3) = (0, 0, 1)$).

Rappelons un résultat bien connu en algèbre linéaire : si A et $B \in M(n, \mathbb{C})$ sont deux matrices hermitiennes, i.e. $A^\dagger = A$ et $B^\dagger = B$, qui commutent, i.e. $[A, B] = 0$, alors elles sont diagonalisables dans une même base. Ce résultat s'étend aux opérateurs hermitiens auto-adjoints qui commutent. En admettant cela et en revenant à l'atome d'hydrogène, nous pouvons prévoir a priori qu'il existe une base hermitienne de Hilbert $(|E, \ell^{(2)}, \ell_z, \alpha\rangle)$ de $L^2(\mathbb{R}^3, \mathbb{C})$, avec

$$\begin{aligned} \widehat{H}|E, \ell^{(2)}, \ell_z, \alpha\rangle &= E|E, \ell^{(2)}, \ell_z, \alpha\rangle, \\ \widehat{L}^2|E, \ell^{(2)}, \ell_z, \alpha\rangle &= \ell^{(2)}|E, \ell^{(2)}, \ell_z, \alpha\rangle, \\ \widehat{L}_z|E, \ell^{(2)}, \ell_z, \alpha\rangle &= \ell_z|E, \ell^{(2)}, \ell_z, \alpha\rangle. \end{aligned}$$

L'indice α a été incorporé ici pour tenir compte de la possible multiplicité du sous-espace propre du triplet d'opérateurs $(\widehat{H}, \widehat{L}^2, \widehat{L}_z)$ pour le triplet de valeurs propres $(E, \ell^{(2)}, \ell_z)$. En fait nous verrons plus loin que, si nous ne tenons pas compte du spin de l'électron, chacun de ces sous-espace, s'il existe, est de dimension 1 et donc que l'indice α est superflu.

D'après les résultats de l'exercice qui suit, le simple fait que $\left[\widehat{L}, \widehat{L}^2 \right] = 0$ et²

$\left[\widehat{L}, \widehat{L} \right] = i\hbar \widehat{L}$ a les conséquences suivantes :

- les seules valeurs propres autorisées pour \widehat{L}^2 sont de la forme $\ell^{(2)} = j(j+1)\hbar^2$, où $2j \in \mathbb{N}$.
- dans le sous-espace vectoriel des fonctions $|\psi\rangle$ telles que $\widehat{L}^2|\psi\rangle = j(j+1)\hbar^2|\psi\rangle$, les seules valeurs propres autorisées pour \widehat{L}_z sont du type $\ell_z = m\hbar$, où $m = -j, -j+1, \dots, j-1, j$.

²i.e. $\left[\widehat{L}_y, \widehat{L}_z \right] = i\hbar \widehat{L}_x$, $\left[\widehat{L}_z, \widehat{L}_x \right] = i\hbar \widehat{L}_y$ et $\left[\widehat{L}_x, \widehat{L}_y \right] = i\hbar \widehat{L}_z$

En particulier nous remarquons que, si j est entier, alors toutes les valeurs de m sont aussi entières et si j est un demi-entier, alors toutes les valeurs de m sont demi-entières.

Exercice 19 On rappelle que le moment cinétique angulaire classique d'une particule, par rapport à l'origine 0 est le vecteur $J := x \times p$. On considère la version quantique de cette observable, c'est à dire l'opérateur

$$\widehat{J} := \widehat{x} \times \widehat{p} = \begin{pmatrix} \widehat{y}\widehat{p}_z - \widehat{z}\widehat{p}_y \\ \widehat{z}\widehat{p}_x - \widehat{x}\widehat{p}_z \\ \widehat{x}\widehat{p}_y - \widehat{y}\widehat{p}_x \end{pmatrix}.$$

On rappelle que $[\widehat{x}^\alpha, \widehat{p}_\beta] = i\hbar\delta_\beta^\alpha$.

1) Démontrer que $\widehat{J} \times \widehat{J} = i\hbar\widehat{J}$.

2) On pose $\widehat{J}^2 := \widehat{J}_x^2 + \widehat{J}_y^2 + \widehat{J}_z^2$ et $\widehat{J}_+ := \widehat{J}_x + i\widehat{J}_y$ et $\widehat{J}_- := \widehat{J}_x - i\widehat{J}_y$. Démontrer que $[\widehat{J}^2, \widehat{J}_+] = [\widehat{J}^2, \widehat{J}_-] = [\widehat{J}^2, \widehat{J}_z] = 0$, $[\widehat{J}_z, \widehat{J}_+] = \hbar\widehat{J}_+$, $[\widehat{J}_z, \widehat{J}_-] = -\hbar\widehat{J}_-$ et $[\widehat{J}_+, \widehat{J}_-] = 2\hbar\widehat{J}_z$. Dans la suite nous allons considérer une représentation sur un espace de Hilbert complexe de ces opérateurs. On désigne par $E_{\Lambda, m}$ l'ensemble des vecteurs propres pour \widehat{J}^2 et \widehat{J}_z avec les valeurs propres $\Lambda\hbar^2$ et $m\hbar$ respectivement. On note $|\Lambda, m\rangle$ un vecteur dans $E_{\Lambda, m}$. Il satisfait donc à $\widehat{J}^2|\Lambda, m\rangle = \Lambda\hbar^2|\Lambda, m\rangle$ et $\widehat{J}_z|\Lambda, m\rangle = m\hbar|\Lambda, m\rangle$. S'il n'y a pas de vecteur propre pour (Λ, m) on convient de poser $E_{\Lambda, m} := \{0\}$.

