

Jean-Noël Roux, Stéphane Rodts et Gabriel Stoltz

Introduction à la physique statistique et à la physique quantique

6 mai 2014

Avant-propos

La mécanique quantique et la physique statistique, deux domaines essentiels de la physique moderne, sont à la base de nombreux développements actuels à la frontière de la science des matériaux, pour lesquels une approche classique macroscopique est inadaptée. Ce cours présente les concepts fondamentaux de ces disciplines sous un aspect pragmatique, celui de la simulation numérique à l'échelle moléculaire, pour laquelle les mathématiques appliquées apportent des contributions décisives. La présentation des techniques de simulation dans une approche multidisciplinaire constitue une voie pédagogique privilégiée pour appréhender concrètement et durablement les notions délicates qui sous-tendent ces théories ; mais c'est également une ouverture vers des outils numériques aujourd'hui incontournables pour des domaines aussi variés que la physique des hautes énergies, l'ingénierie nucléaire, les nanotechnologies et l'électronique, la biologie moléculaire et la pharmacutique, les matériaux poreux et granulaires, les fluides complexes,...

Ce cours se veut une introduction motivée et motivante à la physique contemporaine. Nous insistons sur le terme d'introduction, tant il y a de points que nous n'aurons pas le temps d'évoquer, ou de traiter avec l'attention nécessaire. Cependant, l'étudiant intéressé pourra tout naturellement prolonger, approfondir et s'approprier les éléments théoriques enseignés ici par le biais de stages, dans le milieu académique ou industriel (CEA, EDF, Lafarge, pour ne citer que quelques partenaires privilégiés de l'Ecole des Ponts), d'un Master Recherche (par exemple en Analyse Numérique, en Science des Matériaux, en Physique des Liquides), voire d'une thèse, ou d'une embauche dans une entreprise désireuse de s'ouvrir aux nouvelles possibilités offertes par l'expérimentation numérique à l'échelle microscopique !

L'objectif de la première partie du cours, consacrée à physique statistique, est de montrer comment une approche purement microscopique permet de calculer les grandeurs macroscopiques usuelles de la mécanique et de la thermodynamique : équations d'états, capacités calorifiques, compressibilités, énergies libres, potentiels chimiques... C'est également l'occasion de donner une définition microscopique de concepts thermodynamiques peu intuitifs tels que l'entropie, ainsi que son rôle primordial dans les propriétés macroscopiques familières des matériaux. On y illustre enfin la dynamique des systèmes microscopiques sur un exemple simple, en précisant les relations entre fluctuations spontanées du système et réponses aux perturbations extérieures. Ces enseignements s'appréhendent le mieux de manière pratique : c'est pourquoi, au-delà des séances de travaux pratiques sur ordinateur, des mini-projets orientés, fondés sur une technique de simulation et orientés vers une application physique particulière, viennent compléter la formation.

La seconde partie du cours introduit les concepts fondamentaux de la mécanique quantique (dualité onde-particule, fonction d'onde, quantification, caractère probabiliste et perturbateur de la mesure, relations d'incertitude, indiscernabilité des particules, liaison chimique, théorie des bandes,...), ainsi que leur formalisation mathématique rigoureuse (théorie spectrale des opérateurs auto-adjoints), et les méthodes de simulation numérique associées (pour la résolution pratique de l'équation de Schrödinger). Nous privilégions la compréhension des concepts sur des exemples

VI Avant-propos

simples traités en profondeur, à partir desquels nous donnons des éléments de généralisation.

C'est la seconde année que ce cours a lieu, et c'est donc également la seconde version seulement de ce cours écrit ! Nous prions d'avance le lecteur de nous pardonner les erreurs qui ne doivent pas manquer, et le remercions de nous faire part de toutes les remarques qu'il voudra bien nous faire pour améliorer cet ouvrage.

Marne-la-Vallée, 1er février 2008

Jean-Noël Roux
Stéphane Rodts
Gabriel Stoltz

Table des matières

Partie I Introduction à la physique statistique

1	Préliminaires	3
1.1	Les constituants de la matière et leurs interactions	4
1.2	Interactions des particules	4
1.2.1	Origines, ordres de grandeur	5
1.2.2	Modèles simples	6
1.2.3	Conséquences macroscopiques de la forme des interactions	8
1.3	Physique statistique et simulation moléculaire	9
2	Description microscopique et statistique des grands systèmes	13
2.1	Système, espace des phases, microétats	13
2.2	Évolution des systèmes classiques	15
2.2.1	Équations de Hamilton	15
2.2.2	Pourquoi adopter le formalisme hamiltonien ?	16
2.2.3	Intégration de la dynamique hamiltonienne	18
2.3	Comportement des grands systèmes : observations	19
2.4	Approche statistique des grands systèmes	22
2.4.1	Approche de l'équilibre, moyennes et corrélations en temps	23
2.4.2	Limite thermodynamique, moyennes et corrélations dans l'espace	24
2.4.3	Ensembles statistiques	25
2.5	Quelques observables courantes	26
2.5.1	L'énergie et ses fluctuations	26
2.5.2	Le tenseur des contraintes	27
2.5.3	Microstructure et corrélations de densité	28
	Exercices	32
3	Postulat fondamental et ensemble microcanonique	33
3.1	Le postulat de base et ses conséquences	34
3.1.1	Macroétat d'équilibre	34
3.1.2	Cas du gaz parfait	35
3.1.3	Distribution de Maxwell (cas général)	37
3.2	Entropie, température	38
3.2.1	Définition	38
3.2.2	Propriétés de l'entropie	39
3.2.3	Couplage thermique	40
3.2.4	Les dérivées de la fonction entropie	40
3.2.5	Entropie et fluctuations	41
3.2.6	Discussion	42

3.3	Intégration des équations du mouvement et échantillonnage de la mesure microcanonique	43
3.3.1	Dynamique hamiltonienne et symplecticité.....	43
3.3.2	Approximation et flot numérique	44
3.3.3	Schémas d'intégration usuels pour les équations différentielles ordinaires ...	45
3.3.4	Schémas d'intégration préservant des propriétés de la dynamique Hamiltonienne	48
3.3.5	Comparaison numérique des différents schémas	55
3.3.6	Application à un système physique	56
	Exercices	59
4	L'ensemble canonique	61
4.1	Distribution de Boltzmann-Gibbs	61
4.1.1	Couplage avec un thermostat	61
4.1.2	Distribution de Boltzmann-Gibbs et information manquante	62
4.2	Énergie libre et grandeurs thermodynamiques	64
4.2.1	Définition	64
4.2.2	Dérivées de l'énergie libre	64
4.2.3	Minimisation de l'énergie libre	65
4.2.4	Calculs avec la fonction de partition	65
4.3	Cas de degrés de liberté indépendants	67
4.3.1	Sous-systèmes indépendants	67
4.3.2	Indépendance des positions et des quantités de mouvement pour une collection de particules	67
4.3.3	Gaz parfait, grandeurs idéales et d'excès dans les fluides	68
4.3.4	Cas des oscillateurs harmoniques indépendants	68
4.4	Quelques remarques sur les applications du formalisme canonique	70
4.4.1	Gaz, solides et liquides	70
4.4.2	Approche de l'équilibre et thermalisation	70
	Exercices	72
5	Échantillonnage de la mesure canonique	75
5.1	Le schéma de Metropolis-Hastings	76
5.1.1	L'algorithme de Metropolis	76
5.1.2	Le schéma de Metropolis-Hastings	79
5.1.3	Quelques raffinements	81
5.2	La dynamique de Langevin	82
5.2.1	Brève introduction pratique aux équations différentielles stochastiques	82
5.2.2	La dynamique de Langevin	85
5.3	Tests statistiques de convergence	88
5.3.1	Vitesse de convergence pour les chaînes de Markov	89
5.3.2	Convergence de certaines observables	91
5.4	Application numérique : fluide de Lennard-Jones	93
	Exercices	96
6	Un exemple de propriété dynamique : l'autodiffusion dans les fluides	97
6.1	La diffusion de particules marquées dans un fluide	98
6.1.1	Échelles de temps et nature du mouvement	98
6.1.2	Mouvement brownien et diffusion	99
6.2	Propriétés des fonctions de corrélation	101
6.2.1	Propriétés générales des fonctions de corrélation dépendant du temps	101
6.2.2	Propriétés particulières de la fonction d'autocorrélation des vitesses	103
6.3	Relation entre corrélation et réponse	105
6.3.1	Importance des propriétés de réponse linéaire	105

6.3.2	Réponse linéaire à une perturbation du Hamiltonien	105
6.3.3	Mobilité d'une particule	106
6.4	Remarques finales	107

Partie II Introduction à la physique quantique

7	Introduction historique : la révolution quantique	111
8	La particule quantique isolée	115
8.1	Description quantique de la particule en translation	115
8.1.1	La fonction d'onde	115
8.1.2	L'équation de Schrödinger	116
8.1.3	La constante de Planck	117
8.1.4	Remarques	117
8.2	Conservation des probabilités - dynamique de la particule	117
8.2.1	Equation de Schrödinger et normalisation	117
8.2.2	Forme locale de la conservation des probabilités : <i>Courant</i> et <i>Vitesse</i> de probabilité	118
8.3	La phase de la fonction d'onde	119
8.4	Solutions de l'équation de Schrödinger dans le vide	119
8.4.1	Etude du Hamiltonien quantique de la particule libre	119
8.4.2	Solution générale : Le paquet d'ondes	121
8.5	Quelques solutions particulières	122
8.5.1	Les ondes de De Broglie	122
8.5.2	Le paquet d'onde quasi-monochromatique	124
8.5.3	Le paquet d'ondes gaussien.	125
8.5.4	Superposition d'états	127
	Exercices	129
9	La particule quantique en interaction avec un potentiel	131
9.1	Dynamique de la particule en interaction	131
9.1.1	L'équation de Schrödinger avec potentiel	131
9.1.2	Caractère bien posé de l'équation de Schrödinger	132
9.1.3	Régularité de la fonction d'onde	133
9.1.4	Le terme potentiel et la normalisation	134
9.2	Le paquet d'onde en interaction : la limite classique	135
9.3	Etats stationnaires de la particule quantique	137
9.3.1	Que sont les états stationnaires ?	137
9.3.2	L'équation de Schrödinger indépendante du temps	137
9.3.3	Energie de la particule	138
9.3.4	Régularité des états stationnaires unidimensionnels	138
9.3.5	Minoration de l'énergie des états stationnaires	139
9.3.6	Importance des états stationnaires en mécanique quantique	139
9.4	Une situation purement quantique : la barrière de potentiel	139
9.4.1	La barrière de potentiel	140
9.4.2	Etats stationnaires de la barrière de potentiel	140
9.4.3	Effet tunnel et transmission résonnante	142
9.4.4	Et la limite classique ?	145
9.5	Implémentation numérique	146
9.5.1	Schéma de Crank-Nicholson.	146
9.5.2	Décomposition d'opérateurs.	148
	Exercices	149

10 La particule quantique confinée	151
10.1 Cas des potentiels confinants	152
10.1.1 Puits infini	152
10.1.2 Oscillateur harmonique	153
10.2 Théorie spectrale des Hamiltoniens quantiques - exemple du potentiel coulombien	155
10.2.1 Caractère auto-adjoint	155
10.2.2 Spectre essentiel	157
10.2.3 Spectre ponctuel	157
10.3 Un exemple quasi-analytique : le trou de potentiel	159
10.3.1 Définition	159
10.3.2 Etats d'énergie liés	160
10.3.3 Etats d'énergie de diffusion	162
10.4 Universalité et implications du phénomène de quantification	164
10.4.1 La quantification	164
10.4.2 Caractère général du phénomène de piégeage	164
10.4.3 Quantification et dynamique de la particule	164
10.4.4 Propriétés spectrales des atomes, des molécules, et des corps denses	165
10.4.5 La quantification comme élément stabilisateur de la matière	167
10.4.6 Et la limite classique ?	168
10.4.7 Et ce n'est pas tout...	169
Exercices	170
11 Le formalisme de Dirac	173
11.1 Formalisme vectoriel	173
11.1.1 Grandeurs physiques et observables	173
11.1.2 Notations vectorielles : bra-ket	174
11.1.3 Généralité du formalisme vectoriel	175
11.2 Construction des observables : principe de correspondance	175
11.3 Equation de Schrödinger	176
11.4 La mesure	176
11.4.1 Observable à spectre discret	177
11.4.2 Observable générale	178
11.5 Commutation des observables	179
11.5.1 Crochets de commutation	179
11.5.2 Evolution d'une observable	180
11.5.3 Relations d'incertitude de Heisenberg	181
11.5.4 Commutation et stabilité de l'atome d'hydrogène	183
Exercices	185
12 Le moment cinétique en mécanique quantique :	
Moment orbital et Spin	187
12.1 Le moment cinétique en mécanique classique	187
12.1.1 Moment cinétique d'une particule ponctuelle autour d'un point	187
12.1.2 Rotation d'un objet rigide sur lui-même	189
12.2 Le moment cinétique orbital	190
12.2.1 Relations de commutation	190
12.2.2 Liens avec la mécanique classique	192
12.3 La définition étendue du moment cinétique en mécanique quantique	193
12.4 Quantification du moment cinétique	193
12.4.1 Que quantifier ?	194
12.4.2 Le résultat de quantification	194
12.4.3 Eléments de démonstration	195
12.5 Quantification du moment cinétique orbital	199
12.6 Le spin	203

12.6.1 Définition	203
12.6.2 Espace de spin	203
12.6.3 Manifestations physiques du spin	204
12.6.4 Un exemple d'espace de spin : l'espace du spin $\frac{1}{2}$	205
Exercices	207
13 Le principe de Pauli : fermions et bosons	209
13.1 Système de plusieurs particules identiques	209
13.1.1 Position du problème	209
13.1.2 Description quantique	210
13.1.3 Indiscernabilité des particules et symétrie de la fonction d'onde	210
13.2 Principe de Pauli : spin, fermions et bosons	211
13.3 Système de deux particules sans spin	211
13.3.1 Deux particules discernables sans spin dans le fond d'un puits quantique ..	212
13.3.2 Paire de bosons dans le fond d'un puits quantique	213
13.3.3 Paire de fermions dans le fond d'un puits quantique	215
13.3.4 Principe de Pauli et équation de Schrödinger	217
13.4 Paire d'électrons dans un système à deux états	217
13.4.1 Etats quantiques 'à spin' de l'électron unique	218
13.4.2 Etats quantiques accessibles à la paire d'électrons	218
13.5 Règles de remplissage des niveaux électroniques	220
13.6 Interaction d'échange et magnétisme	221
Exercices	222
14 La liaison chimique : des molécules diatomiques aux corps solides	225
14.1 La molécule diatomique : liaison chimique et liaison ionique	225
14.1.1 Description quantique du système	225
14.1.2 L'approximation des Combinaisons Linéaires d'Orbitales Atomiques (CLOA) ..	226
14.1.3 Cas de la combinaison de deux états : traitement CLOA	227
14.1.4 Orbitales liante et antiliante	228
14.1.5 Hybridation de deux niveaux atomiques proches en énergie : la liaison chimique covalente	230
14.1.6 Hybridation de niveaux atomiques peu compatibles : liaison ionique	231
14.1.7 Compatibilité des orbitales atomiques	233
14.1.8 Exemples	234
14.2 Structure électronique des corps solides : théorie des bandes	235
14.2.1 Approximation CLOA	235
14.2.2 Le cas particulier des solides cristallins	237
14.2.3 propriétés électroniques des solides	241
Exercices	245
Constantes physiques fondamentales	247
A.1 Constantes physiques les plus utilisées	247
A.2 Unités atomiques	247
Analyse numérique des schémas d'intégration	249
B.1 Consistance et ordre	249
B.2 Convergence	250
B.3 Stabilité linéaire en temps long	250

Elements de théorie spectrale	253
C.1 Opérateurs sur des espaces de Banach	253
C.2 Adjoint d'un opérateur	254
C.3 Spectre d'un opérateur	255
C.4 Diagonalisation des opérateurs autoadjoints compacts	258
C.5 Décomposition des opérateurs auto-adjoints quelconques	258
Quelques relations de commutation entre observables	261
Bibliographie	263

Introduction à la physique statistique

Préliminaires

1.1	Les constituants de la matière et leurs interactions	4
1.2	Interactions des particules.....	4
1.2.1	Origines, ordres de grandeur	5
1.2.2	Modèles simples.....	6
1.2.3	Conséquences macroscopiques de la forme des interactions	8
1.3	Physique statistique et simulation moléculaire	9

Bien des systèmes physiques rencontrés dans la nature sont constitués d'un ensemble de particules ou d'objets solides : ainsi les grains de sable, ou les astéroïdes qui forment les anneaux de Saturne. Parmi ceux-ci, il est important de bien différencier les systèmes qui se prêtent naturellement à un traitement par la physique statistique de ceux qui n'en relèvent pas. La mécanique statistique s'applique aux collections d'objets en interaction qui peuvent être décrits dans un formalisme hamiltonien. Comme on le verra, ceci demande en particulier que les interactions entre particules dérivent d'un potentiel, et que la dynamique du système conserve l'énergie mécanique totale.

De la conservation de l'énergie mécanique découle l'existence du phénomène d'agitation thermique : dans un échantillon de matière, atomes et molécules ne cessent de se mouvoir, à des vitesses qui semblent élevées à notre échelle (de l'ordre de 100 m/s). Ceci contraste bien sûr avec l'intuition ordinaire de la matière au repos, et avec l'expérience de la dissipation de l'énergie mécanique. À une échelle plus grande, les systèmes colloïdaux, composés de particules dont la taille est comprise entre le nanomètre et le micron (typiquement de petits objets solides immergés dans un fluide), peuvent parfois être décrits par la physique statistique. On peut en effet parvenir à décrire leurs interactions à la manière des forces interatomiques ou intermoléculaires, à l'aide d'un potentiel, et il subsiste de plus une forme d'agitation thermique, le mouvement brownien.

Macroscopiquement, les collections de particules auxquelles s'appliquent la mécanique statistique obéissent aux lois de la thermodynamique. Le premier principe de la thermodynamique, qui affirme la conservation de l'énergie, dont les échanges prennent la forme soit de travail, soit de chaleur, est la conséquence macroscopique de la conservation de l'énergie mécanique au niveau microscopique. La chaleur est de l'énergie distribuée dans les degrés de liberté microscopiques, de telle sorte qu'elle ne se « voie » pas macroscopiquement.

Adoptant un point de vue statistique, on considérera qu'à grande échelle le système est décrit par son *macroétat*, qui est une loi de probabilité sur les *microétats*, c'est-à-dire sur les valeurs prises par l'ensemble des degrés de liberté microscopiques du système, ce qui constitue une simplification considérable par rapport à la description de tous les degrés de libertés microscopiques. De surcroît, les postulats de base de la physique statistique permettront de déterminer le macroétat correspondant à l'équilibre thermodynamique, et donc les propriétés macroscopiques du système en fonction d'un petit nombre de grandeurs fondamentales (comme la densité ou la température).

Ceci n'est pas possible pour tous les ensembles de particules : ainsi, l'état de N grains de sable (système pour lequel il n'y a pas conservation de l'énergie mécanique et agitation thermique), peut être caractérisé *mécaniquement*, et son état final (densité, force intergranulaires, etc) peut être prédit – mais cet état dépend de tout le détail du processus par lequel les grains de sable ont été rassemblés et déposés dans la boîte !

1.1 Les constituants de la matière et leurs interactions

La matière est constituée de particules chargées, noyaux et électrons, en interaction électromagnétique. La force de Coulomb (ou, dans certains cas, la force de Lorentz) entre particules chargées est l'interaction fondamentale à l'origine de toutes les forces que l'on rencontrera dans le cours. Elle est en effet responsable de l'ensemble des propriétés de la matière ordinaire. Les autres interactions fondamentales sont la gravitation universelle, qui détermine les phénomènes à l'échelle des astres, et les interactions fortes et faibles, qui régissent les échelles inférieures à la taille des noyaux atomiques, qui est de l'ordre du Fermi (10^{-15} m). Nous nous intéressons dans ce cours aux échelles intermédiaires, qui vont de la taille de l'atome (de l'ordre de l'angström, $1 \text{ \AA} = 10^{-10}$ m) aux dimensions macroscopiques pour lesquelles un échantillon de matière peut être modélisé comme un milieu continu. L'électron est une particule de masse $m_e \simeq 9,11 \times 10^{-31}$ kg, portant la charge électrique $e = 1,602 \times 10^{-19}$ C. Les électrons n'obéissent pas à la mécanique classique, c'est pour les décrire que l'on doit recourir à la mécanique quantique, qui remet en cause les notions familières associées au mouvement des particules dans l'espace (trajectoire, position, vitesse...) et qui fait l'objet de la seconde partie de ce cours. Les noyaux, eux, en raison, comme on le verra, de leur masse beaucoup plus élevée (celle du proton étant $m_p = 1,673 \cdot 10^{-27}$ kg, tandis que sa charge est opposée à celle de l'électron) restent décrits par une théorie classique, sauf à très basse température (proche du zéro absolu) ou à très haute densité (comme à l'intérieur des étoiles). La structure et les propriétés de l'atome, édifice neutre et stable dont le volume est constitué en quasi-totalité par le nuage électronique, sont déterminées par la mécanique quantique (la mécanique classique est même incapable d'expliquer la stabilité de l'atome, voir § 11.5.4). Les noyaux sont constitués d'un nombre de protons égal à celui des électrons, auquel correspond le nom d'un élément chimique, et d'un nombre de neutrons variable (très souvent égal ou un peu supérieur à celui des protons). Par définition, deux atomes qui ne diffèrent que par le nombre des neutrons appartiennent à des *isotopes* différents d'un même élément.

Le nombre d'Avogadro, $\mathcal{N}_A = 6,022 \times 10^{23}$ donne une idée de l'énorme quantité d'atomes présents dans un échantillon macroscopique de matière. Ainsi, toute l'eau des océans de la Terre contient moins de 10^{19} moles de H_2O (dont la masse molaire est 18 g/mol, et le volume molaire environ autant de cm^3 dans les conditions terrestres), un nombre inférieur à $\mathcal{N}_A/10000 \dots$

Les atomes se rencontrent pour la plupart sous forme de molécules, ou d'ions (les gaz rares, comme l'argon ou le krypton, étant une exception). Les molécules qui associent plusieurs atomes s'obtiennent par formation de nuages électroniques communs entre différents noyaux, c'est la liaison chimique (voir le chapitre 14). Les interactions des molécules sont supposées connues et modélisées dans cette première partie du cours. La physique statistique traite du passage du niveau microscopique, celui des interactions moléculaires, au niveau macroscopique où l'on décrit le système par la thermodynamique et la mécanique des milieux continus. On suppose donc déjà accompli un certain changement d'échelle, depuis les noyaux et les électrons jusqu'aux molécules ou aux ions et à leurs interactions. Le § 1.2 précise les modèles simples que l'on prendra en exemple.

1.2 Interactions des particules

On verra dans ce cours que toute la physique des systèmes décrits à l'échelle microscopique est déterminée par la forme des interactions élémentaires entre les particules constituant le système. Nous présentons donc dans un premier temps le « contexte physique », puis présentons des modèles simplifiés d'interaction.

1.2.1 Origines, ordres de grandeur

Les interactions moléculaires ou ioniques sont présentées de manière assez générale et pédagogique dans la référence [22], et nous ne donnons ici que quelques très brèves indications, en partie extraites de cet ouvrage.

Les ordres de grandeurs des quantités physiques à l'échelle atomique sont données par les unités atomiques (voir Appendice A.2). L'unité de longueur est le rayon de Bohr $a_0 = 0,529 \text{ \AA}$, qui donne la taille approximative de l'atome d'hydrogène, et l'unité d'énergie est l'énergie électrostatique correspondante, soit $E_{\text{ua}} = \frac{e^2}{4\pi\epsilon_0 a_0} \simeq 27,2 \text{ eV}$. En raison de l'énergie cinétique de l'électron, c'est le double de l'énergie d'ionisation de l'atome d'hydrogène (c'est-à-dire l'énergie qu'il faut apporter pour éloigner à l'infini l'électron de son noyau). Les énergies de liaison entre atomes, molécules ou ions, évaluées de même en rapprochant une paire initialement très éloignée, sont très généralement inférieures à E_{ua} , mais s'élèvent à quelques eV pour les plus fortes. On distingue ainsi les différentes catégories d'interaction suivantes :

- (i) les *liaisons covalentes* mettent en jeu des énergies de l'ordre de quelques eV, soit quelques centaines de kJ par mole.¹ Les liaisons covalentes décident de la nature chimique et de la forme des molécules, fixent très rigidement les distances entre noyaux, qui sont comprises entre 1 et 2 Å et se rompent lors de réactions chimiques : il s'agit d'interactions à très courte portée ;
- (ii) les *liaisons ioniques* impliquent des énergies comparables à celles des liaisons covalentes, mais sont beaucoup plus sensibles aux conditions physiques (voir § 14.1.6) ;
- (iii) les interactions impliquant des *dipôles permanents* sont à longue portée, et on les calcule par les lois de l'électrostatique classique. Les dipôles sont liés à l'asymétrie de la répartition des charges entre les noyaux dans les liaisons comme C-H à l'intérieur des molécules. Les interactions impliquant des dipôles permanents peuvent être attractives ou répulsives selon l'orientation des dipôles. L'interaction d'une charge électronique et d'un dipôle permanent implique des énergies de l'ordre de 1 eV pour les distances interatomiques typiques (et est à l'origine de l'hydratation des ions en solution aqueuse), alors que l'interaction dipôle-dipôle est plus faible, de l'ordre de 0,1 eV. L'effet de ces interactions dipolaires, de même que celui des liaisons ioniques, peut être considérablement réduit en solution, les énergies électrostatiques étant alors divisées par la permittivité relative ϵ du solvant (ϵ vaut environ 80 pour l'eau à 25° C) ;
- (iv) les interactions les plus faibles font intervenir la *polarisation* des molécules, dont l'origine est la capacité du nuage électronique de se déformer sous l'action d'un champ électrique, tel que celui créé à l'approche d'une charge ou d'un dipôle. Cette déformabilité fait que la molécule acquiert un *dipôle induit*, même si elle ne porte pas de dipôle permanent. Les énergies mises en jeu sont typiquement de l'ordre de 10^{-2} à 10^{-1} eV. Le mécanisme de polarisation entraîne également des interactions des dipôles induits. En effet, les fluctuations de la densité d'électrons autour des noyaux de molécules non dipolaires en moyenne créent des moments dipolaires instantanés. Même en l'absence de tout champ électrique permanent, la corrélation des fluctuations des dipôles instantanés qu'elles portent produit des interactions attractives entre de telles molécules. Ce mécanisme est à l'origine des *forces de dispersion*, également appelées forces de London, ou *forces de van der Waals*². Le potentiel attractif est alors de la forme

$$\mathcal{V}(r) = -\frac{C}{r^6}. \quad (1.1)$$

1. 1 eV $\simeq$ 96,5 kJ/mol.

2. On désigne aussi quelquefois sous le nom générique de van der Waals des interactions de dipôles permanents avec des dipôles induits, ou bien encore entre dipôles permanents si les molécules restent libres de leur orientation, ce qui n'est pas le cas lorsqu'elles forment une phase condensée (liquide ou solide). Le potentiel attractif dans ces cas est toujours de la forme (1.1). Dans ce cours nous ne nous intéresserons qu'aux cas les plus simples de particules dénuées de dipôles permanents.

La constante C vaut par exemple environ 10^{-77} J.m^6 pour la molécule de méthane CH_4 et moitié moins pour l'atome d'argon Ar.

Nous avons évoqué les interactions attractives entre molécules (neutres) identiques. Lorsque deux atomes s'approchent jusqu'à interpénétrer leurs nuages électroniques, une répulsion très forte se manifeste (d'origine quantique, liée au principe de Pauli, voir le chapitre 13). La description de cette force, dans les modèles simples, est faite en fixant un rayon effectif des atomes et des molécules (on peut donner plusieurs définitions d'un rayon, voir [22]) de l'ordre de 1 à 2 Å, puis en proposant des potentiels d'interaction répulsifs faisant intervenir ce rayon. On ajuste finalement empiriquement la forme du potentiel en fonction des propriétés thermodynamiques mesurées.

Nous verrons dans la suite du cours que l'énergie des interactions doit être comparée à l'énergie thermique $k_B T$ ($k_B \simeq 1,38 \times 10^{-23} \text{ J/K}$ étant la constante de Boltzmann³), qui vaut environ $4 \times 10^{-21} \text{ J} = 2,5 \times 10^{-2} \text{ eV}$ à la température de 300 K. D'une façon très sommaire, on peut dire que si $k_B T$ reste inférieure à l'énergie d'interaction qui tend à confiner les positions des particules au voisinage d'un minimum du potentiel, alors elles resteront piégées, tandis qu'elles pourront adopter des configurations variées et conserveront une forte mobilité dans le cas contraire.

1.2.2 Modèles simples

Le modèles des sphères dures

Le modèle le plus simple de molécules, déjà considéré par Boltzmann, est celui de billes solides impénétrables, qui interagissent par des chocs élastiques. On le connaît comme le modèle des « sphères dures ». Il se caractérise par un seul paramètre, le diamètre σ des particules. Important historiquement dans l'histoire de la simulation moléculaire (voir § 1.3), ce modèle est une caricature des interactions fortement répulsives entre atomes proches. Sa mise en œuvre dans une simulation numérique du mouvement d'un ensemble de particules est assez différente du cas général d'une loi de force régulière : au lieu de résoudre une équation différentielle, il faut déterminer la séquence des chocs entre lesquels les particules se déplacent à vitesse constante sur des trajectoires rectilignes. En raison de cette particularité on n'en fera pas un objet central d'études dans la partie théorique de ce cours (mais différentes illustrations feront appel aux propriétés de ce modèle). En dépit de la différence au niveau de la technique de simulation, les propriétés physiques des fluides simulés avec cette interaction ou des interactions plus régulières restent toutefois assez proches. Ce modèle demeure utile pour la simulation des fluides, car

- ses propriétés sont très bien connues et répertoriées, et fournissent une base de départ pour aborder celles des systèmes avec interactions plus réalistes dans une démarche perturbative ;
- il décrit assez exactement les interactions effectives entre certains types de particules colloïdales ;
- son étude est instructive en raison même de son apparente singularité : l'énergie interne du fluide de sphères dures se réduit à l'énergie cinétique, comme celle d'un gaz parfait. L'importance des interactions se manifeste au travers d'effets configurationnels et entropiques (rareté des configurations possibles à densité élevée en raison de la gêne stérique).

Formellement, les sphères dures peuvent être considérées comme interagissant par le potentiel de paires donné par :

$$\mathcal{V}_{\text{SD}}(r) = \begin{cases} 0 & \text{si } r \geq \sigma, \\ +\infty & \text{si } r < \sigma. \end{cases} \quad (1.2)$$

Modèles à interactions continues

Pour décrire la répulsion à courte distance, on peut aussi adopter le modèle dit de sphères molles avec potentiel en puissance inverse de la distance :

3. Le Kelvin (K) est l'unité d'énergie que l'on utilise le plus souvent pour exprimer la température. La constante de Boltzmann n'est pas une constante physique fondamentale mais un simple facteur de conversion, et apparaît chaque fois que cette unité est utilisée à la place du Joule (J).

$$\mathcal{V}_{\text{SM}}(r) = \epsilon \left(\frac{\sigma}{r} \right)^n, \quad (1.3)$$

qui fait apparaître, en plus du « diamètre » moléculaire σ , une énergie caractéristique $\epsilon > 0$ et l'exposant n qui détermine la « raideur » du potentiel. On peut s'attendre à retrouver les propriétés des sphères dures dans la limite $n \rightarrow +\infty$, puisque d'après (1.2) et (1.3) on a $\mathcal{V}_{\text{SM}}(r; \epsilon, n) \rightarrow \mathcal{V}_{\text{SD}}(r)$ pour tout ϵ et $r \neq \sigma$.

Ces potentiels purement répulsifs sont incapables de rendre compte de la tendance à la condensation des molécules à basse température, qui se manifeste par exemple lorsqu'un gaz est refroidi dans une enceinte à volume constant : on voit apparaître, en refroidissant, une phase liquide, puis solide (qui coexistent avec la vapeur). De plus on sait que les molécules neutres s'attirent à grande distance, en particulier en raison de l'omniprésente force de van der Waals (1.1). Le *potentiel de Lennard-Jones* combine la répulsion exprimée par (1.3), pour $n = 12$, avec une telle attraction, et s'écrit :

$$\mathcal{V}_{\text{LJ}}(r) = 4\epsilon \left[\left(\frac{\sigma}{r} \right)^{12} - \left(\frac{\sigma}{r} \right)^6 \right]. \quad (1.4)$$

La figure 1.1 représente les potentiels donnés par (1.3) et (1.4). Le potentiel de Lennard-Jones décrit bien les interactions entre les atomes de gaz rares, tels que l'argon, pour lequel les paramètres sont $\epsilon/k_B = 120$ K, et $\sigma = 3,405$ Å.

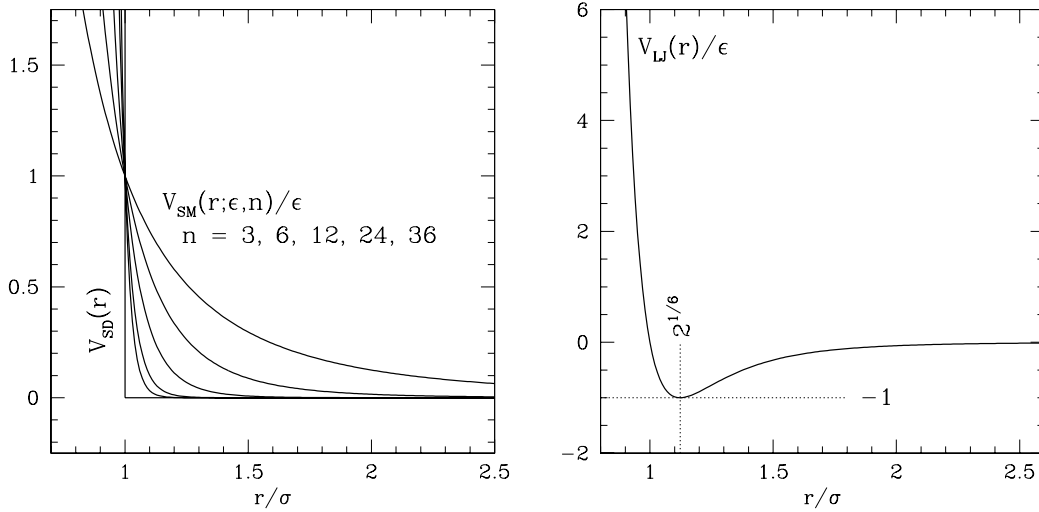


Fig. 1.1. À gauche : potentiels de « sphères molles » (1.3), avec la limite des sphères dures (1.2) pour $n \rightarrow +\infty$. À droite : potentiel de Lennard-Jones (1.4).

Plus généralement, le potentiel de Lennard-Jones donne la forme générale qualitative des interactions entre molécules (non ionisées) pour lesquelles la forme du nuage électronique n'est pas trop anisotrope, avec sa répulsion à courte distance et l'attraction à plus longue portée.

De nombreux raffinements

Pour modéliser plus précisément les interactions des molécules, on a très souvent recours à un modèle qui somme les interactions des atomes. Les atomes formant une même molécule sont considérés comme rigidement liés, ou bien (notamment pour les molécules longues) on admet des degrés de liberté internes (angle de torsion, angle dièdre). Un cas simple d'interaction entre deux molécules diatomiques est illustré par la figure 1.2. Dans cet exemple, chaque molécule est un assemblage rigide, le potentiel intermoléculaire est la somme d'un terme électrostatique, qui résulte par exemple de l'interaction de dipôles permanents, et des interactions entre paires d'atomes, soit

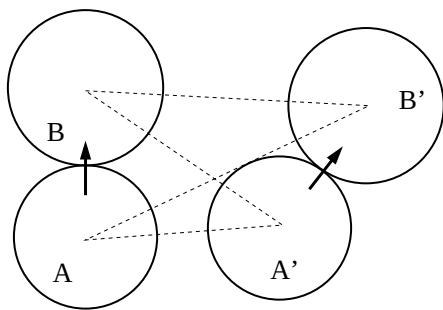


Fig. 1.2. L'interaction entre deux molécules diatomiques AB et A'B' peut être modélisée en sommant les potentiels d'interactions des paires d'atomes n'appartenant pas à la même molécule, donnant lieu aux forces portées par les lignes droites en pointillés. Chaque molécule peut également porter un dipôle permanent, ici schématisé par une flèche.

ici $\mathcal{V}_{AA'} + \mathcal{V}_{AB'} + \mathcal{V}_{BA'} + \mathcal{V}_{BB'}$. Les potentiels interatomiques dans un tel modèle peuvent être choisis sous une forme de Lennard-Jones. Bien souvent, le choix des paramètres d'un modèle de potentiel intermoléculaire se fait en partie par ajustement aux données expérimentales (ou numériques de référence) disponibles. Cette étape semi-empirique est rendue nécessaire par le souci de conserver des interactions faciles à traiter, et en particulier additives par paires.

Certains atomes, molécules ou ions, se prêtent mal à ce type d'approche : ce sont en particulier ceux qui tendent à former des solides covalents, par exemple les atomes de silicium ou de carbone. Les orbitales électroniques ont alors des formes très particulières, une description des atomes par des particules ponctuelles est très mal adaptée, et les interactions sont plus complexes que la simple force de van der Waals (1.1). De tels systèmes se traitent en incluant explicitement des effets d'environnement dans les potentiels interatomiques, empiriquement ou par un calcul de structure électronique pour obtenir la densité électronique locale (méthodes *ab-initio*).

Au niveau de ce cours introductif, on se restreindra au cas simples de particules ponctuelles avec potentiels additifs par paires, fonctions de la seule distance entre voisins et à courte portée (c'est-à-dire que l'on pourra ignorer toutes les interactions des paires de particules plus éloignées l'une de l'autre qu'une certaine distance de coupure).

1.2.3 Conséquences macroscopiques de la forme des interactions

La conséquence la plus remarquable, au niveau macroscopique, de l'interaction de particules modélisées comme ponctuelles avec un potentiel de sphère de type (1.4) est sans doute l'existence de transitions de phase. Le *diagramme de phases* est la représentation, dans l'espace dont les coordonnées sont la densité et la température⁴, des frontières de stabilité des phases gaz, liquide et solide (cristallin) pour les petites molécules et en particulier pour les atomes des gaz rares interagissant par un potentiel de Lennard-Jones. La figure 1.3 représente l'allure typique de ce diagramme pour ces potentiels.

Il est possible de retrouver des aspects qualitatifs du diagramme de phases, et des ordres de grandeurs de certaines propriétés thermodynamiques macroscopiques à partir des paramètres du potentiel d'interaction. Par exemple, en admettant que dans les phases liquide et solide du système de Lennard-Jones, chaque particule interagisse essentiellement avec 12 voisines situées à la distance minimisant $\mathcal{V}_{LJ}(r)$, on peut grossièrement estimer l'énergie potentielle par particule à $e_p = -6\epsilon$ (6 étant le nombre d'interactions de paires par particule, chacune n'étant comptée qu'une seule fois). On peut donc estimer la chaleur latente d'évaporation ou de sublimation à e_p . Quant à la température d'évaporation, elle est de l'ordre de e_p/k_B .

4. On reverra au chapitre 3 pourquoi ce sont les deux seuls variables déterminant l'état d'équilibre d'un système fluide. Les solides admettent aussi le tenseur des déformations comme variable caractérisant localement leur état.

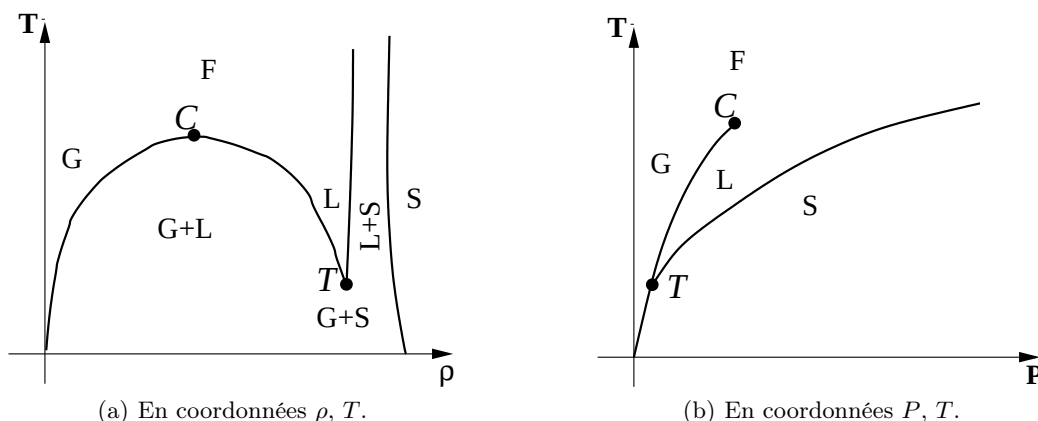


Fig. 1.3. Schéma du diagramme de phase d'un système de Lennard-Jones : (a) dans le plan ρ, T (noter les régions de coexistence, deux phases en équilibre à la même température n'ayant pas la même densité) ; (b) dans le plan P, T . Au point triple T les trois phases sont en équilibre. Au point critique C la différence entre liquide (L) et gaz (G) disparaît, et pour les températures supérieures à la température critique T_C , ou les pressions supérieures à la pression critique P_C on n'a plus qu'une seule phase fluide **F**.

De telles estimations ne peuvent pas, cependant, expliquer les transitions de phase. Même si de simples considérations énergétiques rendent plausibles le passage du solide cristallin à basse température au gaz à température élevée, elles ne peuvent exclure un passage progressif plutôt qu'une transition soudaine à une température de changement de phase, et restent de plus inaptes à rendre compte de l'existence de la phase liquide. Le passage du microscopique au macroscopique révèle aussi des surprises, des phénomènes collectifs impossibles à déduire d'une approche naïve qui ne prend pas en compte la statistique des grands systèmes.

1.3 Physique statistique et simulation moléculaire

La simulation moléculaire, c'est-à-dire la simulation sur ordinateur des modèles de la physique statistique, est un renfort précieux à la théorie et aux expérimentations physiques. On peut lui voir deux grands objectifs :

- (i) la simulation moléculaire est en quelque sorte un *microscope numérique*, qui permet de comprendre le comportement de la matière aux échelles microscopiques d'espace et de temps (pour lesquelles l'expérimentation est difficile) ;
- (ii) la simulation moléculaire est le moyen pratique de calculer les propriétés moyennes des matériaux par les lois prédites par la physique statistique, en particulier dans des régimes thermodynamiques difficilement accessibles à l'expérience (hautes pressions, hautes températures), ou pour tester les caractéristiques de matériaux innovants avant une coûteuse synthèse (médicaments, alliages, ...).