3) Montrer que, pour que $E_{\Lambda, m} \neq \{0\}$, il est nécessaire que $\Lambda, m \in \mathbb{R}$ et $\Lambda \geq 0$.

4) Démontrer que $\widehat{J}_+|\Lambda, m\rangle \in E_{\Lambda, m+1}$ et $\widehat{J}_-|\Lambda, m\rangle \in E_{\Lambda, m-1}$.

5) Démontrer que l'adjoint de \widehat{J}_\pm est \widehat{J}_\mp et que $\widehat{J}_\mp\widehat{J}_\pm = \widehat{J}^2 - \widehat{J}_z^2 \mp \hbar\widehat{J}_z$. Utiliser ce résultat pour calculer $\|\widehat{J}_\pm|\Lambda, m\rangle\|^2$ et en déduire que, si $|\Lambda, m\rangle$ est vecteur propre (et donc en particulier non nul), alors $\Lambda - m^2 \mp m \geq 0$. Dans la suite on fixe (Λ, m) et on suppose que $|\Lambda, m\rangle \neq 0$.

6) Soit $p := \inf\{n \in \mathbb{N} / \widehat{J}_-^n|\Lambda, m\rangle = 0\}$. Démontrer que p existe dans \mathbb{N} et que $\widehat{J}_-^{p-1}|\Lambda, m\rangle \neq 0$ et $\widehat{J}_-^p|\Lambda, m\rangle = 0$. En déduire que $\Lambda = (m-p+1)^2 - (m-p+1)$.

7) Soit $q := \inf\{n \in \mathbb{N} / \widehat{J}_+^n|\Lambda, m\rangle = 0\}$. Démontrer que q existe dans \mathbb{N} et que $\widehat{J}_+^{q-1}|\Lambda, m\rangle \neq 0$ et $\widehat{J}_+^q|\Lambda, m\rangle = 0$. En déduire que $\Lambda = (m+q-1)^2 + (m+q-1)$.

8) Montrer que $m = \frac{1}{2}(p-q)$ et $\Lambda = j(j+1)$, où $j := \frac{1}{2}(p+q) - 1$.

9) En déduire que les valeurs propres autorisées pour \widehat{J}^2 sont de la forme $j(j+1)\hbar^2$, pour $j \in \frac{1}{2}\mathbb{N}$ et que, dans le sous-espace vectoriel des vecteurs propres de \widehat{J}^2 pour $j(j+1)\hbar^2$, les valeurs propres de \widehat{J}_z sont $m\hbar = -j\hbar, -(j-$

1) $\hbar, \dots, (j-1)\hbar, j\hbar^3$.

Adoptons la notation $\psi_{j,m} = |j, m\rangle$. Si nous écrivons $\widehat{J}_z = \widehat{x}\widehat{p}_y - \widehat{y}\widehat{p}_x$ en coordonnées sphériques (r, θ, φ) , nous obtenons

$$\widehat{J}_z = -i\hbar \frac{\partial}{\partial \varphi}.$$

Ainsi, $|j, m\rangle$ est-il solution de

$$-i\hbar \frac{\partial}{\partial \varphi} \psi_{j,m} = m\hbar \psi_{j,m},$$

donc

$$\psi_{j,m}(\mathbf{r}) = \Psi_{j,m}(r, \theta) e^{im\varphi}.$$

Cette fonction n'est périodique de période divisant de 2π que si m est entier. Ainsi, les cas où m est demi-entier sont-ils exclus pour décrire le moment cinétique orbital. Cela force également j à être entier.

Les opérateurs \widehat{J}_z et \widehat{J}^2 s'expriment naturellement en coordonnées sphériques. Nous avons déjà écrit \widehat{J}_z , et nous avons de plus

$$\widehat{J}^2 = -\hbar^2 \left(\frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \sin \theta \frac{\partial}{\partial \theta} + \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \varphi^2} \right).$$

Ces deux opérateurs ne dépendent donc que de φ et θ . Ainsi, les fonctions $\psi_{j,m}$ ne dépendent que de ces deux variables. Nous les noterons $Y_{j,m}(\theta, \varphi)$. Elles s'appellent les *harmoniques sphériques*. À titre d'exemple, voici une table des premières harmoniques sphériques :

$$Y_{0,0} = \frac{1}{\sqrt{4\pi}},$$

$$Y_{1,-1} = -\sqrt{\frac{3}{8\pi}} \sin \theta e^{i\varphi} \quad Y_{1,0} = \sqrt{\frac{3}{4\pi}} \cos \theta \quad Y_{1,1} = \sqrt{\frac{3}{8\pi}} \sin \theta e^{-i\varphi},$$

$$Y_{2,\pm 2} = \sqrt{\frac{15}{32\pi}} \sin^2 \theta e^{\pm 2i\varphi} \quad Y_{2,\pm 1} = \mp \sqrt{\frac{15}{8\pi}} \sin \theta \cos \theta e^{\pm i\varphi} \quad Y_{2,0} = \sqrt{\frac{5}{16\pi}} (3 \cos^2 \theta - 1).$$

Les harmoniques sphériques vont jouer un rôle très important dans la description de l'atome d'hydrogène.