Nous donnons par la suite, à titre culturel, quelques exemples des avancées dans la compréhension de la physique statistique pour lesquelles l'apport des simulations numériques a été décisif. Ces exemples sont centrés sur la simulation des fluides (denses), qui est historiquement le domaine d'émergence de la simulation moléculaire.

Transition de phase entropique

Un des premiers systèmes simulés fut le fluide de sphères dures, dès la fin des années 1950 (on trouvera les références dans [20]). Ceci permit rapidement d'éclaircir un point alors controversé, l'existence ou non d'une phase cristalline pour de telles interactions. En effet, on ne peut pas

justifier l'arrangement cristallin par la minimisation d'une énergie potentielle (comme ce serait le cas avec un potentiel d'interaction continu comme le potentiel de Lennard-Jones), celle-ci étant toujours nulle pour le fluide de sphères dures. C'est en fait *l'entropie* qui est responsable de la cristallisation dans ce cas. Cette grandeur est directement liée au nombre de configurations possibles pour répartir N sphères dans un volume donné, sans recouvrement. La cristallisation des sphères dures manifeste donc qu'au-delà d'une certaine densité l'immense majorité des configurations possibles est constituée d'arrangements réguliers cristallins (de type cubique à faces centrées). La figure 1.4 donne le diagramme de phases des sphères dures, qui ne dépend que de leur densité. D'une façon générale, ce sont souvent les effets entropiques, liées à la statistique dans les grands systèmes, qui réservent des surprises.

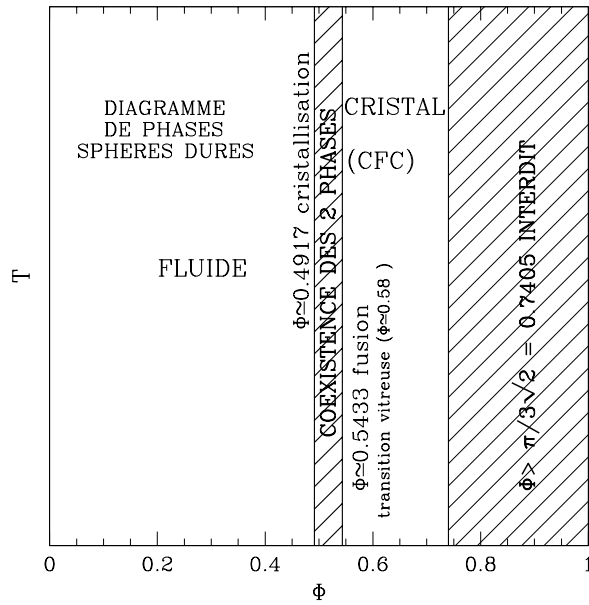


Fig. 1.4. Diagramme de phases des sphères dures, en coordonnées Φ, T , où $\Phi = \pi\rho\sigma^3/6$ est la compacité. La température T n'ayant aucune influence dans ce cas, les frontières sont verticales sur ce graphe. Noter la zone de coexistence entre fluide et cristal.

D'autres formes d'ordres, intermédiaires entre le fluide, dont la structure moyenne est isotrope, et le cristal, apparaissent avec des particules de forme allongée. Ce sont les phases de type cristal liquide (ou *mésophases*). Là aussi, la simulation numérique a révélé l'importance des effets entropiques, puisque des particules en forme de bâtonnets, sans autre interaction que leur mutuelle exclusion stérique, passent à densité croissante par plusieurs transitions de phases. Ainsi, celles⁵ de la figure 1.5 se présentent à basse densité dans une phase fluide ordinaire, isotrope, sans ordre d'orientation ni de position. À densité plus élevée apparaît la phase *nématique*, dans laquelle, les positions restant sans corrélation à longue distance, les grands axes des particules choisissent préférentiellement une orientation commune. Vient ensuite une phase *smectique* dans laquelle apparaît en plus un ordre partiel de positions : les centres des particules se concentrent sur une famille de plans parallèles, mais à l'intérieur de chacune de ces couches on ne trouve pas d'ordre à longue distance. Enfin la phase de plus haute densité est un solide cristallin. Les limites et les zones de coexistence dans le diagramme de phases dépendent du rapport d'aspect des particules

5. Simulations par méthode de Monte-Carlo, dues à Peter Bolhuis et Daan Frenkel, de l'Institut de Physique Atomique et Moléculaire, Amsterdam.

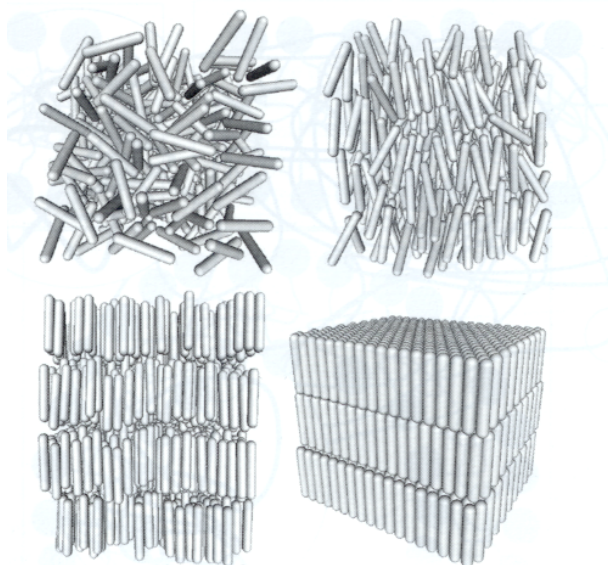


Fig. 1.5. Différentes phases des « sphérocyindres » durs (cylindres fermés par deux hémisphères) : de gauche à droite et de haut en bas, phases isotrope, nématique, smectique, cristalline.

Propriétés dynamiques des fluides

La simulation numérique a également contribué à la compréhension de l'origine microscopique des propriétés dynamiques des fluides, comme la viscosité. Dans les phases denses, solide ou liquide, les mouvements d'une particule sont *a priori* fortement limités par la gêne stérique occasionnée par ses proches voisins. Dans les années 1950 on pensait que l'écoulement d'un fluide, ou bien le mouvement diffusif d'une particule au sein d'un fluide, résultait d'événements localisés dans l'espace et dans le temps, dans lesquels une particule parviendrait exceptionnellement à s'échapper de la « cage » formée par ses voisins. Ce mécanisme est illustré schématiquement par la figure 1.6.

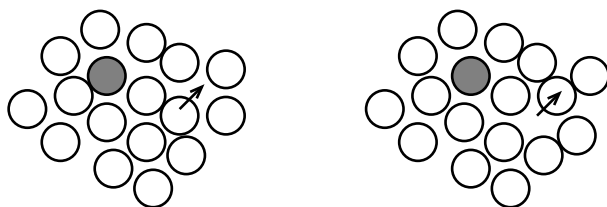


Fig. 1.6. Représentation schématique du modèle de viscosité ou de diffusion moléculaire par sauts. Ici un événement de saut exceptionnel a eu lieu entre les deux configurations représentées, avant (à gauche) et après (à droite) un bref intervalle de temps. Alors que la plupart des particules, comme le disque grisé, oscillaient à l'intérieur de la cage formées par leurs voisines, l'une d'entre elles est parvenue à s'échapper, dans un mouvement brusque, figuré par la flèche.

C'est effectivement le mécanisme de la diffusion moléculaire dans les solides cristallins, dans lesquels les atomes quittent parfois leur site du réseau pour occuper des positions exceptionnelles, dites interstitielles, ou bien sautent brusquement sur un site cristallin voisin, qui se trouve inoccupé (mouvement de lacune), ce qui est fort rare dans les conditions thermodynamiques usuelles.

Toutefois, l'étude des mouvements moléculaires dans les liquides, par la simulation numérique, a révélé que l'on ne pouvait pas isoler des événements de sauts, au moins pour la description de la dynamique des liquides ordinaires : la diffusion est liée à des mouvements collectifs et permanents,

dans lesquels les déplacements présentent une structuration spatiale analogue aux champs de vitesses macroscopiques dans les fluides continus. Conjointement, grâce aux informations nouvelles que l'on a acquises sur la dynamique des liquides, on a pu comprendre comment une version généralisée de l'hydrodynamique macroscopique pouvait s'appliquer jusqu'aux échelles moléculaires. Une manifestation indirecte en est la validité remarquable, dans les liquides, de la relation de Stokes-Einstein jusqu'aux échelles moléculaires. Cette relation concerne normalement les particules colloïdales (de taille mésoscopique), et exprime le coefficient de diffusion D d'une particule sphérique de rayon R en suspension dans un liquide de viscosité η par :

$$D = \frac{k_B T}{6\pi\eta R}. \quad (1.5)$$

L'équation (1.5) fournit une bonne approximation du coefficient de diffusion D , dans les liquides constitués de petites molécules, grossièrement assimilables à des objets sphériques, si on prend pour R une valeur inférieure de 30% au rayon moléculaire.

Description microscopique et statistique des grands systèmes

Dans ce chapitre nous définissons les objets d'étude essentiels du cours de physique statistique. Nous considérons le cas de N particules ponctuelles en interaction, et introduisons les notations et propriétés de base nécessaires à la description statistique de leur mouvement. Nous rappelons aussi quelques observations sur le comportement des grands systèmes et indiquons comment certaines grandeurs macroscopiques peuvent y être mesurées, et comment caractériser leur structure.

2.1	Système, espace des phases, microétats	13
2.2	Évolution des systèmes classiques	15
2.2.1	Équations de Hamilton	15
2.2.2	Pourquoi adopter le formalisme hamiltonien ?	16
2.2.3	Intégration de la dynamique hamiltonienne	18
2.3	Comportement des grands systèmes : observations	19
2.4	Approche statistique des grands systèmes	22
2.4.1	Approche de l'équilibre, moyennes et corrélations en temps	23
2.4.2	Limite thermodynamique, moyennes et corrélations dans l'espace	24
2.4.3	Ensembles statistiques	25
2.5	Quelques observables courantes	26
2.5.1	L'énergie et ses fluctuations	26
2.5.2	Le tenseur des contraintes	27
2.5.3	Microstructure et corrélations de densité	28
	Exercices	32

2.1 Système, espace des phases, microétats

On considère un système formé de N particules (typiquement, des atomes), décrit par une configuration $\mathbf{Q} = (\mathbf{q}_1, \dots, \mathbf{q}_N)$ (avec $\mathbf{q}_i \in \mathbb{R}^d$, où $d = 1, 2, 3$ est la dimension de l'espace ambiant). Chaque particule, de masse m_i , a une impulsion $\mathbf{p}_i = m_i \dot{\mathbf{q}}_i$. On note $\mathbf{P} = (\mathbf{p}_1, \dots, \mathbf{p}_N)$, et M la matrice de masse du système, c'est-à-dire $M = \text{Diag}(m_1, \dots, m_N)$ ¹. Les positions des particules sont des variables à valeurs dans un domaine $\mathcal{D} \subset \mathbb{R}^d$, de volume $|\mathcal{D}|$, alors que les impulsions peuvent prendre toutes les valeurs *a priori*. Ainsi, le système est décrit par le couple $(\mathbf{Q}, \mathbf{P})$, qui évolue dans l'espace des phases $\mathcal{E} = \mathcal{D}^N \times \mathbb{R}^{dN}$.

Définition 2.1 (Microétat ou configuration microscopique). *On appelle microétat (ou configuration microscopique) un élément $(\mathbf{Q}, \mathbf{P}) \in \mathcal{E}$. L'espace des phases $\mathcal{E}$ est par définition l'espace de tous les microétats possibles du système.*

1. Toutefois, les particules auront bien souvent la même masse, car on considèrera des particules identiques.

Le N-uplet $\mathbf{Q}$ évolue dans l'espace des configurations $\mathcal{D}^N$, dont la définition dépend des conditions de bord utilisées :

- si les particules sont libres d'explorer tout l'espace, on a $\mathcal{D} = \mathbb{R}^d$;
- les particules peuvent être assujetties à demeurer à l'intérieur d'un volume fixe. Il faut alors se donner un modèle physique de ce confinement, par exemple avec des forces dérivant d'un potentiel qui croît très rapidement à la traversée de la frontière de $\mathcal{D}$, ou avec des conditions de bord spéculaires (réflexion des impulsions au niveau de l'interface) ;
- si les particules explorent une boîte avec des conditions aux limites périodiques, l'espace des configurations est un tore $\mathbb{T}^{dN}$. C'est très souvent le cas pour la modélisation numérique des solides et des fluides : les conditions aux limites périodiques permettent de représenter plus fidèlement une portion du matériau entouré du même matériau, alors que si on a des conditions de bord libres, ou bien encore faisant intervenir une paroi idéale impénétrable, les effets de surface ne sont pas négligeables, voire dominant le comportement du système.

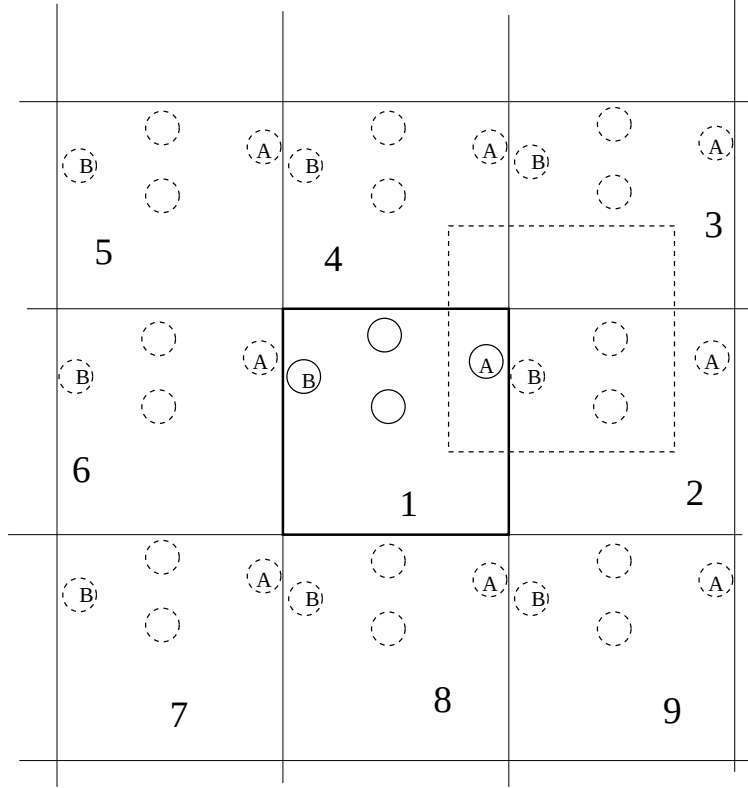


Fig. 2.1. Système bidimensionnel avec conditions aux limites périodiques.

La figure 2.1 illustre les conditions aux limites périodiques dans un espace de dimension $d = 2$. Neuf copies de la cellule de simulation centrale (numérotée 1) y sont représentées. La distance entre les particules marquées A et B est la plus petite entre leurs images les plus proches par l'une des translations qui transforment la cellule de simulation une collection de répliques. Notant A_i la copie de la particule A dans la cellule i , c'est la distance entre A_1 et B_2 (ou entre A_5 et B_4 , ou entre A_7 et B_8) qui détermine la force entre A et B . Cette remarque est importante car nous nous limiterons dans ce cours aux potentiels de courte portée, pour lesquels il suffit d'évaluer l'interaction entre les images les plus proches.² On peut noter aussi que la position des frontières de la cellule de simulation n'a aucune importance, chacune des parties du volume étudié jouant le

2. Pour des potentiels lentement décroissants (en particulier pour les interactions ioniques ou dipolaires) il faut additionner toutes les interactions des A_i et des B_j de toutes les cellules obtenues par périodicité

même rôle. Ainsi on peut tout aussi bien considérer que le domaine de calcul est le carré représenté en pointillés sur la figure 2.1. Dans la pratique, on garde la trace à chaque instant des positions des particules dans une seule cellule, et on veille à prendre en compte les interactions entre images les plus proches.

Remarque 2.1 (Taille minimale de la cellule de simulation). On prendra garde à ce que la cellule de simulation soit suffisamment grande pour qu'une particule n'interagisse qu'une fois avec une autre particule ou son image périodique. Par exemple, pour un domaine de simulation cubique de côté L , il faut que le rayon de coupure r_{cut} du potentiel soit tel que

$$L > 2r_{\text{cut}}.$$

2.2 Évolution des systèmes classiques

2.2.1 Équations de Hamilton

En général, les interactions entre les particules sont décrites par la fonction énergie potentielle $V \equiv V(\mathbf{Q}) = V(\mathbf{q}_1, \dots, \mathbf{q}_N)$. L'énergie totale, somme de l'énergie potentielle et de l'énergie cinétique, s'écrit

$$H(\mathbf{Q}, \mathbf{P}) = V(\mathbf{Q}) + \frac{1}{2} \mathbf{P}^T M^{-1} \mathbf{P}. \quad (2.1)$$

La fonction $H(\mathbf{Q}, \mathbf{P})$ est appelée le *Hamiltonien* du système. Le système est régi par le principe fondamental de la dynamique (loi de Newton), qui prend la forme des équations de Hamilton :

DYNAMIQUE HAMILTONIENNE $\begin{cases} \dot{\mathbf{q}}_i(t) = \nabla_{\mathbf{p}_i} H(\mathbf{Q}, \mathbf{P}) = \frac{\mathbf{p}_i(t)}{m_i}, \\ \dot{\mathbf{p}}_i(t) = -\nabla_{\mathbf{q}_i} H(\mathbf{Q}, \mathbf{P}) = -\nabla_{\mathbf{q}_i} V(\mathbf{Q}). \end{cases} \quad (2.2)$

C'est un système d'équations différentielles ordinaires (EDO). Par la suite, on fera l'hypothèse qu'il y a une et une seule solution à cette EDO. Ceci est le cas par exemple si le champ de force ∇V est globalement lipschitzien (par application du théorème de Cauchy-Lipschitz).³ Un système tel que (2.2), posé dans $\mathbb{R}^{2dN}$ est dit *hamiltonien* si on peut définir des variables $\mathbf{Q} \in \mathbb{R}^{dN}$, $\mathbf{P} \in \mathbb{R}^{dN}$ et une fonction $H(\mathbf{Q}, \mathbf{P})$ telle qu'il puisse s'écrire sous la forme (2.2). Comme toute EDO, la dynamique (2.2) est caractérisée par son flot.

Définition 2.2 (Flot). *On appelle flot de la dynamique hamiltonienne (2.2) la famille d'applications $(\phi_t)_{t \geq 0}$ qui, au microétat $(\mathbf{Q}_0, \mathbf{P}_0) \in \mathcal{E}$ à $t = 0$ font correspondre les microétats $(\mathbf{Q}(t), \mathbf{P}(t)) \in \mathcal{E}$ à tout instant $t \geq 0$.*

On a toujours $\phi_{t+u} = \phi_t \circ \phi_u$ pour $t, u \geq 0$, et la famille $(\phi_t)_{t \geq 0}$ constitue donc un semi-groupe. Cela signifie simplement que si l'on prend $\phi_u(\mathbf{Q}_0, \mathbf{P}_0)$ comme condition initiale, on obtient au bout du temps t la même valeur $(\mathbf{Q}', \mathbf{P}')$ qu'en partant de $(\mathbf{Q}_0, \mathbf{P}_0)$ et en intégrant sur un temps $t + u$. Pour les systèmes hamiltoniens, ϕ_{-t} a un sens (voir (2.4)) et le flot définit un groupe.

(y compris l'interaction de chaque particule chargée avec chacune de ses répliques). Un tel calcul porte le nom de *somme d'Ewald* pour les charges ou *somme d'Ewald et Kornfeld* pour les dipôles.

3. Dans certains cas, on peut évaluer montrer qu'une condition de Lyapounov est vérifiée, *i.e.* on montre qu'une certaine fonction $W(\mathbf{Q}, \mathbf{P})$ décroît au cours du temps, cette décroissance donnant une information sur le comportement de la solution (par exemple, elle reste dans un compact). Un cas classique concerne des dynamiques du type $\dot{q} = -q^3$. Dans ce cas, on n'a pas une condition de Lipschitz globale, mais on montre facilement que la fonction $t \mapsto W(q(t)) = q(t)^2$ décroît au cours du temps, ce qui montre que la solution reste dans le compact $B(0, q(0))$. Au final, comme la solution est bornée, on peut en fait conclure que la force est bornée, et appliquer quand même le théorème de Cauchy-Lipschitz.

2.2.2 Pourquoi adopter le formalisme hamiltonien ?

La forme hamiltonienne (2.2) des équations du mouvement est équivalente à la dynamique suivante, sur les positions et les vitesses :

$$\begin{cases} \dot{\mathbf{q}}_i(t) = \mathbf{v}_i, \\ \dot{\mathbf{v}}_i(t) = -\frac{1}{m_i} \nabla_{\mathbf{q}_i} V(\mathbf{Q}, \mathbf{P}). \end{cases}$$

Cependant, on préfère très souvent l'écriture (2.2) pour un certain nombre de raisons que nous détaillons ci-dessous.

Invariants de la dynamique

On montre tout d'abord que la dynamique (2.2) conserve l'énergie du système. On parle d'ailleurs en mécanique de *système conservatif*. Ceci est une propriété importante de cette dynamique, fondamentale pour la définition de l'ensemble microcanonique (chapitre 3). Elle se prête de surcroît plus aisément au passage de la mécanique classique à la mécanique quantique, comme on le verra dans la seconde partie de ce cours.

Propriété 2.1 (Conservation du Hamiltonien le long d'une trajectoire). *On montre facilement que la dynamique (2.2) est telle que l'énergie totale $H(\mathbf{Q}, \mathbf{P})$ est préservée le long de la trajectoire. En effet,*

$$\frac{dH(\mathbf{Q}(t), \mathbf{P}(t))}{dt} = \partial_{\mathbf{Q}} H(\mathbf{Q}(t), \mathbf{P}(t)) \cdot \dot{\mathbf{Q}}(t) + \partial_{\mathbf{P}} H(\mathbf{Q}(t), \mathbf{P}(t)) \cdot \dot{\mathbf{P}}(t).$$

Utilisant alors (2.2) et la définition (2.1) du Hamiltonien H ,

$$\frac{dH(\mathbf{Q}(t), \mathbf{P}(t))}{dt} = \nabla_{\mathbf{Q}} V(\mathbf{Q}(t)) \cdot M^{-1} \mathbf{P}(t) + M^{-1} \mathbf{P}(t) \cdot (-\nabla_{\mathbf{Q}} V(\mathbf{Q}(t))) = 0.$$

Ceci montre que $H(\mathbf{Q}(t), \mathbf{P}(t)) = H(\mathbf{Q}_0, \mathbf{P}_0)$ pour tout temps t .

D'autres invariants peuvent exister, selon les propriétés du potentiel. En général, la quantité de mouvement totale $\mathbf{P} = \sum_{i=1}^N \mathbf{p}_i$ est conservée, ainsi que le moment cinétique total $\mathbf{L} = \sum_{i=1}^N \mathbf{q}_i \times \mathbf{p}_i$ (si on n'utilise pas de conditions de bord périodiques), pour des formes de l'énergie potentielle invariantes, respectivement, par translation ou par rotation. Voir, à ce sujet, l'exercice 2.1.

Mesure invariante

On peut montrer que la dynamique hamiltonienne conserve la mesure de Lebesgue, au sens où

Propriété 2.2 (Conservation de la mesure de Lebesgue). *Pour tout borélien $B \subset \mathcal{E}$,*

$$\int_{\phi_t(B)} d\mu^{\text{Leb}}(\mathbf{Q}, \mathbf{P}) = \int_B d\mu^{\text{Leb}}(\mathbf{Q}, \mathbf{P}),$$

la mesure μ^{Leb} étant la mesure de Lebesgue et $(\phi_t)_{t \geq 0}$ le flot de la dynamique hamiltonienne.

Pour des raisons de cohérence avec la mécanique quantique, on utilise plutôt la mesure

$$d\mu_0(\mathbf{Q}, \mathbf{P}) = \frac{1}{N! h^{3N}} d\mu^{\text{Leb}}(\mathbf{Q}, \mathbf{P}) = \frac{1}{N! h^{3N}} d\mathbf{q}_1 \dots d\mathbf{q}_N d\mathbf{p}_1 \dots d\mathbf{p}_N. \quad (2.3)$$

Le préfacteur fait apparaître $N!$ et la constante de Planck $h \simeq 6,626 \times 10^{-34} \text{ J.s}$; il conduit à évaluer le volume d'une partie de l'espace des phases par un nombre sans dimension. Ce facteur

d'origine quantique peut être omis sans dommage en général, du moins en présence d'une seule espèce moléculaire.⁴

Pour vérifier la conservation de la mesure (2.3), on montre que, pour tout $t \geq 0$,

$$|\text{Det}(\nabla_{\mathbf{Q}, \mathbf{P}} \phi_t(\mathbf{Q}, \mathbf{P}))| = |\text{Jac } \phi_t(\mathbf{Q}, \mathbf{P})| = 1.$$

Ainsi, le flot définit un changement de variable de Jacobien 1. Cette propriété résulte du fait que le champ de vecteurs associé à la dynamique hamiltonienne est de divergence nulle (voir [19, Lemme VII.3.1] pour une preuve de cette assertion). On calcule en effet

$$\text{div}_{\mathbf{Q}}(\nabla_{\mathbf{P}} H) + \text{div}_{\mathbf{P}}(-\nabla_{\mathbf{Q}} H) = \sum_{i=1}^N (\nabla_{\mathbf{q}_i} \cdot \nabla_{\mathbf{p}_i} H - \nabla_{\mathbf{p}_i} \cdot \nabla_{\mathbf{q}_i} H) = 0.$$

On peut comprendre ce résultat par une analogie avec la mécanique des fluides (voir le cours de cinématique des milieux continus),⁵ en se représentant la dynamique hamiltonienne comme décrivant un écoulement dans l'espace des phases, (2.2) donnant la valeur de la vitesse en chaque point (champ des vitesses en représentation eulérienne). Cette vitesse possède dans ce cas $2dN$ coordonnées, qui sont celles de $(\dot{\mathbf{q}}_i)_{1 \leq i \leq N}$ et $(\dot{\mathbf{p}}_i)_{1 \leq i \leq N}$. La conservation du volume est exprimée par la condition de divergence nulle du champ de vitesse (fluide incompressible), au vu de l'équation de la conservation de la quantité de matière.

Caractère chaotique de la dynamique hamiltonienne

On montre ici sur un exemple que la dynamique hamiltonienne est en générale exponentiellement sensible aux conditions initiales, au sens où une petite perturbation de la donnée initiale grandit exponentiellement au cours du temps. Or, en pratique, il est impossible de connaître exactement les conditions initiales d'un système donné, et de toute façon il existe une erreur numérique intrinsèque des ordinateurs actuels (de l'ordre de 10^{-16}). Ces considérations montrent que la notion de trajectoire d'un système dynamique n'est pas forcément la bonne notion. C'est un point qui aura toute son importance lors de la discussion sur l'échantillonnage de la mesure microcanonique et les schémas numériques d'intégration associés (§ 3.3).

Exemple 2.1. On considère une particule $q \in \mathbb{R}$ soumise au potentiel $V(q) = -\frac{1}{2}\omega^2 q^2$. Dans ce cas, les équations du mouvement

$$\ddot{q} = -\omega q, \quad (q_0, p_0) = (0, 1),$$

ont pour solution $q(t) = \omega^{-1/2} p_0 \exp(\sqrt{\omega} t)$. On voit ainsi qu'une perturbation de l'ordre de ϵ sur l'impulsion initiale p_0 conduit à une perturbation en $\epsilon \omega^{-1/2} \exp(\sqrt{\omega} t)$ au temps t .

La situation présentée dans l'Exemple 2.1 peut sembler caricaturale, mais est en fait très générale : en effet, au voisinage de tout maximum local $\mathbf{Q}_{\max}$ de la surface d'énergie potentielle, on a localement le développement limité

$$V(\mathbf{Q}) = V(\mathbf{Q}_{\max}) + (\mathbf{Q} - \mathbf{Q}_{\max})^T \cdot \nabla^2 V(\mathbf{Q}_{\max}) \cdot (\mathbf{Q} - \mathbf{Q}_{\max}) + O(|\mathbf{Q} - \mathbf{Q}_{\max}|^3),$$

avec la matrice hessienne $\nabla^2 V(\mathbf{Q}_{\max})$ définie négative. Des développements limités de ce type montrent également que la situation rencontrée à l'Exemple 2.1 reste valable au voisinage de tout point-col, où la matrice hessienne du potentiel a au moins une valeur propre négative - ce qui, en pratique, vaut pour une grande partie de l'espace des phases si le système est de grande dimension !

L'analogie avec l'écoulement d'un fluide peut une fois de plus contribuer à rendre intuitive la propriété de sensibilité aux conditions initiales de la dynamique hamiltonienne d'un ensemble

4. Le terme $N!$ qui exprime l'indiscernabilité des particules doit être conservé en présence de plusieurs espèces. Dans le calcul de l'augmentation d'entropie lors d'un mélange, son omission conduit au « paradoxe de Gibbs » – voir le cours « Physique des états de la matière » de première année [8].

5. Le terme « flot » est en fait une traduction un peu impropre de l'anglais *flow*, c'est-à-dire écoulement.

de particules en interaction. Deux trajectoires dans l'espace des phases, partant de points très proches, s'écartent de plus en plus et finissent par explorer des microétats très distants, de même que l'écoulement turbulent d'un fluide conduit à en mélanger les différentes parties. Lorsque l'on matérialise une petite région d'un écoulement turbulent de fluide incompressible, par exemple au moyen d'un traceur coloré, on voit la matière repérée se disperser dans un volume croissant : l'écoulement à divergence nulle rapproche les points selon certaines directions, et les éloigne selon d'autres, d'où une divergence des trajectoires. Il en est de même pour l'écoulement incompressible dans l'espace à $2d \times N$ dimensions $\mathcal{E}$.⁶

Réversibilité en temps

Une autre propriété importante de la dynamique hamiltonienne est sa réversibilité en temps. Si, à l'instant t , on renverse les vitesses, alors la trajectoire antérieure dans $\mathcal{E}$ est parcourue dans l'autre sens. Autrement dit, si on note $\mathcal{S}$ l'application qui change le signe des vitesses alors que les positions restent les mêmes, *i.e.*

$$\mathcal{S}(\mathbf{Q}, \mathbf{P}) = (\mathbf{Q}, -\mathbf{P}),$$

le flot ϕ_t est tel que :

$$\phi_t \circ \mathcal{S} = \mathcal{S} \circ \phi_{-t}. \quad (2.4)$$

Preuve. Soit $(\mathbf{Q}_0, \mathbf{P}_0) \in \mathcal{E}$ donné. Notons $(\mathbf{Q}(t), \mathbf{P}(t))$ la solution de (2.2) partant de la condition initiale $(\mathbf{Q}_0, \mathbf{P}_0)$, et $(\hat{\mathbf{Q}}(t), \hat{\mathbf{P}}(t))$ la solution de (2.2) partant de la condition initiale $\mathcal{S}(\mathbf{Q}_0, \mathbf{P}_0) = (\mathbf{Q}_0, -\mathbf{P}_0)$. On vérifie facilement que $(\hat{\mathbf{Q}}(t), \hat{\mathbf{P}}(t))$ et $\mathcal{S}(\mathbf{Q}(-t), \mathbf{P}(-t)) = (\mathbf{Q}(-t), -\mathbf{P}(-t))$ satisfont (2.2), partant de la même condition initiale $(\mathbf{Q}_0, -\mathbf{P}_0)$. $\square$

2.2.3 Intégration de la dynamique hamiltonienne

On parle quelquefois de *problème à N corps* pour décrire la dynamique hamiltonienne, dans la mesure où l'évolution de la particule i , donnée par

$$m_i \ddot{\mathbf{q}}_i = -\nabla_{\mathbf{q}_i} V(\mathbf{q}_1, \dots, \mathbf{q}_N), \quad (2.5)$$

dépend *a priori* des positions de toutes les autres particules. On a affaire à un système de N équations différentielles ordinaires couplées, et l'information physique importante est le potentiel V . On verra comment intégrer numériquement ces équations au § 3.3.4).

On sait bien sûr que les problèmes à N corps, en général, ne se résolvent pas analytiquement pour $N \geq 3$ (et en pratique, jamais dans un cas vraiment intéressant !). Une résolution numérique des équations du mouvement (2.2) (ou (2.5)) n'est possible que pour des systèmes d'au mieux une centaine de milliers de molécules et sur des temps très courts, de l'ordre de 10^{-9} s. La détermination précise des trajectoires dans l'espace des phases pour le mouvement d'un ensemble de N particules avec N de l'ordre du nombre d'Avogadro $\mathcal{N}_A \simeq 6 \times 10^{23}$ est donc impossible. Fort heureusement, il est possible d'évaluer, dans bien des cas pratiques, des propriétés macroscopiques à partir de simulations numériques d'échantillons assez petits (quelques centaines de particules suffisent parfois), pendant une durée assez brève (cf. les exemples des § 2.3, 3.3.6 et 5.3.2). Insistons une fois de plus sur le fait que la recherche de la *précision* dans le calcul numérique des trajectoires est illusoire et vaine, à cause de la sensibilité aux conditions initiales.

6. On peut trouver dans le livre de Diu *et al.* [10, p. 577] un calcul approché de cet effet, dans le cas d'un gaz de sphères dures, dont le résultat est saisissant. Les collisions de molécules d'air à la pression atmosphérique, modélisées comme des billes élastiques, sont telles qu'elles amplifient les écarts angulaires sur les portions rectilignes de trajectoire entre deux chocs. Les déviations angulaires croissent donc exponentiellement avec le nombre de chocs. Pour qu'une particule soit déviée de 1 radian sous l'effet d'une perturbation infime – celle que cause l'attraction gravitationnelle d'un expérimentateur pesant 70 kg, à 1 mètre de distance – il suffit de 10 collisions environ, et de 50 collisions si la perturbation est due à l'attraction gravitationnelle d'un électron à une distance égale au rayon de l'univers ! 10 collisions, dans l'air à 300 K et à la pression atmosphérique 10^5 Pa, correspondent à un temps de l'ordre de la nanoseconde.

Heureusement, ce qui importe en pratique, pour retrouver les lois macroscopiques, c'est d'étudier des comportements moyens, collectifs. Dans la simulation numérique ce sont les propriétés du flot ϕ_t qui sont importantes, plutôt que celles des trajectoires prises une à une. On adoptera en outre un point de vue statistique : on s'intéressera en particulier à la fréquence avec laquelle sont visitées différentes parties de l'espace des phases. Cette approche statistique a donné lieu à tout le développement historique de la physique statistique (avant les ordinateurs et les simulations numériques). Nous relevons dans le paragraphe suivant quelques observations qualitatives sur le comportement des grands systèmes qui motivent cette démarche et précisons quelques notions de base nécessaires à son développement.

2.3 Comportement des grands systèmes : observations

Grâce à la simulation numérique, il est possible de se livrer à diverses « expériences numériques » en suivant le mouvement d'un ensemble de particules interagissant par un potentiel V . Un cas simple est celui d'un cube avec conditions aux limites périodiques et de côté L , la cellule de base étant définie par $-L/2 \leq x^1, x^2, x^3 \leq L/2$, et contenant N particules. Nous observons par la suite les valeurs prises au cours du temps par les nombres N_1 et $N_2 = N - N_1$ de particules contenues dans chacun des demi-cubes $x^1 \leq 0$ et $x^1 > 0$. Cet exemple illustre le comportement des grandeurs macroscopiques, qui résultent de multiples contributions des degrés de liberté microscopiques dont les fluctuations sont erratiques. Nous utilisons ici le modèle des sphères dures (§ 1.2.2) et leur diamètre σ définit l'unité de longueur. L'unité de temps est σ/u , où u est la valeur moyenne quadratique de l'une des 3 coordonnées de la vitesse des particules. On verra (§ 3.1.3) que cette vitesse est donnée par :

$$u = \frac{k_B T}{m}, \quad (2.6)$$

où m est la masse des particules et T la température. La densité numérique ρ est ici telle que la fraction volumique de sphères (également appelée compacité)

$$\Phi = \pi \rho \sigma^3 / 6, \quad (2.7)$$

soit égale à 0,24. L'observation de l'évolution de grandeurs macroscopique donne des résultats qui deviennent approximativement constants et indépendants des conditions initiales après un certain temps. On s'attend ainsi, dans notre exemple, à ce que la moitié des particules soit contenue, en moyenne, dans chacune des deux moitiés du volume. La figure 2.2 illustre ce comportement, avec une condition initiale délibérément particulière : dans ces calculs, on a choisi $N_1 = N$ et $N_2 = 0$ à $t = 0$. On constate que le rapport $X = N_2/N$ rejoint rapidement $1/2$, et fluctue ensuite en ne s'écartant que légèrement de cette valeur.

On voit également que l'amplitude des fluctuations diminue lorsque N augmente. On peut montrer que, dans la limite $N \rightarrow +\infty$, N_2 se comporte comme une variable gaussienne de moyenne $N/2$ et de variance $\rho k_B T \chi N/4$, χ étant la compressibilité isotherme. On écrit (notation utilisée dans la suite de ce document)

$$N_2 \sim \mathcal{N} \left(\frac{N}{2}, \frac{1}{4} N \rho k_B T \chi \right). \quad (2.8)$$

La variance de N_2 étant proportionnelle au nombre de particules, le rapport écart-type/moyenne est en $O(N^{-1/2})$. Ainsi, $x = \sqrt{N}(X - \frac{1}{2})$ possède la densité de probabilité

$$P(x) = \sqrt{\frac{2}{\pi \rho k_B T \chi}} \exp \left[-\frac{2x^2}{\rho k_B T \chi} \right]. \quad (2.9)$$

La figure 2.3 montre que les résultats de simulation sont bien en accord avec cette prédiction. On a comparé la variance de N_2 à la prédiction (2.8), en utilisant l'équation d'état connue des sphères dures [15] qui donne, pour la fraction volumique de sphères $\Phi = 0,24$, $\chi \equiv \rho \left(\frac{\partial P}{\partial \rho} \right)_T \simeq \frac{0,199}{\rho k_B T}$.

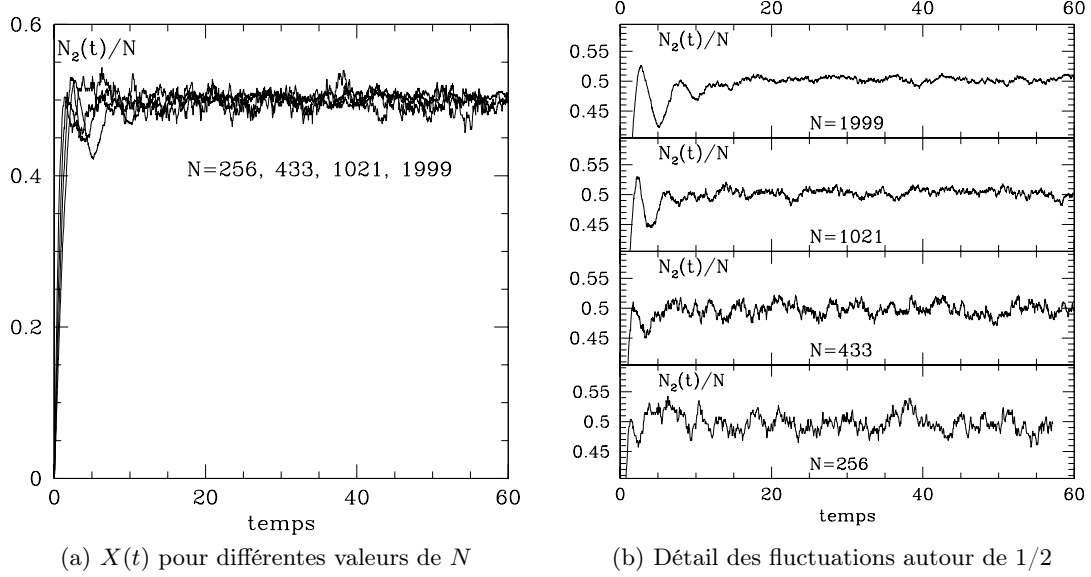


Fig. 2.2. Évolution dans le temps de la fraction du nombre de particules contenue dans une moitié de la cellule de simulation, partant de zéro, pour 4 échantillons avec des valeurs de N différentes.

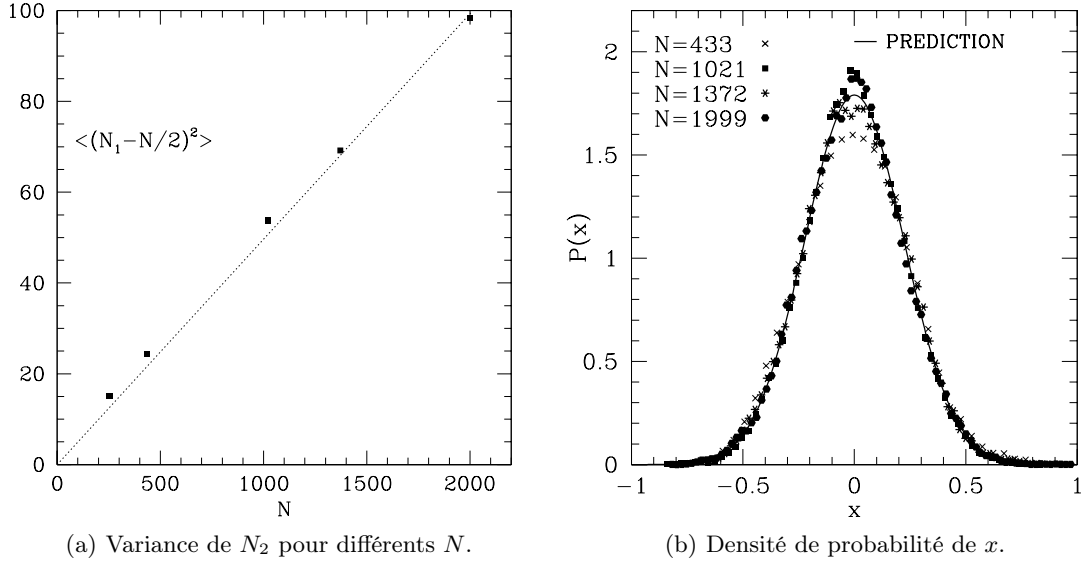


Fig. 2.3. Pour 4 valeurs de N , variance de N_2 (a) et densité de probabilité de $x = N_2/N$ (b), comparées à la prédiction (2.9), avec la valeur connue $\rho k_B T \chi = 0,199$ dans ces conditions thermodynamiques. La droite pointillée sur le graphe (a) est $N \mapsto \frac{1}{4} N \sqrt{\rho k_B T \chi}$.

Les données statistiques sur les fluctuations de N_2 ont été acquises une fois écoulé un temps nécessaire à l'oubli de l'état initial particulier (ici, $t = 15$), quand on peut estimer que les oscillations initiales de $N_2(t)$ sont amorties. Il semble alors que l'on ait atteint un état stationnaire, au sens statistique du terme, c'est-à-dire que la succession des valeurs prises au cours du temps par la grandeur observée s'apparente à la trajectoire d'un processus aléatoire stationnaire : connaissant la valeur à l'instant t , la valeur à l'instant $t' > t$ est aléatoire et distribuée suivant une loi qui ne dépend que de $t' - t$. Ceci peut se vérifier par des tests statistiques adéquats.