³l'usage est alors de noter $|j, m\rangle$ ce que nous avons noté ici $|\Lambda, m\rangle = |j(j+1), m\rangle$.

7.5 L'atome d'hydrogène

Nous allons maintenant étudier le comportement d'un électron dans un potentiel coulombien. Vu les symétries de ce système, il est commode de passer en coordonnées sphériques. Nous utilisons l'expression suivante du laplacien :

$$\Delta = \frac{1}{r} \frac{\partial^2}{\partial r^2} r - \frac{1}{r^2 \hbar^2} \widehat{J}^2.$$

L'équation de Schrödinger stationnaire s'écrit donc

$$\left(-\frac{\hbar^2}{2m} \frac{1}{r} \frac{\partial^2}{\partial r^2} r + \frac{\widehat{J}^2}{2mr^2} + V(x) \right) \psi(x) = E\psi(x).$$

Il est important de remarquer que cet hamiltonien commute à \widehat{J}_z et à \widehat{J} . Nous affirmons que $\widehat{H}, \widehat{J}_z, \widehat{J}^2$ forment un ECOOC (*exercice* : le vérifier).

Cherchons les solutions sous la forme

$$\psi_{j,m}(x) = R_j(r) Y_{j,m}(\theta, \varphi).$$

On a $V(r) = -\frac{q^2}{4\pi\epsilon_0 r} = -\frac{e^2}{r}$, où la seconde égalité définit e . Comme $\widehat{J}^2 Y_{j,m} = j(j+1)\hbar^2 Y_{j,m}$, l'équation stationnaire devient

$$\left(-\frac{\hbar^2}{2m} \frac{1}{r} \frac{\partial^2}{\partial r^2} r + \frac{j(j+1)\hbar^2}{2mr^2} - \frac{e^2}{r} \right) R_j(r) = ER_j(r).$$

Afin de simplifier l'équation, nous allons nous placer dans un système d'unités adapté à notre problème. Tout d'abord, définissons la *constante de structure fine* sans dimension,

$$\alpha = \frac{e^2}{c\hbar} \simeq \frac{1}{137}.$$

Elle est construite comme le rapport de deux vitesses, e^2/\hbar et c . L'unité de longueur du problème est le rayon de Bohr

$$a_1 = \frac{\hbar^2}{me^2} \simeq 0,53.$$

C'est l'ordre de grandeur de la taille d'un atome. Enfin, l'unité d'énergie est

$$E_I = \frac{1}{2} m \frac{e^4}{\hbar^2} = \frac{1}{2} mc^2 \alpha^2 \simeq 13,6 \text{ eV},$$

qui est l'énergie d'ionisation de l'hydrogène.

En posant maintenant $r = \rho a_1$ et $E = -\epsilon E_I$, l'équation devient

$$\left(\frac{1}{\rho} \frac{d^2}{d\rho^2} \rho - \frac{j(j+1)}{\rho^2} + \frac{2}{\rho} - \epsilon \right) R_j(r) = 0.$$

Posons $R(\rho) = e^{-\sqrt{\epsilon}\rho} F(\rho)$. L'équation en F s'écrit alors

$$F''(\rho) + 2 \left(\frac{1}{\rho} - \sqrt{\epsilon} \right) F'(\rho) + \left(-\frac{j(j+1)}{\rho^2} + \frac{2}{\rho} (1 - \sqrt{\epsilon}) \right) F(\rho) = 0.$$

Cherchons la solution sous la forme $F(\rho) = \rho^s (c_0 + c_1 \rho + \dots)$. Lorsque ρ tend vers 0, l'équation se comporte comme

$$F''(\rho) + \frac{2}{\rho} F'(\rho) - \frac{j(j+1)}{\rho^2} F(\rho) = 0$$

et on doit avoir $s(s+1) = j(j+1)$, c'est-à-dire $s = j$ ou $s = -j - 1$. Seule la solution $s = j$ est acceptable pour des raisons de normalisabilité. Posons donc

$$F(\rho) = \rho^j f(\rho).$$

L'équation satisfaite par f s'écrit alors

$$f''(\rho) + 2 \left(\frac{j+1}{\rho} - \sqrt{\epsilon} \right) f'(\rho) + \frac{2}{\rho} (1 - (j+1)\sqrt{\epsilon}) f(\rho) = 0.$$

On cherche f sous la forme d'une série entière : $f(\rho) = \sum c_n \rho^n$. On trouve la relation de récurrence

$$c_{n+1} = -2 \frac{1 - (j+n+1)\sqrt{\epsilon}}{(n+1)(2j+n+2)} c_n.$$

De la même façon que pour l'oscillateur harmonique, cette série doit s'arrêter après un nombre fini de termes pour que la solution soit de carré intégrable. Il doit donc exister n' tel que

$$\epsilon = \frac{1}{(n' + j + 1)^2}.$$

Les polynômes $f(\rho)$ s'appellent polynômes de Laguerre. La fonction d'onde radiale s'écrit alors

$$R_j(\rho) = e^{-\sqrt{\epsilon}\rho} \rho^j (c_0 + c_1\rho + \cdots + c_{n'}\rho^{n'}).$$

On appelle $n = n' + j + 1$ le *nombre quantique principal*.