Il est intéressant de mesurer la *fonction d'autocorrélation* entre les valeurs de N_2 à des instants différents, que l'on définit comme :

$$C(t) = \langle (N_2(t) - \langle N_2 \rangle) (N_2(0) - \langle N_2 \rangle) \rangle.$$

La notation $\langle \cdot \rangle$ représente une moyenne sur le choix de l'origine des temps, pour une longue trajectoire donnée (éventuellement, cette moyenne pourrait également être prise sur plusieurs trajectoires partant de conditions initiales différentes). La fonction $C(t)$, normalisée par sa valeur à $t = 0$ (c'est-à-dire la variance de N_2), est représentée sur la figure 2.4. Le temps de corrélation de

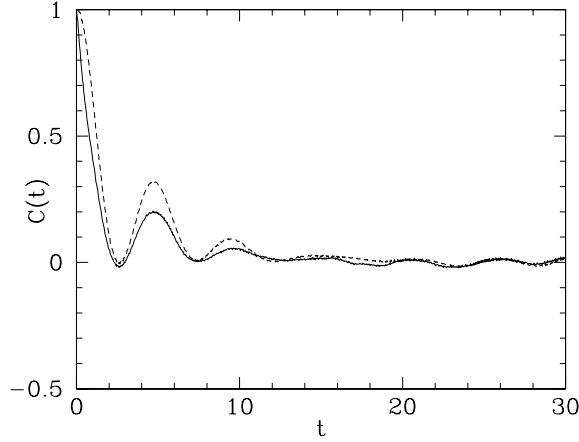


Fig. 2.4. Fonctions d'autocorrélation dépendant du temps, normalisées à leur valeurs initiales, de N_2 (courbe continue) et de $\rho_{\mathbf{k}}$ (courbe en pointillés). Cette dernière quantité est définie par (2.11) et (2.10), pour un vecteur d'onde $\mathbf{k}$ tel que $|\mathbf{k}| = \frac{2\pi}{L}$. Ici, $N = 1999$ et $\Phi = 0,24$.

N_2 est donc assez long, puisque $C(t)$ est toujours oscillant pour $t = 30$, et donne une indication du temps de mesure nécessaire pour bien échantillonner la distribution statistique des valeurs prises par N_2 : il faut en principe moyenner sur une durée grande devant le temps de corrélation de N_2 , pour que la distribution des valeurs observées cesse d'être biaisée par l'état initial au début des mesures.

En fait, cette fonction pourrait être évaluée à partir de la connaissance macroscopique de la thermodynamique et de l'hydrodynamique du fluide de sphères dures. On peut en effet faire l'hypothèse que les fluctuations de la densité à l'échelle de la cellule de simulation à $t = 0$ vont, en moyenne, régresser comme d'il s'agissait d'un petit écart macroscopique à l'équilibre global de l'échantillon de fluide étudié.⁷ Une telle approche permet de calculer $C(t)$, à partir d'une décomposition de $N_2(t)$ en série de Fourier. Les longueurs d'onde les plus faibles ne sont pas bien décrites par l'approche macroscopique, mais la décroissance vers zéro de $C(t)$ aux temps longs est dominée par les longueurs d'onde de l'ordre de la taille L de la cellule de simulation. La plus grande est précisément L , et sur la figure 2.4 on a aussi représenté la fonction d'autocorrélation de $\rho_{\mathbf{k}}$,

$$F(\mathbf{k}, t) = \frac{\langle \rho_{\mathbf{k}}(t) \rho_{-\mathbf{k}}(0) \rangle}{N}, \quad (2.10)$$

normalisée par sa valeur initiale. Le coefficient $\rho_{\mathbf{k}}$ est la composante de Fourier du champ de densité, pour un vecteur d'onde $\mathbf{k}$ tel que $|\mathbf{k}| = \frac{2\pi}{L}$:

$$\rho_{\mathbf{k}} = \sum_{i=1}^N \exp[-i\mathbf{k} \cdot \mathbf{q}_i]. \quad (2.11)$$

7. Cette hypothèse est appelée le *postulat d'Onsager pour la régression des fluctuations*.

On appelle *facteur de structure* la quantité $S(\mathbf{k}) \equiv F(\mathbf{k}, 0)$:

$$S(\mathbf{k}) = \frac{\langle \rho_{\mathbf{k}} \rho_{-\mathbf{k}} \rangle}{N}. \quad (2.12)$$

Cette quantité est reliée aux fonctions de corrélation de densité dans l'espace réel (voir § 2.5.3). La dépendance en temps de $F(\mathbf{k}, t)$, dont la forme est prédite par l'hydrodynamique⁸ (voir [20]), coïncide bien pour les temps assez grands avec celle de la fonction d'autocorrélation de N_2 .

Enfin, une remarque importante concerne la réversibilité en temps des trajectoires, propriété que le système de sphères dures partage avec les systèmes hamiltoniens du § 2.2. Pour les simulations dont nous rapportons ici les résultats, cette propriété signifie en principe que si l'on renverse les vitesses de toutes les particules à l'instant t , alors, à l'instant $2t$, on doit retrouver la configuration de départ ($t = 0$), qui était telle que toutes les particules se situent dans la moitié de la cellule de simulation. Or il n'en est rien si t est suffisamment grand ($t = O(1)$) : on observe simplement des fluctuations analogues à celles que l'on voit sur la figure 2.3(b). Ceci s'explique une fois encore par l'accumulation d'erreurs numériques et l'instabilité des trajectoires.

2.4 Approche statistique des grands systèmes

Plutôt que de se confronter avec l'énorme quantité d'informations que constitue la donnée de toutes les coordonnées de $(\mathbf{Q}, \mathbf{P})$, on s'intéresse lorsque l'on souhaite retrouver le comportement macroscopique du système à partir de l'approche microscopique, à certaines fonctions de $(\mathbf{Q}, \mathbf{P})$ seulement, à valeurs dans $\mathbb{R}^m$ avec $m \ll 2dN$.

Définition 2.3 (Observable). *On appelle observable une fonction $A : \mathcal{E} \rightarrow \mathbb{R}^m$ ($m < 2dN$) des microétats $(\mathbf{Q}, \mathbf{P}) \in \mathcal{E}$ du système.*

L'approche statistique permet une description réduite du système en ne s'intéressant qu'aux valeurs *moyennes* des observables (éventuellement à leur fonctions de corrélations en temps), et non pas à la distribution de toutes leurs valeurs possibles. On se donne donc des lois de probabilité sur $\mathcal{E}$, qui correspondent à l'état du système au sens macroscopique de la thermodynamique.

Définition 2.4 (Macroétat). *On appelle macroétat une mesure de probabilité $d\rho_N$ sur l'espace des phases $\mathcal{E}$. L'état macroscopique du système est alors caractérisé par les valeurs moyennes $\langle A \rangle$ des observables A , selon*

$$\langle A \rangle = \int_{\mathcal{E}} A(\mathbf{Q}, \mathbf{P}) d\rho_N(\mathbf{Q}, \mathbf{P}). \quad (2.13)$$

Ces moyennes sont appelées moyennes d'ensemble.

On rappelle qu'une mesure de probabilité est une mesure de Radon positive de masse totale égale à 1. Ceci signifie que

$$\int_{\mathcal{E}} d\rho_N(\mathbf{Q}, \mathbf{P}) = 1,$$

et, pour tout borélien $B \in \mathcal{B}(\mathcal{E})$,

$$0 \leq \int_B d\rho_N(\mathbf{Q}, \mathbf{P}) \leq 1.$$

Une telle mesure de probabilité peut admettre une densité f_N , auquel cas

$$d\rho_N(\mathbf{Q}, \mathbf{P}) = f_N(\mathbf{Q}, \mathbf{P}) d\mathbf{Q} d\mathbf{P}, \quad f_N \geq 0, \quad \int_{\mathcal{E}} f_N(\mathbf{Q}, \mathbf{P}) d\mathbf{Q} d\mathbf{P} = 1,$$

⁸. La période des oscillations de $F(k, t)$ sur la figure 2.4 est celle des ondes de compression – ondes acoustiques longitudinales – de longueur d'onde L .

mais ce n'est pas forcément le cas. On pourrait imaginer que $d\rho_N(\mathbf{Q}, \mathbf{P})$ est une masse de Dirac concentrée en un point, ou plus généralement, une mesure restreinte à une sous-variété de l'espace. Un tel exemple sera présenté au § 3.1.1.

Notons également que le point de vue probabiliste est naturel en pratique pour un observateur qui ne dispose que d'une information partielle sur le mouvement des N particules et à qui les fluctuations, comme celles de N_2 dans l'exemple du § 2.3, semblent erratiques. Plus fondamentalement, la sensibilité aux conditions initiales interdit de prévoir les trajectoires dans $\mathcal{E}$, en dépit du déterminisme exprimé par (2.2), car les valeurs initiales des $2dN$ coordonnées de $(\mathbf{Q}, \mathbf{P})$ ne sont jamais connues qu'avec une précision finie. Dans la pratique usuelle du calcul numérique, la précision relative sur un nombre réel (codé sur 8 octets) est de l'ordre de 10^{-16} . En fait, pour certains systèmes dynamiques chaotiques analogues à (2.2), on peut montrer⁹ qu'il est impossible de distinguer :

- (i) les trajectoires exactes des équations déterministes pour une certaine condition initiale ;
- (ii) les trajectoires des équations déterministes calculées avec des erreurs relatives de 10^{-16} ;
- (iii) les trajectoires exactes d'un modèle modifié, un terme de bruit d'ordre relatif 10^{-16} étant ajouté dans les équations déterministes.

Ce résultat est appelé *lemme de poursuite* dans [11], car la trajectoire fautive que l'on calcule est en fait une vraie trajectoire pour une certaine condition initiale tellement proche de celle que l'on a prise qu'on ne la distingue pas avec le niveau de précision dont on dispose. Pour de tels systèmes dynamiques chaotiques, la distinction entre évolution déterministe ou aléatoire n'est donc qu'affaire de point de vue.

2.4.1 Approche de l'équilibre, moyennes et corrélations en temps

On observe en général lorsque l'on prépare un système dans un état particulier, comme dans l'exemple précédent où toutes les particules sont initialement dans la moitié du volume accessible, que cette situation de départ est oubliée après un certain temps, au-delà duquel tout semble indiquer que le macroétat du système, c'est-à-dire la mesure de probabilité $d\rho_N(\mathbf{Q}, \mathbf{P})$ des microétats, est indépendante du temps. Un tel macroétat est par définition un état d'équilibre au sens thermodynamique.

Dans la pratique, lorsque l'on simule le mouvement de N particules, on repère un instant t_0 où l'influence de l'état initial est devenue négligeable (on peut par exemple tester la stationnarité de certaines observables, considérées comme processus aléatoires, voir § 5.3) et on remplace la moyenne d'ensemble (2.13) par une moyenne en temps

$$\langle A \rangle = \frac{1}{\tau} \int_{t_0}^{t_0+\tau} A(\phi_t(\mathbf{Q}, \mathbf{P})) dt, \quad (2.14)$$

dans laquelle on doit prendre τ assez grand pour avoir un échantillon représentatif de la statistique des fluctuations au cours du temps. La propriété (2.14) est la définition pratique de la moyenne selon le macroétat d'équilibre. Dire que les systèmes macroscopiques approchent spontanément d'un état d'équilibre, c'est dire qu'il existe un macroétat ρ_N^{eq} tel que, pour toute observable A , (2.14) coïncide avec (2.13) pour τ, t_0 assez grands. On fait en sorte une hypothèse d'*ergodicité*.

Une indication sur les échelles de temps concernées est fournie par les fonctions de corrélation dépendant du temps. Si A et B sont deux observables, on appelle fonction de corrélation de A et de B la quantité

$$C_{AB}(t) = \langle A(0)B(t) \rangle - \langle A \rangle \langle B \rangle, \quad (2.15)$$

où $B(t) = B(\phi_t(\mathbf{Q}_0, \mathbf{P}_0))$, la moyenne étant prise par exemple sur toutes les réalisations de la dynamique, partant de conditions initiales $(\mathbf{Q}_0, \mathbf{P}_0)$ distribuées selon $d\rho_N$. Si on admet que les valeurs prises par les différentes observables sont indépendantes à des temps éloignés (du fait de

9. Nous recommandons la lecture de l'article d'Adrien Douady [11] qui illustre ces propriétés sur un modèle simple.

l'oubli des conditions initiales), alors $C_{AB}(t)$ tend vers zéro pour $t \rightarrow +\infty$. Une condition nécessaire pour que les valeurs prises à t et à $t + \tau$ puissent être considérées comme indépendantes est que l'on ait $\tau \gg \tau_r$, où τ_r est le temps caractéristique de décroissance de C_{AB} . On doit donc prendre $\tau \gg \tau_r$ dans (2.14).

2.4.2 Limite thermodynamique, moyennes et corrélations dans l'espace

C'est une constatation d'expérience que les observables ayant un sens macroscopique, comme la pression ou l'énergie interne, et qui font intervenir un grand nombre de degrés de liberté du système (à l'instar du nombre N_2 dont l'évolution est rapportée au § 2.3) ne s'écartent que peu de leur valeur moyenne, avec des fluctuations qui tendent à diminuer, en valeur relative, comme $O(N^{-1/2})$ lorsque le nombre de particules N augmente. Bien souvent, de telles observables A peuvent s'écrire, au moins approximativement,

$$A = \sum_{k=1}^p A_k, \quad (2.16)$$

où A_k représente la contribution d'un sous-système, c'est-à-dire un sous-ensemble des degrés de liberté du système entier, correspondant à une liste donnée de molécules, ou bien à celles des molécules qui à l'instant considéré se trouvent dans une partie donnée du volume entier du système. Les systèmes que l'on étudie sont constitués d'un grand nombre de particules identiques, et il est possible si on procède à un découpage du volume de prendre des sous-domaines de même taille et de même forme. On s'attend alors à ce que chacune des p observables partielles A_k « joue le même rôle », et par conséquent, si on les considère comme des variables aléatoires, elles soient *équidistribuées*. Par exemple, pour le nombre N_2 de molécules dans la moitié du volume accessible on peut découper la moitié de la cellule en p petites cellules identiques, et écrire N_2 comme la somme des nombres de particules contenues dans chacune de ces petites cellules.

L'approche consistant à définir un sous-système par la liste des molécules qui lui sont attribuées est très commode quand on a affaire à des degrés de liberté indépendants. On a alors $\mathcal{E} = \mathcal{E}_1 \times \mathcal{E}_2 \times \cdots \times \mathcal{E}_p$, et le microétat est ainsi un p -uplet $(\mathbf{Q}, \mathbf{P}) = (\mathbf{Q}_k, \mathbf{P}_k)_{1 \leq k \leq p}$ avec $(\mathbf{Q}_k, \mathbf{P}_k) \in \mathcal{E}_k$ (pour $1 \leq k \leq p$). On peut également décomposer le Hamiltonien comme une somme de Hamiltoniens partiels :

$$H(\mathbf{Q}, \mathbf{P}) = \sum_{k=1}^p H_k(\mathbf{Q}_k, \mathbf{P}_k). \quad (2.17)$$

Il est immédiat dans ce cas que le mouvement de chacun des $(\mathbf{Q}_k, \mathbf{P}_k)$ est indépendant de tous les autres. On aura alors indépendance statistique entre les fluctuations de chacune des observables partielles. L'application du Théorème de la Limite Centrale à (2.16), somme de variables indépendantes équidistribuées, assure alors (sachant que le nombre p augmente proportionnellement à N) que A est une variable gaussienne de variance

$$\langle (A - \langle A \rangle)^2 \rangle \propto N.$$

Toutefois, le cas d'un Hamiltonien se décomposant en une somme de termes faisant intervenir des degrés de liberté indépendants est exceptionnel et correspond aux rares systèmes physiques, comme le gaz parfait, pour lesquels la physique statistique fournit des prédictions exactes.

En général, une décomposition de la forme (2.17) est impossible à écrire exactement, mais on peut obtenir une partition approximative en découplant le volume total en sous-parties, et en faisant une hypothèse d'indépendance à grande distance. En effet, si les sous-systèmes sont eux-mêmes assez grands, deux sous-systèmes adjacents ne sont corrélés que par l'intermédiaire des degrés de liberté des molécules situées à la proximité de la frontière, dont la contribution décroît relativement lorsque la taille des sous-systèmes augmente. On peut alors formellement s'attendre à ce qu'un résultat du type Théorème de la Limite Centrale s'applique, et donc à ce que les fluctuations relatives par rapport aux valeurs moyennes soient en $O(N^{-1/2})$. Une indication utile

pour estimer grossièrement l'importance des fluctuations résiduelles dans un échantillon de taille finie est celle de la *longueur de corrélation*, longueur caractéristique de décroissance vers zéro de l'amplitude des corrélations entre les observables qui dépendent des degrés de liberté des molécules au voisinage d'un point donné dans le volume occupé par le système. Pour considérer que des sous-systèmes associés à un découpage du volume sont approximativement indépendants, une condition nécessaire est que leur taille soit grande devant la longueur de corrélation.

On appelle *limite thermodynamique* la limite dans laquelle le nombre de particules N tend vers l'infini, ainsi que le volume $V = |\mathcal{D}|$ du système, tandis que la densité numérique $\rho = N/V$ est maintenue constante. Dans la limite thermodynamique, les observables du type (2.16) qui résultent de la somme de contributions locales peuvent être considérées comme fixées, et on retrouve alors la description déterministe de la thermodynamique pour les systèmes macroscopiques. De telles grandeurs, qui ont une moyenne croissant proportionnellement à N , sont appelées grandeurs *extensives*. Les grandeurs caractérisant l'état thermodynamique du système et possédant une limite thermodynamique finie sont dites *intensives* (pression, température, densité par exemple).

En général, il est très difficile de prédire le comportement des systèmes physiques dans la limite thermodynamique. Les considérations heuristiques qui précèdent ne sont que des arguments de plausibilité de la *régression des fluctuations* dans cette limite (c'est-à-dire que les fluctuations des observables caractérisant l'état thermodynamique du système deviennent négligeables). Il convient toutefois de vérifier ces prédictions, ce dont on peut s'assurer par l'expérience et la simulation numérique. Les exceptions sont associées à des phénomènes physiques particuliers, comme les transitions de phase du second ordre.

2.4.3 Ensembles statistiques

Jusqu'ici, nous avons supposé que l'on observait l'évolution dans le temps d'un système unique, et que l'on pouvait éventuellement donner le sens de probabilités aux fréquences d'occurrence des microétats. Pour motiver et justifier le recours à une approche probabiliste, nous avons brièvement invoqué l'apparence erratique des fluctuations des grandeurs observables, et le caractère chaotique de la dynamique hamiltonienne d'un grand nombre de particules, c'est-à-dire son imprévisibilité pratique du fait de la sensibilité aux conditions initiales et à la loi d'interaction. La physique statistique s'est construite sur un *postulat fondamental* (voir § 3.1), et qui permet de prédire les macroétats d'équilibre (c'est-à-dire la mesure sur $\mathcal{E}$ correspondant à l'échantillonnage par la dynamique hamiltonienne dans la limite des temps longs). Grâce à ce postulat de départ, on a deux possibilités pour la détermination numérique des moyennes des observables :

- soit on simule directement le mouvement des particules avec la dynamique hamiltonienne, c'est la méthode de la *dynamique moléculaire* (voir chapitre 3) ;
- soit on cherche à tirer au sort des microétats selon la loi de probabilité prédite par le postulat, ce qui conduit aux *méthodes de Monte-Carlo* (voir chapitre 5).

Traditionnellement, la physique statistique fait appel à la notion d'*ensemble statistique*, qui est une manière de se représenter un macroétat ou une mesure de probabilité sur $\mathcal{E}$. L'idée d'ensemble statistique renvoie à une collection fictive de systèmes identiques, chacun représentant un tirage au sort selon une certaine loi de probabilité. On distingue différents types d'ensemble selon le modèle que l'on se donne pour les interactions du système étudié avec le reste de l'univers. En fin de compte, les propriétés thermodynamiques telles que l'équation d'état d'un fluide, par exemple, ne dépendent pas du récipient qui en contient un échantillon macroscopique. Aussi s'attend-on à ce que les différents ensembles statistiques, qui sont des constructions abstraites et commodes pour passer de l'échelle microscopique à l'échelle macroscopique, conduisent aux mêmes résultats physiques dans la limite thermodynamique.

À chacun des ensembles statistiques usuels correspond une mesure de probabilité d'équilibre différente sur $\mathcal{E}$. Les ensembles usuels sont caractérisés par les paramètres de contrôle qui sont utilisés pour décrire l'état du système. Comme on le verra, les macroétats d'équilibre correspondant à ces différents ensembles possèdent certaines propriétés de maximisation de l'entropie. On distingue ainsi les cas usuels suivants :

- (i) la situation d'un système isolé maintenu dans un volume fixe $\mathcal{D}$ (avec $V = |\mathcal{D}|$), composé de N particules et dont l'énergie totale E est constante, est décrite par l'*ensemble microcanonique* ou ensemble (N, V, E) . On verra que la mesure sur $\mathcal{E}$ qui lui correspond n'est pas la plus commode d'utilisation (chapitre 3) ;
- (ii) l'*ensemble canonique* décrit la situation d'un système de N particules dans un récipient de volume V fixé pouvant échanger de l'énergie avec un autre système appelé *thermostat*, qui fixe sa température T . On parle donc aussi d'ensemble (N, V, T) (voir chapitre 4). La température caractérise la propension du système à céder de l'énergie ;
- (iii) l'*ensemble grand-canonique*, ou ensemble (μ, V, T) , décrit la situation d'un système, en contact avec un thermostat, et constitué par des particules se trouvant dans un volume V donné. Ces particules sont cependant en nombre variable, des échanges ayant lieu avec un certain réservoir de particules. Le *potentiel chimique* μ mesure la propension du système à absorber des particules ;
- (iv) l'*ensemble isotherme-isobare*, enfin, ou ensemble (N, P, T) , décrit la situation du système en contact avec un thermostat et avec un réservoir de volume qui lui impose une certaine pression P .

Ces différents ensembles statistiques font intervenir des échanges (d'énergie, de volume, de particules) avec un système réservoir, supposé très grand, de telle sorte que la quantité d'énergie qu'il fournit au système étudié (ou le nombre de particules, ou le volume) est infime par rapport à sa taille et n'affecte pas son état. Le thermostat est un réservoir d'énergie par exemple.

Beaucoup d'observables, en particulier celles que l'on peut écrire comme somme de contributions locales (à la manière de (2.16)), ont, à des termes de fluctuation en $O(N^{-1/2})$ près, la même moyenne dans les différents ensembles statistiques, pourvu que le sens macroscopique du macroétat considéré soit le même. Toutefois leur variance, c'est-à-dire

$$\text{Var}(A) = (\Delta A)^2 = \langle (A - \langle A \rangle)^2 \rangle, \quad (2.18)$$

est différente. Par exemple, la variance de l'énergie totale est identiquement nulle dans l'ensemble microcanonique (puisque la valeur de $H(\mathbf{Q}, \mathbf{P})$ est une constante du mouvement), alors qu'elle n'est pas nulle dans les ensembles canonique, grand-canonique et isotherme-isobare, en raison des échanges d'énergie avec le thermostat.

2.5 Quelques observables courantes

2.5.1 L'énergie et ses fluctuations

L'observable *énergie* est donnée par la fonction Hamiltonien H . C'est une constante du mouvement dans l'ensemble microcanonique. Sa variance est alors nulle. Dans les ensembles canonique ou isotherme-isobare, cette variance est liée aux capacités calorifiques, selon ¹⁰

$$c_v = \frac{1}{N} \left(\frac{\partial \langle H \rangle}{\partial T} \right)_V = \frac{\text{Var}(H)_{\text{NVT}}}{N k_B T^2} \quad (2.19)$$

pour la capacité calorifique par particule à volume constant, évaluée dans l'ensemble canonique, ou

$$c_p = \frac{1}{N} \left(\frac{\partial \langle H \rangle}{\partial T} \right)_P = \frac{\text{Var}(H)_{\text{NPT}}}{N k_B T^2} \quad (2.20)$$

pour la capacité calorifique par particule à pression constante, évaluée dans l'ensemble isotherme-isobare. Dans le cas de l'ensemble microcanonique, on peut lier c_v aux fluctuations de l'énergie cinétique totale

10. Ces formules seront établies dans les chapitres suivants.

$$K = \sum_{i=1}^N \frac{p_i^2}{m_i},$$

selon

$$\frac{1}{N} \text{Var}(K)_{\text{NVE}} = \frac{3k_B^2 T^2}{2} \left(1 - \frac{3k_B}{2c_v}\right). \quad (2.21)$$

2.5.2 Le tenseur des contraintes

Considérons un état stationnaire (invariant dans le temps) et homogène (uniforme dans l'espace) d'un système de N particules de masse m , maintenues à l'intérieur du domaine $\mathcal{D} \subset \mathbb{R}^3$, de volume $|\mathcal{D}|$, et interagissant par un potentiel de paires $\mathcal{V}(r)$ qui ne dépend que des distances $r_{ij} = |\mathbf{q}_j - \mathbf{q}_i|$ entre les positions $\mathbf{q}_i$ ($1 \leq i \leq N$) des particules. On cherche à évaluer la valeur du tenseur des contraintes (de Cauchy) $\underline{\sigma}$. Pour simplifier, supposons que nous nous intéressons à un échantillon de forme cubique et de côté L , avec des conditions aux limites périodiques. Ainsi, $(x^1, x^2, x^3) \in [-L/2, L/2]^3$. Nous admettons de plus que la fonction potentiel $\mathcal{V}(r)$ est de portée finie r_c (*i.e.* $\mathcal{V}(r) = 0$ si $r \geq r_c$), et nous nous limitons¹¹ à $r_c < L/2$.

Fixant une direction $\alpha \in \{1, 2, 3\}$, on peut évaluer le flux $\mathbf{J} = (J_1, J_2, J_3)$ de quantité de mouvement, transportée en direction des x^α croissants, au travers du plan $x^\alpha = x_0^\alpha$ (flux *sortant* de la région $x^\alpha < x_0^\alpha$). Par définition du tenseur des contraintes,

$$J_\beta(x_0^\alpha) = -L^2 \sigma_{\beta\alpha}, \quad 1 \leq \alpha \leq 3.$$

Ce flux a deux origines physiques distinctes. Le mouvement des particules qui traversent la surface $x^\alpha = x_0^\alpha$ lui apporte une contribution $J_\beta^{(1)}$, à laquelle s'ajoute un autre terme $J_\beta^{(2)}$, dû aux forces que les particules situées d'un côté de la surface ($x^\alpha < x_0^\alpha$) exercent sur les particules situées de l'autre côté ($x^\alpha > x_0^\alpha$).

La stationnarité de la quantité de mouvement moyenne contenue dans une tranche quelconque $x_1^\alpha < x^\alpha < x_0^\alpha$ montre que

$$L^2 \sigma_{\beta\alpha} = -\frac{1}{L} \int_{-L/2}^{L/2} J_\beta(x^\alpha) dx^\alpha.$$

Dans cette intégrale, le terme $J_\beta^{(1)}(x^\alpha)$ ne contribue que pour les valeurs de x^α correspondant à la coordonnée q_i^α du centre d'une particule i . Le flux de quantité de mouvement correspondant est alors

$$\int_{-L/2}^{L/2} J_\beta^{(1)}(x^\alpha) dx^\alpha = \sum_{i=1}^N m v_i^\alpha \mathbf{v}_i.$$

Pour le second terme, notons d'abord que

$$\mathbf{J}^{(2)}(x_0^\alpha) = \sum_{\substack{i \text{ t.q. } x_0^\alpha - r_c < q_i^\alpha < x_0^\alpha \\ j \text{ t.q. } x_0^\alpha < q_j^\alpha < x_0^\alpha + r_c}} \mathbf{F}_{ij},$$

$\mathbf{F}_{ij}$ désignant la force exercée par la particule i sur la particule j . Dans le calcul de l'intégrale de $\mathbf{J}^{(2)}$ chaque paire (i, j) telle que $q_i^\alpha < q_j^\alpha$ contribue pour les x^α compris entre q_i^α et q_j^α . On a donc, en symétrisant les rôles de i et j ,

$$\int_{-L/2}^{L/2} \mathbf{J}^{(2)}(x^\alpha) dx^\alpha = \frac{1}{2} \sum_{i \neq j} \mathbf{F}_{ij} \mathbf{r}_{ij} \cdot \mathbf{e}_\alpha,$$

où la somme est étendue à tous les couples (i, j) avec $i \neq j$, et $\mathbf{e}_\alpha$ est le vecteur unitaire le long de l'axe des x^α . En rassemblant les différentes contributions, et en notant que le volume de la cellule est $L^3 = |\mathcal{D}|$, on trouve :

11. Cette condition assure que les conditions de bord périodiques ont un sens.

$$\underline{\underline{\sigma}} \cdot \mathbf{n} = -\frac{1}{|\mathcal{D}|} \left\langle \sum_{i=1}^N m \mathbf{v}_i \otimes \mathbf{v}_i + \frac{1}{2} \sum_{i \neq j} \mathbf{F}_{ij} \otimes \mathbf{r}_{ij} \right\rangle \cdot \mathbf{n}$$

pour tout vecteur $\mathbf{n}$ orthogonal à l'une des faces du cube, de sorte que les deux tenseurs du second ordre coïncident :

$$\underline{\underline{\sigma}} = -\frac{1}{|\mathcal{D}|} \left\langle \sum_{i=1}^N (m \mathbf{v}_i \otimes \mathbf{v}_i) + \frac{1}{2} \sum_{i \neq j} (\mathbf{F}_{ij} \otimes \mathbf{r}_{ij}) \right\rangle.$$

Dans cette formule, le vecteur $\mathbf{r}_{ij}$ désigne la différence $\mathbf{q}_j - \mathbf{q}_i$ entre les positions des copies les plus proches de i et de j par la périodicité des conditions aux limites. Comme la force dérive du potentiel $\mathcal{V}$, fonction de la distance $r_{ij} = |\mathbf{r}_{ij}|$, on peut également écrire :

$$\underline{\underline{\sigma}} = -\frac{1}{|\mathcal{D}|} \left\langle \sum_{i=1}^N (m \mathbf{v}_i \otimes \mathbf{v}_i) - \frac{1}{2} \sum_{i \neq j} \left(\frac{\mathcal{V}'(r_{ij})}{r_{ij}} \mathbf{r}_{ij} \otimes \mathbf{r}_{ij} \right) \right\rangle. \quad (2.22)$$

La pression moyenne $P = -\text{tr} \underline{\underline{\sigma}}/3$ est alors

$$P = \frac{1}{3|\mathcal{D}|} \left\langle \sum_{i=1}^N m \mathbf{v}_i^2 - \frac{1}{2} \sum_{i \neq j} r_{ij} \mathcal{V}'(r_{ij}) \right\rangle \quad (2.23)$$

Rappelons que dans un fluide simple à l'équilibre, on a toujours $\underline{\underline{\sigma}} = -P \underline{\underline{\mathbf{1}}}$. On montre également plus généralement que

$$P = \frac{1}{3|\mathcal{D}|} \left\langle \sum_{i=1}^N \frac{\mathbf{p}_i^2}{m_i} - \mathbf{q}_i \cdot \nabla_{\mathbf{q}_i} V(\mathbf{Q}) \right\rangle$$

On notera que ces formules peuvent se mettre sous la forme de sommes de contributions locales dans la mesure où le potentiel d'interaction de paire a une portée r_c . Par ailleurs, comme on normalise la moyenne par le volume de la cellule de simulation, on a une grandeur *intensive*.

En l'absence d'écoulement (vitesse moyenne des particules nulle), chaque coordonnée $1 \leq \alpha \leq 3$ de la vitesse d'une particule $1 \leq i \leq N$ est telle que (voir les chapitres suivants)

$$\langle (v_i^\alpha)^2 \rangle = \frac{k_B T}{m}. \quad (2.24)$$

Le premier terme de la pression donne donc $\rho k_B T$. Le gaz parfait est le fluide sans interactions pour lequel $\mathcal{V} = 0$, on retrouve donc son équation d'état $P = \rho k_B T$.

2.5.3 Microstructure et corrélations de densité

La macroétat $d\rho_N$, que l'on suppose symétrique vis-à-vis des permutations des N particules, contient toute l'information statistique sur le système, ce qui est énorme. Beaucoup d'observables d'intérêt pratique peuvent être moyennées sans faire appel à une connaissance aussi détaillée.

Densité à un corps

Si les observables que l'on cherche à calculer sont de la forme $A_1 = \sum_{i=1}^N \phi_1(\mathbf{q}_i)$, il n'est besoin, pour évaluer la moyenne $\langle A \rangle$, que de connaître la loi marginale de la position d'une particule, obtenue en prenant la marginale de $d\rho_N$ par rapport à toutes les variables $\mathbf{q}_2, \dots, \mathbf{q}_N, \mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_N$. On suppose que $d\rho_N$ est dominée par la mesure de Lebesgue,

$$d\rho_N(\mathbf{Q}, \mathbf{P}) = f_N(\mathbf{Q}, \mathbf{P}) d\mathbf{Q} d\mathbf{P},$$

et que la densité f_N est régulière. Multipliant par N et exploitant la symétrie, on définit la *densité à un corps* $\rho^{(1)}(\mathbf{r})$ comme

$$d\rho^{(1)}(\mathbf{r}) = N \int f_N(\mathbf{r}, \mathbf{q}_2, \dots, \mathbf{q}_N, \mathbf{p}_1, \dots, \mathbf{p}_N) d\mathbf{q}_2 \dots d\mathbf{q}_N d\mathbf{P}.$$

On a alors $\int_{\mathcal{D}} \rho^{(1)}(\mathbf{r}) d\mathbf{r} = N$ et $\langle A_1 \rangle = \int_{\mathcal{D}} \phi_1(\mathbf{r}) \rho^{(1)}(\mathbf{r}) d\mathbf{r}$. Dans un système homogène (les fluides simples), $\rho^{(1)}(\mathbf{r})$ est constante et coïncide avec sa moyenne spatiale, qui est simplement la densité (numérique) $\rho = N/V$. Dans un solide, $\rho^{(1)}(\mathbf{r})$ présente des maxima bien marqués au voisinage des sites du réseau (cristallin ou amorphe) des positions moyennes des particules.

Densité de paires

On peut généraliser cette approche à des densités à n corps. Toutefois, en pratique, on s'intéresse essentiellement à la densité à 2 corps, qui peut être mesurée expérimentalement par des techniques de diffusion de rayonnement ou de particules. Par ailleurs, un certain nombre d'observables résultent de la somme de contributions de paires de particules (comme l'énergie potentielle dans le cas d'un potentiel d'interactions par paires), et sont donc de la forme $A_2 = \sum_{1 \leq i, j \leq N, i \neq j} \phi_2(\mathbf{q}_i, \mathbf{q}_j)$. Leurs moyennes font apparaître la *densité à deux corps* $\rho^{(2)}(\mathbf{r}_1, \mathbf{r}_2)$, densité de la loi marginale de la paire de positions $(\mathbf{q}_1, \mathbf{q}_2)$, multipliée par $N(N-1)$ (pour raisons de symétries, tenant compte du nombre total de paires de $\mathcal{D} \times \mathcal{D}$) :

$$\langle A_2 \rangle = \int_{\mathcal{D} \times \mathcal{D}} \rho^{(2)}(\mathbf{r}_1, \mathbf{r}_2) \phi_2(\mathbf{r}_1, \mathbf{r}_2) d\mathbf{r}_1 d\mathbf{r}_2,$$

avec

$$\rho^{(2)}(\mathbf{r}_1, \mathbf{r}_2) = N(N-1) \int f_N(\mathbf{r}_1, \mathbf{r}_2, \mathbf{q}_3, \dots, \mathbf{q}_N, \mathbf{p}_1, \dots, \mathbf{p}_N) d\mathbf{q}_3 \dots d\mathbf{q}_N d\mathbf{P}.$$

On a $\int_{\mathcal{D} \times \mathcal{D}} \rho^{(2)}(\mathbf{r}_1, \mathbf{r}_2) d\mathbf{r}_1 d\mathbf{r}_2 = N(N-1)$. Dans le cas d'un système homogène, $\rho^{(2)}(\mathbf{r}_1, \mathbf{r}_2)$ ne dépend que de $\mathbf{r}_{12} = \mathbf{r}_2 - \mathbf{r}_1$, que l'on notera simplement $\mathbf{r}$. Si le système est de plus isotrope, alors la densité à deux corps $\rho^{(2)}(\mathbf{r}_1, \mathbf{r}_2)$ est une fonction scalaire de la norme $r = |\mathbf{r}|$ du vecteur $\mathbf{r} = \mathbf{r}_2 - \mathbf{r}_1$. On définit la *fonction de corrélation de paires* $g(r)$ par

$$\rho^{(2)}(\mathbf{r}_1, \mathbf{r}_2) = \rho^2 g(|\mathbf{r}_2 - \mathbf{r}_1|),$$

où ρ est la densité du système. La fonction $g(r)$ décrit les corrélations entre positions des particules, prises deux à deux. Si une particule est placée à l'origine des coordonnées, alors, à la distance r , la densité de particules n'est pas ρ mais $\rho g(r)$, c'est-à-dire que le nombre moyen de particules entre les distances r_1 et r_2 d'une particule de référence n'est pas $\int_{r_1}^{r_2} 4\pi r^2 \rho dr$, mais $\int_{r_1}^{r_2} 4\pi r^2 \rho g(r) dr$. Dans un fluide $g(r)$ tend vers 1 lorsque $r \rightarrow +\infty$. Si r est inférieur au diamètre du nuage électronique, à l'intérieur duquel la pénétration du centre d'une autre molécule est interdite par une très forte répulsion, on a $g(r) = 0$. Entre ces deux limites, $g(r)$ présente des oscillations caractéristiques de la structure d'un liquide.

La figure 2.5 présente ainsi la fonction $g(r)$ dans le liquide de sphères dures pour la fraction solide $\Phi = 0,49$. On note que $g(r)$ est discontinue en $r = \sigma$. Le premier maximum de $g(r)$ manifeste l'accumulation de particules dans une couche plus dense à une certaine distance caractéristique des premières voisines, qui compense partiellement leur exclusion autour du centre de la particule origine. Les oscillations, rapidement amorties, à plus grande distance traduisent une organisation locale en couches successives, de moins en moins marquée à mesure que l'on s'éloigne de la particule centrale. L'amplitude de ces oscillations est une fonction croissante de la densité (ou décroissante de la température).

Avec un potentiel fortement répulsif, mais fini, comme dans le cas du fluide de Lennard-Jones, plutôt qu'une répulsion « de cœur dur » (c'est-à-dire une stricte exclusion stérique à courte distance), $g(r)$ passe très rapidement de zéro à une valeur maximale correspondant à la distance des particules proches voisines lorsque r augmente, mais sans discontinuité, comme on le voit sur la figure 2.6.

Dans un solide cristallin, on peut également définir une fonction de corrélation de paires $g(r)$ analogue, à condition de moyenner $\rho^{(2)}(\mathbf{r}_1, \mathbf{r}_2)$ sur la position $\mathbf{r}_1$ et sur les orientations de $\mathbf{r}_{12}$:

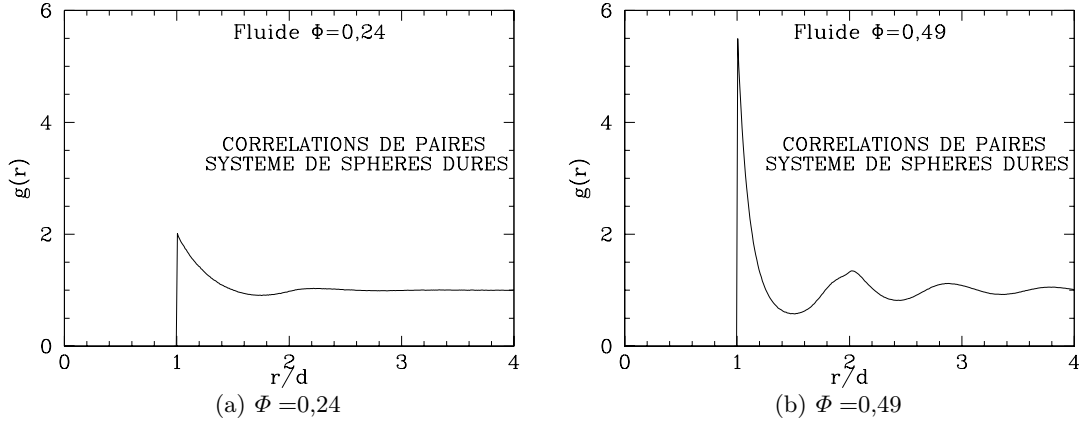


Fig. 2.5. Fonction de corrélation de paires dans le fluide de sphères dures pour deux valeurs de la fraction volumique, la plus élevée ($\Phi=0,49$) correspondant à la coexistence avec le cristal)

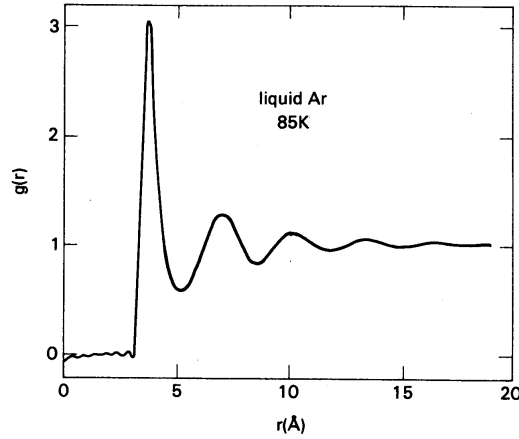


Fig. 2.6. Fonction de corrélation de paires de l'argon liquide au voisinage du point triple (résultat expérimental, cité dans [20])

$$g(r) = \frac{1}{4\pi\rho^2|\mathcal{D}|} \int_{\mathcal{D}} \int_{\mathbb{S}} \rho^{(2)}(\mathbf{r}, \mathbf{r} + \mathbf{R}(r, \Omega)) d\Omega d\mathbf{r}, \quad (2.25)$$

$\mathbf{R}(r, \Omega)$ désignant ici le vecteur de norme r et d'orientation $\Omega \in \mathbb{S}$. Cette fonction est représentée sur la figure 2.7, pour un cristal cubique à faces centrées de sphères dures. Les pics correspondent aux distances entre sites du réseau, et leur largeur est liée aux vibrations des atomes autour de leur position moyenne.

Calcul de quantités moyennes

Un exemple d'observable dont on peut évaluer la moyenne avec la fonction de corrélation de paires est l'énergie potentielle d'un système avec interactions de paires :

$$E_p = 2\pi\rho N \int_0^{+\infty} g(r) r^2 \mathcal{V}(r) dr,$$

tandis que l'énergie cinétique moyenne est d'après (2.6) égale à $3Nk_B T/2$. La pression, d'après (2.23) (et en utilisant (2.6)), est donnée par

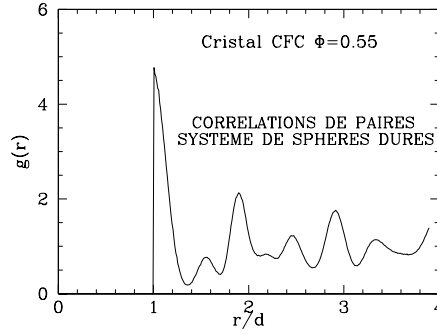


Fig. 2.7. Fonction de corrélation de paires dans le cristal de sphères dures pour la compacité $\Phi = 0,55$.

$$P = \rho k_B T - \frac{2\pi}{3} \rho \int_0^{+\infty} r^3 g(r) \mathcal{V}'(r) dr. \quad (2.26)$$

Le premier terme de (2.26) est la pression du gaz parfait, dans lequel les molécules sont sans interaction. Les d'interactions répulsives, $\mathcal{V}'(r) < 0$, tendent logiquement à augmenter la pression, et les interactions attractives, $\mathcal{V}'(r) > 0$, à la diminuer.

Facteur de structure

On définit aussi la *fonction de structure* (ou facteur de structure), fonction du vecteur d'onde $\mathbf{k}$ (dont la dimension est l'inverse d'une longueur), par

$$S(\mathbf{k}) = 1 + \rho \int_{\mathbb{R}^3} [g(r) - 1] \exp(-i\mathbf{k} \cdot \mathbf{r}) d\mathbf{r},$$

lorsque $|g(r) - 1|$ décroît suffisamment vite vers zéro. Le facteur de structure $S(\mathbf{k})$ est la transformée de Fourier de $\delta_{\mathbf{r}=0} + \rho(g(r) - 1)$, c'est-à-dire de la densité de paires « avec répétition »

$$\tilde{\rho}^{(2)}(\mathbf{r}_1, \mathbf{r}_2) = \rho^{(2)}(\mathbf{r}_1, \mathbf{r}_2) + \rho^{(1)}(\mathbf{r}_1) \delta(\mathbf{r}_2 - \mathbf{r}_1). \quad (2.27)$$

L'isotropie entraîne que $S(\mathbf{k})$ ne dépend que de $k = |\mathbf{k}|$, et

$$S(k) = 1 + 4\pi\rho \int_0^{+\infty} r \frac{\sin(kr)}{k} [g(r) - 1] dr. \quad (2.28)$$

Cette définition coïncide avec (2.12). Le facteur de structure est directement accessible par des techniques expérimentales de diffusion de rayonnement X ou de neutrons. Il vérifie

$$\lim_{k \rightarrow 0} S(k) = \rho k_B T \chi, \quad (2.29)$$

Cette formule, qui relie corrélations de densité et réponse à une sollicitation extérieure, fait intervenir la compressibilité isotherme, $\chi = \frac{1}{\rho} \left(\frac{\partial \rho}{\partial P} \right)_T$, qui apparaît dans les relations (2.8) et (2.9).

Les expériences de diffusion de rayons X ou de neutrons permettent de mesurer directement la fonction de structure, et c'est ainsi que la fonction de corrélation de paires représentée sur la figure 2.6 a été obtenue.¹²

12. Les petites oscillations qui apparaissent sur le graphe aux faibles distances, là où $g(r)$ devrait exactement s'annuler, sont des artefacts du calcul numérique de la transformée de Fourier inverse dans (2.28).

Exercices

Exercice 2.1 (Invariants de la dynamique hamiltonienne). Montrer que la quantité de mouvement totale est conservée si on prend des conditions de bord périodiques ou si le système n'est pas confiné, lorsque le potentiel d'interaction entre les particules est invariant par translation, au sens où, pour tout $\mathbf{r} \in \mathbb{R}^d$,

$$V(\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_N) = V(\mathbf{q}_1 + \mathbf{r}, \mathbf{q}_2 + \mathbf{r}, \dots, \mathbf{q}_N + \mathbf{r}).$$

Montrer que le moment cinétique total n'est pas conservé pour des conditions de bord périodiques, mais qu'il l'est si le système n'est pas confiné, à condition que le potentiel soit invariant par rotation, au sens où, quelle que soit la rotation $\mathcal{R} \in SO_3(\mathbb{R})$,

$$V(\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_N) = V(\mathcal{R}\mathbf{q}_1, \mathcal{R}\mathbf{q}_2, \dots, \mathcal{R}\mathbf{q}_N).$$

Dans les deux cas, on pourra commencer par traiter le cas simple d'un potentiel de paire ne dépendant que de la distance.

Exercice 2.2. On considère deux systèmes hamiltoniens du type (2.2), constitué chacun du même nombre N de particules, avec le même potentiel d'interaction. La masse des particules est m_1 dans le premier et m_2 dans le second. Comparer les flots Φ_t^1 et Φ_t^2 correspondants. Deux tels systèmes peuvent être, en pratique, l'un de molécules de méthane CH_4 , l'autre de méthane deutérié, CD_4 (le deutérium est l'isotope de l'hydrogène comportant un neutron dans le noyau). Qu'en déduit-on sur leurs propriétés thermodynamiques ?

Exercice 2.3 (Cas des potentiels en puissance inverse). On considère la dynamique (2.2), avec le potentiel d'interaction de paire des « sphères molles », donné par (1.3).

- (1) Quelle est la condition sur l'exposant n pour que l'énergie reste une grandeur extensive ?
- (2) En général, dans un fluide, les grandeurs thermodynamiques et les moyennes des observables ne dépendent que de la densité ρ et de l'énergie par particule, $e = E/N$. Montrer dans ce cas que $P/(\rho k_B T)$, $e/(k_B T)$ et, généralement, toute observable sans dimension, ainsi que la fonction de corrélation de paire fonction de la distance exprimée en unités de $\rho^{-1/3}$, ne dépendent que de la combinaison $\rho e^{-3/n}$.

Pour les besoins de l'exercice on définira la température T à partir de l'énergie cinétique moyenne :

$$\langle E_c \rangle = \frac{3}{2} N k_B T.$$

Postulat fondamental et ensemble microcanonique

3.1	Le postulat de base et ses conséquences	34
3.1.1	Macroétat d'équilibre	34
3.1.2	Cas du gaz parfait	35
3.1.3	Distribution de Maxwell (cas général)	37
3.2	Entropie, température	38
3.2.1	Définition	38
3.2.2	Propriétés de l'entropie	39
3.2.3	Couplage thermique	40
3.2.4	Les dérivées de la fonction entropie	40
3.2.5	Entropie et fluctuations	41
3.2.6	Discussion	42
3.3	Intégration des équations du mouvement et échantillonnage de la mesure microcanonique	43
3.3.1	Dynamique hamiltonienne et symplecticité	43
3.3.2	Approximation et flot numérique	44
3.3.3	Schémas d'intégration usuels pour les équations différentielles ordinaires	45
3.3.4	Schémas d'intégration préservant des propriétés de la dynamique hamiltonienne	48
3.3.5	Comparaison numérique des différents schémas	55
3.3.6	Application à un système physique	56
	Exercices	59

Nous avons souligné au chapitre 2 la pertinence d'une description statistique des fluctuations des grandeurs observables associées au mouvement conservatif d'une collection de particules en interaction. La notion d'équilibre thermodynamique, associée à la stationnarité des processus modélisant l'évolution dans le temps des observables d'un système isolé, fait naturellement apparaître un *macroétat*, qui est une mesure de probabilité sur l'espace des phase $\mathcal{E}$. Or, jusqu'ici, nous ne savons évaluer que des moyennes sur une trajectoire (ce qui suppose un calcul numérique), au lieu de moyennes d'ensemble (en référence à la notion d'« ensemble statistique » évoquée au § 2.4.3).

Le *postulat fondamental* sur lequel s'est construite la physique statistique donne le macroétat d'équilibre, et permet de prédire directement certains résultats sans passer par le calcul numérique. De plus, c'est à partir de la description statistique, en termes de macroétats, que l'on peut définir les notions essentielles d'entropie, de température et de transfert de chaleur, grandeurs dont les valeurs macroscopiques ne s'obtiennent pas en général par la moyenne d'une observable appropriée (elles n'ont pas de sens pour un microétat).

Ce chapitre énonce ce postulat de base, et en donne certaines conséquences immédiates (§ 3.1). Il rappelle ensuite comment on retrouve l'entropie et la température (§ 3.2), et enfin discute de la mise en œuvre numérique de la simulation par la technique de la dynamique moléculaire, qui

consiste à intégrer les équations du mouvement (§ 3.3). On verra que le choix d'un algorithme approprié est dicté par des considérations de stabilité, du fait de la nécessité d'un échantillonnage par intégration sur des temps très longs.

3.1 Le postulat de base et ses conséquences

3.1.1 Macroétat d'équilibre

La mesure d'équilibre doit être invariante par le flot ϕ_t de la dynamique hamiltonienne, et rester concentrée sur la sous-variété de $\mathcal{E}$ correspondant aux invariants de la dynamique (ou constantes du mouvement) : énergie E , éventuellement quantité de mouvement totale et moment cinétique (selon les conditions aux limites utilisées). On suppose par la suite que seule l'énergie est conservée dans le mouvement. On note

$$\mathcal{M}(E) = \{(\mathbf{Q}, \mathbf{P}) \in \mathcal{E}, \quad H(\mathbf{Q}, \mathbf{P}) = E\}$$

la sous-variété de $\mathcal{E}$ des états d'énergie égale à E . Notons que les sous-variétés $\{\mathcal{M}(E), E \in \mathbb{R}\}$ forment une partition disjointe de $\mathcal{E}$, au sens où

$$\mathcal{E} = \bigcup_{E \in \mathbb{R}} \mathcal{M}(E), \quad \mathcal{M}(E) \cap \mathcal{M}(E') = \emptyset \text{ si } E \neq E'.$$

Le postulat fondamental stipule que la sous-variété correspondant à la valeur constante E de l'énergie, est entièrement explorée par le système dans son mouvement, et que le macroétat d'équilibre s'écrit

$$d\rho_{\text{NVE}}(\mathbf{Q}, \mathbf{P}) = \frac{1}{\Omega(N, V, E)} \delta_{H(\mathbf{Q}, \mathbf{P}) - E} = \frac{1}{\Omega(N, V, E)} \frac{d\sigma_{\mathcal{M}(E)}(\mathbf{Q}, \mathbf{P})}{|\nabla H(\mathbf{Q}, \mathbf{P})|}, \quad (3.1)$$

$d\sigma_{\mathcal{M}(E)}(\mathbf{Q}, \mathbf{P})$ étant la mesure de Lebesgue induite sur la sous-variété $\mathcal{M}(E)$, et $\Omega(N, V, E)$ une constante de normalisation (de sorte que $d\rho_{\text{NVE}}$ est bien une mesure de probabilité) :

$$\Omega(N, V, E) = \int_{\mathcal{M}(E)} \frac{d\sigma_{\mathcal{M}(E)}(\mathbf{Q}, \mathbf{P})}{|\nabla H(\mathbf{Q}, \mathbf{P})|}.$$

L'aire de l'hypersurface $\mathcal{M}(E)$ dans $\mathcal{E}$ est appelée *fonction de partition microcanonique*. On peut comprendre la définition (3.1) comme la densité de probabilité uniforme par rapport à la mesure de Lebesgue dans le domaine compris entre les hypersurfaces $\mathcal{M}(E)$ et $\mathcal{M}(E + \Delta E)$, dans la limite $\Delta E \rightarrow 0$. On dit traditionnellement que « tous les microétats de même énergie sont équiprobables ».

Ainsi, le postulat fondamental revient à faire l'hypothèse d'*ergodicité* suivante :

ERGODICITÉ SELON UNE TRAJECTOIRE HAMILTONIENNE

Hypothèse 3.1. La moyenne par rapport à la mesure microcanonique d'une observable A sur $\mathcal{E}$ est donnée par

$$\lim_{T \rightarrow +\infty} \frac{1}{T} \int_0^T A(\phi_t(\mathbf{Q}_0, \mathbf{P}_0)) dt = \int_{\mathcal{E}} A(\mathbf{Q}, \mathbf{P}) d\rho_{\text{NVE}}(\mathbf{Q}, \mathbf{P}), \quad (3.2)$$

où $E = H(\mathbf{Q}_0, \mathbf{P}_0)$, la mesure $d\rho_{\text{NVE}}$ est donnée par (3.1), et (ϕ_t) est le flot de la dynamique hamiltonienne (2.2).

On ne peut montrer rigoureusement une telle égalité que dans quelques cas (les systèmes intégrables, *i.e.* qui peuvent être résolus analytiquement). En revanche, il existe des contre-exemples simples à cette propriété... qui montrent donc qu'il faut être prudent, et éventuellement modifier la définition de la sous-variété explorée $\mathcal{M}(E)$ en la séparant en ses différentes composantes connexes (voir la Remarque 3.1).

Remarque 3.1 (Un cas de non-ergodicité). On considère le potentiel unidimensionnel

$$V(q) = h(q+1)^2(q-1)^2. \quad (3.3)$$

Ce potentiel a deux minima globaux $V(\pm 1) = 0$ en ± 1 , et un maximum local $V(0) = h$ en 0. On a $V'(1) = V'(0) = V'(-1) = 0$ et $V''(-1) = V''(1) = 8h$, $V''(0) = -4h$. Remarquons que les sous-variétés $\mathcal{M}(E)$ pour $E < h$ sont formées de deux composantes connexes. Partant du voisinage d'un des deux minima d'énergie, on n'explore qu'un des deux puits, et on ne peut donc pas explorer toute la surface d'énergie $\{H(q, p) = E < h\}$ (voir la figure 3.1).

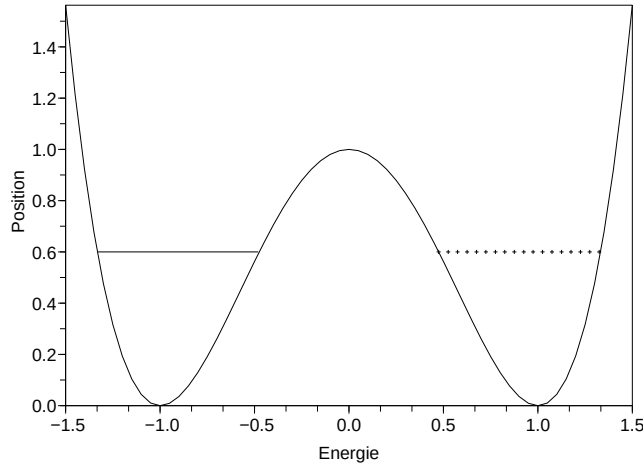


Fig. 3.1. Positions accessibles dans la surface d'énergie $H(q, p) = 0.6$, pour le potentiel (3.3) avec $h = 1$ et une particule de masse $m = 1$. Noter qu'il y a deux composantes connexes et que si la dynamique hamiltonienne est initialement dans une des composantes, elle y reste.

Remarque 3.2 (Non-ergodicité et comportement des grands systèmes.). On constate toutefois en pratique qu'en présence d'un grand nombre de particules, les situations telles que celles de la Remarque 3.1 tendent à disparaître. Même si, en moyenne, l'énergie cinétique d'une particule est insuffisante pour qu'elle puisse franchir certaines barrières de potentiels, les échanges avec le reste du système peuvent lui autoriser des fluctuations suffisantes pour autoriser de tels mouvements. Dans les grands systèmes, les violations manifestes du postulat fondamental sont en général associées à des phénomènes physiques particuliers, qui méritent une étude spécifique. Ce sont des barrières *entropiques* (et non énergétiques comme dans la Remarque 3.1) qui empêchent un échantillonnage correct.

Un exemple célèbre concerne les transitions de phase, dans lesquelles un choix s'opère de la direction d'un certain « paramètre d'ordre » – comme l'orientation de molécules allongées dans le cas de la transition isotrope/nématique citée au § 1.3. Le temps de relaxation des fluctuations de cette orientation augmente alors exponentiellement avec la taille du système. La perte d'ergodicité survient aussi dans le cas de la *transition vitreuse*, qui est une augmentation brutale du temps de relaxation des fluctuations de la structure d'un fluide lorsqu'il est refroidi et comprimé trop vite pour que les particules puissent s'organiser sur le motif cristallin correspondant à la phase thermodynamiquement stable.

3.1.2 Cas du gaz parfait

En général on ne sait pas calculer $\Omega(N, V, E)$, ni évaluer la mesure d'une partie de $\mathcal{E}$ selon (3.1). Une exception est le cas du gaz parfait, pour lequel les molécules sont sans interaction. Dans ce

cas, pour un gaz confiné dans un volume $V = |\mathcal{D}|$ donné,

$$H(\mathbf{Q}, \mathbf{P}) = \sum_{i=1}^N \frac{\mathbf{p}_i^2}{2m}.$$

On a alors

$$\Omega(N, V, E) = \frac{V^N}{h^{3N} N!} \Sigma(3N, \sqrt{2mE}), \quad (3.4)$$

où on a noté $\Sigma(n, R)$ l'aire de la sphère de rayon R dans $\mathbb{R}^n$:

$$\Sigma(n, R) = \frac{n\pi^{n/2} R^{n-1}}{\Gamma(1 + \frac{n}{2})}, \quad \Gamma(x) = \int_0^{+\infty} t^{x-1} \exp(-t) dt.$$

La formule de Stirling montre que $\Gamma(x) \sim \sqrt{2\pi} x^{x-\frac{1}{2}} \exp(-x)$ lorsque $x \rightarrow +\infty$.

Distribution des positions

La mesure (3.1) assure l'uniformité de la densité de $\mathbf{Q} = (\mathbf{q}_i)_{1 \leq i \leq N}$ dans l'espace des configurations V^N . Ainsi, autour d'un point $\mathbf{Q}$ de V^N tel que chaque $\mathbf{q}_i$ appartienne à une moitié donnée du volume accessible, la probabilité de présence est la même qu'au voisinage de $\mathbf{Q}$ tel que chaque volume moitié contienne environ la moitié des particules. Cela n'est pas contradictoire avec l'homogénéité macroscopique de la distribution des molécules dans le volume accessible, car les microétats qui contreviennent nettement à une telle homogénéité correspondent à un volume très faible dans l'espace des configurations $\mathcal{D}^N$. En effet, les valeurs de $\mathbf{Q}$ telles que chaque particule soit dans une moitié donnée du domaine accessible ne constituent que la fraction 2^{-N} du volume total dans l'espace des configurations $\mathcal{D}^N$, négligeable devant le nombre $C_N^{N/2}$ des configurations ayant exactement $N/2$ particules dans chaque moitié du volume.

Distribution des impulsions

D'après le macroétat d'équilibre pour le gaz parfait, on a une expression de la densité de probabilité de la quantité de mouvement d'une particule donnée :

$$P(\mathbf{p}) = V h^{-3} \frac{\Omega(N-1, V, E - \frac{\mathbf{p}^2}{2m})}{\Omega(N, V, E)}.$$

En utilisant (3.4) et la formule de Stirling (lorsque $N \rightarrow +\infty$),

$$P(\mathbf{p}) = \frac{1}{(2\pi m k_B T)^{3/2}} \exp \left[-\frac{\mathbf{p}^2}{2m k_B T} \right], \quad (3.5)$$

formule dans laquelle T (la température) est fixée par l'énergie selon

$$E = \frac{3}{2} N k_B T. \quad (3.6)$$

La densité (3.5) définit la *distribution de Maxwell* pour la quantité de mouvement d'une particule. On retrouve ainsi la formule (2.6) de la vitesse quadratique moyenne.

Cas des sphères dures

Dans le cas du modèle de sphères dures, (3.1) nous donne directement aussi un résultat d'uniformité de la probabilité dans la partie de l'espace des configurations telle qu'aucune paire de particules ne s'interpénètre, la distribution des impulsions étant la même que pour le gaz parfait.

Dans (3.4), il faut remplacer $V^N = |\mathcal{D}|^N$ par la mesure Q , sensiblement plus petite, du domaine de $\mathcal{D}^N$ dans lequel les sphères ne s'interpénètrent pas. On a :

$$\Omega(N, V, E) = \frac{Q}{h^{3N} N!} \Sigma(3N, \sqrt{2mE}). \quad (3.7)$$

Une approximation de Q est, pour les faibles densités,

$$Q \simeq V^N (1 - 4\Phi)^N, \quad (3.8)$$

où la compacité Φ définie par (2.7) fait intervenir le volume exclu¹ autour du centre de chaque sphère.

Notons aussi qu'en raison de la factorisation (3.7) de la fonction de partition microcanonique, la distribution de Maxwell pour les vitesses ou les impulsions s'applique directement aussi dans le cas des sphères dures.

3.1.3 Distribution de Maxwell (cas général)

Pour un système avec des interactions à distance, l'énergie totale E se partage entre énergie cinétique E_c dépendant de $\mathbf{P}$ et énergie potentielle E_p , dépendant de $\mathbf{Q}$.

La distribution de Maxwell s'applique quelles que soient les interactions des particules, et n'est pas limitée au gaz parfait. On verra au chapitre 4 que si on considère l'ensemble canonique, ce résultat est immédiat. Dans le cas microcanonique, on peut s'en convaincre en notant qu'aussi bien l'énergie potentielle que l'énergie cinétique sont des variables extensives, sommes de multiples contributions locales, et qu'elles fluctuent très peu. Si on les considère comme fixées à leurs valeurs les plus probables, on est ramené au cas précédent.

Remarquons également, toujours dans la limite des grands systèmes, l'indépendance statistique entre impulsions et positions des particules, et l'indépendance entre différentes coordonnées de $\mathbf{P}$. Cette indépendance contraste avec les corrélations des positions des particules voisines, décrites en partie par la fonction de corrélation de paires introduite au § 2.5.3.

Ordres de grandeur

Pour une température de l'ordre de 300 K, selon (3.5) ou (2.6), la vitesse totale en moyenne quadratique

$$\sqrt{\left\langle \frac{\mathbf{p}^2}{m^2} \right\rangle} = \sqrt{\frac{3k_B T}{m}}$$

est de l'ordre de 600 m/s pour la molécule d'eau H_2O (masse molaire 18 g/mol), 500 m/s pour la molécule d'azote N_2 (masse molaire 28 g/mol), 470 m/s pour la molécule d'oxygène O_2 (masse molaire 32 g/mol).²

Limite classique de la mécanique quantique pour un ensemble de particules

Les phénomènes de petite échelle échappent à la mécanique classique et doivent être traités par la physique quantique. Anticipant sur la seconde partie de ce cours, citons la fameuse *relation d'indétermination de Heisenberg*. Pour une particule donnée, on ne saurait déterminer exactement

1. En fait le centre d'une sphère est exclu de l'intérieur de la sphère dont le *rayon* est σ , c'est-à-dire le *diamètre* d'une sphère dure. Le volume exclu total est donc $4\Phi V$ si V est le volume du système.

2. Ces valeurs élevées des vitesses instantanées ne signifient pas que, dans un gaz ou un liquide, les molécules parcourent rapidement des distances importantes. En effet, le mouvement des molécules, après un certain temps τ d'oubli de la vitesse initiale, sous l'effet des interactions, apparaît comme erratique et est bien décrit comme un mouvement brownien, de sorte que la distance parcourue pendant le temps $t \gg \tau$ est aléatoire et d'ordre $\sqrt{t}$. Ce phénomène est étudié au chapitre 6.

sa coordonnée q_α et sa quantité de mouvement dans la même direction, p_α . Les incertitudes Δq_α et Δp_α sur les composantes q_α et p_α de la position et de la quantité de mouvement d'une particule sont liées par

$$\Delta q_\alpha \Delta p_\alpha \geq \frac{\hbar}{2}. \quad (3.9)$$

En raison de ces effets quantiques, la notion même de trajectoire classique est dénuée de sens à petite échelle. Pour qu'un traitement classique reste justifié, il faut que l'on puisse décrire le mouvement de la particule avec, simultanément

- Δq_x très petit devant la distance typique $\rho^{-1/3}$ entre deux particules voisines ;
- Δp_x très petit devant la valeur typique $mk_B T$ de p_α .

Une condition nécessaire est donc $mk_B T \rho^{-1/3} \gg \hbar$. On a coutume de l'exprimer en recourant à la *longueur d'onde thermique* (ou de de Broglie), donnée par

$$\Lambda = \hbar \sqrt{\frac{2\pi}{mk_B T}}. \quad (3.10)$$

Ainsi, le régime classique correspond à

$$\rho^{-1/3} \gg \Lambda. \quad (3.11)$$

La longueur Λ est de l'ordre d'une fraction d'Å pour les petites molécules (eau, azote, dioxyde de carbone, ...) et les atomes (sauf l'hydrogène). Comme elle varie en $m^{-1/2}$ et $T^{-1/2}$, il est cohérent de négliger les effets quantiques pour les noyaux atomiques et les molécules, sauf à basse température, ou à densité très élevée (cas des étoiles naines). Les électrons, en revanche, doivent être traités par la mécanique quantique.

3.2 Entropie, température

3.2.1 Définition

Pour un système isolé, dont la statistique est décrite dans l'ensemble microcanonique, on définit l'entropie par la formule de Boltzmann :

$$S(N, V, E) = k_B \ln \Omega(N, V, E). \quad (3.12)$$

La température T est alors donnée par :

$$\frac{1}{T} = \left(\frac{\partial S}{\partial E} \right)_{N, V}. \quad (3.13)$$

Cas du gaz parfait

Pour le gaz parfait, un calcul (que l'on pourra refaire en exercice) dans lequel on ne garde que les termes dominants (d'ordre N) donne, en utilisant la formule de Stirling et en notant $e = E/N$ l'énergie par particule :

$$\frac{S}{Nk_B} = \frac{5}{2} - \ln \left(\rho h^3 \left(\frac{3}{4\pi m e} \right)^{3/2} \right). \quad (3.14)$$

On en tire la température par (3.13) :

$$T = \frac{2e}{3k_B},$$

et on retrouve l'énergie du gaz parfait

$$E = \frac{3}{2} N k_B T, \quad (3.15)$$

résultat que l'on avait anticipé avec les formules (2.6) et (3.6). Remplaçant e par T dans (3.14), l'expression de l'entropie en fonction de la densité et de la température fait réapparaître la longueur d'onde thermique :

$$\frac{S}{Nk_B} = \frac{5}{2} - \ln(\rho\Lambda^3). \quad (3.16)$$

3.2.2 Propriétés de l'entropie

L'entropie définie par (3.12) est bien, on peut le vérifier, celle de la thermodynamique macroscopique. Bien que le lien entre physique statistique microscopique et thermodynamique macroscopique classique soit assez peu détaillé dans ce cours (nous renvoyons aux références [8, 23, 2, 25] pour des présentations plus développées), nous indiquons ici quelques propriétés fondamentales et utiles de l'entropie, en signalant au passage, lorsqu'il y a lieu, la cohérence avec la thermodynamique macroscopique.

Extensivité de l'entropie et robustesse du calcul

L'entropie du gaz parfait est extensive (*i.e.* proportionnelle à N). Il en est de même, nous l'admettrons, pour l'entropie d'un système avec des interactions, la fonction de partition Ω étant dominée par les termes exponentiels de l'ordre de $\exp(\alpha N)$ avec une constante α . Cette extensivité entraîne par (3.13) l'intensivité de la température.

On peut considérer, pour justifier l'extensivité de l'entropie, un argument de dénombrement de configurations accessibles. Soit un découpage d'un système macroscopique en p sous-systèmes identiques, identifiés par le volume qu'ils occupent ou bien par les particules qu'ils contiennent (comme dans le § 2.4.2). Ils occupent chacun un volume V/p , et contiennent en moyenne N/p particules et l'énergie E/p . Si on néglige les interactions entre sous-systèmes différents, ainsi que les fluctuations de l'énergie et du volume ou du nombre de particules dans un sous-système (car elles n'affectent qu'une proportion négligeable de l'ensemble des degrés de liberté), on peut remplacer $\Omega(N, V, E)$ par $[\Omega(N/p, V/p, E/p)]^p$. On en déduit que l'entropie totale est proportionnelle à p , et donc à N .

On notera qu'une erreur d'un facteur N^γ dans le calcul de Ω (où γ est constant), erreur énorme si N est de l'ordre du nombre d'Avogadro, est sans conséquence sur le calcul de l'entropie, proportionnelle au logarithme de Ω . De même, on peut remplacer la surface $\{H(\mathbf{Q}, \mathbf{P}) = E\}$, dans le calcul de Ω , par le volume tout entier donné par $\{H(\mathbf{Q}, \mathbf{P}) \leq E\}$, sans en affecter l'évaluation de la partie extensive (dominante) de l'entropie. On le vérifiera en exercice dans le cas du gaz parfait.

Enfin, la suppression du facteur h^{-3N} dans la mesure de l'espace des phases décale la valeur de l'entropie d'une constante $3N \ln h$. Sans considération de mécanique quantique, l'entropie d'un système classique n'est déterminée qu'à une constante près.

Extensions de la définition

Sous la forme (3.12), l'entropie n'est définie *a priori* que pour le macroétat de l'équilibre microcanonique. Par extension, on a coutume d'associer une valeur de l'entropie à tous les macroétats similaires, c'est-à-dire pour des mesures de Lebesgue induites sur des sous-variétés $\{(\mathbf{Q}, \mathbf{P}) \in \mathcal{E} \mid A_1(\mathbf{Q}, \mathbf{P}) = A_1^0, \dots, A_p(\mathbf{Q}, \mathbf{P}) = A_p^0\}$. Le plus souvent, les observables A_i ont un sens macroscopique, et/ou sont associées à certaines fractions macroscopiques du système. On écrit $S(N, V, E, A^0)$ sans plus de précision pour l'entropie d'un tel macroétat, en se fondant sur la définition (3.12) de l'entropie comme logarithme du volume dans l'espace des phases. La définition la plus générale de l'entropie est celle de la théorie de l'information (voir le chapitre 4).

3.2.3 Couplage thermique

Considérons deux systèmes macroscopiques pouvant échanger de l'énergie, mais faiblement couplés, et isolés du milieu extérieur, contenant respectivement N_1 et N_2 particules, dans des volumes fixés V_1 et V_2 . Leurs énergies E_1 et E_2 varient en raison des échanges, mais la somme $E = E_1 + E_2$ est fixée. La fonction de partition microcanonique, peut se factoriser, en très bonne approximation, comme

$$\Omega(N, V, E) = A \Omega_1(N_1, V_1, \bar{E}_1) \Omega_2(N_2, V_2, E - \bar{E}_1), \quad (3.17)$$

où A est un terme borné par $(N_1 N_2)^a$ (avec une certain $a < +\infty$ indépendant des N_i) et $\bar{E}_i$ désigne la valeur la plus probable de E_i .³ L'énergie $\bar{E}_1$ réalise donc le maximum du produit $\Omega_1(N_1, V_1, \bar{E}_1) \Omega_2(N_2, V_2, E - \bar{E}_1)$, ou de son logarithme, qui n'est autre que la somme des entropies des deux sous-systèmes. On a ainsi

$$\frac{\partial S_1(N_1, V_1, \bar{E}_1)}{\partial E_1} + \frac{\partial S_1(N_2, V_2, E - \bar{E}_1)}{\partial E_2} = 0,$$

c'est-à-dire :

$$\frac{1}{T_1} = \frac{1}{T_2}.$$

L'équilibre thermique (c'est-à-dire l'équilibre vis-à-vis des échanges d'énergie au moyen d'un petit hamiltonien de couplage) entre deux systèmes impose l'égalité de leurs températures. On retrouve là une propriété caractéristique de la température en thermodynamique.

3.2.4 Les dérivées de la fonction entropie

La démarche conduisant à la définition de la température, et établissant l'égalité des températures entre deux systèmes échangeant de l'énergie par un faible couplage peut être reproduite pour le volume, ou bien encore pour le nombre de particules, pourvu que l'on s'en donne une définition convenable, comme des observables. Leur point commun est d'être des grandeurs qui se conservent dans les échanges, de sorte que si deux sous-systèmes échangent du volume on a $V = V_1 + V_2$ constant, et s'ils échangent des particules c'est $N = N_1 + N_2$ qui est constant. Lorsqu'il s'agit de deux parties d'un même corps, il est normal de considérer que de tels transferts impliquent aussi des échanges d'énergie. En plus de l'égalité des températures, la maximisation de (3.17) par rapport au partage du volume, N_1 et N_2 constants, donne alors :

$$\left(\frac{\partial S_1}{\partial V_1} \right)_{N_1, E_1} = \left(\frac{\partial S_2}{\partial V_2} \right)_{N_2, E_2}.$$

Si on identifie S à l'entropie de la thermodynamique macroscopique des fluides, on sait que

$$\left(\frac{\partial S}{\partial V} \right)_{N, E} = \frac{P}{T}, \quad (3.18)$$

où P est la pression. L'équilibre thermodynamique vis-à-vis des échanges de volume conduit donc à l'égalité des pressions, en plus de celle des températures. La formule (3.18), appliquée au gaz parfait, dont l'entropie est donnée par (3.16), donne bien

$$P = \rho k_B T. \quad (3.19)$$

3. Cette évaluation de Ω à partir des valeurs les plus probables des observables macroscopiques E_1 et E_2 est similaire à celle que l'on a évoquée pour justifier la distribution de Maxwell pour les quantités de mouvement dans le cas général : on avait alors argué que l'on pouvait en première approximation considérer l'énergie cinétique et l'énergie potentielle comme fixées.

Si on identifie plutôt les deux systèmes macroscopiques par leur localisation spatiale, fixant exactement leurs volumes V_1 et V_2 mais autorisant des échanges de particules, alors la maximisation de l'entropie totale donne l'égalité des températures de des *potentiels chimiques* μ , car

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

Pour comprendre que la pression définie avec (3.18) est bien la pression habituelle, on peut concevoir un dispositif idéal (appelé « piston adiabatique » dans la présentation traditionnelle de la thermodynamique) qui est une paroi sur laquelle on exerce une force constante, correspondant à une pression imposée P , mais dépourvue d'entropie. La maximisation de l'entropie totale, alors égale à celle du système, impose

$$\frac{dS(N, V, E - PV)}{dV} = 0 = \left(\frac{\partial S}{\partial V}\right)_{E,N} + \left(\frac{\partial S}{\partial E}\right)_{V,N} \frac{\partial E}{\partial V} = \left(\frac{\partial S}{\partial V}\right)_{E,N} - \frac{P}{T}.$$

Notons qu'en réécrivant les dérivées de l'entropie évaluées ci-dessus sous la forme

$$dE = TdS - PdV,$$

on retrouve la décomposition d'une variation d'énergie du système en somme du travail $\delta W = -PdV$ et de la chaleur $\delta Q = \frac{dS}{T}$ reçus. On reverra au passage ces notions dans le chapitre 4.

3.2.5 Entropie et fluctuations

L'entropie est jusqu'ici calculée comme intégrale de la mesure microcanonique sur certains ensembles, correspondant à des valeurs imposées de certaines observables macroscopiques. Dans ce calcul on peut considérer les observables macroscopiques extensives comme fixées à leur valeur la plus probable, dont le voisinage immédiat concentre la distribution de probabilité. En raffinant légèrement ce calcul, on peut évaluer la densité de probabilité d'une telle variable A au voisinage de sa valeur la plus probable $\bar{A}$. L'entropie $S(A)$ est maximale pour $A = \bar{A}$, et on peut écrire

$$S(A) = S(\bar{A}) - \frac{1}{2} \frac{\partial^2 S}{\partial A^2}(\bar{A}) \cdot (A - \bar{A})^2,$$

ce qui correspond à

$$\Omega(A) \propto \exp \left[-\frac{1}{2k_B} \frac{\partial^2 S}{\partial A^2}(\bar{A}) \cdot (A - \bar{A})^2 \right],$$

où la dérivée seconde est évaluée pour $A = \bar{A}$. La quantité A est donc distribuée selon une loi gaussienne de moyenne $\bar{A}$ et de variance

$$\text{Var}(A) = \left(\frac{\partial^2 (S/k_B)}{\partial A^2} \right)^{-1}.$$

Si A est extensive ($\bar{A} \propto N$), alors on a bien $\text{Var}(A) \propto N$, et donc $\text{Var}(A)/N$ est intensive.

L'entropie S doit être maximale pour la valeur $\bar{A}$ la plus probable de l'observable macroscopique A , et donc sa dérivée seconde doit être négative. Dans le cas contraire, il se produirait une évolution spontanée tendant à augmenter l'entropie, et l'état de départ caractérisé par la valeur $\bar{A}$ serait instable.

On peut raisonner de la même façon avec p observables scalaires, formant un p -uplet $(A_k)_{1 \leq k \leq p}$. On relie alors [23] les covariances de ces variables à la matrice hessienne de l'entropie. La stabilité requiert qu'elle soit définie négative.

Preuve de la formule (2.21).

L'énergie E étant fixée, on peut écrire l'entropie fonction de l'énergie cinétique comme la somme de l'entropie associée aux quantités de mouvement et de l'entropie associée aux degrés de liberté de position, les énergies correspondantes étant E_c et l'énergie potentielle $E_p = E - E_c$.

On écrit ensuite l'entropie totale S comme la somme d'un terme $S_c(E_c)$ associé aux quantités de mouvement et d'un terme $S_p(E_p)$ associé aux variables positions, et le développement au second ordre donne :

$$\begin{aligned} S &= S_c(\bar{E}_c) + S_p(\bar{E}_p) - \frac{1}{2} \frac{\partial(1/T)}{\partial E_c} (E_c - \bar{E}_c)^2 - \frac{1}{2} \frac{\partial(1/T)}{\partial E_p} (E_p - \bar{E}_p)^2 \\ &= S_c(\bar{E}_c) + S_p(\bar{E}_p) - \frac{1}{2} \left[\frac{\partial(1/T)}{\partial E_c} + \frac{\partial(1/T)}{\partial E_p} \right] (E_c - \bar{E}_c)^2. \end{aligned}$$

La capacité calorifique à volume constant par particule c_v se décompose en une partie liée à l'énergie cinétique, $c_v^{(\text{id})}$ (on note id pour « idéale », c'est la valeur de c_v dans le gaz parfait ou gaz idéal) et une partie provenant de l'énergie potentielle, que nous notons $c_v^{(\text{ex})}$ (pour *en excès* relativement au gaz parfait). La valeur $c_v^{(\text{id})} = \frac{1}{N} \frac{\partial E_c}{\partial T} = \frac{3}{2} k_B$ étant connue, on peut réécrire :

$$\frac{\partial(1/T)}{\partial E_c} = -\frac{2}{3Nk_B T^2} \quad \text{et} \quad \frac{\partial(1/T)}{\partial E_p} = -\frac{1}{Nc_v^{(\text{ex})} T^2} = -\frac{1}{(Nc_v - \frac{3}{2}Nk_B) T^2}.$$

On a donc pour la densité de probabilité de E_c , avec $\bar{E}_c = \frac{3}{2}Nk_B T$

$$P(E_c) \propto \exp \left[- \left(\frac{1}{3Nk_B T^2} + \frac{1}{2Nc_v - 3Nk_B T^2} \right) (E_c - \frac{3}{2}Nk_B T)^2 \right],$$

dont on déduit la formule cherchée (2.21).

3.2.6 Discussion

On verra au chapitre 4 que la mesure canonique permet de remédier à certains inconvénients de l'approche microcanonique, liés à la contrainte de l'énergie totale fixée. En effet, cette contrainte impose que la mesure d'équilibre dans $\mathcal{E}$ est concentrée sur une sous-variété. Il existe ainsi des corrélations entre les degrés de liberté $\mathbf{Q}, \mathbf{P}$ du système même lorsqu'ils sont en principe indépendants, c'est-à-dire avec un Hamiltonien de la forme (2.17).

Bien que l'on ait pu rapidement retrouver certaines notions de la thermodynamique macroscopique, des questions d'apparence délicate doivent être soulevées. Ce cours n'entend nullement clore des débats anciens et complexes à ce propos, et ces difficultés connues n'ont pas empêché le développement de la physique statistique et le succès des simulations numériques. On pourra se reporter aux références [2, 10, 25].

Ergodicité et degrés de libertés indépendants

Ainsi, précisément dans le cas d'un Hamiltonien présentant une décomposition selon (2.17), on peut noter que le postulat fondamental est en toute rigueur injustifié. En fait, chacune des énergies des sous-systèmes est, en principe, une constante du mouvement, et l'ergodicité est manifestement violée. Ainsi, chaque molécule d'un gaz parfait devrait conserver indéfiniment la même quantité de mouvement et la même énergie cinétique, et on voit bien mal comment sa vitesse pourrait échantillonner la distribution de Maxwell au cours du temps...

C'est un dilemme classique de la physique statistique : avec des degrés de liberté indépendants, les moyennes selon les macroétats d'équilibre se calculent aisément, mais le système ne peut pas se mettre à l'équilibre ! En pratique, on suppose que de très faibles interactions existent, que l'on néglige lorsque l'on écrit le hamiltonien du système, mais qui, à long terme, parviennent à en thermaliser les différentes parties. Ainsi, dans le cas du gaz parfait, on admet l'occurrence de rares collisions entre les molécules.

Irréversibilité

Une autre classe de difficultés, plus générale et plus sérieuse, apparaît lorsque l'on cherche à retrouver la notion thermodynamique d'irréversibilité, qui est associée à l'augmentation de l'entropie de la thermodynamique. Nous avons invoqué une condition d'entropie maximale en discutant, par exemple, de la répartition de l'énergie entre deux systèmes couplés. Par ailleurs, nous avons défini l'entropie comme logarithme d'un certain volume dans $\mathcal{E}$ (admettons que l'on donne une petite « épaisseur » aux sous-variétés sur lesquelles on concentre la mesure, quitte à la faire tendre vers zéro par la suite). La dynamique hamiltonienne conserve les volumes dans $\mathcal{E}$. Par conséquent l'entropie, logarithme du volume sur lequel on répartit uniformément la densité de probabilité, est constante. Comment peut-on donc atteindre l'état d'entropie maximale, si on part d'un état d'entropie inférieure, dans lequel, par exemple, deux parties macroscopiques du système sont à des températures différentes ?

Une explication possible est à rechercher dans l'emploi de deux notions d'entropie différentes, une entropie microscopique liée aux détails des mouvements dans l'espace des phases, et une entropie macroscopique, qui serait insensible aux détails fins de la répartition de la mesure dans $\mathcal{E}$. Si, à $t = 0$, la mesure est uniforme dans un domaine régulier $A \subset \mathcal{E}$, alors elle sera uniforme à $t > 0$ dans $\Phi_t(A)$, qui est une partie de $\mathcal{E}$ de même volume, mais, en raison du caractère chaotique de la dynamique, de forme beaucoup plus complexe, avec des feuilletés et des replis. De façon imagée, si A est une boule pleine, alors $\Phi_t(A)$ ressemble davantage à une boulette de papier froissé. Vue à grande échelle, on peut ainsi lui attribuer un volume plus grand, et on peut considérer que c'est un tel volume qui doit être pris en compte pour évaluer l'entropie macroscopique. Ainsi, cette dernière peut augmenter.