Une solution est donc déterminée par trois nombres quantiques : le nombre quantique principal $n = 1, 2, \dots$, les nombres j et m qui satisfont $0 \leq j < n$ et $-j \leq m \leq j$. L'énergie de la solution $\psi_{n,j,m}$ est $E_n = -E_I/n^2$. On a

$$\psi_{n,j,m} = Y_{j,m}(\theta, \varphi) e^{-\frac{r}{na_1}} \left(\frac{r}{a_1}\right)^j \left[c_0 + c_1 \frac{r}{a_1} + \cdots + c_{n-j-1} \left(\frac{r}{a_1}\right)^{n-j-1} \right].$$

Chapitre 8

Le point de vue de Heisenberg

8.1 La mécanique des matrices

à rédiger

8.2 Les relations d'incertitudes d'Heisenberg

Soient a et b deux observables et \hat{A} et \hat{B} les opérateurs auto-adjoints hermitiens correspondants. Rappelons que, si l'on observe un très grand nombre de systèmes physiques tous « physiquement identiques », c'est à dire décrits par le même vecteur $|\psi\rangle \in \mathcal{H}$, les valeurs moyennes de a et de b sont données par

$$\langle a \rangle = \langle \psi | \hat{A} | \psi \rangle \quad \text{et} \quad \langle b \rangle = \langle \psi | \hat{B} | \psi \rangle,$$

mais que ces deux espérances mathématiques correspondent à des lois de probabilité en général différentes. La relation d'Heisenberg précise justement cette différence, en imposant une relation entre les variances de chacune des observables, pour les lois de probabilités correspondantes. Définissons les variances :

$$\begin{aligned} (\Delta a)^2 &:= \langle (a - \langle a \rangle)^2 \rangle = \langle \psi | (\hat{A}')^2 | \psi \rangle \\ &\text{et} \\ (\Delta b)^2 &:= \langle (b - \langle b \rangle)^2 \rangle = \langle \psi | (\hat{B}')^2 | \psi \rangle, \end{aligned}$$

où $\hat{A}' := \hat{A} - \langle a \rangle$ et $\hat{B}' := \hat{B} - \langle b \rangle$.

Nous partons^{1 2} du fait que, $\forall \lambda \in \mathbb{R}$,

$$\begin{aligned}
0 &\leq \|(\widehat{A}' + i\lambda\widehat{B}')|\psi\rangle\|^2 \\
&= \langle\psi|(\widehat{A}' + i\lambda\widehat{B}'), (\widehat{A}' + i\lambda\widehat{B}')|\psi\rangle \\
&= \langle\psi|, (\widehat{A}' - i\lambda\widehat{B}')(\widehat{A}' + i\lambda\widehat{B}')|\psi\rangle \\
&= \langle\psi|(\widehat{A}')^2|\psi\rangle + \lambda^2\langle\psi|(\widehat{B}')^2|\psi\rangle + \lambda\langle\psi|i(\widehat{A}'\widehat{B}' - \widehat{B}'\widehat{A}')|\psi\rangle.
\end{aligned}$$

Donc le discriminant de ce polynome doit être négatif, c'est à dire

$$\Delta = \langle\psi|i[\widehat{A}', \widehat{B}']|\psi\rangle^2 - 4\langle\psi|(\widehat{A}')^2|\psi\rangle\langle\psi|(\widehat{B}')^2|\psi\rangle \leq 0.$$

D'où l'on déduit que

$$\langle\psi|(\widehat{A}')^2|\psi\rangle\langle\psi|(\widehat{B}')^2|\psi\rangle \geq \frac{1}{4}\langle\psi|i[\widehat{A}, \widehat{B}]|\psi\rangle^2.$$

Cela entraîne l'inégalité suivante sur les écarts-type :

$$(\Delta a)(\Delta b) \geq \frac{1}{2}|\langle\psi|, [\widehat{A}, \widehat{B}]|\psi\rangle|. \quad (8.1)$$

C'est la relation d'incertitude d'Heisenberg. Elle entraîne notamment que, si \widehat{A} et \widehat{B} ne commutent pas, alors *il n'existe pas d'état physique $|\psi\rangle$ pour lequel l'écart-type de a et celui de b soient simultanément nuls.*

L'exemple le plus important est celui des observables \widehat{x}^α et \widehat{p}_β : comme $[\widehat{x}^\alpha, \widehat{p}_\beta] = i\hbar\delta_\beta^\alpha$, on en déduit que

$$(\Delta x^\alpha)(\Delta p_\beta) \geq \frac{\hbar}{2}\delta_\beta^\alpha.$$

8.3 Reformulation des principes de la mécanique quantique

à rédiger

¹Il faut se souvenir que, pour tout opérateur U , $\langle\psi|(U|\psi\rangle) = (\langle\psi|U^\dagger)|\psi\rangle$ ou encore, en notant $\langle\psi|, U|\psi\rangle := \langle\psi|(U|\psi\rangle)$ et $\langle\psi|U, |\psi\rangle := (\langle\psi|U)|\psi\rangle$, $\langle\psi|, U|\psi\rangle = \langle\psi|U^\dagger, |\psi\rangle$. C'est pourquoi nous avons $\langle\psi|, U|\psi\rangle \neq \langle\psi|U, |\psi\rangle$ en général, sauf si $U^\dagger = U$. De plus, il est clair que, si U et V sont des opérateurs auto-adjoints hermitiens, alors $(U + iV)^\dagger = U - iV$. Enfin rappelons que si $U^\dagger = U$ on note $\langle\psi|U|\psi\rangle = \langle\psi|, U|\psi\rangle = \langle\psi|U, |\psi\rangle$.