En thermodynamique, l'augmentation de l'entropie est associée à l'irréversibilité. Les transformations réversibles sont celles qui maintiennent constante l'entropie d'un système isolé.

3.3 Intégration des équations du mouvement et échantillonnage de la mesure microcanonique

3.3.1 Dynamique hamiltonienne et symplecticité

Nous avons présenté au § 2.2 les équations de Hamilton et certaines des propriétés du flot $(\phi_t)_{t \in \mathbb{R}}$ qui en découlent. Nous avons insisté sur l'importance que l'on doit attacher aux propriétés statistiques des ensembles de trajectoires plutôt qu'aux trajectoires individuelles, d'où l'intérêt de bien connaître les caractéristiques du flot. Avant de s'attaquer à la résolution numérique approchée des équations du mouvement, nous allons compléter la liste des propriétés de la dynamique hamiltonienne (voir § 2.2.2), en mentionnant en particulier une propriété structurelle très importante, la symplecticité, qui, comme on le verra, détermine le choix d'un schéma de discrétisation approprié.

Posant $y = (\mathbf{Q}, \mathbf{P}) \in \mathbb{R}^{2Nd}$, on peut réécrire la dynamique de Newton sous la forme d'une équation différentielle ordinaire :

$$\frac{dy}{dt} = J \cdot \nabla H = J \cdot \begin{pmatrix} \nabla_{\mathbf{Q}} H(\mathbf{Q}, \mathbf{P}) \\ \nabla_{\mathbf{P}} H(\mathbf{Q}, \mathbf{P}) \end{pmatrix}, \quad (3.20)$$

où J est la matrice réelle carrée de taille $2dN$

$$J = \begin{pmatrix} 0 & I_{dN} \\ -I_{dN} & 0 \end{pmatrix}.$$

On remarque que J est antisymétrique et orthogonale (*i.e.* $J^T = -J = J^{-1}$), et donc de déterminant égal à 1.

Définition 3.1 (Symplecticité). *Pour un ouvert $U \subset \mathcal{E}$, une application $g : U \rightarrow \mathbb{R}^{2dN}$ est symplectique si la matrice jacobienne $g'(\mathbf{Q}, \mathbf{P})$ vérifie*

$$[g'(\mathbf{Q}, \mathbf{P})]^T \cdot J \cdot g'(\mathbf{Q}, \mathbf{P}) = J \quad (3.21)$$

pour tout $(\mathbf{Q}, \mathbf{P}) \in U$.

Rappelons qu'étant donnée une application $g(\mathbf{Q}, \mathbf{P}) = (g_1(\mathbf{Q}, \mathbf{P}), \dots, g_{2dN}(\mathbf{Q}, \mathbf{P}))^T$, sa matrice jacobienne $g'(\mathbf{Q}, \mathbf{P})$ est définie par

$$g'(\mathbf{Q}, \mathbf{P}) = \begin{pmatrix} \frac{\partial g_1}{\partial q_1} & \cdots & \frac{\partial g_1}{\partial q_{2dN}} \\ & \ddots & \\ \frac{\partial g_{2dN}}{\partial q_1} & \cdots & \frac{\partial g_{2dN}}{\partial q_{2dN}} \end{pmatrix}.$$

On montre alors

Théorème 3.1 (Symplecticité du flot hamiltonien). *Soit $H(\mathbf{Q}, \mathbf{P})$ une fonction de $C^2(U)$ ($U \subset \mathcal{E}$ ouvert). Alors, pour tout t fixé tel que ϕ_t existe, le flot ϕ_t est une transformation symplectique.*

Preuve. Soit le flot ϕ_t de (3.20). On remarque tout d'abord que

$$\frac{d}{dt} \left(\frac{\partial \phi_t(y)}{\partial y} \right) = \frac{\partial}{\partial y} \left(\frac{d\phi_t(y)}{dt} \right) = \frac{\partial}{\partial y} (J \cdot \nabla H(\phi_t(y))) = J \cdot \nabla^2 H(\phi_t(y)) \cdot \frac{\partial \phi_t(y)}{\partial y}.$$

Ainsi, notant $\psi : t \mapsto \frac{\partial \phi_t(y)}{\partial y}$, on a

$$\frac{d}{dt} (\psi(t)^T \cdot J \cdot \psi(t)) = \psi(t)^T \cdot \nabla^2 H(\phi_t(y)) \cdot J^T J \cdot \psi(t) + \psi(t)^T \cdot \nabla^2 H(\phi_t(y)) \cdot J^2 \cdot \psi(t) = 0,$$

en utilisant $J^T = -J$. Ceci montre que $\psi(y)^T \cdot J \cdot \psi(y) = \psi(0)^T \cdot J \cdot \psi(0) = J$, ce qui permet de conclure à la symplecticité. $\square$

En fait, on peut même montrer qu'une EDO est localement hamiltonienne si et seulement si son flot est symplectique. La symplecticité permet de retrouver directement la propriété de conservation du volume dans l'espace des phases. En effet, si g est une application symplectique, le déterminant de sa matrice jacobienne vaut ± 1 d'après (3.21). Cette conservation du volume explique de manière heuristique les bonnes propriétés numériques de certains schémas d'intégration : en effet, les trajectoires partant du voisinage d'un point ne peuvent pas diverger dans toutes les directions, sans quoi, le volume balayé dans l'espace des phase augmenterait. La réversibilité en temps est également importante de ce point de vue : partant du voisinage d'un même point, si on a divergence de toutes les trajectoires dans un sens du temps, on a convergence dans l'autre... ce qui est absurde !

3.3.2 Approximation et flot numérique

On cherche ici à intégrer numériquement (2.2) (ou (3.20)), afin d'obtenir des moyennes d'observables, ou bien des fonctions de corrélation dépendant du temps. Etant donné un pas de temps Δt , une intégration numérique d'une EDO

$$\frac{dy}{dt} = f(y). \quad (3.22)$$

consiste à approximer la solution exacte $y(t_n) = \Phi_{\Delta t}^n(y^0)$ au temps $t_n = n\Delta t$ par une solution approchée y^n . Cette solution est obtenue par itération d'une méthode numérique $\Phi_{\Delta t}$. On a ainsi, pour une méthode à un pas (le calcul de y^{n+1} ne demande que la connaissance de y^n),

$$y(t_{n+1}) = \phi_{\Delta t}(y(t_n)), \quad y^{n+1} = \Phi_{\Delta t}(y^n).$$

Or, la solution exacte est telle que

$$y(t_{n+1}) = y(t_n) + \int_0^{\Delta t} f(y(t_n + s)) ds. \quad (3.23)$$

Une approximation du flot $\phi_{\Delta t}$ par un flot numérique $\Phi_{\Delta t}$ revient donc à trouver une approximation convenable de (3.23).

La convergence d'une méthode numérique est souvent caractérisée sur une trajectoire, *i.e.* pour une condition initiale donnée y^0 , et en intégrant sur un temps fixé T . On va voir cependant que la notion de convergence sur une trajectoire n'est pas la bonne notion : on va lui préférer la convergence d'un ensemble de trajectoires et la préservation des propriétés particulières du flot hamiltonien. On parlera ainsi d'intégration numérique géométrique pour caractériser les schémas préservant des invariants particuliers de la dynamique hamiltonienne. En particulier, comme on cherche à échantillonner des sous-variétés qui sont des surfaces de niveau du Hamiltonien, il est important de bien préserver l'énergie du système au cours du temps, et ce, pour des temps très longs – sans quoi la sous-variété échantillonnée aura peu de rapport avec la sous-variété associée à la dynamique (2.5).

Avant de se lancer dans l'étude des schémas numériques, précisons tout d'abord un critère très important en pratique : le coût de calcul. En dynamique moléculaire, le principal coût d'une itération de la méthode numérique est l'évaluation de la force $-\nabla V(q)$. On peut s'en convaincre aisément en prenant l'exemple d'un système comme le fluide de Lennard-Jones, où, en l'absence d'hypothèse supplémentaires, le calcul des interactions demande d'effectuer une boucle sur toutes les paires de particules, ce qui représente un coût de calcul en $O(N^2)$.

3.3.3 Schémas d'intégration usuels pour les équations différentielles ordinaires

Comme les équations de Newton peuvent se mettre sous la forme d'une EDO, on peut utiliser tous les schémas d'intégration usuels pour intégrer numériquement (3.20). Nous allons présenter successivement les schémas d'Euler explicite et implicite, ainsi que des schémas plus fins tels que les schémas de Runge-Kutta d'ordres 2 et 4. Ce dernier schéma est généralement un très bon schéma, et constitue l'algorithme par défaut de nombreux codes d'intégration d'EDO générales. L'étude de ces schémas demande de considérer :

- l'*ordre* des schémas numériques, qui traduit la précision de la trajectoire calculée, et la petitesse des erreurs d'intégration locales ;
- la *stabilité* (éventuellement la stabilité linéaire), qui traduit la robustesse du schéma par rapport à l'accumulation des erreurs d'intégration locales et d'arrondi.

Les notions d'ordre, de stabilité, et de stabilité linéaire, sont rappelées dans l'Appendice B. On peut d'ores et déjà sentir que la seconde notion, celle de stabilité, sera très importante car on souhaite calculer des moyennes sur de longues trajectoires, en échantillonnant correctement le macroétat d'équilibre microcanonique.

Remarque 3.3. Le lecteur prendra garde qu'il n'y a pas de contradiction pour un schéma numérique à être à la fois consistant (convergence trajectorielle à temps d'intégration T fixé, dans la limite $\Delta t \rightarrow 0$) et (linéairement) instable (explosion de la norme pour $T \rightarrow +\infty$ à Δt fixé).

Remarque 3.4. Il est évidemment souhaitable qu'un schéma numérique soit (linéairement) stable. Ce n'est toutefois pas une condition suffisante pour être un bon schéma en dynamique moléculaire ! On souhaite en effet que le schéma conserve bien l'énergie du système, et donc qu'il ne soit pas dissipatif.

Schéma d'Euler explicite.

Le schéma d'Euler explicite s'écrit pour une EDO générale de la forme (3.22),

$$y^{n+1} = y^n + \Delta t f(y^n).$$

Ceci correspond à une approximation de l'intégrale apparaissant dans (3.23) par une formule de quadrature où l'on n'utilise que l'information à l'extrémité gauche de l'intervalle en temps. Pour les équations de Newton (2.5), cela conduit au schéma

$$\begin{cases} q^{n+1} = q^n + \Delta t M^{-1} p^n, \\ p^{n+1} = p^n - \Delta t \nabla V(q^n). \end{cases} \quad (3.24)$$

On a ainsi

$$\Phi_{\Delta t}^{\text{Euler}}(q, p) = (q + \Delta t M^{-1} p, p - \Delta t \nabla V(q)).$$

On vérifie facilement que le schéma d'Euler explicite est d'ordre 1 (voir l'Exemple B.1). En effet,

$$\begin{aligned} |y(\Delta t) - \Phi_{\Delta t}^{\text{Euler}}(y(0))| &= \left| \int_0^{\Delta t} f(y(s)) ds - f(y(0)) \Delta t \right| \\ &\leq \int_0^{\Delta t} |f(y(s)) - f(y(0))| ds \\ &\leq \int_0^{\Delta t} s |y'(\theta(s))| ds, \end{aligned}$$

avec $\theta(s) \in]0, s[$, en utilisant une inégalité des accroissements finis. On en déduit que la différence entre la solution exacte intégrée sur un pas de temps Δt et l'approximation numérique donnée par le schéma est en $O(\Delta t^2)$, ce qui signifie en effet que le schéma est d'ordre 1 (voir la Définition B.2). Retenons également de ce calcul que la vérification de l'ordre d'un schéma consiste en pratique à effectuer des développements de Taylor, et à étudier quels termes se compensent. On montre également facilement que le schéma d'Euler explicite est stable (au sens de la Définition B.3), c'est-à-dire robuste à des erreurs d'arrondi et/ou sur les conditions initiales, à condition d'intégrer sur un temps T fini et donné *a priori*. Finalement, ce schéma est donc convergent, au sens de la Définition B.4.

En revanche, le schéma d'Euler explicite est toujours linéairement instable en temps long (Rappelons que cette notion est définie au § B.3). On peut s'en convaincre facilement dans le cas d'une particule de masse 1 soumise au potentiel harmonique $V(q) = \frac{1}{2} \omega^2 q^2$ en dimension 1, auquel cas on peut réécrire (3.24) sous la forme matricielle

$$y^{n+1} = \begin{pmatrix} q^n \\ p^n \end{pmatrix} = A(\Delta t) y^n \quad (3.25)$$

avec

$$A(\Delta t) = \begin{pmatrix} 1 & \Delta t \\ -\omega^2 \Delta t & 1 \end{pmatrix}.$$

La matrice $A(\Delta t)$ a pour valeurs propres $\lambda_{\pm} = 1 \pm i\omega \Delta t$. Comme $|\lambda_{\pm}| = \sqrt{1 + \omega^2 \Delta t^2} > 1$ si $\Delta t > 0$, on en déduit bien que le schéma est inconditionnellement linéairement instable. En particulier, partant d'une condition initiale (q^0, p^0) donnée, on a divergence de l'énergie :

$$H(q^n, p^n) \sim (1 + \omega^2 \Delta t^2)^n H(q^0, p^0) \rightarrow +\infty$$

lorsque $n \rightarrow +\infty$. Ainsi, on n'échantillonne pas du tout la surface d'énergie constante $\{H(q, p) = H(q^0, p^0)\}$. L'expérience numérique montre également que l'énergie diverge au cours du temps pour des potentiels non-linéaires généraux.

Schéma d'Euler implicite.

Le schéma d'Euler implicite s'écrit pour une EDO générale de la forme (3.22),

$$y^{n+1} = y^n + \Delta t f(y^{n+1}).$$

Ceci correspond à une approximation de l'intégrale apparaissant dans (3.23) par une formule de quadrature où l'on n'utilise que l'information à l'extrémité droite de l'intervalle en temps. En pratique, cela demande de résoudre un problème de point fixe non-linéaire, ce qui est plus coûteux qu'une itération simple d'un schéma explicite. Toutefois, l'expérience montre qu'on peut utiliser des pas de temps Δt plus grands qu'avec un schéma explicite.

Pour les équations de Newton (2.5), cela conduit à l'algorithme

$$\begin{cases} q^{n+1} = q^n + \Delta t M^{-1} p^{n+1}, \\ p^{n+1} = p^n - \Delta t \nabla V(q^{n+1}). \end{cases} \quad (3.26)$$

On vérifie facilement que le schéma d'Euler implicite est d'ordre 1 et convergent. Concernant la stabilité linéaire en temps long, on peut réécrire (3.26) sous la forme matricielle (3.25) avec

$$A(\Delta t) = \frac{1}{1 + \omega^2 \Delta t^2} \begin{pmatrix} 1 & \Delta t \\ -\omega^2 \Delta t & 1 \end{pmatrix}.$$

Il est en effet remarquable que dans le cas linéaire la méthode d'Euler implicite peut être réécrite de manière explicite⁴. La matrice $A(\Delta t)$ a pour valeurs propres $\lambda_{\pm} = (1 \pm i\omega\Delta t)^{-1}$. Comme $|\lambda_{\pm}| = (1 + \omega^2 \Delta t^2)^{-\frac{1}{2}} < 1$ si $\Delta t > 0$, on en déduit que le schéma est inconditionnellement linéairement stable. Cependant, on a convergence de l'énergie vers 0 lorsque $n \rightarrow +\infty$, puisque

$$H(q^n, p^n) \sim (1 + \omega^2 \Delta t^2)^{-n} H(q^0, p^0).$$

Pour un potentiel général, on s'attend encore à avoir un schéma dissipatif, ne se comportant pas bien sur des temps longs. L'expérience numérique confirme cette conjecture.

Schéma de Runge-Kutta d'ordre 4.

Les méthodes d'intégration numérique de Runge-Kutta sont très utilisées pour l'intégration des EDO, particulièrement dans le cas d'EDO non-autonomes $\dot{y} = f(t, y)$. Le schéma de Runge-Kutta d'ordre 4 suivant (présenté dans le cas d'une EDO autonome (3.22)) est très utilisé en pratique :

$$\begin{cases} Y^n &= y^n \\ Y^{n+1} &= y^n + f(Y^n) \Delta t/2 \\ Y^{n+2} &= y^n + f(Y^{n+1}) \Delta t/2 \\ Y^{n+3} &= y^n + f(Y^{n+2}) \Delta t \end{cases}$$

et on pose alors

$$y^{n+1} = y^n + \Delta t \frac{f(Y^n) + 2f(Y^{n+1}) + 2f(Y^{n+2}) + f(Y^{n+3})}{6}. \quad (3.27)$$

Ce schéma est d'ordre 4 (d'où son nom), et demande 4 évaluations de la force par pas de temps. Concernant la stabilité linéaire, on a la formulation matricielle (3.25) avec

$$A(\Delta t) = \begin{pmatrix} 1 - \frac{(\omega\Delta t)^2}{2} + \frac{(\omega\Delta t)^4}{24} & \Delta t \left(1 - \frac{(\omega\Delta t)^2}{6}\right) \\ -\omega^2 \Delta t \left(1 - \frac{(\omega\Delta t)^2}{6}\right) & 1 - \frac{(\omega\Delta t)^2}{2} + \frac{(\omega\Delta t)^4}{24} \end{pmatrix}.$$

4. C'est un bon exercice que de mener ce calcul...

On en déduit que les valeurs propres sont

$$\lambda_{\pm} = 1 - \frac{(\omega\Delta t)^2}{2} + \frac{(\omega\Delta t)^4}{24} \pm i\Delta t \left(1 - \frac{(\omega\Delta t)^2}{6}\right),$$

et donc que

$$|\lambda_{\pm}| = \sqrt{1 - \frac{(\omega\Delta t)^6}{72} + \frac{(\omega\Delta t)^8}{576}}.$$

Ceci montre que pour $\omega\Delta t > 0$ assez petit (numériquement, $\omega\Delta t \leq 2,828427$), on a $|\lambda_{\pm}| < 1$. Ceci montre que le schéma de Runge-Kutta d'ordre 4 est conditionnellement stable en temps long, mais, en analogie avec le schéma d'Euler implicite, qu'il est également dissipatif dans sa zone de stabilité, et ne permet donc pas un échantillonnage correct de la sous-variété d'énergie constante.

3.3.4 Schémas d'intégration préservant des propriétés de la dynamique Hamiltonienne

Nous avons vu aux § 2.2 et 3.3.1 que la dynamique hamiltonienne possède certaines propriétés qui la distinguent d'une EDO dans son type le plus général. Nous introduisons ici des schémas numériques dont le flot vérifie des propriétés analogues au flot de la dynamique hamiltonienne.

L'analyse numérique de ces schémas repose sur l'analyse d'erreur rétrograde. Alors que l'analyse numérique traditionnelle cherche des bornes *a priori* sur les erreurs commises, en interprétant le schéma numérique comme une approximation de la solution exacte, l'idée originale de l'analyse d'erreur rétrograde est d'interpréter le schéma numérique comme la solution exacte d'un problème approché, et d'étudier ce dernier. On montre alors que les schémas symplectiques préservent exactement une énergie approchée, ce qui les conduit à préserver approximativement l'énergie du système sur des temps exponentiellement longs⁵ (voir Théorème 3.2). La conservation approchée de l'énergie garantit alors que l'on échantillonne bien une sous-variété proche de la sous-variété cible. On parle d'*intégration géométrique*, c'est-à-dire d'une intégration numérique qui préserve bien les propriétés géométriques/structurelles fondamentales.

Définition 3.2 (Symplecticité d'un schéma numérique). *Une méthode numérique à un pas est symplectique si l'application*

$$\Phi_{\Delta t} : y^0 \mapsto y^1 = \Phi_{\Delta t}(y^0)$$

est symplectique lorsque la méthode est appliquée à un système hamiltonien régulier.

Là encore, la notion de symplecticité est liée à la conservation du volume dans l'espace des phases. Ceci explique que pour une méthode symplectique, les trajectoires ne peuvent pas converger vers une trajectoire donnée (ce qui impliquerait une diminution du volume dans l'espace des phases, ce qui est le cas du schéma d'Euler implicite et de Runge-Kutta d'ordre 4 pour Δt assez petit), ou au contraire diverger (ce qui impliquerait une augmentation du volume dans l'espace des phases, ce qui est le cas du schéma d'Euler explicite). Cette préservation du volume explique heuristiquement la très bonne stabilité des schémas symplectiques, précisée par le Théorème 8.1 de [19] :

Théorème 3.2 (Conservation approchée de l'énergie en temps long pour des méthodes symplectiques). *On considère un Hamiltonien $H : \mathcal{E} \rightarrow \mathbb{R}$ analytique et une méthode numérique $\Phi_{\Delta t}$ symplectique, d'ordre p . Si la solution numérique reste dans un compact $K \subset \mathcal{E}$, alors il existe $h > 0$ tel que*

$$H(y^n) = H(y^0) + O(\Delta t^p)$$

pour des temps exponentiellement longs $n\Delta t \leq e^{h/\Delta t}$.

5. On peut montrer également la conservation des invariants et une dérive linéaire de l'erreur sur la trajectoire sur des temps exponentiellement longs pour des systèmes complètement intégrables. Le lecteur (très motivé, car le sujet est technique!) pourra, en guise d'introduction, consulter le cours de Master de Philippe Chartier et Erwan Faou [7], et ensuite se plonger avec délices dans les chapitres IX et X de [19].

Ce résultat est d'énoncé technique, et pose de fortes contraintes sur le système. On a toutefois des résultats analogues sous des hypothèses un peu plus faibles. Retenons cependant que ce genre de résultat mathématique est peut-être sous-optimal, au sens où on n'a jamais observé en pratique une dérive de l'énergie pour un schéma symplectique! On motivera par la suite ce résultat en présentant des résultats élémentaires d'analyse rétrograde. En anticipant, on peut dire que les bonnes propriétés des schémas symplectiques sont la conséquence du fait qu'une *énergie approchée est conservée de manière exacte, ce qui signifie que l'énergie exacte est conservée de manière approchée*.

Outre la notion de symplecticité, le flot hamiltonien jouit d'autres propriétés, en particulier la réversibilité en temps, et la symétrie. Ces propriétés sont également importantes pour les schémas numériques, et mêmes suffisantes dans certains cas pour assurer le bon comportement en temps long.⁶

Définition 3.3 (Réversibilité en temps d'un schéma numérique). *Définissant l'opérateur de réversibilité en temps $\mathcal{S}$ par $\mathcal{S}(q, p) = (q, -p)$, on dit qu'un schéma numérique est réversible en temps si*

$$\mathcal{S} \circ \Phi_{\Delta t} = \Phi_{-\Delta t} \circ \mathcal{S}.$$

Cette propriété est satisfaite pour la grande majorité des schémas numériques associés à (2.5), et en tout cas, tous les schémas que nous allons rencontrer satisfont cette propriété, et ne s'avère donc pas être une propriété discriminante. Ceci n'est pas le cas pour la symétrie en temps :

Définition 3.4 (Symétrie d'un schéma numérique). *On dit qu'un schéma numérique est symétrique lorsque*

$$\Phi_{\Delta t} \circ \Phi_{-\Delta t} = \text{Id}.$$

On vérifie facilement qu'aucun schéma du § 3.3.3 ne vérifie cette condition (voir Exercice 3.2). En revanche le flot hamiltonien vérifie ces deux propriétés (la symétrie provenant du fait que le flot hamiltonien est un groupe).

Exemple 3.1. Vérifions à titre d'exemple que le schéma d'Euler explicite est réversible en temps mais pas symétrique. Rappelons que

$$\Phi_{\Delta t}^{\text{Euler}}(q, p) = (q + \Delta t M^{-1} p, p - \Delta t \nabla V(q)).$$

On a ainsi

$$\Phi_{\Delta t}^{\text{Euler}}(q, -p) = (q - \Delta t M^{-1} p, -p - \Delta t \nabla V(q)) = \mathcal{S}(q - \Delta t M^{-1} p, p + \Delta t \nabla V(q)) = \mathcal{S}(\Phi_{-\Delta t}^{\text{Euler}}(q, p)),$$

ce qui montre la réversibilité en temps. En revanche,

$$\begin{aligned} \Phi_{-\Delta t}^{\text{Euler}} \circ \Phi_{\Delta t}^{\text{Euler}}(q, p) &= \Phi_{-\Delta t}^{\text{Euler}}(q + \Delta t M^{-1} p, p - \Delta t \nabla V(q)) \\ &= (q + \Delta t^2 M^{-1} \nabla V(q) p + \Delta t (\nabla V(q + \Delta t M^{-1} p) - \nabla V(q))) \\ &= (q, p) + O(\Delta t^2), \end{aligned}$$

ce qui montre que le schéma n'est pas symétrique.

6. Attention, il existe des schémas symétriques qui ne sont pas symplectiques, et des schémas symplectiques qui ne sont pas symétriques... On pourra se reporter au chapitre XI de [19] pour l'étude des propriétés des schémas réversibles en temps : on montre que, pour des systèmes complètement intégrables ou quasi-intégrables, la symétrie suffit à assurer la conservation des invariants sur des temps exponentiellement longs.

Schémas d'Euler symplectiques.

Pour un Hamiltonien séparable (2.1), on peut écrire $H(q, p) = H_1(q, p) + H_2(q, p)$, avec $H_1(q, p) = \frac{1}{2}p^T M^{-1}p$ et $H_2(q, p) = V(q)$. Le champ de force se décompose selon

$$\begin{cases} \dot{q} = M^{-1}p, \\ \dot{p} = 0, \end{cases}, \quad \begin{cases} \dot{q} = 0, \\ \dot{p} = -\nabla V(q). \end{cases}$$

Chacun des deux systèmes hamiltoniens ci-dessous a un flot que l'on peut calculer exactement. Ce sont respectivement

$$\phi_t^1(q, p) = (q + t M^{-1}p, p), \quad \phi_t^2(q, p) = (q, p - t \nabla V(q)).$$

On utilise alors une formule de Lie-Trotter.⁷ Une approximation numérique de (2.5) obtenue par cette méthode de décomposition est par exemple

$$(q^{n+1}, p^{n+1}) = \phi_{\Delta t}^2 \circ \phi_{\Delta t}^1(q^n, p^n).$$

Les applications ϕ_t^2, ϕ_t^1 sont symplectiques pour tout temps t car ce sont les flots de dynamiques hamiltoniennes. Par ailleurs, on vérifie facilement par une formule de dérivation des applications composées que la composition de deux applications symplectiques est encore symplectique. Plus explicitement, on obtient le schéma dit "Euler symplectique A"

$$\begin{cases} q^{n+1} = q^n + \Delta t M^{-1}p^n, \\ p^{n+1} = p^n - \Delta t \nabla V(q^{n+1}). \end{cases} \quad (3.28)$$

Si on compose les flots dans l'autre sens, on obtient encore un schéma symplectique, dit "Euler symplectique B"

$$\begin{cases} q^{n+1} = q^n + \Delta t M^{-1}p^{n+1}, \\ p^{n+1} = p^n - \Delta t \nabla V(q^n). \end{cases} \quad (3.29)$$

Ces deux schémas sont tous deux explicites et d'ordre 1, réversibles mais pas symétriques.

Pour l'étude de la stabilité linéaire, on considère à nouveau l'oscillateur harmonique avec $m = 1$. Dans ce cas particulier, on a la formulation matricielle (3.25) avec respectivement

$$A_{\text{Euler A}}(\Delta t) = \begin{pmatrix} 1 & \Delta t \\ -\omega^2 \Delta t & 1 - (\omega \Delta t)^2 \end{pmatrix}, \quad A_{\text{Euler B}}(\Delta t) = \begin{pmatrix} 1 - (\omega \Delta t)^2 & \Delta t \\ -\omega^2 \Delta t & 1 \end{pmatrix}. \quad (3.30)$$

Les valeurs propres de ces matrices sont

$$\lambda_{\pm} = 1 - \frac{(\omega \Delta t)^2}{2} \pm i \omega \Delta t \sqrt{1 - \frac{(\omega \Delta t)^2}{4}}.$$

On a donc bien $|\lambda_{\pm}| = 1$ à condition que $\omega \Delta t < 2$. Ceci montre que les schémas d'Euler symplectique sont conditionnellement linéairement stables en temps long. Comme les valeurs propres sont de module 1, on s'attend heuristiquement à ce qu'il n'y ait pas de dérive de l'énergie dans un sens ou dans l'autre.

Remarque 3.5. Insistons sur l'importance de la condition de stabilité linéaire en temps long $\omega \Delta t < 2$. Cette condition montre que s'il existe des fréquences hautes dans le système, le pas de temps d'intégration devra être très petit! La très bonne stabilité des schémas et la conservation de l'énergie sur des temps longs vient donc avec une contrainte sur le pas de temps maximal que l'on peut utiliser...

⁷. Avec des mots simples (ou pour ceux qui ont joué au serpent sur leur portable) : pour faire un pas en diagonale, on fait un pas sur le côté et un pas tout droit. ... C'est l'exemple le plus simple de décomposition.

Schéma de Störmer-Verlet.

Le schéma de Störmer-Verlet est le schéma le plus utilisé en pratique pour l'intégration des équations (2.5). Il a été découvert et redécouvert plusieurs fois au cours du temps, en particulier par Störmer (1907) en astronomie, et par Verlet (1967) en dynamique moléculaire. Mais, comme l'a remarqué Feynman, on en trouve déjà des traces chez Newton (1687) ! Une très bonne introduction sur l'historique de ce schéma est donnée dans [18].

La procédure conduisant à l'obtention des schémas d'Euler symplectiques est quelque peu arbitraire, dans la mesure où il faut décider d'un certain ordre dans la composition. On peut symétriser cette opération par la composition

$$\Phi_{\Delta t} = \phi_{\Delta t/2}^2 \circ \phi_{\Delta t}^1 \circ \phi_{\Delta t/2}^2,$$

qui conduit au schéma de Störmer-Verlet :

ALGORITHME DE STÖRMER-VERLET

Partant d'une configuration initiale (q^0, p^0) ,

$$\begin{cases} p^{n+1/2} = p^n - \frac{\Delta t}{2} \nabla V(q^n), \\ q^{n+1} = q^n + \Delta t M^{-1} p^{n+1/2}, \\ p^{n+1} = p^{n+1/2} - \frac{\Delta t}{2} \nabla V(q^{n+1}). \end{cases} \quad (3.31)$$

On vérifie que ce schéma est d'ordre 2 et qu'il est symétrique, alors que les schémas d'Euler symplectiques A et B ne le sont pas. L'intérêt de la symétrisation est qu'on peut gagner un ordre de précision, en n'augmentant pas le nombre d'évaluations totales de la force (quitte à stocker les forces calculées au pas précédent pour la première mise à jour des vitesses sur un demi-pas de temps).

Concernant la stabilité linéaire en temps long, on a, toujours dans le cas de l'oscillateur harmonique avec des particules de masse $m = 1$, la formulation matricielle (3.25) avec

$$A(\Delta t) = A_{\text{Euler A}}(\Delta t/2) A_{\text{Euler B}}(\Delta t/2) = \begin{pmatrix} 1 - \frac{(\omega \Delta t)^2}{2} & \Delta t \\ -\omega^2 \Delta t \left(1 - \frac{(\omega \Delta t)^2}{4}\right) & 1 - \frac{(\omega \Delta t)^2}{2} \end{pmatrix},$$

où $A_{\text{Euler A}}, A_{\text{Euler B}}$ sont données par (3.30). On montre facilement que les valeurs propres de $A(\Delta t)$ sont toutes deux de module 1 si et seulement si $\omega \Delta t < 2$.

Remarque 3.6 (Formulation sans les impulsions). Un calcul simple montre que la solution de (3.31) vérifie

$$M \frac{q^{n+1} - 2q^n + q^{n-1}}{\Delta t^2} = -\nabla V(q^n).$$

On reconnaît ainsi une discrétisation par différences finies centrées de l'équation du mouvement

$$M \ddot{q} = -\nabla V(q).$$

Finalement, le schéma de Störmer-Verlet peut être construit très simplement ! En revanche, cette construction simple ne permet pas de mettre en évidence les propriétés structurelles profondes du schéma numérique, et il serait difficile de montrer que l'énergie est préservée de manière approchée sur des temps longs avec cette manière de faire.

Remarque 3.7 (Algorithme de Verlet généralisé). Dans le cas où le Hamiltonien n'est pas sous la forme séparable (2.1), la dynamique hamiltonienne s'écrit en général

$$\begin{cases} \dot{q}(t) = \nabla_p H(q, p), \\ \dot{p}(t) = -\nabla_q H(q, p). \end{cases}$$

Dans ce cas, le schéma de Verlet est

$$\begin{cases} p^{n+1/2} = p^n - \frac{\Delta t}{2} \nabla_q H(p^{n+1/2}, q^n), \\ q^{n+1} = q^n + \frac{\Delta t}{2} \left[\nabla_p H(p^{n+1/2}, q^n) + \nabla_p H(p^{n+1/2}, q^{n+1}) \right], \\ p^{n+1} = p^{n+1/2} - \frac{\Delta t}{2} \nabla_q H(p^{n+1/2}, q^{n+1}). \end{cases}$$

Le schéma n'est alors plus un schéma explicite. On pourra se reporter à [19, 24] pour plus de précisions.

Conservation exacte d'une énergie approchée : un exemple.

Dans le cas de l'oscillateur harmonique, on peut préciser la bonne conservation de l'énergie pour le schéma de Verlet. On cherche une matrice $C(\Delta t) \in \mathbb{R}^{2 \times 2}$ diagonale telle que

$$\forall y^0 \in \mathbb{R}^2, \quad (y^n)^T C(\Delta t) y^n = (y^0)^T C(\Delta t) y^0,$$

c'est-à-dire $C(\Delta t)$ diagonale telle que

$$A(\Delta t)^T C(\Delta t) A(\Delta t) = C(\Delta t).$$

Par exemple,

$$C(\Delta t) = \begin{pmatrix} \omega^2 \left(1 - \frac{(\omega \Delta t)^2}{4} \right) & 0 \\ 0 & 1 \end{pmatrix}$$

convient. Ceci montre que

$$\forall n \geq 0, \quad H_{\Delta t}(q^n, p^n) = H_{\Delta t}(q^0, p^0),$$

où le Hamiltonien modifié

$$H_{\Delta t}(q, p) = H(q, p) - (\omega \Delta t)^2 q^2 / 4.$$

Ainsi, on a bien conservation exacte d'une énergie approchée à $O(\Delta t^2)$ près, ce qui conduit au final à une conservation de l'énergie exacte à $O(\Delta t^2)$ pour tout temps.

Eléments d'analyse rétrograde

Rappelons que l'idée originale de l'analyse rétrograde est d'interpréter le schéma numérique comme la solution exacte d'un problème modifié (au lieu de l'interprétation usuelle en tant que solution approchée d'un problème exact). C'est un résultat profond de la théorie de l'intégration géométrique que le système hamiltonien modifié est encore hamiltonien... Ainsi, le schéma numérique conserve bien une énergie approchée, ce qui permet de montrer qu'il conserve de manière approchée l'énergie exacte. On a vu un tel résultat dans le cas de l'oscillateur harmonique, mais comment généraliser ce résultat ?

La manière de procéder est d'écrire un développement du flot modifié en puissances de Δt . Considérons la dynamique

$$\dot{y}(t) = f(y(t)), \tag{3.32}$$

de flot ϕ_t , et un schéma numérique associé $y^{n+1} = \Phi_{\Delta t}(y^n)$, approximation d'ordre p . On cherche une équation modifiée

$$\dot{z}(t) = f_{\Delta t}(z(t)), \quad (3.33)$$

dont le flot ψ_t est une bonne approximation de $\Phi_{\Delta t}$, au sens où $\Phi_{\Delta t}$ est une approximation numérique d'ordre supérieur ou égal à $p + 1$ de ψ_t . Pour ce faire, on développe $f_{\Delta t}$ en puissances de Δt , et on identifie les termes dans les développements des flots.

Premier exemple : le schéma d'Euler explicite

Prenons un exemple. Dans le cas du schéma d'Euler explicite, $p = 1$ et $\Phi_{\Delta t}(y) = y + \Delta t f(y)$. On écrit $f_{\Delta t}$ sous la forme

$$f_{\Delta t}(y) = f(y) + \Delta t F_1(y) + \frac{\Delta t^2}{2} F_2(y) + \dots$$

Il s'agit donc de chercher F_1 tel que $\Phi_{\Delta t}$ soit une approximation d'ordre 2 de ψ_t . Pour ce faire, on développe la solution de (3.33) selon

$$\begin{aligned} z(t) &= z(0) + \Delta t \dot{z}(0) + \frac{\Delta t^2}{2} \ddot{z}(0) + \dots \\ &= z(0) + \Delta t [f(z(0)) + \Delta t F_1(z(0)) + \dots] + \frac{\Delta t^2}{2} f'_{\Delta t}(z(0)) \dot{z}(0) \\ &= z(0) + \Delta t f(z(0)) + \Delta t^2 \left[\frac{1}{2} f'(z(0)) f(z(0)) + F_1(z(0)) \right] + O(\Delta t^3). \end{aligned}$$

Ainsi,

$$\Phi_{\Delta t}(z) - \psi_t(y) = O(\Delta t^3)$$

si on choisit

$$F_1(y) = -\frac{1}{2} f'(y) f(y).$$

Ceci montre que le schéma d'Euler est une approximation d'ordre 2 de la dynamique

$$\dot{y}(t) = f(y(t)) - \frac{\Delta t}{2} f'(y(t)) f(y(t)).$$

On peut itérer cette construction pour construire une dynamique modifiée qui dont le flot est d'ordre arbitrairement proche du flot du schéma numérique.⁸

Dans le cas où la dynamique exacte (3.32) est hamiltonienne, le terme correctif est

$$\Delta t F_1(y) = -\frac{1}{2} \begin{pmatrix} 0 & M^{-1} \\ -\nabla_q^2 V(q) & 0 \end{pmatrix} \begin{pmatrix} M^{-1} p \\ -\nabla_q V(q) \end{pmatrix} = \frac{1}{2} \left(M^{-1} \nabla_q V(q) \cdot M^{-1} p \right).$$

On ne peut pas réécrire ce terme comme dérivant d'un Hamiltonien selon

$$F_1(q, p) = \begin{pmatrix} \nabla_p H_1(q, p) \\ -\nabla_q H_1(q, p) \end{pmatrix}.$$

Si cela était possible, on aurait

$$F_1(q, p) = \begin{pmatrix} F_1^1(q, p) \\ F_1^2(q, p) \end{pmatrix}, \quad \nabla_q F_1^1 + \nabla_p F_1^2 = 0,$$

ce qui n'est pas le cas. Ceci montre que le système modifié n'est pas de type hamiltonien.

8. En fait, on a omis de préciser la constante qui apparait dans le terme $O(\Delta t^3)$ ci-dessus. Une analyse fine montrerait que cette constante diverge avec l'augmentation de l'ordre, et on ne peut donc pas atteindre un ordre quelconque si facilement... En fait, on ne peut pas écrire en général $f_{\Delta t}$ comme une série convergente en puissances de Δt .

Second exemple : le schéma d'Euler symplectique

On peut effectuer le même genre de calculs pour le schéma d'Euler symplectique (3.28). On a dans ce cas

$$\begin{cases} q^{n+1} = q^n + \Delta t M^{-1} p^n, \\ p^{n+1} = p^n - \Delta t \nabla_q V(q^n) - \Delta t^2 \nabla_q^2 V(q^n) M^{-1} p^n + O(\Delta t^3). \end{cases}$$

Or, on a vu ci-dessus que la solution de la dynamique modifiée partant de $y^n = (q^n, p^n)$ est

$$\begin{pmatrix} q^{n+1} \\ p^{n+1} \end{pmatrix} = \begin{pmatrix} q^n \\ p^n \end{pmatrix} + \Delta t \begin{pmatrix} M^{-1} p^n \\ -\nabla V(q^n) \end{pmatrix} + \Delta t^2 \left(F_1(y^n) + \frac{1}{2} f'(y^n) f(y^n) \right) + \dots$$

La comparaison avec le schéma d'Euler symplectique montre que le choix suivant permet d'assurer que le flot numérique du schéma d'Euler symplectique coïncide avec le flot de la dynamique modifiée à l'ordre 2 :

$$F_1(q, p) = \frac{1}{2} \begin{pmatrix} M^{-1} \nabla_q V(q) \\ -\nabla_q^2 V(q) \cdot M^{-1} p \end{pmatrix} = \begin{pmatrix} \nabla_p H_1(q, p) \\ -\nabla_q H_1(q, p) \end{pmatrix},$$

avec

$$H_1(q, p) = \frac{1}{2} p^T M^{-1} \nabla_q V(q).$$

Dans ce cas, la dynamique modifiée est encore hamiltonienne, et est associée au Hamiltonien

$$H^{\Delta t}(q, p) = H(q, p) + \Delta t H_1(q, p).$$

On vient ainsi de montrer que

$$H^{\Delta t}(q^n, p^n) = H^{\Delta t}(q^0, p^0) + O(\Delta t^2),$$

alors que

$$H(q^n, p^n) = H(q^0, p^0) + O(\Delta t).$$

On conserve donc mieux une énergie approchée. Une fois de plus, ce processus peut être itéré. La figure 3.2 illustre numériquement ce résultat.

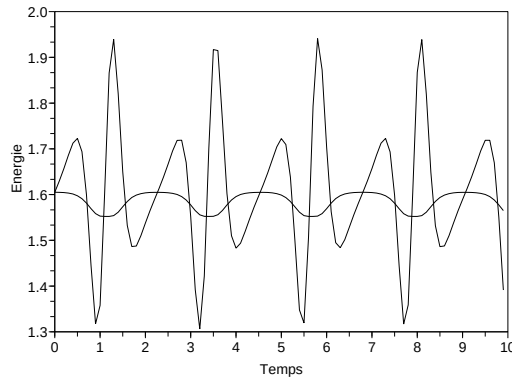


Fig. 3.2. Evolution temporelle de H et de $H^{\Delta t}$ (ligne plus forte) pour le potentiel double-puits, pour le schéma d'Euler symplectique avec un pas de temps $\Delta t = 0,05$. Notons que les variations de $H^{\Delta t}$ sont plus faibles que celles de H .

3.3.5 Comparaison numérique des différents schémas

On illustre en figure 3.3 l'erreur en fonction du pas de temps Δt dans le cas d'un oscillateur harmonique de pulsation $\omega = 1$, pour un temps d'intégration $T = 1$, et la condition initiale $(q^0, p^0) = (0, 1)$, l'erreur étant alors $e_T(\Delta t) = |q^{T/\Delta t} - \sin(T)|$. On vérifie bien que la relation $-\log(e_T(\Delta t)) = F(-\log(\Delta t))$ est affine avec un pente proche de 4 dans le cas du schéma de Runge-Kutta⁹ d'ordre 4, et proche de 2 dans le cas du schéma de Störmer-Verlet.