²Nous utilisons le fait que, si U et V sont deux opérateurs auto-adjoints hermitiens, alors $i[U, V]$ est aussi auto-adjoint hermitien.

8.4 Le spin

8.4.1 Fonctions propres pour \widehat{L}_z

Revenons sur les calculs effectués à propos du moment cinétique angulaire. Nous considérons l'opérateur \widehat{L}_z agissant sur $L^2(\mathbb{R}^3, \mathbb{C}) \cap \mathcal{C}^\infty(\mathbb{R}^3 \setminus \{0\}, \mathbb{C})$ et nous recherchons les vecteurs propres $\phi \in L^2(\mathbb{R}^3, \mathbb{C}) \cap \mathcal{C}^\infty(\mathbb{R}^3 \setminus \{0\}, \mathbb{C})$ pour \widehat{L}_z avec la valeur propre ℓ_z . Cela signifie que :

$$\widehat{L}_z \phi = (\widehat{x}\widehat{p}_y - \widehat{y}\widehat{p}_x) \phi = \frac{\hbar}{i} \left(x \frac{\partial \phi}{\partial y} - y \frac{\partial \phi}{\partial x} \right) = \ell_z \phi. \quad (8.2)$$

Considérons alors la famille de fonctions $(\phi_\theta)_{\theta \in \mathbb{R}}$, définie par

$$\phi_\theta(x, y, z) := \phi(x \cos \theta - y \sin \theta, x \sin \theta + y \cos \theta, z), \quad \forall (x, y, z) \in \mathbb{R}^3 \setminus \{0\}.$$

Alors, en utilisant (8.2) on obtient, notant $x_\theta := x \cos \theta - y \sin \theta$ et $y_\theta := x \sin \theta + y \cos \theta$, que

$$\begin{aligned} \frac{d}{d\theta} (\phi_\theta(x, y, z)) &= -y_\theta \frac{\partial \phi}{\partial x}(x_\theta, y_\theta, z) + x_\theta \frac{\partial \phi}{\partial y}(x_\theta, y_\theta, z) \\ &= \left(\frac{i}{\hbar} \widehat{L}_z \phi \right) (x_\theta, y_\theta, z) \\ &= \frac{i}{\hbar} \ell_z \phi(x_\theta, y_\theta, z) \\ &= \frac{i}{\hbar} \ell_z \phi_\theta(x, y, z). \end{aligned}$$

Nous en concluons que

$$\frac{d\phi_\theta}{d\theta} = \frac{i\ell_z}{\hbar} \phi_\theta.$$

Nous pouvons intégrer cette équation différentielle ordinaire à valeurs dans $\mathcal{C}^\infty(\mathbb{R} \times (\mathbb{R}^3 \setminus \{0\}))$: nous trouvons que

$$\phi_\theta = e^{\frac{i\theta\ell_z}{\hbar}} \phi_0 = e^{\frac{i\theta\ell_z}{\hbar}} \phi.$$

Nous remarquons alors qu'il découle de la définition de ϕ_θ que nous devons avoir $\phi_{2\pi} = \phi$, ce qui n'est possible que si $\frac{2\pi\ell_z}{\hbar} \in 2\pi\mathbb{Z}$, c'est à dire

$$\ell_z = \hbar m, \quad \text{où } m \in \mathbb{Z}.$$

Nous retrouvons ainsi une condition nécessaire sur ℓ_z analogue à celle obtenue à l'exercice (19), à la différence qu'ici m est nécessairement entier et ne peut pas être un demi-entier. Donc si ϕ est en même temps un vecteur propre pour \widehat{L}^2 , cela ne peut être qu'avec une valeur propre de la forme $j(j+1)\hbar^2$, où j est un entier positif.

8.4.2 Retour sur le moment cinétique angulaire

Rappelons à nouveau les résultats de l'Exercice (19). On considère un espace de Hilbert hermitien \mathcal{H} et des opérateurs auto-adjoints hermitiens \widehat{J}_x , \widehat{J}_y et \widehat{J}_z agissant de \mathcal{H} dans lui-même. On note

$$\widehat{\vec{J}} = \begin{pmatrix} \widehat{J}_x \\ \widehat{J}_y \\ \widehat{J}_z \end{pmatrix} \quad \text{et} \quad \widehat{J}^2 := (\widehat{J}_x)^2 + (\widehat{J}_y)^2 + (\widehat{J}_z)^2$$

et on suppose que ces opérateurs satisfont les relations

$$\left[\widehat{\vec{J}}, \widehat{J}^2 \right] = 0 \quad \text{et} \quad \widehat{\vec{J}} \times \widehat{\vec{J}} = i\hbar \widehat{\vec{J}}.$$

Alors \mathcal{H} se décompose³ en une somme directe $\bigoplus_{a \in A} \mathcal{H}_a$, où chaque sous-espace \mathcal{H}_a satisfait les propriétés suivantes :

- \mathcal{H}_a est *stable* par $\widehat{\vec{J}}$, i.e. $\forall \phi \in \mathcal{H}_a$, $\widehat{J}_x \phi$, $\widehat{J}_y \phi$, $\widehat{J}_z \phi \in \mathcal{H}_a$
- \mathcal{H}_a est un sous-espace propre de \widehat{J}^2 pour une valeur propre $j_a(j_a+1)\hbar^2$, avec $2j_a \in \mathbb{N}$
- \mathcal{H}_a est *irréductible*, i.e. les seuls sous-espaces de \mathcal{H}_a stables par $\widehat{\vec{J}}$ sont \mathcal{H}_a et $\{0\}$.
- Enfin, on peut décomposer chaque sous-espace

$$\mathcal{H}_a = \bigoplus_{m=-j_a}^{j_a} \mathbb{C} |a, j_a, m\rangle,$$

³ici, soit \mathcal{H} est de dimension finie et alors il n'y a pas de difficulté particulière, soit \mathcal{H} est de dimension infinie et alors, nous devons supposer que \mathcal{H} est séparable et que les composantes de $\widehat{\vec{J}}$ sont des opérateurs compacts (cf. H. Brezis, *Analyse fonctionnelle*, Masson) pour obtenir ce résultat. Alors la somme est hilbertienne, ce qui signifie que A est un ensemble dénombrable et $\bigoplus_{a \in A} \mathcal{H}_a$ est dense dans \mathcal{H} . Dans le cas où les opérateurs ne sont pas compacts, un résultat similaire s'obtient, mais son énoncé, en terme de mesure spectrale, est plus compliqué (cf. T. Kato, *Perturbation theory for linear operators*, Springer).

où $\widehat{J}_z|a, j_a, m\rangle = m\hbar|a, j_a, m\rangle$ et⁴ $(\widehat{J}_x \pm i\widehat{J}_y)|a, j_a, m\rangle \in \mathbb{C}|a, j_a, m \pm 1\rangle$. Ce résultat permet de décomposer \mathcal{H} par blocs irréductibles, qui sont tous de dimension finie. Chaque bloc est isomorphe à un espace de la forme $\bigoplus_{m=-j}^j \mathbb{C}|j, m\rangle$, pour un certain j tel que $2j \in \mathbb{N}$ et l'action des opérateurs $\widehat{\vec{J}}$ sur la base peut être explicitée.

Observations du moment cinétique — Le moment cinétique angulaire pour une particule chargée peut s'observer en présence d'un champ magnétique : tout d'abord on définit l'opérateur *aimantation* d'une particule chargée en mouvement de rotation comme étant le vecteur

$$\widehat{\vec{\mu}} := \gamma \widehat{\vec{J}},$$

où γ est une constante réelle, qui vaut $\gamma_0 = -\frac{q}{2m_e}$ pour un électron de charge $-q$ et de masse m_e dans un potentiel central. Si ce système est soumis à un champ magnétique \vec{B} , alors l'énergie d'interaction entre la particule et le champ magnétique est

$$\widehat{H}_M = -\widehat{\vec{\mu}} \cdot \vec{B}.$$

Si par exemple \vec{B} est constant, de la forme $B(0, 0, 1)$, alors

$$\widehat{H}_M = -\widehat{\mu}_z B = -\gamma B \widehat{J}_z.$$

Imaginons à présent un atome d'hydrogène baignant dans un champ magnétique constant. Alors l'hamiltonien gouvernant le système est la somme $\mathcal{H}_H + \mathcal{H}_M$, où \mathcal{H}_H est l'hamiltonien standard pour l'atome d'hydrogène, sans champ magnétique (il contient l'énergie cinétique de l'électron et l'énergie d'interaction électrostatique avec le noyau.) Puisque les deux hamiltoniens commutent, les niveaux d'énergie de l'hamiltonien total sont de la forme $E_n - m\gamma B\hbar$, où E_n , pour $n \in \mathbb{N}$, est une des valeurs propres de \mathcal{H}_H et où $m \in \{-j, -j+1, \dots, j-1, j\}$. Ainsi les niveaux E_n se subdivisent en familles de niveaux d'énergie dont la valeur reste proche de E_n (si B est petit), mais les niveaux d'énergie restent discrets. On parle de *clivage* des niveaux d'énergie de l'atome. On peut remarquer à ce propos que, en faisant des calculs dans un cadre classique, on trouverait que l'énergie de l'atome

⁴en posant $\widehat{J}_+|a, j_a, m\rangle = \hbar\sqrt{(j_a(j_a+1) - m(m+1))}|a, j_a, m+1\rangle$ et $\widehat{J}_-|a, j_a, m\rangle = \hbar\sqrt{(j_a(j_a+1) - m(m-1))}|a, j_a, m-1\rangle$, on peut supposer sans perte de généralité que $\| |a, j_a, m\rangle \| = 1$.

serait aussi altérée par le champ magnétique, mais d'une façon continue. Les clivages peuvent être mesurés par des expériences de spectroscopie, puisque les photons émis ou absorbés par l'atome doivent avoir une énergie égale à $(E_n - m\gamma B\hbar) - (E_{n'} - m'\gamma B\hbar)$.

De tels clivages en présence d'un champ magnétique ont été observés très tôt. Donc tout concorde... sauf que les clivages pour l'hydrogène et les atomes alcalins sont en nombre *pair*. Cet effet a été constaté entre 1896 et 1903 par Pieter Zeeman et porte le nom de ce physicien.

Ou est le problème ?