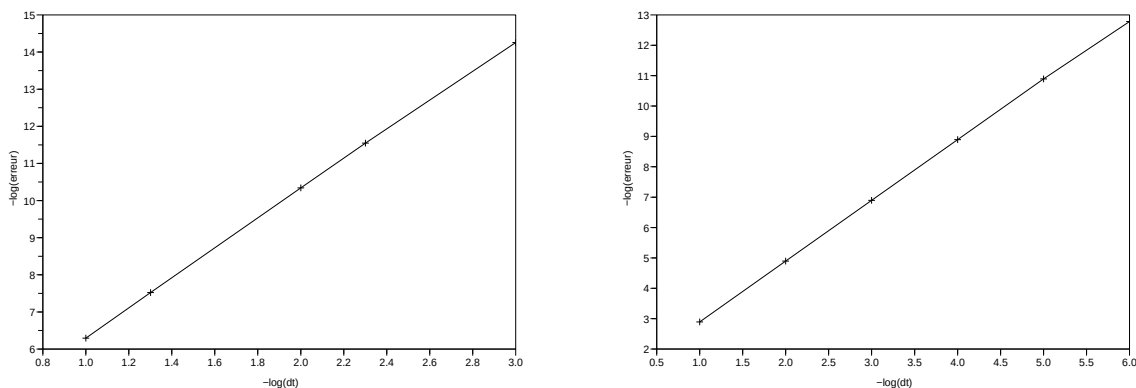


Fig. 3.3. À gauche : tracé de $-\log(e_T(\Delta t))$ en fonction de $-\log(\Delta t)$ pour le schéma de Runge-Kutta d'ordre 4 dans le cas de l'oscillateur harmonique. À droite : cas du schéma de Störmer-Verlet.

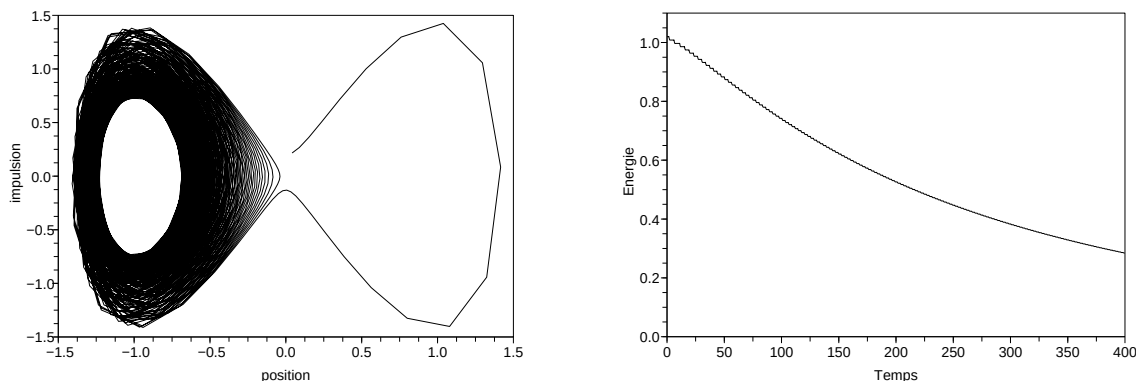


Fig. 3.4. Intégration en temps long du potentiel double puits par une méthode de Runge-Kutta d'ordre 4.

On compare ensuite une intégration sur un temps assez long ($T = 400$) de la dynamique hamiltonienne pour une particule de masse $m = 1$ dans le potentiel double puits (3.3), pour $h = 1$ et $\Delta t = 0,2$, partant de $(q, p) = (0, 1)$. Les figures 3.4 et 3.5 présentent les résultats obtenus pour le schéma de Runge-Kutta d'ordre 4 et de Störmer-Verlet, respectivement. On voit que dans le cas du schéma de Verlet, l'énergie oscille autour de l'énergie initiale $E = 1$, et le portrait de phase montre que la trajectoire numérique est localisée autour de la sous-variété des configurations d'énergie égale à 1 (on peut le voir en traçant une trajectoire de référence avec un pas d'intégration très petit). En revanche, le schéma de Runge-Kutta d'ordre 4, quoique plus précis, a un comportement en temps long moins bon du fait de sa dissipativité. La trajectoire numérique est rapidement piégée dans le puits de gauche, et l'énergie décroît au cours du temps.

9. C'est d'ailleurs pour cela qu'on se limite à des pas de temps plus grands que 10^{-4} , car dans ce cas, l'erreur est de l'ordre de 10^{-16} , soit la précision machine...

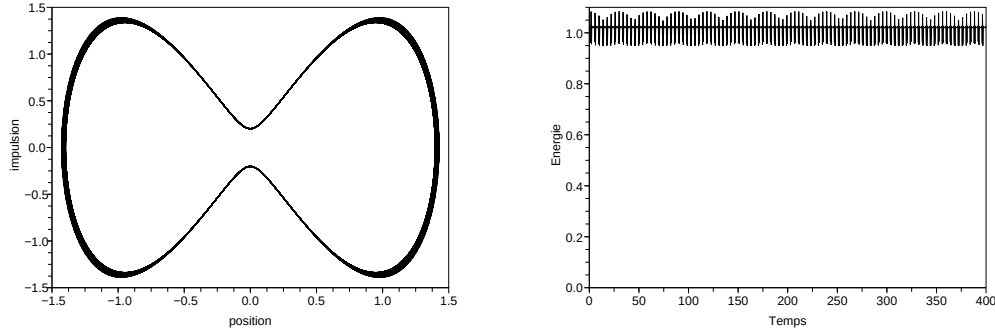


Fig. 3.5. Intégration en temps long du potentiel double puits par le schéma de Störmer-Verlet.

En conclusion, on voit bien que la notion d'ordre local, liée à une convergence trajectorielle en temps court, n'est pas la plus pertinente dans le contexte de la dynamique moléculaire, où on cherche à calculer des moyennes en temps long selon l'hypothèse d'ergodicité du postulat fondamental.

3.3.6 Application à un système physique

Adimensionnalisation et choix du pas de temps

Il est commode de mettre en œuvre les algorithmes précédents dans un code en travaillant en unités réduites, ne serait-ce qu'afin de manipuler des nombres qui ne soient ni trop petits, ni trop grands, ou de vérifier plus rapidement les ordres de grandeurs des résultats (si on obtient une quantité moyenne de l'ordre de 10^{20} en unités réduites, il faut s'inquiéter). On peut également se servir plusieurs fois du même calcul pour obtenir des résultats sur des systèmes différents (décrits par les mêmes fonctions potentielles, mais avec des paramètres différents!). On considère :

- l'unité de masse m , par exemple la masse de la fraction majoritaire des atomes simulés ;
- l'unité d'énergie, donnée par le minimum du potentiel d'interaction liant typique entre deux molécules (par exemple, ϵ pour le potentiel de Lennard-Jones (1.4)) ;
- l'unité de longueur l , donnée par la distance d'interaction typique (par exemple, la distance $2^{1/6}\sigma$ entre deux molécules liées par le potentiel de Lennard-Jones).

On en déduit des unités dérivées, comme l'unité de temps t_0 . Pour ce faire, on écrit que l'énergie cinétique typique est en $m(\sigma/t_0)^2 \sim \epsilon$, et donc $t_0 \sim \sigma(m/\epsilon)^{1/2}$.

Pour un fluide ou solide d'argon, de masse molaire $M = 39,95$ g/mol, les interactions entre les atomes sont données par le potentiel de Lennard-Jones (1.4), avec comme paramètres $\epsilon/k_B = 120$ K et $\sigma = 3,405$ Å. On en déduit que l'unité de masse est $m = 6,64 \times 10^{-26}$ kg, l'unité de longueur est $l = 2^{1/6}\sigma = 3,82$ Å, et donc l'unité de temps $t_0 = 2,42 \times 10^{-12}$ s.

Le pas de temps est choisi suffisamment petit pour que l'intégration numérique soit stable sur des temps longs, mais suffisamment grand pour que l'on puisse effectivement intégrer sur des temps longs. En pratique, il est bon de calculer la fréquence fondamentale du système, notée ω , et de choisir Δt tel que $\omega\Delta t < 2$, typiquement, $\omega\Delta t \sim 0,1$. Pour le potentiel de Lennard-Jones utilisé pour l'argon, une fréquence fondamentale peut être obtenue par un développement limité au voisinage de la position d'équilibre : $\omega^2 = V''(\epsilon)/m$, soit $\omega \sim 2,93 \times 10^{13}$ s⁻¹. Notons que la connaissance de la fréquence fondamentale permet également de déduire un temps caractéristique d'évolution du problème.

Système de Lennard-Jones

On présente un cas test réel : un fluide ou un solide d'argon. En pratique, on utilise un potentiel de Lennard-Jones tronqué¹⁰ à un certain rayon de coupure r_c , typiquement $r_c \sim 2,5 - 3\sigma$. On considère ensuite un domaine de simulation de la forme $[0, L]^3$, avec $L > 2r_c$ (sinon, l'utilisation de conditions de bord périodiques n'a aucun sens, voir Remarque 2.1).

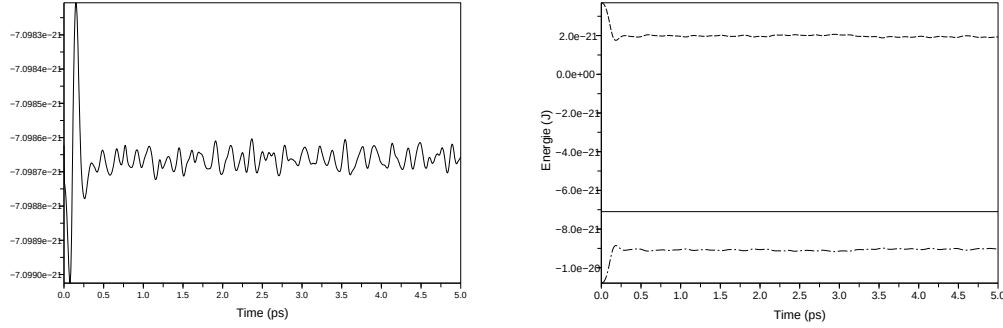


Fig. 3.6. Gauche : évolution de l'énergie moyenne par particule pour un solide d'argon en structure CFC à la densité $\rho = 1456 \text{ kg/m}^3$, pour des vitesses tirées initialement selon une maxwellienne à une température $T_{\text{init}} = 180 \text{ K}$. Droite : évolutions conjointes de l'énergie totale (ligne pleine), l'énergie cinétique (courbe du haut en traits interrompus) et de l'énergie potentielle (courbe du bas en traits interrompus).

Il est commode de répartir initialement les atomes selon un réseau régulier, par exemple cubique à face centrées (CFC) pour des solides, ou cubique simple si on veut simuler un fluide. En effet, cette dernière structure est moins stable, et est donc plus facile à faire fondre. Le paramètre important est la taille a de la cellule élémentaire du réseau cristallin. Dans l'exemple des figures 3.6 et 3.7, on a considéré un cube de 512 cellules élémentaires, pour un réseau CFC de paramètre $a = 4,5 \text{ \AA}$ (soit 1024 particules). On tire ensuite des vitesses selon une distribution de Maxwell à une certaine température T_{init} , et on intègre la dynamique hamiltonienne avec un schéma de Störmer-Verlet avec un pas de temps $\Delta t = 5 \times 10^{-15} \text{ s}$, utilisant un rayon de coupure $r_c = 9 \text{ \AA}$.

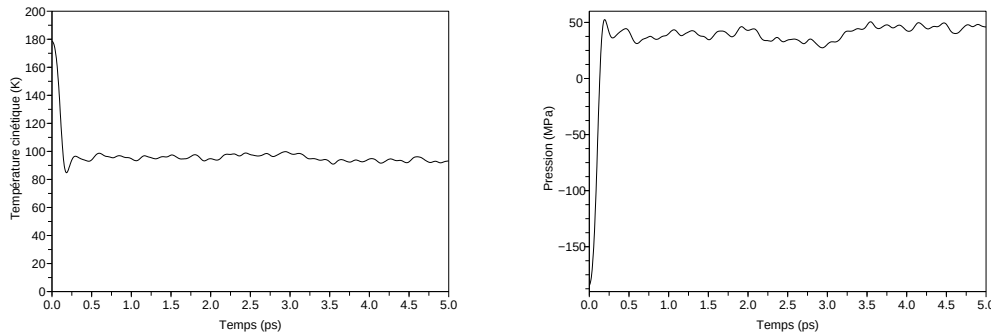


Fig. 3.7. Mêmes paramètres que pour la figure 3.6. Gauche : évolution de la température cinétique en fonction du temps. Droite : pression instantanée.

10. Idéalement, cette troncature s'effectue de manière à ce que le potentiel reste $C^1([0, +\infty[)$, par le biais d'une spline dans un domaine $[r_{\text{mod}}, r_c]$. Mais tous les codes ne font pas les choses aussi proprement, et souvent on tronque brutalement le potentiel...

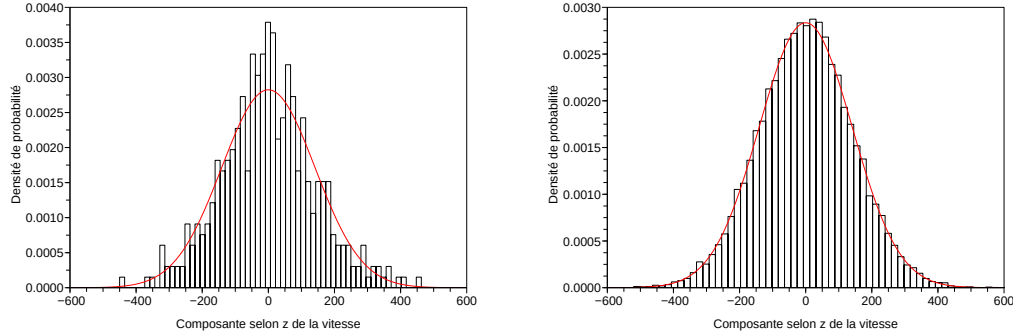


Fig. 3.8. Distribution des vitesses v_z au temps $t = 5 \times 10^{-14}$ s pour les mêmes conditions que pour la figure 3.6. Gauche : $N = 1296$ atomes. Droite : $N = 16000$ atomes. La courbe en ligne pleine est la maxwellienne de même variance que la distribution empirique des vitesses.

On montre en figure 3.6 que l'énergie est bien conservée pour le pas de temps choisi (erreur de l'ordre de 0,01%). On voit également que les différentes composantes de l'énergie s'équilibrent rapidement pour se stabiliser vers des valeurs à peu près constantes. La figure 3.7 présente l'évolution de la température cinétique et de la pression en fonction du temps. Il faut moyenner ces quantités sur la trajectoire pour obtenir une approximation des moyennes d'ensemble (voir le § 5.3 pour plus de précisions sur le calcul des barres d'erreur et les tests de convergence). Enfin, la figure 3.8 présente des distributions de vitesses à un instant donné. Comme on s'y attendait, cette distribution des vitesses est bien approchée par une distribution de Maxwell, et ce, d'autant mieux que le nombre d'atomes simulés est grand.

Exercices

Exercice 3.1. Vérifier que les schémas de Runge-Kutta (3.27) et le schéma de Störmer-Verlet (3.31) sont respectivement d'ordres 2 et 4.

Exercice 3.2. Montrer que tous les schémas numériques de ce chapitre sont réversibles en temps, mais que seul le schéma de Verlet est symétrique.

Exercice 3.3. Montrer que, dans le cas d'un oscillateur harmonique, les schémas d'Euler symplectiques préservent exactement l'énergie à un terme $O(\Delta t)$ près.

Exercice 3.4. Montrer que la composition de schémas symplectiques est encore symplectique.

Exercice 3.5. On change de coordonnées dans l'espace des phases : les $\mathbf{q}_i$, $\mathbf{p}_i$ sont remplacés par les $\mathbf{q}'_i = \mathbf{f}_i(\mathbf{Q}, \mathbf{P})$, et les $\mathbf{p}'_i = \mathbf{g}_i(\mathbf{Q}, \mathbf{P})$, les $\mathbf{f}_i$, $\mathbf{g}_i$ définissant un difféomorphisme régulier dans $\mathcal{E}$, et telle que les équations de Hamilton décrivent l'évolution des nouvelles variables avec le hamiltonien $H'(\mathbf{Q}', \mathbf{P}') = H(\mathbf{Q}, \mathbf{P})$. Montrer que le jacobien de ce changement de variables vaut 1. On pourra montrer pour ce faire qu'on a en fait une transformation symplectique.

Exercice 3.6. On considère une collection de N oscillateurs harmoniques identiques, dans un espace de dimension 1 pour simplifier. Le Hamiltonien du système s'écrit :

$$H(q, p) = \sum_{i=1}^N \frac{1}{2} \frac{p_i^2}{m} + \frac{1}{2} K q_i^2.$$

Évaluer la fonction de partition microcanonique, l'entropie et la température pour la valeur $E = Ne$ de l'énergie totale. Montrer que l'une quelconque des quantités de mouvement est distribuée selon la loi de Maxwell si N est assez grand.

Exercice 3.7 (Pression osmotique d'une solution diluée de molécules neutres). Un récipient de volume V est rempli d'un liquide, dans lequel sont dissoutes des molécules d'une espèce chimique différente, en faible proportion. La concentration (nombre de molécules de soluté par unité de volume) est c .

- (1) On suppose la solution diluée, c'est-à-dire que c est assez faible pour que l'énergie de la solution puisse être considérée comme indépendante des positions des molécules du soluté. Justifier cette approximation en considérant le volume v de la région du solvant dont l'état est affecté par la présence d'une molécule de soluté.
- (2) L'une des parois du récipient est en fait semi-perméable, elle laisse passer les molécules de solvant mais pas celles du soluté. De l'autre côté de cette paroi se trouve un compartiment contenant du solvant pur. En raisonnant sur l'entropie, montrer que la pression est plus grande dans la solution que dans le solvant pur, la différence étant par définition la *pression osmotique* Π , ressentie par la paroi semi-perméable, et qui vaut :

$$\Pi = ck_{\text{B}}T$$

sous l'hypothèse de solution diluée.

Indications : on pourra chercher une forme factorisée de la fonction de partition microcanonique. On identifiera Π à partir de

$$\Pi = T \left(\frac{\partial S}{\partial V} \right)_E.$$

L'ensemble canonique

Le recours à l'ensemble canonique, représentatif d'un système dont les échanges avec le milieu extérieur sont des transferts d'énergie avec un thermostat, est motivé par le manque de commodité d'utilisation de la mesure microcanonique, et aussi par la proximité avec l'expérience : il est beaucoup plus pertinent de modéliser les objets physiques de notre univers quotidien en considérant que leur température est fixée, plutôt que leur énergie totale.

4.1	Distribution de Boltzmann-Gibbs	61
4.1.1	Couplage avec un thermostat	61
4.1.2	Distribution de Boltzmann-Gibbs et information manquante	62
4.2	Énergie libre et grandeurs thermodynamiques	64
4.2.1	Définition	64
4.2.2	Dérivées de l'énergie libre	64
4.2.3	Minimisation de l'énergie libre	65
4.2.4	Calculs avec la fonction de partition	65
4.3	Cas de degrés de liberté indépendants	67
4.3.1	Sous-systèmes indépendants	67
4.3.2	Indépendance des positions et des quantités de mouvement pour une collection de particules	67
4.3.3	Gaz parfait, grandeurs idéales et d'excès dans les fluides	68
4.3.4	Cas des oscillateurs harmoniques indépendants	68
4.4	Quelques remarques sur les applications du formalisme canonique	70
4.4.1	Gaz, solides et liquides	70
4.4.2	Approche de l'équilibre et thermalisation	70
	Exercices	72

4.1 Distribution de Boltzmann-Gibbs

La mesure de probabilité d'équilibre sur $\mathcal{E}$ dans l'ensemble canonique est appelée la distribution de Boltzmann-Gibbs, et sa forme constitue un des résultats les plus fondamentaux de la physique statistique. Nous indiquons ici deux manières de l'obtenir : la première (§ 4.1.1) est fondée sur le couplage idéal avec un thermostat, traité comme au § 3.2.3 ; la seconde (§ 4.1.2) fait appel à l'entropie définie selon la théorie de l'information.

4.1.1 Couplage avec un thermostat

Nous avons admis au § 3.2.3 que s'il y a couplage thermique entre deux systèmes macroscopiques, on peut évaluer la fonction de partition microcanonique de la réunion des deux comme le

produit des fonctions de partitions revenant à chacun, et à cette factorisation correspond l'évaluation de l'entropie totale comme somme des entropies des deux systèmes (voir (3.17)).

Nous nous replaçons ici dans ce cadre, et supposons que le second système, que l'on appelle thermostat, est très grand, de sorte que son couplage avec le premier, qui est le système physique que l'on souhaite étudier (et que l'on l'appellera simplement « le système » dorénavant), n'affecte son état que de façon négligeable. Les variations d'entropie du thermostat seront alors traitées au premier ordre dans les variations de son énergie. Ainsi, quelle que soit la valeur de E , on écrit pour l'entropie du thermostat, en vertu de la définition (3.13) de la température :

$$S_{\text{th}} = S_0 - \frac{E}{T},$$

S_0 désignant l'entropie du thermostat lorsqu'il prend toute l'énergie disponible, et T sa température. À une constante additive près, l'entropie totale vaut donc $S(N, V, E) - \frac{E}{T}$. On obtient ainsi la densité de probabilité $p(E)$ d'une valeur donnée de l'énergie du système par (3.12) :

$$p(E) \propto \exp\left(-\frac{E}{k_B T}\right).$$

Il sera commode dans toute la suite du cours d'utiliser la notation :

$$\beta = \frac{1}{k_B T}. \quad (4.1)$$

L'équiprobabilité pour les microétats de même énergie de la réunion du système et du thermostat entraîne l'équiprobabilité des microétats du système de même énergie E . On en tire donc la forme suivante de la densité de probabilité (macroétat d'équilibre canonique), par rapport à la mesure $\mu_0(\mathbf{Q}, \mathbf{P})$ introduite en (2.3).

DISTRIBUTION DE BOLTZMANN-GIBBS, MACROÉTAT D'ÉQUILIBRE CANONIQUE

La mesure canonique a pour densité de probabilité

$$\rho_{\text{NVT}}(\mathbf{Q}, \mathbf{P}) d\mu_0 = Z^{-1} e^{-\beta H(\mathbf{Q}, \mathbf{P})} d\mu_0, \quad (4.2)$$

La constante de normalisation Z est la *fonction de partition canonique* :

$$Z = \int_{\mathcal{E}} e^{-\beta H(\mathbf{Q}, \mathbf{P})} d\mu_0. \quad (4.3)$$

Le macroétat d'équilibre est donc déduit ici de l'approche microcanonique pour la réunion du système avec le thermostat. La mesure de Boltzmann-Gibbs sur $\mathcal{E}$ qu'il définit s'avère très commode d'utilisation. Dans la pratique on peut effectivement vérifier que le macroétat d'équilibre d'une petite partie macroscopique d'un grand système, si son volume et le nombre de particules qu'elle contient sont fixés, est bien décrit par cette distribution.

4.1.2 Distribution de Boltzmann-Gibbs et information manquante

Entropie et défaut d'information

Dans l'approche de la théorie de l'information ¹ l'entropie est définie, vis-à-vis d'un observateur, comme le manque d'information sur le système observé. On définit ainsi l'entropie $S[\rho_N]$ d'un macroétat possédant une densité ρ_N par rapport à la mesure μ_0 par

1. On trouvera dans la référence [2] une présentation claire et détaillée de cette approche. En revanche, la référence [25] évite d'y avoir recours, en raison du rôle qu'elle fait jouer à un observateur humain, et définit l'entropie comme (à la constante k_B près) « logarithme du volume d'espace de phases exploré par le système dans son mouvement ».

$$S[\rho_N] = -k_B \int_{\mathcal{E}} \rho_N(\mathbf{Q}, \mathbf{P}) \ln \rho_N(\mathbf{Q}, \mathbf{P}) d\mu_0. \quad (4.4)$$

La détermination du macroétat d'entropie maximale est alors motivée par l'exploitation de l'information disponible sans introduire de choix ou de biais injustifié.

La définition (4.4) se justifie d'abord pour les distributions de probabilités discrètes, auquel cas on a affaire à l'entropie de la théorie de l'information (ou entropie de Shannon)

$$S = -k_B \sum_i p_i \ln p_i. \quad (4.5)$$

On prend ensuite une limite continue, le choix de la mesure μ_0 dans $\mathcal{E}$ étant naturel car c'est une mesure invariante pour la dynamique hamiltonienne. Pour motiver la définition (4.5), on peut montrer (voir par exemple [2]) qu'elle est compatible avec les propriétés de base que l'on doit attribuer à une fonction des p_i ayant le sens du manque d'information, ou de l'information moyenne acquise par l'observation de l'un des événements discrets possibles. La propriété essentielle est l'additivité, qui exprime qu'une information acquise en deux temps, correspondant par exemple à deux niveaux de précision des observations, est la somme de l'information acquise lors de l'étape grossière et de l'information moyenne acquise lors de l'étape la plus fine. Une autre propriété est que l'entropie doit être maximale lorsque tous les événements sont équiprobables. La constante k_B fixe conventionnellement l'unité d'information.

La définition (4.4) généralise celle que nous avons utilisée au chapitre 3, avec laquelle elle coïncide lorsque les macroétats possèdent une densité uniforme par rapport à la mesure μ_0 dans $A \subset \mathcal{E}$ de mesure positive finie. En effet, on a alors

$$S = -k_B \int_A \frac{1}{\mu_0(A)} \ln \left[\frac{1}{\mu_0(A)} \right] d\mu_0 = k_B \ln \mu_0(A).$$

Maximisation sous contrainte

Faisons l'hypothèse que l'énergie du système étudié est connue en moyenne, c'est-à-dire que la valeur de

$$\langle H \rangle = \int_{\mathcal{E}} \rho_N(\mathbf{Q}, \mathbf{P}) H(\mathbf{Q}, \mathbf{P}) d\mu_0 \quad (4.6)$$

est imposée. La maximisation de $S[\rho_N]$ sous les contraintes (4.6) et

$$\int_{\mathcal{E}} \rho_N d\mu_0 = 1, \quad (4.7)$$

conduit à la condition d'optimalité suivante, qui doit être satisfaite pour toute variation $\delta\rho_N$ de ρ_N :

$$- \int_{\mathcal{E}} (\ln \rho_N(\mathbf{Q}, \mathbf{P}) + 1) \delta\rho_N(\mathbf{Q}, \mathbf{P}) d\mu_0 + \beta \int_{\mathcal{E}} H(\mathbf{Q}, \mathbf{P}) \delta\rho_N(\mathbf{Q}, \mathbf{P}) d\mu_0 + \gamma \int_{\mathcal{E}} \delta\rho_N(\mathbf{Q}, \mathbf{P}) d\mu_0 = 0,$$

où on a introduit les multiplicateurs de Lagrange β et γ associés respectivement aux contraintes (4.6) et (4.7). Ceci entraîne

$$\ln \rho_N = -\beta H - \gamma - 1,$$

et on en déduit donc le résultat (4.2), en reconnaissant dans β le paramètre défini en (4.1), tandis que la fonction de partition $Z = \exp(\gamma + 1)$ normalise la distribution.

4.2 Énergie libre et grandeurs thermodynamiques

4.2.1 Définition

L'énergie libre F (appelée parfois énergie libre de Helmholtz s'il y a une ambiguïté possible avec d'autres potentiels thermodynamiques) est classiquement définie en thermodynamique comme²

$$F(N, V, T) = E(N, V, S) - TS,$$

et il est commode de la considérer³ comme une fonction de T , tandis que E est fonction de S . L'énergie libre F est directement reliée à la fonction de partition Z . Pour le voir, écrivons l'entropie, évaluée par (4.4) pour la densité ρ_N de Boltzmann-Gibbs. On trouve :

$$S = k_B (\beta \langle H \rangle - \ln Z),$$

et on a donc, l'énergie interne macroscopique E s'identifiant à la moyenne de l'observable H ,

$$F = -k_B T \ln Z(N, V, T). \quad (4.8)$$

On peut remarquer que dans la première approche que nous avons employée pour obtenir la distribution de Boltzmann-Gibbs (§ 4.1.1), l'entropie du système étudié se confondait avec l'entropie microcanonique correspondant à la valeur la plus probable de l'énergie du système. Ceci entraîne une légère différence pour l'énergie libre F , qui disparaît dans la limite $N \rightarrow +\infty$. Dans la pratique, on utilisera la formule (4.8), et on gardera donc la seconde définition de l'entropie (§ 4.1.2), qui est la plus générale.

4.2.2 Dérivées de l'énergie libre

Les dérivées de l'énergie libre par rapport à ses variables naturelles V et T se déduisent des dérivées de l'énergie interne, ou de l'entropie. On a ainsi :

$$\left(\frac{\partial F}{\partial T} \right)_V = \left(\frac{\partial E}{\partial T} \right)_V - S - T \left(\frac{\partial S}{\partial T} \right)_V = \left(\frac{\partial E}{\partial S} \right)_V \left(\frac{\partial S}{\partial T} \right)_V - S - T \left(\frac{\partial S}{\partial T} \right)_V = -S, \quad (4.9)$$

tandis que

$$\left(\frac{\partial F}{\partial V} \right)_T = \left(\frac{\partial E}{\partial V} \right)_T - T \left(\frac{\partial S}{\partial V} \right)_T = \left(\frac{\partial E}{\partial V} \right)_S + \left(\frac{\partial E}{\partial S} \right)_V \left(\frac{\partial S}{\partial V} \right)_T - T \left(\frac{\partial S}{\partial V} \right)_T = -P. \quad (4.10)$$

Ces relations permettent, en supposant que l'on parvienne à calculer Z et donc $F = -k_B T \ln Z$, d'en déduire l'entropie, l'énergie interne et la pression.

Le *potentiel chimique* μ est défini comme dérivée de l'énergie libre par rapport au nombre de particules N ; ainsi, lorsque N est grand,

$$\mu = \left(\frac{\partial F}{\partial N} \right)_{V,T} \simeq F(N+1, V, T) - F(N, V, T).$$

On s'attend en effet à une variation d'énergie libre entre deux états d'équilibre à la même température dans le même volume lorsque l'on introduit une particule supplémentaire.

2. Cela signifie que $F = E - S \frac{\partial E}{\partial S}$ est – au signe près selon les conventions – la transformée de Legendre de E par rapport à la variable S .

3. La thermodynamique s'est développée, historiquement, en mettant l'accent sur l'énergie plus que sur l'entropie, et c'est pourquoi on a pris l'habitude d'écrire E fonction de S et F fonction de T plutôt que S/k_B fonction de E et $(S/k_B) - \beta E$ fonction de β .

4.2.3 Minimisation de l'énergie libre

Au chapitre 3 nous avons discuté de la répartition de l'énergie (plus généralement, de toute grandeur conservée, c'est-à-dire qui ne varie que par des échanges) entre différentes parties d'un système macroscopique, et caractérisé l'équilibre par le maximum de l'entropie. Dans le cas d'un système couplé avec un thermostat, la même démarche appliquée à la réunion du système et du thermostat conduit à maximiser la somme de leurs entropies, qui est à une constante près

$$S - \frac{E}{T} = -\frac{F}{T},$$

T désignant ici la température du thermostat. Il faut donc *minimiser l'énergie libre* (la température étant fixée).

Une application est l'équilibre des phases (comme dans les régions de coexistence des diagrammes de phase des figures 1.3a et 1.4). Deux phases en équilibre sont à la même température, et elles peuvent échanger, en plus de la chaleur, du volume et des particules. On a donc égalité des potentiels chimiques et égalité des pressions. Pour chaque phase, le potentiel chimique peut être pris comme fonction de la densité et de la température, ou bien de la pression et de la température. Les indices se rapportant aux deux phases en présence, on a ainsi la condition

$$\mu_1(P, T) = \mu_2(P, T),$$

qui détermine la ligne de coexistence dans le plan P, T (comme sur la figure 1.3b).

Une autre conséquence remarquable de la minimisation de l'énergie libre est la compétition entre énergie et entropie. À basse température, c'est l'énergie E qui tend à dominer dans l'énergie libre $F = E - TS$. Minimiser l'énergie libre, c'est alors essentiellement minimiser l'énergie. Pour les systèmes usuels composés d'un grand nombre de particules, l'énergie potentielle est souvent minimale pour des configurations ordonnées à basse température.⁴ Des considérations entropiques entrent quand même en jeu cependant, et elles peuvent même déterminer le changement de phase, comme on l'a évoqué au chapitre 1. En revanche, à haute température, le terme entropique devient prépondérant, et minimiser l'énergie libre revient alors à maximiser l'entropie. C'est en phase gazeuse et non dans les phases condensées liquide ou solide⁵ que l'entropie est la plus grande, parce que les particules accèdent à un domaine beaucoup plus étendu.

4.2.4 Calculs avec la fonction de partition

À la différence de la fonction de partition microcanonique $\Omega(N, V, E)$, la fonction de partition canonique Z , fonction de (N, V, T) , permet des calculs commodes.

Dérivation par rapport à la température

La dérivée de la fonction de partition par rapport à β est

$$\frac{\partial Z}{\partial \beta} = - \int_{\mathcal{E}} H(\mathbf{Q}, \mathbf{P}) \exp[-\beta H(\mathbf{Q}, \mathbf{P})] d\mu_0.$$

Comme

$$\langle H \rangle = \frac{1}{Z} \int_{\mathcal{E}} H(\mathbf{Q}, \mathbf{P}) \exp[-\beta H(\mathbf{Q}, \mathbf{P})] d\mu_0,$$

4. .

5. On considère ici implicitement un système maintenu dans une enceinte de volume constant, et dont le volume est assez grand pour permettre le changement de phase entre liquide et vapeur. À basse température il y a condensation et on observe un équilibre liquide-vapeur ou solide-vapeur. Ce sont dans les conditions isothermes et isobares – auquel cas il faut minimiser l'enthalpie libre $G(N, P, T) = F + PV$, également appelée énergie libre de Gibbs – que l'on a une seule phase.

on a alors

$$E = \langle H \rangle = -\frac{\partial \ln Z}{\partial \beta}. \quad (4.11)$$

Une seconde dérivation donne

$$\frac{\partial^2 \ln Z}{\partial \beta^2} = \langle H^2 \rangle - \langle H \rangle^2 = \text{Var}(H) = -\frac{\partial \langle H \rangle}{\partial \beta},$$

ce qui établit la formule (2.19) pour la capacité calorifique à volume constant par particule c_v :

$$c_v = \frac{1}{N} \left(\frac{\partial E}{\partial T} \right)_V = \frac{\text{Var}(H)}{N k_B T^2}.$$

Dérivation par rapport à un paramètre du Hamiltonien

Supposons que le Hamiltonien dépende d'un certain paramètre λ . Par exemple, des forces extérieures sont appliquées sur les particules, et λ apparaît dans l'expression du champ de force correspondant. On a donc un Hamiltonien $H \equiv H(\lambda, q)$. Un calcul analogue à (4.11) donne alors :

$$\left(\frac{\partial F}{\partial \lambda} \right)_T = \left\langle \frac{\partial H}{\partial \lambda} \right\rangle_\lambda, \quad (4.12)$$

où la notation $\langle \cdot \rangle_\lambda$ renvoie à une moyenne canonique pour la mesure $Z_\lambda^{-1} e^{-\beta H(\lambda, q)}$. Une action extérieure sur le système entraîne une variation de l'énergie des microétats, égale au travail des forces extérieures. La relation (4.12) signifie donc que la variation d'énergie libre d'un système couplé à un thermostat lors de l'application de sollicitations extérieures est égale à la moyenne des travaux microscopiques pour chaque macroétat visité (paramétré par λ), et on identifie ainsi au second membre, une fois intégré sur un intervalle de λ , le *travail* fourni au système. Notant W ce travail on a ainsi

$$\left(\frac{\partial F}{\partial \lambda} \right)_T = \frac{dW}{d\lambda}.$$

Au sens de la thermodynamique, on a ici une transformation isotherme *réversible*, c'est-à-dire dans laquelle le système reste en équilibre interne et en équilibre avec le milieu extérieur.

Le travail des forces extérieures est égal à la variation d'énergie libre, qui prend donc le sens d'une énergie potentielle macroscopique pour les évolutions à température constante. Dans les problèmes de statique macroscopiques, à température constante, on utilise souvent des énergies potentielles (par exemple élastiques), qui, du point de vue de la thermodynamique, sont en fait des *énergies libres*. En particulier, les forces généralisées qui dérivent de tels potentiels macroscopiques sont donc pour une part dues à des variations d'entropie.

Travail et chaleur

La variation de l'énergie interne $\langle H \rangle$ est, par un calcul analogue à (4.11) :

$$\frac{\partial \langle H \rangle}{\partial \lambda} = \left\langle \frac{\partial H}{\partial \lambda} \right\rangle - \beta \left[\left\langle H \frac{\partial H}{\partial \lambda} \right\rangle - \langle H \rangle \left\langle \frac{\partial H}{\partial \lambda} \right\rangle \right].$$

On reconnaît dans le premier terme le travail reçu $\frac{dW}{d\lambda}$, qui est la moyenne des variations d'énergies des microétats. Le second terme, dû à la variation des poids statistiques $\exp(-\beta H)$ des différentes microétats, est la *chaleur reçue* par le système. Notant Q la chaleur on a alors, en utilisant (4.12) et $TdS = dE - dF$,

$$\frac{dE}{d\lambda} = \frac{dW}{d\lambda} + \frac{dQ}{d\lambda}, \quad \text{avec} \quad \frac{dQ}{d\lambda} = T \frac{dS}{d\lambda},$$

pour une transformation isotherme réversible. Ceci est conforme à la thermodynamique macroscopique, qui décrit l'évolution de l'énergie en fonction du travail et de la chaleur.

Réponse linéaire statique

Le résultat précédent pour la variation de $\langle H \rangle$ peut être réécrit pour une observable A quelconque. On le trouve en général présenté sous une forme légèrement différente, avec un hamiltonien dépendant de λ sous la forme

$$H(\lambda, \mathbf{Q}, \mathbf{P}) = H_0(\mathbf{Q}, \mathbf{P}) - \lambda B(\mathbf{Q}, \mathbf{P}), \quad (4.13)$$

où B est une observable donnée. On a donc $-\frac{\partial H}{\partial \lambda} = B$ dans les formules précédentes. Au premier ordre dans le paramètre λ , la variation de la moyenne de A fait intervenir sa covariance⁶ avec B , évaluée pour $\lambda = 0$:

$$\langle A \rangle_\lambda = \langle A \rangle_0 + \lambda \beta \text{Cov}|_{\lambda=0}(A, B) + O(\lambda^2). \quad (4.14)$$

4.3 Cas de degrés de liberté indépendants

4.3.1 Sous-systèmes indépendants

Pour un hamiltonien de la forme (2.17), somme d'énergies associées à des degrés de liberté différents, on a une forme tensorisée de la densité de probabilité canonique :

$$\rho_{\text{NVT}}(\mathbf{Q}, \mathbf{P}) = \prod_{k=1}^p \rho_{\text{NVT}}^{(k)}(\mathbf{Q}_k, \mathbf{P}_k),$$

où

$$\rho_{\text{NVT}}^{(k)}(\mathbf{Q}_k, \mathbf{P}_k) = \frac{1}{Z_k} \exp[-\beta H_k(\mathbf{Q}_k, \mathbf{P}_k)], \quad Z_k = \int \exp[-\beta H_k(\mathbf{Q}_k, \mathbf{P}_k)] d\mathbf{Q}_k d\mathbf{P}_k.$$

La fonction de partition globale s'écrit elle aussi $Z = \prod_k Z_k$. Une telle forme de la distribution de probabilité est caractéristique de l'indépendance statistique entre les observables dépendant des $\mathbf{Q}_k, \mathbf{P}_k$ et des $\mathbf{Q}_l, \mathbf{P}_l$ si $l \neq k$.

4.3.2 Indépendance des positions et des quantités de mouvement pour une collection de particules

En fait, on peut pour tout Hamiltonien séparable (somme des énergies cinétiques associées à chacune des coordonnées des $3N$ quantités de mouvements $\mathbf{p}_i$ d'une part, et de l'énergie potentielle $V(\mathbf{Q})$ d'autre part), tensoriser la mesure de probabilité entre les impulsions et les positions selon

$$d\rho_{\text{NVT}}(\mathbf{Q}, \mathbf{P}) = \frac{1}{N!} \left[\prod_{\substack{1 \leq i \leq N \\ 1 \leq \alpha \leq 3}} \frac{1}{\Lambda} \exp\left(-\beta \frac{p_{i,\alpha}^2}{2m}\right) \right] \times \frac{1}{Z_q} \exp[-\beta V(\mathbf{Q})] d\mathbf{Q} d\mathbf{P},$$

où Λ est la longueur d'onde thermique introduite en (3.10). Ainsi, chacune des composantes $p_{i,\alpha}$ est indépendante des autres et indépendante des positions, et les quantités de mouvement sont distribuées selon la loi de Maxwell (3.5), dont on établit ainsi la validité générale pour des particules classiques quelles que soient leurs interactions. La fonction de partition Z s'écrit :

$$Z = \frac{Z_q}{N! \Lambda^{3N}}, \quad Z_q = \int \exp[-\beta V(\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_N)] d\mathbf{q}_1 d\mathbf{q}_2, \dots, d\mathbf{q}_N.$$

L'intégrale de configuration Z_q ne se calcule pas analytiquement en général, et il faut recourir à des méthodes numériques approchées pour calculer les moyennes d'ensemble pour la mesure canonique (voir chapitre 5).

6. La covariance des observables A et B est $\text{Cov}(A, B) = \langle AB \rangle - \langle A \rangle \langle B \rangle$. Ici, les moyennes sont prises par rapport à la mesure de Boltzmann associée au Hamiltonien non perturbé $H(0, \cdot, \cdot)$.

4.3.3 Gaz parfait, grandeurs idéales et d'excès dans les fluides

Le formalisme canonique permet une détermination simple et directe des propriétés thermodynamiques du gaz parfait, pour lequel $Z_q = V^N$, et ainsi⁷

$$Z_{\text{id}}(N, V, T) = \frac{V^N}{N! \Lambda^{3N}}.$$

Ceci donne l'énergie libre (par (4.8))

$$F(N, V, T) = Nk_B T [\ln(\rho \Lambda^3) - 1], \quad (4.15)$$

puis, par dérivation, l'entropie, l'énergie interne et la pression, pour lesquelles on retrouve les résultats (3.16), (3.15) et (3.19). Le potentiel chimique du gaz parfait vaut :

$$\mu = k_B T \ln(\rho \Lambda^3). \quad (4.16)$$

Il est négatif et nettement plus grand que $k_B T$ en valeur absolue, en raison de l'inégalité $\rho \Lambda^3 \ll 1$. Ceci provient de l'augmentation d'entropie résultant de l'ajout d'une particule.

En général, l'énergie libre et ses dérivées, pour un système de N particules avec des interactions, peuvent s'écrire comme une somme d'une contribution « idéale » qui est la valeur de la grandeur considérée dans le gaz parfait, et d'une contribution d'« excès ». Cette séparation se déduit directement de l'écriture de la fonction de partition comme produit :

$$Z = Z_{\text{id}} \cdot \frac{Z_q}{V^N},$$

et on écrit par exemple $F = F^{\text{id}} + F^{\text{ex}}$, avec F^{id} donnée par (4.15) et

$$F^{\text{ex}} = -k_B T \ln \left[\frac{Z_q}{V^N} \right].$$

Lorsque l'on recherche les propriétés macroscopiques du système en mesurant des moyennes d'observables dans un calcul numérique, on peut se dispenser de calculer la contribution connue du gaz parfait (contribution cinétique) pour se concentrer sur la contribution inconnue de la partie configurationnelle de l'énergie libre. Ainsi, la méthode de Monte-Carlo pour échantillonner la distribution de Boltzmann-Gibbs de l'ensemble canonique (voir chapitre 5) ne s'intéresse qu'à la simulation de la loi de probabilité marginale par rapport aux positions. Pour calculer par exemple c_v , plutôt que d'utiliser la formule (2.19), on utilise la valeur connue $c_v = \frac{3}{2}k_B$ pour les gaz parfaits monoatomiques, et on évalue seulement

$$c_v^{\text{ex}} = \frac{\text{Var}(E_p)}{Nk_B T^2},$$

en échantillonnant les fluctuations de la seule énergie potentielle E_p .

4.3.4 Cas des oscillateurs harmoniques indépendants

Motivation : solide à basse température

À température suffisamment basse, les N particules en interaction que nous étudions restent au voisinage d'un minimum de l'énergie potentielle. C'est en effet le terme d'énergie qui domine dans la minimisation de l'énergie libre. Soit $\mathbf{Q}_0$ le vecteur des positions qui minimisent le potentiel $V(\mathbf{Q})$. On peut donc écrire, en développant le potentiel $V(\mathbf{Q})$ au voisinage de $\mathbf{Q}_0$ au second ordre dans le vecteur des déplacements $\mathbf{U} = \mathbf{Q} - \mathbf{Q}_0$,

⁷. La notation id renvoie au gaz parfait, également appelé gaz idéal – la traduction anglaise est *ideal gas*.

$$V(\mathbf{Q}_0 + \mathbf{U}) = V(\mathbf{Q}_0) + \frac{1}{2} \mathbf{U} \cdot \mathbf{K} \cdot \mathbf{U} + O(|\mathbf{U}|^3),$$

où la matrice symétrique $\mathbf{K} = \nabla^2 V(\mathbf{Q}_0)$ est appelée *matrice des raideurs* (ou *matrice dynamique* en physique des solides). On suppose par la suite que $\mathbf{K}$ est définie positive (*i.e.* on se place au voisinage d'un minimum strict). Par un changement de variable diagonalisant $\mathbf{K}$, on peut réécrire le Hamiltonien comme⁸

$$H(\tilde{\mathbf{Q}}, \tilde{\mathbf{P}}) = \sum_{k=1}^{3N} \left(\frac{\tilde{p}_k^2}{2m} + \frac{1}{2} K_k \tilde{q}_k^2 \right),$$

c'est-à-dire comme somme des hamiltoniens de $3N$ oscillateurs harmoniques (à une dimension) indépendants. À chacun est associé une raideur K_k , et une pulsation propre

$$\omega_k = \sqrt{\frac{K_k}{m}}.$$

Un cas particulier très important en pratique est celui des cristaux, arrangements périodiques pour lesquels la manière appropriée de diagonaliser la matrice $\mathbf{K}$ a été étudiée en détail en physique des solides (voir l'Exercice 4.12 pour un exemple simplifié de système cristallin).

Résultats

La fonction de partition se factorise⁹ en $Z = \prod_k Z_k$, avec $Z_k = \frac{k_B T}{\hbar \omega_k}$. L'énergie libre est alors :

$$F = -k_B T \sum_{k=1}^{3N} \ln \left(\frac{k_B T}{\hbar \omega_k} \right). \quad (4.17)$$

Limite classique de la mécanique quantique pour un oscillateur harmonique

Il est utile d'étudier la condition pour qu'on puisse négliger les effets quantiques pour un oscillateur harmonique. Avec $q \sim \sqrt{k_B T / K}$ et $p \sim \sqrt{k_B T m}$, on peut ignorer les incertitudes Δq et Δp liées par la condition de Heisenberg (3.9) pourvu que l'on ait $\hbar \ll k_B T \sqrt{\frac{m}{K}}$, soit

$$\hbar \omega \ll k_B T. \quad (4.18)$$

Capacité calorifique

Avec une famille d'oscillateurs harmoniques indépendants, une quantité ne dépendant pas des pulsations propres est la *capacité calorifique*. En effet celle-ci est simplement égale au nombre d'oscillateurs indépendants, multiplié par la constante de Boltzmann k_B . La capacité calorifique à volume constant est en effet

$$C_v = T \frac{\partial S}{\partial T} = -T \frac{\partial^2 F}{\partial T^2},$$

On obtient alors

$$C_v = 3N k_B. \quad (4.19)$$

8. Dans cette transformation, la matrice qui fait passer des p_i^α aux $\tilde{p}_k$ est la transposée de celle qui transforme les u_i^α en $\tilde{q}_k$. On vérifie (exercice) que les $\tilde{q}_k$ et les $\tilde{p}_k$ sont des variables conjuguées pour le Hamiltonien.

9. Dans le cas de particules confinées au voisinage des sites dont les positions sont données par $\mathbf{Q}_0$, on considère que les *sites* sont discernables et on ne divise pas par $N!$ dans le calcul de la fonction de partition.

Ce résultat constitue la *loi de Dulong et Petit* pour la capacité calorifique des solides. Les recherches pour expliquer les écarts systématiques à cette loi au tournant du 20^{ème} siècle (on observe des capacités calorifiques plus faibles à basse température) conduisirent à l'introduction, d'abord arbitraire, d'effets de quantification de l'énergie des oscillateurs, qui ultérieurement s'expliquèrent par la mécanique quantique. D'après (4.18), ces écarts sont d'autant plus manifestes que la température est basse, ou bien que les raideurs sont grandes. Ainsi, le diamant, matériau dont les modules élastiques sont très élevés, présente toujours à 300 K des écarts notables par rapport à la loi de Dulong et Petit.

4.4 Quelques remarques sur les applications du formalisme canonique

Les applications de la distribution de Boltzmann-Gibbs et les problèmes que l'on peut traiter dans l'ensemble canonique sont innombrables et ce bref chapitre n'en donne qu'un échantillon assez réduit. Quelques prolongements sont suggérés dans les exercices. En guise de conclusion sur le sujet, nous formulons ici quelques observations générales.

4.4.1 Gaz, solides et liquides

Nous avons rencontré deux systèmes dont les propriétés se calculent simplement, en particulier dans l'ensemble canonique. Ce sont le gaz parfait, pour lequel les particules n'ont pas d'interaction, et la collection d'oscillateurs harmoniques indépendants, à laquelle se ramènent les systèmes dont l'énergie potentielle est une fonction quadratique (positive) des variables. Ces deux systèmes exactement solubles fournissent des bases efficaces pour étudier les propriétés des gaz et des solides réels. Les interactions dans un gaz, l'écart à l'harmonicité dans un solide, peuvent être traités perturbativement dans certains régimes, c'est-à-dire que l'on peut utiliser des développements asymptotiques des grandeurs étudiées dans les petits paramètres qui expriment la différence avec le modèle dont la solution exacte est connue (voir par exemple l'Exercice 4.11).

Il n'en est rien pour les liquides, et c'est pourquoi leur étude s'appuie aussi fortement sur la simulation numérique.

Remarquons toutefois que le recours à l'approche numérique se justifie aussi pleinement quand il s'agit de propriétés de systèmes « simples », tels qu'un cristal quasi-harmonique, en présence de défauts, ou lors de perturbations de forte amplitude – toutes circonstances qui sollicitent le système hors du domaine de validité des hypothèses des modèles solubles ! Une application traditionnelle de la simulation par dynamique moléculaire est ainsi l'étude de l'effet d'un bombardement d'un réseau cristallin par des particules ou des rayonnements susceptibles de fournir une énergie suffisante pour déplacer des atomes hors de leur site d'équilibre. Les parois métalliques destinées à contenir le rayonnement de matériaux fissiles subissent de telles agressions, et l'estimation de la perte éventuelle de leur capacité de confinement est un objectif important pour la sûreté des installations nucléaires, par exemple.

4.4.2 Approche de l'équilibre et thermalisation

Le gaz parfait, les oscillateurs harmoniques, et généralement tous les systèmes dans lesquels on peut isoler des ensembles de degrés de liberté indépendants, sont simples du point de vue de leurs propriétés statiques d'équilibre. Il faut remarquer qu'ils sont cependant singuliers du point de vue de l'approche de l'équilibre et de la dynamique. En effet, il n'y a aucun transfert d'énergie entre les composantes indépendantes. On retrouve donc les mêmes difficultés que dans l'approche microcanonique, et on peut invoquer de même de petits termes de couplage pour rendre le modèle dynamiquement cohérent.

Une question supplémentaire se pose dans l'ensemble canonique : le fonctionnement dynamique du couplage avec le thermostat, qui, lui aussi, s'opère par un Hamiltonien que l'on n'écrit pas. En fait, on a coutume de considérer, dans les questions de dynamique, que les échanges avec

le thermostat se font sur des temps très longs, de sorte que leur influence ne se fait pas sentir pendant la durée d'étude du système, à l'échelle de ses temps de relaxation propres. On admet cependant que le thermostat a joué son rôle de préparation de l'état initial. Cela revient à considérer des configurations initiales distribuées selon la mesure canonique, qui évoluent ensuite par la dynamique hamiltonienne du système découplé du thermostat (ce qui est bien cohérent, car la dynamique hamiltonienne préserve la mesure canonique).

Exercices

Exercice 4.1. Que vaut la capacité calorifique molaire à volume constant d'un gaz de molécules diatomiques (comme l'azote et l'oxygène de l'air) ?

Exercice 4.2. On considère un système de N particules interagissant par un potentiel de paire, dans un volume V à température T , avec des conditions aux limites périodiques.

- (1) Au moyen d'un changement de variables dans l'espace des phases $\mathcal{E}$, montrer que l'effet d'un changement de volume sur la distribution de Boltzmann-Gibbs se ramène à la variation d'un paramètre du hamiltonien. On veillera à utiliser des changements de variable qui conservent la mesure μ_0 .
- (2) Dédurre la formule (2.23) pour la pression, à partir de la dérivée de l'énergie libre par rapport au volume.

Exercice 4.3 (Potentiels en puissance inverse). On se replace dans le cas de l'Exercice 2.3. Établir les mêmes conclusions à partir de la distribution de Boltzmann-Gibbs dans l'ensemble canonique.

Exercice 4.4. Établir la formule (2.19) sans utiliser la fonction de partition ni la forme de la mesure canonique, avec une approche similaire à la preuve de la formule (2.21) (chapitre 3, page 42).

Exercice 4.5 (Calcul du potentiel chimique par la méthode d'insertion de Widom). On étudie un système de N particules dans l'ensemble canonique, et on cherche à évaluer le potentiel chimique d'excès μ^{ex} .

- (1) Écrire μ^{ex} en fonction du rapport des intégrales de configuration Z_q à $N+1$ et à N particules
- (2) On place au hasard une $N+1^{\text{ème}}$ particule, avec une probabilité uniforme dans le volume V occupé par le système de N particules à l'équilibre. On note W l'énergie potentielle d'interaction de cette particule supplémentaire avec les N premières. Montrer que $\exp[-\beta\mu^{\text{ex}}]$ est égale à la moyenne sur la position de cette particule test de la moyenne selon la distribution de l'ensemble canonique des N particules de $\exp[-\beta W]$.

Exercice 4.6 (Fréquence des collisions dans un système de sphères dures). Montrer que le nombre moyen de collisions par particule et par unité de temps dans un système de sphères dures est

$$n_c = 2\rho\sigma^2 \sqrt{\frac{\pi k_B T}{m}} g(\sigma_+),$$

$g(\sigma_+) = \lim_{r \rightarrow \sigma, r > \sigma} g(r)$ étant l'amplitude de la discontinuité de la fonction de corrélation de paires en $r = \sigma$.

Exercice 4.7 (Pression du fluide de sphères dures (1)). Pour les sphères dures on ne peut pas exprimer la pression à partir de la formule (2.23), car on n'a pas de potentiel différentiable pour calculer des forces. On peut néanmoins obtenir une formule fondée sur le bilan de quantité de mouvement dans les chocs.

- (1) Calculer la loi du choc, qui donne les vitesses après le choc en fonction des vitesses antérieures au choc. Les collisions conservent la quantité de mouvement et l'énergie cinétique.
- (2) Montrer que l'on a :

$$P = \rho k_B T \left[1 + \frac{2\pi\rho}{3} \sigma^3 g(\sigma_+) \right],$$

où $g(\sigma_+)$ est défini dans l'Exercice 4.6.

Exercice 4.8 (Pression du fluide de sphères dures (2)). On approche le comportement des sphères dures par celui des sphères molles, c'est-à-dire interagissant par le potentiel (1.3), dans la limite $n \rightarrow +\infty$. À cette fin, on définira la fonction $y(r; n)$ par :

$$g(r; n) = y(r; n)e^{-\beta\mathcal{V}(r; n)}.$$

On rappelle ici que ces fonctions sont relatives au système de sphères molles avec une valeur donnée de l'exposant n . Dans la limite des sphères dures ($n \rightarrow +\infty$), on supposera que $y(r; n)$ tend vers une fonction régulière y , à la différence de $e^{-\beta\mathcal{V}(r; n)}$ qui tend vers la fonction de Heaviside $H(r - \sigma)$. Discuter rapidement cette hypothèse et retrouver le résultat de l'Exercice 4.7.

Exercice 4.9 (Équilibre d'un gaz isotherme sous gravité). Un gaz parfait étant en équilibre dans le champ de la pesanteur, donner sa densité ρ et sa pression P en fonction de l'altitude z .

Exercice 4.10 (Réponse linéaire à un champ de forces extérieur). À un système de N particules en interaction, traité dans l'ensemble canonique, on applique un champ de forces extérieures qui dérive du potentiel $\varphi(\mathbf{r})$, supposé petit, et dont les effets seront traités au premier ordre. Le hamiltonien est donc de la forme (4.13), avec $\lambda = 1$ et

$$B(\mathbf{Q}, \mathbf{P}) = \sum_{i=1}^N \varphi(\mathbf{q}_i).$$

- (1) En utilisant (4.14), calculer au premier ordre la variation de $\langle A \rangle$ à l'équilibre induite à température constante par l'application du champ extérieur $\varphi(\mathbf{r})$, dans le cas où l'observable A est de la forme :

$$A(\mathbf{Q}) = \sum_{i=1}^N \psi(\mathbf{q}_i),$$

avec une fonction ψ quelconque. En déduire une expression de la variation du champ de densité à un corps $\rho^{(1)}(\mathbf{r})$ faisant intervenir la densité à deux corps $\rho^{(2)}$ introduite en (2.27), du système non perturbé.

- (2) Montrer que pour un système homogène (fluide) on peut exprimer ce résultat avec les transformées de Fourier de φ et de la fonction de corrélation de paires, faisant ainsi apparaître le facteur de structure $S(k)$ (on pourra admettre pour simplifier que la perturbation s'annule en-dehors d'une certaine région).
- (3) Dans l'hypothèse où le potentiel appliqué varie lentement dans l'espace, écrire au premier ordre la variation du champ de densité du système fluide qui en résulte à l'équilibre, par une approche purement macroscopique (statique des fluides), faisant apparaître la compressibilité isotherme χ .
- (4) En déduire la relation (2.29) entre facteur de structure et compressibilité isotherme.

Exercice 4.11 (Deuxième coefficient du viriel). On appelle *développement du viriel* l'écriture de la pression d'un fluide comme série entière de la densité, sous la forme :

$$P(\rho, T) = \rho k_B T (1 + b_2(T)\rho + b_3(T)\rho^2 + \dots),$$

et le coefficient $b_i(T)$ est le i -ème *coefficient du viriel*. On se propose d'évaluer $b_2(T)$ dans le cas d'un système de particules interagissant par un potentiel de paire ne dépendant que de la distance, $\mathcal{V}(r)$.

- (1) Écrire l'intégrale de configuration Z_q en fonction des quantités f_{ij} associées aux paires de particules, définies par :

$$f_{ij} = \exp[-\beta\mathcal{V}(r_{ij})] - 1,$$

en adoptant l'approximation suivante dans l'intégrande de l'intégrale :

$$\prod_{1 \leq i < j \leq N} (1 + f_{ij}) \simeq 1 + \sum_{1 \leq i < j \leq N} f_{ij}.$$

Discuter cette approximation et l'utiliser pour évaluer l'énergie libre d'excès F^{ex} .

- (2) En déduire la pression d'excès et la valeur du deuxième coefficient du viriel $b_2(T)$. Que vaut-il pour les sphères dures ? Vérifier que l'on retrouve l'approximation donnée par (3.8).
- (3) Dans le cas général, montrer que le développement de la pression jusqu'au second ordre en ρ correspond à une certaine approximation de basse densité pour la fonction de corrélation de paires $g(r)$ dans la formule (2.26).

Exercice 4.12 (Modes propres d'une chaîne de ressorts). On considère un système unidimensionnel formé d'une chaîne de particules de masse m liées à leurs voisins par des ressorts de raideur K , avec des conditions aux limites périodiques. Trouver le changement de coordonnées qui permet d'écrire le hamiltonien du système sous forme d'une somme de hamiltoniens associés à des oscillateurs harmoniques indépendants. On cherchera les coordonnées de ces oscillateurs en considérant des modulations sinusoïdales du déplacement le long de la chaîne, avec des longueurs d'onde sous-multiples de la longueur de la chaîne.

Échantillonnage de la mesure canonique

5.1	Le schéma de Metropolis-Hastings	76
5.1.1	L'algorithme de Metropolis	76
5.1.2	Le schéma de Metropolis-Hastings	79
5.1.3	Quelques raffinements	81
5.2	La dynamique de Langevin	82
5.2.1	Brève introduction pratique aux équations différentielles stochastiques	82
5.2.2	La dynamique de Langevin	85
5.3	Tests statistiques de convergence	88
5.3.1	Vitesse de convergence pour les chaînes de Markov	89
5.3.2	Convergence de certaines observables	91
5.4	Application numérique : fluide de Lennard-Jones	93
	Exercices	96

Ce chapitre présente des méthodes d'échantillonnage de la mesure canonique. L'objectif est d'engendrer, par une méthode numérique adéquate, une suite de configurations microscopiques $(q^n, p^n)_{n \geq 0}$ telles que

$$\lim_{N_{\text{iter}} \rightarrow +\infty} \frac{1}{N_{\text{iter}}} \sum_{n=0}^{N_{\text{iter}}-1} A(q^n, p^n) = \int_{\mathcal{E}} A(q, p) d\rho_{\text{can}}(q, p), \quad (5.1)$$

avec

$$d\rho_{\text{can}}(q, p) = Z^{-1} e^{-\beta H(q, p)} dq dp, \quad Z = \int_{\mathcal{E}} e^{-\beta H(q, p)} dq dp. \quad (5.2)$$

Remarquons que l'échantillonnage des impulsions ne pose pas de problème (car les impulsions peuvent être considérées comme des variables aléatoires gaussiennes indépendantes), et la vraie difficulté est d'échantillonner la partie configurationnelle de la mesure canonique, c'est-à-dire la mesure

$$d\pi_{\text{can}}(q, p) = \tilde{Z}^{-1} e^{-\beta V(q)} dq, \quad \tilde{Z} = \int_{\mathcal{D}^N} e^{-\beta V(q)} dq,$$

où $\mathcal{D}^N$ est l'ensemble des configurations accessibles. Cette tâche n'est pas triviale car le nombre de degrés de libertés du système est usuellement très grand, et la mesure $d\pi_{\text{can}}$ très concentrée (*i.e.* elle prend des valeurs non négligeables sur un ensemble de très petite taille). Des méthodes usuelles d'intégration numérique ou de quadrature (éléments finis, différences finies, etc) ne peuvent pas être utilisées du fait de la grande dimension de l'espace. En effet, le nombre de points à considérer est bien trop grand : si on utilise par exemple des différences finies, il y a $K^{N_{\text{dim}}}$ points à considérer, le nombre de degrés de libertés valant $N_{\text{dim}} \sim 10^{3-6}$ dans la pratique. De même, des méthodes usuelles d'échantillonnage en probabilités (algorithme de rejet, fonction d'importance) ne peuvent

généralement pas être utilisées en tant que telles car on ne dispose pas d'une bonne approximation *a priori* de la mesure à échantillonner.

Nous présentons dans ce chapitre des techniques couramment employées dans la pratique quotidienne de la simulation moléculaire pour calculer des moyennes, utilisant une chaîne de Markov (en particulier le schéma de Metropolis-Hastings), ou des équations différentielles stochastiques (surtout la dynamique de Langevin) convenablement discrétisées. Dans les deux cas, on peut montrer que la mesure canonique est préservée par la dynamique. On peut même obtenir des théorèmes de convergence, du type de la Loi des Grands Nombres, pour justifier des limites telles que (5.1). Ainsi, contrairement au cas microcanonique, l'ergodicité est ici un résultat qu'on peut prouver rigoureusement. Cependant, on verra aussi que, même si l'ergodicité est assurée d'un point de vue théorique, il se peut qu'en pratique les temps de simulations associées soient prohibitivement longs pour observer la convergence de la moyenne (5.1).

Finissons par un point d'avertissement important : les dynamiques que nous présentons dans ce chapitre ne sont pas des dynamiques physiques ! On doit simplement les considérer comme des moyens commodes d'échantillonner une mesure de probabilité.

5.1 Le schéma de Metropolis-Hastings

Le schéma de Metropolis et ses raffinements permettent d'échantillonner une mesure de probabilité connue à sa constante de normalisation près. Ceci est heureux en pratique car les fonctions de partitions sont difficiles à calculer. Le cadre théorique qui sous-tend ces algorithmes est celui de la théorie des chaînes de Markov. Une chaîne de Markov est un processus aléatoire à temps discret dont le résultat à l'itération $n + 1$ ne dépend que du résultat à l'itération n , et non des itérations précédentes (incrément sans mémoire).

Une chaîne de Markov est déterminée de manière unique par son noyau de transition, qui donne la probabilité de passer d'une configuration q à une configuration q' . La suite des configurations visitées $(q^0, q^1, \dots, q^n, \dots)$ est une suite de variables aléatoires *corrélées*. En particulier, on ne peut pas appliquer les théorèmes de convergence usuels pour les variables aléatoires indépendantes et identiquement distribuées (loi des grands nombres, théorème de la limite centrale, etc). Cependant, il existe des théorèmes analogues dans le cas des chaînes de Markov, qui permettent de conclure à la convergence de l'échantillonnage (voir § 5.3).

5.1.1 L'algorithme de Metropolis

L'algorithme historique de Metropolis *et al.* [26] utilise une fonction de proposition $T(q, q')$ qui permet, à partir d'une configuration donnée q , de proposer une nouvelle configuration q' en perturbant la configuration q . La fonction $T(q, \cdot)$ est la densité d'une mesure de probabilité :

$$T(q, \cdot) \geq 0, \quad \int_{\mathcal{D}^N} T(q, q') dq' = 1.$$

Pour l'algorithme de Metropolis originel, cette fonction de proposition doit être symétrique, au sens où

$$T(q, q') = T(q', q).$$

ALGORITHME DE METROPOLIS

Algorithme 5.1. Pour $n \geq 0$,

- (1) proposer une nouvelle configuration $\tilde{q}^{n+1}$ à partir de la configuration q^n , avec une fonction de proposition symétrique $T(q^n, \tilde{q}^{n+1})$;
- (2) accepter la nouvelle configuration avec une probabilité

$$p = \min \left(1, \frac{\pi_{\text{can}}(\tilde{q}^{n+1})}{\pi_{\text{can}}(q^n)} \right) = \min \left(1, e^{-\beta[V(\tilde{q}^{n+1}) - V(q^n)]} \right),$$

et poser $q^{n+1} = \tilde{q}^{n+1}$ dans ce cas ; sinon $q^{n+1} = q^n$.

Remarque 5.1 (Accepter avec une probabilité p). En pratique, pour décider si on effectue une action avec une probabilité $0 \leq p \leq 1$, on tire un nombre aléatoire U uniformément entre 0 et 1 (on notera $U \sim \mathcal{U}[0, 1]$), et on effectue l'action si $U \leq p$. Ainsi, si p est petit, il est peu probable que $U \leq p$. La génération d'un nombre aléatoire selon $\mathcal{U}[0, 1]$ se fait par exemple en Scilab avec la commande **rand**.

On peut interpréter cet algorithme de la manière suivante : si la nouvelle configuration a une énergie inférieure à la configuration actuelle, on l'accepte à coup sûr ; sinon, on l'accepte avec une certaine probabilité, d'autant plus grande que la nouvelle énergie est proche de la précédente. Notons qu'il est important d'assurer qu'on peut accepter des énergies plus grandes, car cela permet de franchir des barrières énergétiques (éventuellement en plusieurs coups) et de passer d'un minimum local du potentiel à un autre. Remarquons également que lorsqu'une proposition est rejetée, on garde la configuration précédente, qui compte donc plusieurs fois dans la moyenne (5.1) (sans quoi, ladite moyenne serait fausse!).

On peut considérer des fonctions de proposition non symétriques dans la généralisation l'algorithme proposée par Hastings [21] (voir § 5.1.2). La preuve de convergence de l'Algorithme 5.1 est d'ailleurs faite dans le cadre plus général du § 5.1.2.

Exemple 5.1 (Échantillonnage d'un potentiel double puits). On considère une particule astreinte à se mouvoir dans le potentiel unidimensionnel (3.3), à savoir

$$V(q) = h(q+1)^2(q-1)^2.$$

On va échantillonner la mesure canonique (5.2) pour $\beta = 1$, et différentes valeurs de h , en utilisant le schéma de Metropolis. Il faut en premier lieu décider quelle fonction de proposition symétrique T nous allons utiliser. Le choix le plus simple est d'utiliser des incréments indépendants et aléatoirement distribués selon une loi uniforme ou gaussienne par exemple. Notant $\mathcal{N}(a, b)$ la loi d'une variable gaussienne centrée en a et de variance b , et $\mathcal{U}[a, b]$ la loi uniforme sur $[a, b]$, on peut considérer

$$q' = q + G, \quad G \sim \mathcal{N}(0, \sigma^2) \quad \text{et donc} \quad T(q, q') = \left(\sigma \sqrt{2\pi} \right)^{-1} \exp \left(-\frac{|q' - q|^2}{2\sigma^2} \right) = T(q', q), \quad (5.3)$$

ou

$$q' = q + U, \quad U \sim \mathcal{U}(-\alpha, \alpha) \quad \text{et donc} \quad T(q, q') = \frac{1}{2\alpha} \mathbf{1}_{|q' - q| \leq \alpha} = T(q', q). \quad (5.4)$$

Partant de $q^0 = 0$, une réalisation de l'Algorithme 5.1 avec la fonction de proposition (5.4) donne, pour $h = 1$, $\alpha = 0,4$ et $N_{\text{iter}} = 10^4$ itérations, des positions distribuées de manière à peu près égale dans les deux puits, avec des transitions nombreuses entre les puits (voir la figure 5.1). L'histogramme des positions visitées semble indiquer que les positions sont asymptotiquement correctement réparties. En revanche, si la barrière énergétique h pour passer d'un puits à un autre est plus grande (par exemple $h = 4$), alors l'algorithme de Metropolis reste longtemps coincé dans un des puits, ce qui met en péril la convergence numérique de l'échantillonnage. On pourrait

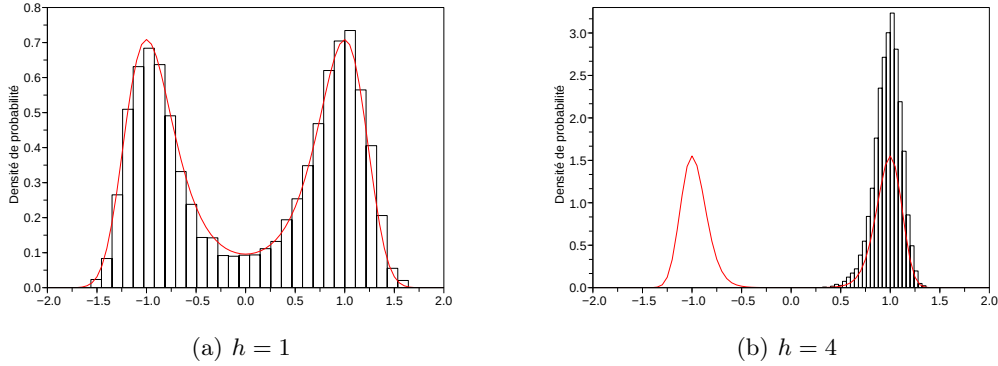


Fig. 5.1. Comparaison entre la densité de probabilité théorique (courbe) et empirique (histogramme) des configurations visitées (algorithme de Metropolis), pour deux valeurs de h .

remédier à ce problème en considérant des incréments moins locaux (qui permettraient de passer directement d'un puits à un autre, comme $\alpha \sim 3$). Cet exemple montre tout le savoir-faire parfois nécessaire pour assurer un échantillonnage correct. Ceci est d'autant plus vrai que la dimension du système est grande.

Exemple 5.2 (Fluide de Lennard-Jones). On considère un fluide de Lennard-Jones, dont les positions des N particules sont notées $q = (q_1, \dots, q_N)$. Partant d'une configuration q^n donnée, on peut créer une nouvelle configuration de plusieurs manières différentes. Par exemple, on peut choisir une particule au hasard (*i.e.* tirer un entier $\{1, \dots, N\}$ au hasard) et modifier aléatoirement sa position d'un déplacement tiré selon une loi uniforme ou gaussienne. Si on déplace par exemple une particule par un mouvement gaussien standard (en dimension 3) multiplié par un facteur σ , on a la fonction de proposition

$$T(q^n, q^{n+1}) = \sum_{j=1}^N \frac{1}{N} \left[\prod_{k \neq j} \delta_{q_k^n}(q_k^{n+1}) \right] \cdot \left(\frac{1}{\sigma\sqrt{2\pi}} \right)^3 \exp \left(-\frac{|q_j^{n+1} - q_j^n|^2}{2\sigma^2} \right).$$

En effet, on a une probabilité $1/N$ de choisir un indice j au hasard (ce qui donne le premier terme de la somme, qui exprime le fait que parmi les positions q_k^{n+1} , seul q_j^{n+1} est différent de q_j^n), et ensuite on effectue un déplacement gaussien pour cette particule, d'où le second terme. On peut imaginer bien d'autres modifications, par exemple effectuer un déplacement pour chaque particule selon un ordre défini (et pas aléatoire comme ci-dessus), en acceptant ou rejetant à chaque fois la nouvelle configuration ainsi produite; ou en déplaçant toutes les particules avant de rejeter ou accepter, auquel cas

$$T(q^n, q^{n+1}) = \left(\frac{1}{\sigma\sqrt{2\pi}} \right)^{3N} \exp \left(-\frac{\sum_{j=1}^N |q_j^{n+1} - q_j^n|^2}{2\sigma^2} \right).$$

Bien sûr, on peut considérer des magnitudes de déplacements σ_i dépendant de la particule en question (si jamais certaines ont des "longueurs d'interaction" plus grandes que les autres, on pourra réduire σ_i).

Remarque 5.2 (Choix des paramètres de la fonction de proposition et taux de rejet). La *doxa* en simulation moléculaire est que les paramètres des fonctions de proposition doivent être choisis afin que le taux d'acceptation dans l'algorithme de Metropolis soit de l'ordre de 0,5. C'est une règle d'usage, fondée sur une constatation empirique, et qui ne peut être justifiée mathématiquement que dans certains cas limites analytiques. Mais en pratique, on se rend compte que cette prescription

permet en effet d'atteindre une bonne performance de l'algorithme en terme de variance (*i.e.* permet de réduire au mieux l'erreur statistique). L'idée est que les erreurs statistiques sont liées à la trop grande corrélation entre les configurations échantillonnées. Cette corrélation peut avoir deux origines pour l'algorithme de Metropolis :

- (i) un taux de rejet trop grand conduit à considérer souvent $q^{n+1} = q^n$, ce qui est, en un sens, une corrélation maximale. Ceci arrive si les déplacements proposés sont trop grands, conduisant à des configurations atypiques ;
- (ii) des déplacements trop petits, ce qui conduit certes à taux de rejet très petit, mais les nouvelles configurations ressemblent alors énormément aux précédentes.

Il s'agit donc de trouver un bon compromis entre le fait d'accepter d'une nouvelle configuration et le fait d'avoir des nouvelles configurations assez différentes des précédentes, d'où cette règle du moitié rejet/moitié acceptation.

5.1.2 Le schéma de Metropolis-Hastings

Le schéma de Metropolis-Hastings a été proposé par Hastings en 1970 [21]. On ne fait plus l'hypothèse que la fonction de proposition est symétrique, mais on tient compte du biais engendré par l'asymétrie de la proposition en modifiant la probabilité d'acceptation/rejet.

ALGORITHME DE METROPOLIS-HASTINGS

Algorithme 5.2. Pour $n \geq 0$,

- (1) proposer une nouvelle configuration $\tilde{q}^{n+1}$ à partir de la configuration q^n , avec une probabilité $T(q^n, \tilde{q}^{n+1})$;
- (2) accepter la nouvelle configuration avec une probabilité

$$p = r(q^n, \tilde{q}^{n+1}) = \min \left(1, \frac{\pi_{\text{can}}(\tilde{q}^{n+1})T(\tilde{q}^{n+1}, q^n)}{\pi_{\text{can}}(q^n)T(q^n, \tilde{q}^{n+1})} \right),$$

et poser $q^{n+1} = \tilde{q}^{n+1}$ dans ce cas ; sinon $q^{n+1} = q^n$.

On vérifie que l'on retrouve bien l'algorithme de Metropolis originel dans le cas d'une fonction de proposition symétrique. La chaîne de Markov correspondante est caractérisée par le noyau de transition

$$P(q, dq') = r(q, q')T(q, q') dq' + \left(1 - \int_{\mathcal{D}^N} r(q, Q)T(q, Q) dQ \right) \delta_q,$$

qui donne la probabilité, partant de q , d'arriver dans un voisinage qd' d'une autre configuration $q' \in \mathcal{D}^N$. On peut donc voir le noyau de transition comme une famille de mesures de probabilité indexée par $q \in \mathcal{D}^N$. La première partie du noyau de transition provient des pas acceptés de q vers q' : on a une probabilité $T(q, q')$ de proposer la configuration q' partant de q , et ensuite une probabilité $r(q, q')$ d'accepter ce mouvement. La seconde partie du membre de droite résume tous les mouvements rejetés, pour lesquels on reste à la position initiale q . La probabilité de rejet est la probabilité complémentaire de la probabilité d'acceptation, qui est la somme des probabilités élémentaires d'acceptation $r(q, q')T(q, q')$ d'un mouvement de q vers q' . On vérifie ainsi que

$$\int_{\mathcal{D}^N} P(q, dq') = 1.$$

On définit ensuite récursivement, pour un borélien $B \in \mathcal{B}(\mathcal{D}^N)$, la probabilité d'arriver en B en exactement n pas, partant de q :

$$P^n(q, B) = \int_{\mathcal{D}^N} P(q, dq') P^{n-1}(q', B).$$

On dit que la chaîne de Markov est *irréductible* si pour tout $q \in \mathcal{D}^N$ et tout borélien $B \in \mathcal{B}(\mathcal{D}^N)$ de mesure de Lebesgue strictement positive, il existe $n \geq 1$ tel que $P^n(q, B) > 0$. Cela traduit une propriété d'accessibilité globale de toutes les régions de l'espace des phases (on peut atteindre un voisinage de n'importe quel point en un nombre fini de pas). Dans le cas du schéma de Metropolis-Hastings, cette accessibilité est liée à un choix judicieux de la fonction de proposition T .

Exemple 5.3. Pour le fluide de Lennard-Jones, on peut vérifier facilement que des déplacements aléatoires d'une ou plusieurs particules permettent de passer d'une configuration à n'importe quelle autre en un nombre fini de pas (comme dans un jeu de taquin, on déplace des particules pour placer q_1 dans la position souhaitée, puis on fait de même pour q_2 , etc). Chaque déplacement aléatoire a une probabilité strictement positive (même si elle est très petite!) d'être observé, ce qui donne au final une probabilité strictement positive de passer d'une configuration à une autre.

On montre que le schéma de Metropolis-Hastings converge au sens du théorème suivant.

Théorème 5.1 (Convergence du schéma de Metropolis-Hastings). *Supposons que la chaîne de Markov engendrée par le schéma de Metropolis-Hastings soit irréductible. Alors, pour toute observable $A \in L^1(\pi_{\text{can}})$, et pour presque tout $q^0 \in \mathcal{D}^N$,*

$$\lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{i=1}^n A(q^i) \rightarrow \langle A \rangle = \int_{\mathcal{D}^N} A(q) \pi_{\text{can}}(q) dq \quad \text{p.s.}$$

Remarquons donc que l'on peut montrer que la dynamique de Metropolis-Hastings est ergodique, contrairement au cas microcanonique, où l'ergodicité est une hypothèse. La preuve de ce résultat utilise un certain nombre de concepts avancés de la théorie des chaînes de Markov. Le lecteur intéressé pourra se référer à [27], et nous nous contentons de donner les éléments principaux de la preuve.

Preuve. La preuve de la convergence se décompose en deux parties : (i) on montre tout d'abord que la mesure π_{can} est une mesure invariante de la dynamique de Metropolis-Hastings ; (ii) comme elle est par ailleurs irréductible (et apériodique, car on a une probabilité non nulle de rester sur place) et admet une mesure de probabilité invariante, on obtient la propriété d'ergodicité le long d'une trajectoire recherchée.

La preuve de convergence se réduit donc à montrer l'invariance de la mesure cible π_{can} . Pour ce faire, on vérifie que pour toute fonction $f \in L^1(\pi_{\text{can}})$,

$$\int_{\mathcal{D}^N} \int_{\mathcal{D}^N} f(q') \pi_{\text{can}}(q) P(q, dq') dq = \int_{\mathcal{D}^N} f(q') \pi_{\text{can}}(q') dq'.$$

Ceci traduit en effet de manière faible (en dualité contre des fonctions intégrables) que, partant de π_{can} et faisant un pas de la dynamique de Metropolis-Hastings avec le noyau P (terme $\pi_{\text{can}}(q)P(q, dq')$), on retrouve bien π_{can} .

On remarque en premier lieu que, notant $a(q, q') = T(q, q')r(q, q')$ la probabilité d'accepter un mouvement de q vers q' , on a la propriété dite de « bilan détaillé »

$$\pi_{\text{can}}(q)a(q, q') = \pi_{\text{can}}(q')a(q', q). \quad (5.5)$$

Alors,

$$\int_{\mathcal{D}^N} \int_{\mathcal{D}^N} f(q') \pi_{\text{can}}(q) a(q, q') dq' dq = \int_{\mathcal{D}^N} \int_{\mathcal{D}^N} f(q') \pi_{\text{can}}(q') a(q', q) dq' dq.$$

Par ailleurs,

$$\begin{aligned} & \int_{\mathcal{D}^N} \int_{\mathcal{D}^N} f(q') \pi_{\text{can}}(q) \left(1 - \int_{\mathcal{D}^N} r(q, Q) T(q, Q) dQ \right) \delta_q(dq') dq = \\ & = \int_{\mathcal{D}^N} f(q') \pi_{\text{can}}(q') \left(1 - \int_{\mathcal{D}^N} a(q', Q) dQ \right) dq'. \end{aligned}$$

En sommant les deux contributions précédentes,

$$\int_{\mathcal{D}^N} \int_{\mathcal{D}^N} f(q') \pi_{\text{can}}(q) P(q, dq') dq = \int_{\mathcal{D}^N} f(q') \pi_{\text{can}}(q') dq'.$$

Ceci conclut la preuve de l'invariance de la mesure. $\square$

Remarque 5.3 (Invariance de la mesure). On retiendra surtout de la preuve précédente qu'une condition nécessaire pour la convergence de la chaîne de Markov est que la mesure cible soit invariante. Ceci peut s'interpréter de deux manières :

- (i) si on avait des conditions initiales canoniquement distribuées, *i.e.* $q^0 \sim d\pi_{\text{can}}$, alors l'ensemble des q^1 donnés par une itération de la chaîne de Markov est encore distribué selon $d\pi_{\text{can}}$;
- (ii) on a également une interprétation trajectorielle : si on considère l'ensemble des configurations visitées le long d'une trajectoire suffisamment longue, elles forment globalement un échantillon distribué selon $d\pi_{\text{can}}$, et ce, bien qu'elles soient corrélées.

Remarque 5.4 (Bilan détaillé). La propriété de bilan détaillé (5.5) est également une propriété importante, qui exprime une certaine réversibilité de la chaîne de Markov : étant dans une configuration q plus probable que la configuration q' , on a une probabilité plus grande d'aller en q partant de q' que le contraire. Ceci est lié à un flux de probabilité assurant la conservation de la mesure invariante.

Exemple 5.4 (Diffusion brownienne ajustée). On présente ici un schéma de Metropolis-Hastings introduit dans [30] (MALA, *Metropolized Adjusted Langevin Algorithm*). Pour une particule dans un potentiel unidimensionnel V , et partant d'une configuration q , on propose une nouvelle configuration q' par une dynamique de gradient perturbée :

$$q' = q - \alpha \nabla V(q) + \sqrt{\frac{2\alpha}{\beta}} U, \quad U \sim \mathcal{N}(0, 1). \quad (5.6)$$

La configuration proposée diffère de la configuration originelle à cause du terme $\alpha \nabla V(q)$ qui tend à ramener les configurations vers des minima locaux de l'énergie potentielle (donc des configurations plus probables), le terme aléatoire permettant de ne pas systématiquement diminuer l'énergie. La probabilité d'atteindre q' partant de q est ainsi

$$T(q, q') = \sqrt{\frac{\beta}{4\pi\alpha}} \exp \left(-\beta \frac{|q' - q + \alpha \nabla V(q)|^2}{4\alpha} \right).$$

Notons que $T(q, q') \neq T(q', q)$ en général, à cause du terme de force $-\nabla V$. En pratique, l'intensité α des perturbations proposées est choisie afin d'obtenir un taux d'acceptation/rejet proche de 0,5, afin que l'algorithme soit le plus efficace possible. Ceci peut être justifié théoriquement dans certains cas. Nous verrons en § 5.3 selon quels critères on peut numériquement chercher un paramètre α optimal. On présente en figure 5.2 quelques résultats d'échantillonnages obtenus avec cet algorithme pour le potentiel (3.3), pour $\alpha = 0, 2$ et $n = 10^4$ itérations de la chaîne de Markov.