8.4.3 Le spin d'un électron

On ne peut pas expliquer un clivage en un nombre pair des niveaux d'énergie par un calcul utilisant l'équation de Schrödinger comme nous avons fait au début de ce chapitre. En effet, en supposant que $\vec{B} = B(0, 0, 1)$, l'énergie d'interaction de l'électron dans l'atome d'hydrogène avec \vec{B} devrait provenir de la composante « verticale » (en z) du moment cinétique angulaire de cet électron et donc des valeurs propres de l'opérateur $\widehat{L}_z = \widehat{x}\widehat{p}_y - \widehat{y}\widehat{p}_x$, agissant sur l'« espace des phases » de l'électron, c'est à dire $L^2(\mathbb{R}^3, \mathbb{C})$. Or nous avons déjà calculé au début de ce chapitre que les valeurs propres de \widehat{L}_z sont nécessairement de la forme $m\hbar$, où $m = -j, -j + 1, \dots, j - 1, j$ est entier (et donc j aussi est un entier). Et nous voyons bien que, pour chaque valeur entière de j , il y a $2j + 1$ valeurs possibles pour m , c'est à dire un nombre impair !

Il s'agit là d'un problème qui a occupé les physiciens pendant beaucoup de temps. La solution, élaborée en 1925-26 par Wolfgang Pauli et par George-Eugène Uhlenbeck et Samuel Abraham Goudsmit, nécessite de prendre encore plus au sérieux le point de vue, inspiré des idées de Heisenberg, que nous avons vu au début de ce chapitre. A savoir : l'espace des états physiques est un espace de Hilbert dont l'essence est totalement inobservable et on n'observe que les valeurs propres d'opérateurs agissant sur cet espace de Hilbert et qui sont caractérisés par des relations de commutations. Ainsi les résultats de l'Exercice (19) sont bien plus qu'un simple exercice d'algèbre destiné à rendre plus élégants des calculs que l'on pourrait réaliser à partir des expressions explicites de l'action de \vec{J} sur $L^2(\mathbb{R}^3, \mathbb{C})$. En effet ils montrent

que les opérateurs \widehat{J}^2 et \widehat{J}_z sont a priori autorisés à admettre respectivement des valeurs propres $j(j+1)\hbar^2$ et $m\hbar$, où j et m sont soit des entiers, soit des demi-entiers. Mais alors, pour une valeur demi-entière de j , le nombre de valeurs de m compatibles est pair.

Il faut donc admettre que de tels opérateurs soient observables dans un atome d'hydrogène. Plus précisément l'effet Zeeman est explicable en admettant que l'électron est doté d'une sorte de « moment cinétique angulaire », pour lequel \widehat{J}_z ne possède que deux valeurs propres, c'est à dire correspondant à $j = \frac{1}{2}$. Ce « moment cinétique angulaire » est appelé le *spin*. Il n'existe pas d'analogue en mécanique classique de ce degré de liberté purement quantique, et en particulier le spin ne correspond pas à la rotation d'une charge dans l'espace.

Examinons un exemple d'espace de Hilbert \mathcal{H} sur lequel agit \widehat{J} avec $j = \frac{1}{2}$. Il s'agit du « bloc » non trivial le plus simple pouvant apparaître dans la décomposition d'un espace de Hilbert sur lequel agirait \widehat{J} avec $\widehat{J} \times \widehat{J} = i\hbar\widehat{J}$. (On parle de *représentation irréductible la plus simple* du groupe Spin). Cet espace est engendré par deux vecteurs $|\frac{1}{2}, -\frac{1}{2}\rangle$ (correspondant à $j = \frac{1}{2}$ et $m = -\frac{1}{2}$) et $|\frac{1}{2}, \frac{1}{2}\rangle$ (correspondant à $j = m = \frac{1}{2}$). Nous noterons, pour plus de simplicité

$$|\uparrow\rangle := |\frac{1}{2}, \frac{1}{2}\rangle \quad \text{et} \quad |\downarrow\rangle := |\frac{1}{2}, -\frac{1}{2}\rangle.$$

Rappelons aussi la notation $\widehat{J}_\pm := \widehat{J}_x \pm i\widehat{J}_y$. Alors nous avons $\widehat{J}^2|\uparrow\rangle = \frac{3\hbar^2}{4}|\uparrow\rangle$, $\widehat{J}^2|\downarrow\rangle = \frac{3\hbar^2}{4}|\downarrow\rangle$ et $\widehat{J}_z|\uparrow\rangle = \frac{\hbar}{2}|\uparrow\rangle$ et $\widehat{J}_z|\downarrow\rangle = -\frac{\hbar}{2}|\downarrow\rangle$. De plus il est possible de poser, sans perte de généralité, $\hbar|\uparrow\rangle = \widehat{J}_+|\downarrow\rangle$. Alors nous avons $\| |\downarrow\rangle \| = \| |\uparrow\rangle \|$, ce qui entraîne que nous pouvons normaliser ces deux vecteurs simultanément. Enfin ces deux vecteurs sont hermitiens orthogonaux, puisqu'ils sont vecteurs propres de \widehat{J}_z pour deux valeurs propres différentes.

Ainsi l'espace \mathcal{H} , muni de la base hermitienne $(|\uparrow\rangle, |\downarrow\rangle)$, est isomorphe à \mathbb{C}^2 . Nous pouvons donc identifier les vecteurs « états quantiques »

$$|\uparrow\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad |\downarrow\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix},$$

et alors les opérateurs \widehat{J}^2 , \widehat{J}_z et \widehat{J}_\pm s'identifient aux matrices

$$\widehat{J}^2 = \frac{3\hbar^2}{4} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \widehat{J}_z = \frac{\hbar}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

$$\widehat{J}_+ = \hbar \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad \widehat{J}_- = \hbar \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}.$$

Nous en déduisons que

$$\widehat{J}_x = \frac{\hbar}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \widehat{J}_y = \frac{\hbar}{2} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}.$$

On écrit souvent que $\widehat{J}_x = \frac{\hbar}{2}\sigma_x$, $\widehat{J}_y = \frac{\hbar}{2}\sigma_y$ et $\widehat{J}_z = \frac{\hbar}{2}\sigma_z$, où l'on a introduit les *matrices de Pauli*⁵

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad \text{et} \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Appendice : produits tensoriels

Définitions

Soient E et F deux espaces vectoriels sur un corps \mathbb{K} (qui est souvent \mathbb{R} dans les applications à la géométrie différentielle et \mathbb{C} dans les applications à la physique quantique).

Supposons dans un premier temps que E et F sont de dimension finie et considérons (e_1, \dots, e_n) , une base de E , et (f_1, \dots, f_p) , une base de F . Alors l'espace vectoriel engendré par les vecteurs notés $e_i \otimes f_j$, pour $1 \leq i \leq n$ et $1 \leq j \leq p$ est appelé *produit tensoriel* de E et F et est noté $E \otimes F$. Ainsi un vecteur $z \in E \otimes F$ se décompose de façon unique selon

$$z = \sum_{i=1}^n \sum_{j=1}^p z^{ij} e_i \otimes f_j,$$

⁵ces matrices admettent plusieurs interprétations. L'une d'elles est que l'application $\mathbb{H} \longrightarrow \mathbb{R}_+ \cdot SU(2)$, $t + ix + jy + kz \longmapsto t - i(x\sigma_x - y\sigma_y - z\sigma_z)$ est un isomorphisme de corps entre les quaternions et $\mathbb{R}_+ \cdot SU(2) := \{r \cdot a/r \in [0, \infty[, a \in SU(2)\}$.

où les coefficients z^{ij} sont dans \mathbb{K} . On remarque qu'à toute forme bilinéaire $B : E \times F \longrightarrow \mathbb{K}$ correspond une unique forme linéaire $\ell_B : E \otimes F \longrightarrow \mathbb{K}$, définie par

$$\ell_B \left(\sum_{i=1}^n \sum_{j=1}^p z^{ij} e_i \otimes f_j \right) := \sum_{i=1}^n \sum_{j=1}^p z^{ij} B(e_i, f_j).$$

Réciproquement il est facile de voir à partir de cette relation que toute forme linéaire sur $E \otimes F$ définit une unique forme bilinéaire sur $E \times F$. Cette propriété peut être utilisée pour donner une autre définition de $E \otimes F$, indépendante du choix d'une base.

Définition 8 *Etant donné deux espaces vectoriels E et F sur \mathbb{K} . Il existe un espace vectoriel sur \mathbb{K} , unique à un isomorphisme près, que nous noterons $E \otimes F$ et une application bilinéaire $P : E \times F \longrightarrow E \otimes F$, unique au même isomorphisme près, tels que pour toute application bilinéaire $B : E \times F \longrightarrow \mathbb{K}$, il existe une unique application linéaire $L : E \otimes F \longrightarrow \mathbb{K}$ telle que $B = L \circ P$.*

L'espace $E \otimes F$ est appelé produit tensoriel de E par F et on note $P(x, y) = x \otimes y, \forall x \in E, \forall y \in F$.

Remarque — Il faut prendre garde au fait que le produit tensoriel est très différent du produit cartésien. Ainsi, si $\dim E = n$ et $\dim F = p$, alors $\dim(E \times F) = n + p$ et $\dim(E \otimes F) = np$.

Exemple — Le produit tensoriel $E^* \otimes E^*$ s'identifie à l'espace vectoriel des formes bilinéaires sur E . Plus généralement, si (e_1, \dots, e_n) est une base de E et si $(\theta^1, \dots, \theta^n)$ est la base duale (de E^*), alors, on a (les signes sommes $\sum_{i,j=1}^n$ étant sous-entendus, à chaque fois qu'un indice apparaît une fois en haut et une fois en bas)

$$\begin{aligned} E \otimes E &= \{x = x^{ij} e_i \otimes e_j / x^{ij} \in \mathbb{K}\}, \\ E^* \otimes E^* &= \{x = x_{ij} \theta^i \otimes \theta^j / x_{ij} \in \mathbb{K}\}, \\ E \otimes E^* &= \{x = x^i_j e_i \otimes \theta^j / x^i_j \in \mathbb{K}\}, \\ E^* \otimes E &= \{x = x_i^j \theta^i \otimes e_j / x_i^j \in \mathbb{K}\}. \end{aligned}$$

Dans le cas où E et F sont des espaces vectoriels de dimension infinie, il est utile dans ce qui nous concerne de supposer que E et F sont des espaces de

Banach. Alors nous pouvons adapter la définition précédente en demandant que : l'application P soit continue et que les applications B et L soient continues.