5.1.3 Quelques raffinements

De nombreux travaux ont proposé des améliorations des algorithmes de type Metropolis ou Metropolis-Hastings. En fait, ce qu'on cherche à améliorer, ce sont les fonctions de proposition T utilisées. Ainsi,

- (i) on peut proposer de nouveaux mouvements adaptés à la physique du système. Par exemple, pour des fluides de Lennard-Jones ou des mélanges de fluides, il peut être intéressant de combiner des mouvements locaux (du type des algorithmes précédents), et des mouvements non-locaux, comme, par exemple, déplacer une seule particule en tirant sa nouvelle position complètement aléatoirement dans le domaine de simulation (ou seulement parmi des zones de l'espace à densité plus faible), changer le type d'une particule, casser une liaison chimique et en recréer une autre ailleurs, etc ;

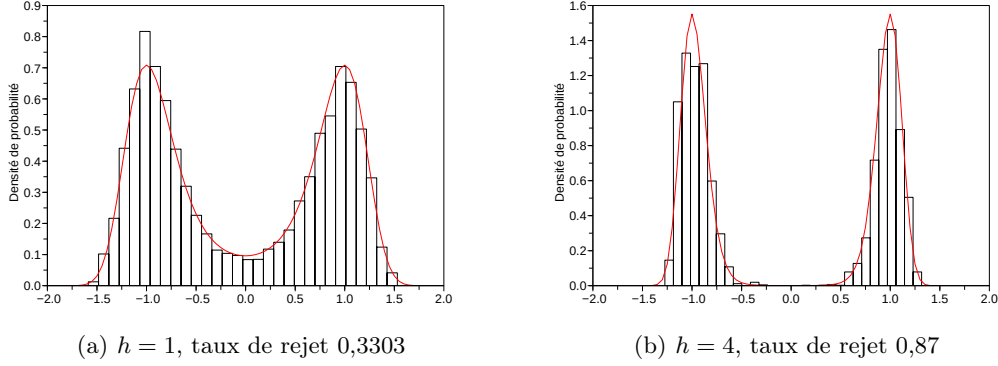


Fig. 5.2. Comparaison entre la densité de probabilité théorique (courbe) et empirique (histogramme) des configurations visitées (algorithme de Metropolis-Hastings, version MALA avec $\alpha = 0.2$).

- (ii) on peut également utiliser des manières plus abstraites de proposer des nouveaux mouvements : un cas typique est celui où l'on simule deux répliques du système à des températures différentes, une température haute (on s'attend donc à ce que la chaîne de Markov des configurations échantillonnées à cette température converge plus rapidement), et une température basse ; les configurations proposées à la température basse étant celles échantillonnées par la température haute. Cette technique, connue sous le nom de *parallel tempering* permet de lutter contre la métastabilité¹ du système (voir l'Exercice 5.3).

5.2 La dynamique de Langevin

On présente dans cette section une manière de modifier la dynamique hamiltonienne pour faire en sorte qu'elle puisse explorer tous les niveaux d'énergie, et avec la bonne fréquence (c'est-à-dire de manière cohérente avec la mesure canonique). On utilise pour ce faire des processus stochastiques, qui sont des dynamiques en temps continu (comme la dynamique hamiltonienne), au contraire des approches de type chaînes de Markov présentées ci-dessus. Cela dit, numériquement, la discrétisation des processus stochastiques conduit naturellement à considérer des chaînes de Markov.

5.2.1 Brève introduction pratique aux équations différentielles stochastiques

Quelques propriétés du mouvement brownien

On rappelle rapidement quelques propriétés élémentaires du mouvement brownien afin de présenter des discrétisations d'équations différentielles stochastiques (EDS) simples. Considérons le processus stochastique q_t donné par

$$dq_t = \sigma dW_t, \quad q_0 = q^0, \quad (5.7)$$

où W_t est un mouvement brownien standard de dimension 1 (pour simplifier). Le mouvement brownien standard est tel que

$$\forall 0 \leq t_0 \leq \dots \leq t_n, \quad W_{t_{i+1}-t_i} \sim \mathcal{N}(0, t_{i+1} - t_i),$$

les variables $(W_{t_{i+1}-t_i})_{i=0, \dots, n-1}$ étant par ailleurs indépendantes. On peut donc voir l'EDS (5.7) comme la limite lorsque $\Delta t \rightarrow 0$ du schéma numérique (il sera suffisant, pour les besoins de ce cours, d'avoir cette image en tête) :

1. C'est à dire une convergence numérique très lente, voir la § 5.3 pour plus de précisions.

$$q^{n+1} = q^n + \sigma\sqrt{\Delta t} G^n, \quad (5.8)$$

où q^n est une approximation de $q_{n\Delta t}$, les $(G^n)_{n \in \mathbb{N}}$ étant des variables aléatoires gaussiennes standard indépendantes. Notons que l'incrément du mouvement brownien est distribué selon une gaussienne de variance Δt , ce qui revient en pratique à ajouter une variable aléatoire gaussienne multipliée par un facteur $\sqrt{\Delta t}$.

On déduit facilement de (5.8) que q^n est une variable gaussienne centrée et q^0 et de variance $n\sigma^2\Delta t$ (et en fait, dans ce cas simple, q^n et $q_{n\Delta t}$ sont distribués selon les mêmes lois). La

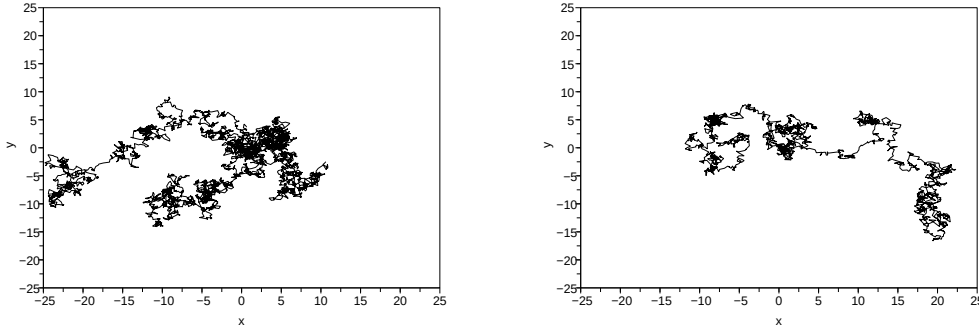


Fig. 5.3. Deux réalisations d'un mouvement brownien (5.7) selon la discrétisation (5.8) pour $\Delta t = 0,1$ et $n\Delta t = 100$.

figure 5.3 présente deux réalisations (*i.e.* deux trajectoires) de (5.7), avec la discrétisation (5.8), dans le cas d'un mouvement brownien de dimension 2. Un mouvement brownien de dimension d est $W_t = (W_{1,t}, \dots, W_{d,t})$, où chacun des $W_{i,t}$ est un mouvement brownien standard de dimension 1, ces différents mouvements browniens étant par ailleurs indépendants.

Evolution en temps de la loi de q_t

Nous avons caractérisé q_t par sa loi (gaussienne centrée en x^0 , de variance $\sigma^2 t$), ce qui était facile car nous en avons une expression analytique explicite. De tels calculs ne peuvent toutefois pas être menés analytiquement en général, et il faut savoir comment évolue la loi des solutions d'une EDS. Pour cela, il est commode de considérer un ensemble (infini) de réalisations de l'EDS (5.7), partant de configurations initiales distribuées selon une mesure $d\pi^0 = \phi(q) dq$. Les configurations sont alors distribuées au temps t selon la loi de densité $\psi(t, \cdot)$.²

Nous allons à présent établir formellement l'équation aux dérivées partielles (EDP) satisfaite par ψ . Nous omettons volontairement de préciser les notions de régularité de fonctions et les espaces fonctionnels adéquats. L'espérance (la moyenne) d'une fonction A est, au temps t ,

$$\mathbb{E}(A(q_t)) = \int_{\mathcal{D}^N} A(q) \psi(t, q) dq.$$

Pour simplifier, on se place dans le cas où $q_t \in \mathbb{R}$ (on indiquera ensuite comment les résultats sont modifiés dans le cas d'un espace de dimension supérieure), auquel cas $G^n \sim \mathcal{N}(0, 1)$. On regarde comment $\mathbb{E}(A(q^n))$ varie d'une itération à l'autre. Pour ce faire, on utilise la discrétisation (5.8), et on effectue un développement de Taylor :

2. Nous admettrons que, pour $t > 0$, la loi de la variable aléatoire q_t est effectivement à densité par rapport à la mesure de Lebesgue. Cela provient de la propriété régularisante de l'équation d'évolution parabolique satisfaite par cette loi.

$$\begin{aligned}
\mathbb{E}(A(q^{n+1})) &= \mathbb{E}\left(A(q^n + \sigma\sqrt{\Delta t} G^n)\right) \\
&= \mathbb{E}\left(A(q^n) + \sigma\sqrt{\Delta t} A'(q^n) G^n + \frac{\sigma^2 \Delta t}{2} A''(q^n)(G^n)^2 + O(\Delta t^{3/2})\right) \\
&= \mathbb{E}(A(q^n)) + \sigma\sqrt{\Delta t} \mathbb{E}(A'(q^n))\mathbb{E}(G^n) + \frac{\sigma^2 \Delta t}{2} \mathbb{E}(A''(q^n))\mathbb{E}((G^n)^2) + O(\Delta t^{3/2}) \\
&= \mathbb{E}(A(q^n)) + \frac{\sigma^2 \Delta t}{2} \mathbb{E}(A''(q^n)) + O(\Delta t^{3/2}),
\end{aligned}$$

en utilisant, pour passer de la deuxième à la troisième ligne, l'indépendance des variables aléatoires q^n et G^n et la linéarité de l'espérance. Ainsi,

$$\int_{\mathcal{D}^N} A(q) \psi((n+1)\Delta t, q) dq = \int_{\mathcal{D}^N} \left[A(q) + \frac{\sigma^2 \Delta t}{2} A''(q) \right] \psi(n\Delta t, q) dq + O(\Delta t^{3/2}). \quad (5.9)$$

Une intégration par parties donne (on suppose que les termes de bords valent 0 ou se compensent) :

$$\int_{\mathcal{D}^N} A(q) \left[\frac{\psi((n+1)\Delta t, q) - \psi(n\Delta t, q)}{\Delta t} - \frac{\sigma^2}{2} \partial_q^2 \psi(n\Delta t, q) \right] dq = O(\sqrt{\Delta t}).$$

Lorsque $\Delta t \rightarrow 0$ avec $n\Delta t \rightarrow t$, on obtient :

$$\int_{\mathcal{D}^N} A(q) \left[\partial_t \psi(t, q) - \frac{\sigma^2}{2} \partial_q^2 \psi(t, q) \right] dq = 0.$$

Cette équation étant valable quelle que soit la fonction A , on en déduit que

$$\partial_t \psi = \frac{\sigma^2}{2} \partial_q^2 \psi.$$

Dans le cas multidimensionnel, on montre de la même manière que

$$\partial_t \psi = \frac{\sigma^2}{2} \Delta \psi, \quad \psi(t=0, \cdot) = \phi(\cdot).$$

Cette EDP est l'équation de Fokker-Planck associée à l'EDS (5.7). On a ainsi une simple équation de diffusion. Si on part initialement d'une mesure de Dirac en x^0 , on sait que l'on trouve au temps t une gaussienne de variance $\sigma^2 t$, centrée en x^0 . Rappelons en effet que la solution élémentaire de l'équation de la chaleur

$$\partial_t \psi = \frac{\sigma^2}{2} \Delta \psi, \quad \psi(t=0, \cdot) = \delta_0,$$

en dimension $3N$ est bien

$$\psi(t, q) = \left(\frac{1}{2\pi\sigma^2 t} \right)^{3N/2} \exp\left(-\frac{q^2}{2\sigma^2 t}\right),$$

comme on peut le vérifier directement.

Une équation différentielle stochastique modèle

La dynamique brownienne ci-dessus était une diffusion libre sur tout l'espace. On souhaiterait la modifier afin d'obtenir un processus qui échantillonne la mesure canonique. Une manière de faire est d'ajouter un terme de dérive, qui ramène la dynamique vers les zones importantes de la mesure (*i.e.* vers des valeurs pas trop grandes de V) au lieu de laisser la diffusion s'échapper. On peut penser par exemple à une dynamique de gradient perturbée par un mouvement brownien, telle que

$$dq_t = -\nabla V(q_t) dt + \sigma dW_t, \quad (5.10)$$

où W_t est un mouvement brownien standard de dimension dN , où d est la dimension de l'espace physique ambiant. Une discrétisation de cette EDS est par exemple le schéma

$$q^{n+1} = q^n - \nabla V(q^n) \Delta t + \sigma \sqrt{\Delta t} G^n, \quad (5.11)$$

où les G^n sont des vecteurs gaussiens standard indépendants et identiquement distribués (i.i.d.) de dimension dN .

Notant encore $\psi(t, \cdot)$ la loi de q_t , on montre que l'équation de Fokker-Planck associée à l'EDS (5.10) est

$$\partial_t \psi = \nabla \cdot (\psi \nabla V) + \frac{\sigma^2}{2} \Delta \psi = \mathcal{L} \psi.$$

On peut en effet motiver cette équation par un travail similaire à celui fait ci-dessus pour la diffusion libre. Il apparaît un terme supplémentaire de la forme

$$- \int_{\mathcal{D}^N} \Delta t \nabla A(q) \cdot \nabla V(q) \psi(n \Delta t, q) dq$$

dans (5.9), qu'on transforme à l'aide d'une intégration par parties pour éliminer l'action des dérivées sur la fonction A .

On peut alors vérifier que, si

$$\sigma^2 = \frac{2}{\beta}, \quad (5.12)$$

la mesure canonique est une solution stationnaire de l'équation de Fokker-Planck, au sens où $\mathcal{L} \pi_{\text{can}} = 0$. Ceci montre bien en effet que si on part d'une variable distribuée initialement selon π_{can} et qu'on la fait évoluer selon (5.10), on aura $\psi(t, \cdot) = \pi_{\text{can}}$ pour tout $t \geq 0$. La vérification de la propriété $\mathcal{L} \pi_{\text{can}} = 0$ ne pose pas de problème particulier :

$$\mathcal{L} (e^{-\beta V}) = \nabla \cdot (e^{-\beta V} \nabla V) + \frac{\sigma^2}{2} \Delta (e^{-\beta V}) = \left(1 - \frac{\beta \sigma^2}{2}\right) \nabla \cdot (e^{-\beta V} \nabla V) = 0,$$

avec le choix (5.12). La relation (5.12), fondamentale pour assurer l'échantillonnage de la bonne mesure, est appelée relation de *fluctuation-dissipation*, dans la mesure où elle détermine la taille typiques des fluctuations aléatoires à imposer (ces fluctuations apportant de l'énergie au système) pour compenser le terme de dissipation $-\nabla V$ et ce, afin que l'énergie moyenne du système soit correcte.

Concernant le schéma numérique (5.11), si le potentiel V vérifie certaines conditions (par exemple, si ∇V est globalement lipschitzien) et que le pas de temps Δt est assez petit, on peut montrer que les positions visitées par le processus stochastique sont distribuées selon une mesure $d\pi_{\Delta t}$ proche de la mesure $d\pi_{\text{can}}(q) = Z^{-1} e^{-\beta V(q)}$. Ce biais peut être corrigé en ajoutant une étape de Metropolis (voir la figure 5.4, et comparer également (5.11) avec l'Algorithme 5.2 avec la fonction de proposition (5.6)). On verra au § 5.3 que, sans correction de type Metropolis, il faudra distinguer dans le calcul de moyenne canonique un terme d'erreur statistique, et un terme de biais lié au pas de temps.

5.2.2 La dynamique de Langevin

En 1827, le botaniste R. Brown a découvert au microscope le mouvement incessant et irrégulier de petites particules en suspension dans l'eau. Le mouvement de ces particules, nommées alors *browniennes* est une conséquence de l'agitation thermique à petite échelle. C'est Albert Einstein, qui, en 1905, a donné la première explication théorique claire de ce phénomène, et c'est pour représenter la position d'une particule brownienne qu'un processus stochastique a été construit pour la première fois, par N. Wiener en 1923. Le modèle de Langevin est un modèle phénoménologique dans lequel on analyse l'effet du fluide sur la particule brownienne en considérant deux forces : une

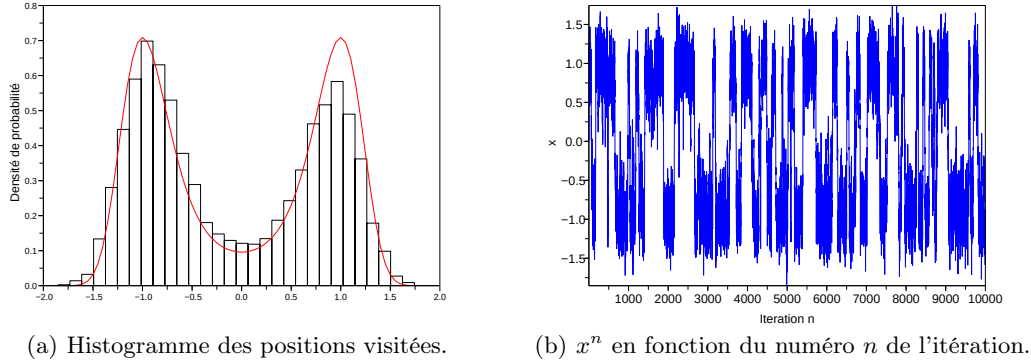


Fig. 5.4. Réalisation de l'EDS (5.10) avec la discrétisation (5.11) pour $\Delta t = 0,05$, $n\Delta t = 50$, $\beta = 1$ pour le potentiel (3.3).

force de frottement visqueux, et une force fluctuante, censée représenter les impacts incessants des molécules du fluide sur la particule brownienne³. La dynamique de Langevin est ainsi donnée par l'EDS suivante :

$$\begin{cases} dq_{i,t} = \frac{p_{i,t}}{m} dt, \\ dp_{i,t} = -\nabla_{q_i} V(q_{i,t}) dt - \xi \frac{p_{i,t}}{m_i} dt + \sigma dW_{i,t}, \end{cases} \quad (5.13)$$

où les $(W_{i,t})_{t \geq 0}$ sont des mouvements browniens standard indépendants de dimension 3. Les paramètres ξ and σ sont respectivement la friction et l'amplitude des fluctuations aléatoires. Ils sont liés par la relation de fluctuation-dissipation, qui assure que la mesure canonique est une solution stationnaire de l'équation de Fokker-Planck associée à (5.13) :

$$\sigma^2 = \frac{2\xi}{\beta}. \quad (5.14)$$

La calcul qui mène à cette relation est analogue au calcul qui mène à (5.12). Il suffit de noter que l'équation de Fokker-Planck associée à (5.13) est

$$\partial_t \psi = (-M^{-1}p \cdot \nabla_q + \nabla_q V \cdot \nabla_p) \psi + \xi \nabla_p (M^{-1}p \psi) + \frac{\sigma^2}{2} \Delta_p \psi,$$

et que $\exp(-\beta H)$ en est une solution stationnaire.

Régimes limites

Notons que si $\xi = 0$ (et donc $\sigma = 0$), on retrouve la dynamique hamiltonienne usuelle (voir la figure 5.5 pour une illustration). On peut également montrer qu'on retrouve (5.10) dans la limite $m \rightarrow 0$. Cependant, en pratique, la dynamique de Langevin est souvent plus efficace pour échantillonner la mesure canonique que sa limite de fortes frictions (5.10) (voir par exemple [5]).

3. En fait ceci n'a pas grand sens : les petites molécules qui constituent la particule brownienne sont tout aussi agitées que celles du fluide. De plus, la séparation d'échelles de temps, sur laquelle se fonde le modèle, entre grosses particules (lentes) et molécules du fluide (rapides) ignore le couplage entre le mouvement des particules browniennes et ceux du fluide à la même échelle : ainsi la vitesse moyenne dans un volume de fluide égal à celui de la particule varie lentement, elle aussi. Le modèle de Langevin, bien qu'ayant joué un rôle important pour le traitement de différents phénomènes à une échelle mésoscopique en espace et en temps, n'est pas une description correcte de la dynamique des particules browniennes, pour laquelle il a initialement été inventé. Fort heureusement, il coïncide dans ce cas, pour des temps assez longs, avec le modèle d'Einstein, qui consiste à écrire (5.10), et est pleinement justifié pour les particules colloïdales en suspension. Et quelle que soit la validité physique de l'équation de Langevin, cela ne remet bien sûr pas en cause l'usage qui en est fait ici, où on fait appel à une dynamique fictive pour échantillonner une mesure.

Mise en œuvre numérique

Une discrétisation de $dp_{i,t} = -\nabla_{q_i} V(q_{i,t}) dt - \xi \frac{p_{i,t}}{m_i} dt + \sigma dW_{i,t}$ est par exemple

$$p_i^{n+1} = p_i^n - \nabla_{q_i} V(q_i^n) \Delta t - \xi \frac{p_i^n}{m_i} \Delta t + \sigma \sqrt{\Delta t} G_i^n,$$

avec $(G_i^n)_{i,n}$ des vecteurs gaussiens standard indépendants. On peut ainsi discrétiser l'EDS (5.13) par le schéma suivant, qui consiste à faire un demi-pas de temps en impulsions, un pas de temps en positions, et un demi-pas de temps en impulsions :

$$\begin{cases} p_i^{n+1/2} = p_i^n - \nabla_{q_i} V(q_i^n) \frac{\Delta t}{2} - \xi \frac{p_i^n}{m_i} \frac{\Delta t}{2} + \frac{1}{2} \sigma \sqrt{\Delta t} G_i^n, \\ q_i^{n+1} = q_i^n + \Delta t \frac{p_i^{n+1/2}}{m_i}, \\ p_i^{n+1} = p_i^{n+1/2} - \nabla_{q_i} V(q_i^{n+1}) \frac{\Delta t}{2} - \xi \frac{p_i^{n+1}}{m_i} \frac{\Delta t}{2} + \frac{1}{2} \sigma \sqrt{\Delta t} G_i^n. \end{cases}$$

Pour des questions de stabilité numérique, on a implicité le second demi-pas de temps en impulsions, mais on peut remettre le schéma sous forme explicite :

$$\begin{cases} p_i^{n+1/2} = p_i^n + \frac{\Delta t}{2} \left(-\nabla_{q_i} V(q_i^n) - \xi \frac{p_i^n}{m_i} + \frac{\sigma}{\sqrt{\Delta t}} G_i^n \right), \\ q_i^{n+1} = q_i^n + \Delta t \frac{p_i^{n+1/2}}{m_i}, \\ p_i^{n+1} = \frac{1}{1 + \frac{\xi \Delta t}{2m_i}} \left(p_i^{n+1/2} - \frac{\Delta t}{2} \nabla_{q_i} V(q_i^{n+1}) + \sigma \frac{\sqrt{\Delta t}}{2} G_i^n \right). \end{cases} \quad (5.15)$$

Notons que si $\xi = 0$ (et donc $\sigma = 0$), on retrouve le schéma de Verlet usuel. Notons également que

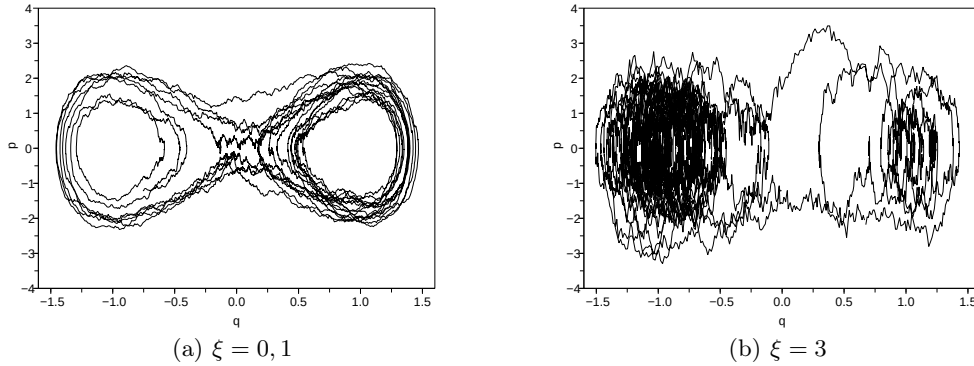


Fig. 5.5. Trajectoire d'une réalisation de la dynamique de Langevin avec la discrétisation (5.15) pour $\Delta t = 0.005$, $n\Delta t = 5$ pour le potentiel (3.3). Pour ξ petit (a), les trajectoires sont (localement) proches des trajectoires hamiltoniennes. Lorsque ξ augmente (b) – ainsi, en raison de (5.14), que l'amplitude du bruit σ – le caractère erratique de la dynamique est plus apparent.

l'on emploie deux fois le même nombre aléatoire dans la première et la dernière équations car on a, dans une certaine mesure, “coupé en deux” le terme aléatoire. Dans le cas $\nabla V = 0$ et $\xi = 0$, on alors

$$p_i^{n+1} = p_i^n + \sigma \sqrt{\Delta t} G_i^n,$$

qui est bien une discrétisation consistante de $dp_{i,t} = \sigma dW_t$. On aurait également pu intégrer la première et la dernière équation en employant des nombres aléatoires indépendants, mais auquel cas il aurait fallu tenir compte du fait qu'on intègre pendant un temps $\Delta t/2$. On aurait ainsi des mises à jour des demi-pas de temps de la forme

$$p_i^{n+1/2} = p_i^n - \nabla_{q_i} V(q^n) \frac{\Delta t}{2} - \xi \frac{p_i^n}{m_i} \frac{\Delta t}{2} + \sigma \sqrt{\frac{\Delta t}{2}} G_i^n,$$

et

$$p_i^{n+1} = p_i^{n+1/2} - \nabla_{q_i} V(q^{n+1}) \frac{\Delta t}{2} - \xi \frac{p_i^{n+1}}{m_i} \frac{\Delta t}{2} + \sigma \sqrt{\frac{\Delta t}{2}} G_i^{n+1/2},$$

avec $G_i^n, G_i^{n+1/2}$ indépendants.

Convergence

On peut montrer, pour toute friction ξ (à condition que σ soit donné par (5.14)) que l'EDS (5.13) admet la mesure canonique comme mesure invariante, et que, partant d'une condition initiale quelconque et moyennant quelques hypothèses sur le potentiel d'interaction V (voir [5]), on a

$$\lim_{T \rightarrow +\infty} \frac{1}{T} \int_0^T A(q_t, p_t) dt = \int_{\mathcal{E}} A(q, p) d\rho_{\text{can}}(q, p).$$

En pratique, on discrétise convenablement (5.13) (avec le schéma (5.15) par exemple), et on calcule une approximation de la moyenne d'ensemble par

$$\lim_{N_{\text{iter}} \rightarrow +\infty} \frac{1}{N_{\text{iter}}} \sum_{n=0}^{N_{\text{iter}}-1} A(q^n, p^n) = \int_{\mathcal{E}} A(q, p) d\rho_{\text{can}}(q, p) + O(\Delta t).$$

Notons en effet que, contrairement au schéma de Metropolis-Hastings qui échantillonne exactement la mesure canonique, la mesure échantillonnée par la dynamique de Langevin n'est pas exactement la mesure canonique, et l'erreur est usuellement de l'ordre de $O(\Delta t)$. Si le pas de temps est suffisamment petit, cette erreur est difficilement mesurable, eu égard aux fluctuations statistiques des quantités que l'on cherche à calculer (voir § 5.3).

5.3 Tests statistiques de convergence

Le Théorème 5.1 montre que, sous des hypothèses souvent facilement vérifiées en pratique, le schéma de Metropolis-Hastings est ergodique. On a également des résultats analogues pour la dynamique de Langevin. Toutefois, en pratique, il est possible que cette convergence soit lente, et que l'algorithme reste coincé pour des temps longs dans certaines régions de l'espace des configurations. Ces régions sont appelées des *états métastables*, et un des grands défis de la simulation moléculaire actuelle est d'échantillonner au mieux la mesure canonique malgré la présence de ces états métastables. Pour mieux comprendre la métastabilité, on présente au § 5.3.1 des résultats de convergence plus précis pour les chaînes de Markov, qui sont, rappelons le, le bon cadre théorique pour la étude la convergence des méthodes numérique.

Pour évaluer la qualité de la convergence, on a souvent recours à des conditions nécessaires de convergence : on pourra demander ainsi que les moyennes empiriques sur une trajectoire d'un certain nombre d'observables aient convergé. Cette approche est détaillée au § 5.3.2, mais est insuffisante pour caractériser la convergence. On ne sait décider que la non-convergence, et il faudra alors se dire que si rien n'indique qu'on n'a pas convergé, alors on admettra que l'on a convergé... La question de la convergence de l'échantillonnage est éminemment difficile et non encore résolue. Le lecteur pourra en juger par lui-même au cours des séances de travaux pratiques...

5.3.1 Vitesse de convergence pour les chaînes de Markov

Erreur statistique

Sous des hypothèses techniques que nous ne détaillons pas ici (voir par exemple [27]), on peut montrer qu'une chaîne de Markov (q^n, p^n) ergodique pour la mesure canonique vérifie également une sorte de théorème de la limite centrale :

$$\sqrt{N} \left| \hat{A}_N - \langle A \rangle_{\rho_{\text{can}}} \right| \rightarrow \mathcal{N}(0, \sigma^2), \quad (5.16)$$

lorsque $N \rightarrow +\infty$, où

$$\hat{A}_N = \frac{1}{N} \sum_{n=0}^{N-1} A(q^n, p^n),$$

est la moyenne trajectorielle⁴ de l'observable A , et

$$\langle A \rangle_{\rho_{\text{can}}} = \int_{\mathcal{E}} A(q, p) \rho_{\text{can}}(q, p) dq dp$$

sa moyenne d'ensemble. La quantité σ^2 est la variance de la chaîne de Markov, et peut se décomposer comme la somme d'une variance intrinsèque (égale à la variance qui résulterait d'une moyenne de variables aléatoires indépendantes), et d'une variance additionnelle provenant des corrélations entre les configurations successives :

$$\sigma^2 = \text{Var}_{\rho_{\text{can}}}(A) + 2 \sum_{n=1}^{+\infty} \mathbb{E}_{\rho_{\text{can}}} [(A(q^0, p^0) - \langle A \rangle_{\rho_{\text{can}}})(A(q^n, p^n) - \langle A \rangle_{\rho_{\text{can}}})], \quad (5.17)$$

où

$$\text{Var}_{\rho_{\text{can}}}(A) = \mathbb{E}_{\rho_{\text{can}}} [(A - \langle A \rangle_{\rho_{\text{can}}})^2].$$

Ainsi, on peut dire que l'erreur statistique commise en approchant $\langle A \rangle_{\rho_{\text{can}}}$ par $\hat{A}_N$ est de l'ordre de $\sigma/\sqrt{N}$. Il est donc important en pratique de bien estimer σ , afin d'obtenir des estimations précises de la moyenne d'ensemble, ou au moins de quantifier correctement les erreurs. On peut par exemple obtenir un intervalle de confiance à 95% selon

$$I_N(A) = \left[\hat{A}_N - 1.96 \frac{\sigma}{\sqrt{N}}, \hat{A}_N + 1.96 \frac{\sigma}{\sqrt{N}} \right].$$

Exemple 5.5 (Exemple analytique : importance des termes de corrélations). Considérons l'observable $A(q) = q$ et la chaîne de Markov unidimensionnelle

$$q^{n+1} = (1 - \Delta t)q^n + \sqrt{2\beta^{-1}\Delta t}G^n, \quad (5.18)$$

où les variables aléatoires $(G^n)_{n \geq 0}$ sont des gaussiennes standard i.i.d. Pour simplifier, on suppose que

$$q^0 \sim \mathcal{N}\left(0, \frac{1}{\beta(1 - \Delta t/2)}\right)$$

qui est la mesure invariante de la chaîne (voir Exemple 5.6). La variance (5.17) peut être calculée analytiquement dans ce cas car $\mathbb{E}(q^n) = \mathbb{E}(q^0) = 0$ pour tout $n \geq 0$, et $\mathbb{E}[q^n q^0] = (1 - \Delta t)^n \mathbb{E}[(q^0)^2]$. Ainsi,

$$\sigma^2 = \frac{2}{\beta \Delta t}.$$

Il est instructif de comparer cette variance à

$$\mathbb{E}[q^2] = \beta^{-1}(1 - \Delta t/2)^{-1},$$

qui est la variance obtenue pour des variables i.i.d. distribuées selon la mesure invariante précisée ci-dessus. On note que l'effet des corrélations augmente significativement la variance, et ce, d'autant plus que Δt est petit.

4. Dans cette section, on notera, pour simplifier, N le nombre d'itérations, noté précédemment N_{iter} . On fera attention à ne pas le confondre avec le nombre de particules formant le système moléculaire...

Éléments de démonstration de (5.17)

Donnons une preuve formelle de la formule (5.17). Quitte à remplacer A par $A - \langle A \rangle_{\rho_{\text{can}}}$, on peut supposer que $\langle A \rangle_{\rho_{\text{can}}} = 0$. On a alors

$$\text{Var}_{\rho_{\text{can}}}(\hat{A}_N) = \frac{1}{N^2} \mathbb{E}_{\rho_{\text{can}}} \left[\sum_{i,j=1}^N A(q^i) A(q^j) \right].$$

Supposons que la chaîne soit stationnaire (pour cela, il faut attendre un temps assez long, dit temps de thermalisation, pour que la chaîne soit à l'équilibre), c'est-à-dire que

$$\mathbb{E}_{\rho_{\text{can}}} [A(q^i) A(q^j)] = \mathbb{E}_{\rho_{\text{can}}} [A(q^0) A(q^{|i-j|})].$$

Alors,

$$N \text{Var}_{\rho_{\text{can}}}(\hat{A}_N) = \mathbb{E}_{\rho_{\text{can}}}(A^2) + 2 \sum_{k=1}^N \left(1 - \frac{k}{N}\right) \mathbb{E}_{\rho_{\text{can}}} [A(q^0) A(q^k)].$$

Supposant que

$$\sum_{k=1}^{+\infty} |\mathbb{E}_{\rho_{\text{can}}} [A(q^0) A(q^k)]| < +\infty,$$

une application du théorème de convergence dominée montre que

$$N \text{Var}_{\rho_{\text{can}}}(\hat{A}_N) \longrightarrow \sum_{k=1}^{+\infty} \mathbb{E}_{\rho_{\text{can}}} [A(q^0) A(q^k)],$$

ce qui donne (5.17).

Biais

On a vu que les chaînes de Markov obtenues par discrétisation d'équations différentielles stochastiques ayant la mesure canonique pour mesure stationnaire, peuvent ne plus admettre la mesure canonique comme mesure invariante, mais plutôt une mesure canonique approchée, notée $\rho_{\text{can}}^{\Delta t}$. Dans ce cas, on s'attend à ce que

$$\sqrt{N} \left| \hat{A}_N - \langle A \rangle_{\rho_{\text{can}}^{\Delta t}} \right| \rightarrow \mathcal{N}(0, \sigma^2),$$

la variance étant toujours donnée dans par la formule (5.17), quitte à remplacer ρ_{can} par $\rho_{\text{can}}^{\Delta t}$. Dans ce cas, on peut dire que l'erreur commise en approchant $\langle A \rangle_{\rho_{\text{can}}}$ par $\hat{A}_N$ est de la somme d'une erreur statistique, toujours de l'ordre de $\sigma/\sqrt{N}$, et d'un biais systématique, de l'ordre de

$$|\langle A \rangle_{\rho_{\text{can}}^{\Delta t}} - \langle A \rangle_{\rho_{\text{can}}} = \mathcal{O}(\Delta t^p),$$

où p dépend de la précision du schéma numérique. Pour les schémas présentés dans ce chapitre, $p = 1$. En pratique, le terme de biais est souvent largement négligeable devant l'erreur statistique.

Exemple 5.6. Une situation dans laquelle on peut démontrer rigoureusement une telle estimation est celle de l'Exemple 5.5, pour lequel on peut identifier la mesure invariante effectivement échantillonnée par la dynamique numérique. La chaîne de Markov (5.18) est une discrétisation consistante de l'EDS $dq_t = -q_t + \sqrt{2\beta^{-1}} dW_t$, qui a pour mesure invariante

$$\pi_{\text{exact}} \sim \mathcal{N}(0, \beta^{-1}),$$

alors que la dynamique discrétisée a pour mesure invariante

$$\pi_{\text{num}}^{\Delta t} \sim \mathcal{N}\left(0, \frac{1}{\beta(1 - \Delta t/2)}\right).$$

Pour montrer ce résultat, on commence par remarquer que si on part d'une distribution gaussienne centrée, on conserve une distribution gaussienne centrée. On peut donc écrire $q^n \sim \mathcal{N}(0, a_n)$ à la n -ième itération. On cherche la limite de a_n lorsque $n \rightarrow +\infty$. On écrit pour cela une équation sur la variance a_n :

$$\mathbb{E}((q^{n+1})^2) = \mathbb{E}\left((1 - \Delta t)^2 (q^n)^2 + 2(1 - \Delta t)\sqrt{2\Delta t\beta^{-1}} q^n G^n + \frac{2\Delta t}{\beta} (G^n)^2\right),$$

et donc

$$a_{n+1} = (1 - \Delta t)^2 a_n + \frac{2\Delta t}{\beta}.$$

On vérifie alors facilement que

$$a_n \longrightarrow a_\infty = \frac{1}{\beta(1 - \Delta t/2)}.$$

lorsque $n \rightarrow +\infty$.

5.3.2 Convergence de certaines observables

Pour les systèmes réels, on ne peut pas savoir si on a vraiment convergé ou non vers la loi cible. Tout ce qu'on peut faire, c'est vérifier que certaines moyennes représentatives ont convergé.

Utilisation de réalisations indépendantes

On peut penser à tester la convergence de la moyenne d'une observable le long d'une trajectoire. Ceci n'est toutefois qu'une condition nécessaire de convergence, pas suffisante en pratique. En effet, lorsque la dynamique d'échantillonnage reste coincée dans un état métastable, seule une exploration locale est faite. Ainsi, dans le cas de la figure 5.1b, si l'observable en question était $A(q, p) = q$, on aurait une moyenne fautive (autour de 1), mais une variance plus petite que dans la situation de la figure 5.1a !

Il faut ajouter à la convergence le long d'une trajectoire une mesure de la convergence sur les différentes réalisations possibles. Pour cela, on considère M réalisations indépendantes du processus d'échantillonnage, partant de M conditions initiales indépendantes⁵ $(q^{1,0}, p^{1,0}), \dots, (q^{M,0}, p^{M,0})$. On note $(q^{m,n}, p^{m,n})_{n=0, \dots, N-1}$ les configurations échantillonnées partant de la condition initiale $(q^{m,0}, p^{m,0})$. Pour une observable donnée A , on calcule la moyenne trajectorielle pour chacune des réalisations :

$$\hat{A}_N^m = \frac{1}{N} \sum_{n=0}^{N-1} A(q^{m,n}, p^{m,n}),$$

la moyenne empirique globale :

$$\hat{A}_N = \frac{1}{M} \sum_{m=1}^M \hat{A}_N^m,$$

et la variance entre les moyennes trajectoires :

$$\sigma_{M,N}^2(A) = \frac{N}{M} \sum_{m=1}^M \left(\hat{A}_N^m - \hat{A}_N \right)^2.$$

Sous de bonnes hypothèses,

$$\sigma^2 = \lim_{M \rightarrow +\infty} \lim_{N \rightarrow +\infty} \sigma_{M,N}^2.$$

En effet, (5.16) montre que pour N assez grand, on a $N \text{Var} \left(\hat{A}_N^m - \langle A \rangle_{\rho_{\text{can}}} \right) \simeq \sigma^2$.

5. Obtenir de telles conditions initiales est déjà un défi en soi ! Usuellement, on thermalise le système avec dynamique de Langevin par exemple, et on choisit les conditions initiales parmi les configurations visitées, en veillant à les prendre suffisamment espacées afin de limiter la corrélation. On peut éventuellement effectuer cette dynamique à une température plus élevée pour augmenter encore plus la décorrélation.

Moyenne par blocs

Dans la pratique, il est souvent plus commode de considérer une seule longue trajectoire pour calculer les moyennes canonique. Il est encore possible d'obtenir une estimation de l'erreur statistique par une procédure dite “moyennes par blocs”, qui consiste à décomposer une longue trajectoire en morceaux de taille croissante (voir [14] pour une référence dans le monde de la physique statistique computationnelle, ou [16] dans la littérature en statistique).

Plus précisément, considérons une trajectoire $(q^i)_{i \geq 1}$ de longueur $N = nm$, m étant le nombre de blocs et n le nombre de points dans un bloc. Les variables aléatoires

$$A_n^k = \frac{1}{n} \sum_{i=(k-1)n+1}^{kn} A(q_i), \quad k = 1, \dots, m$$

sont approximativement distribuées selon $\mathcal{N}(\langle A \rangle_{\rho_{\text{can}}}, \sigma^2/n)$, où σ^2 est donnée par (5.17). Un estimateur de σ^2 est donc

$$\sigma_{n,m}^2 = \frac{n}{m-1} \sum_{k=1}^m (A_n^k - A_{nm})^2,$$

$A_{nm} \equiv A_{nm}^1$ étant la moyenne sur toute la trajectoire. On peut également obtenir une estimation de l'erreur statistique commise en approchant σ^2 par $\sigma_{n,m}^2$, voir [14] pour plus de détails. Dans tous les cas, la méthode n'est valide que si la longueur de corrélation typique est (beaucoup) plus petite que la taille des blocs. En pratique, on augmente progressivement la taille des blocs, ce qui fait que la variance estimée augmente (car on prend de plus en plus compte de la corrélation) et on regarde à quelle valeur la variance se stabilise.

Quelques exemples

Exemple 5.7 (Exemple analytique). On reprend la chaîne de Markov (5.18). La figure 5.6 présente des estimations de la variance σ^2 , obtenues en utilisant plusieurs répliques indépendantes, ou des moyennes par blocs, dans le cas $\beta = 1$ et $\Delta t = 0,05$. Rappelons que la valeur analytique de la variance est $2/(\beta \Delta t) = 40$.

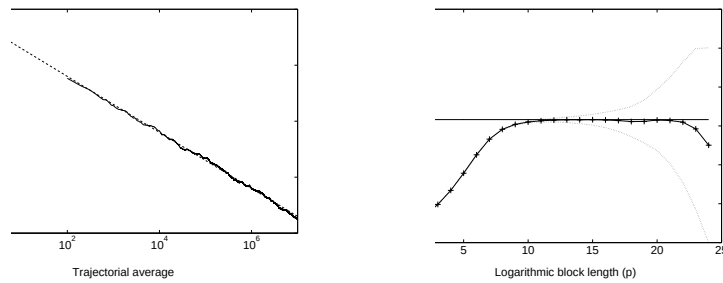


Fig. 5.6. Gauche : Variance $\sigma_{N,M}^2$ en fonction de la longueur N de la trajectoire, avec $M = 100$ trajectoires indépendantes. L'approximation linéaire (ligne en traits interrompus) donne $\sigma_{N,M}^2 \simeq 40N^{-1}$. Droite : Estimation de la déviation standard par moyenne par blocs, en fonction de la taille $n = 2^p$ des blocs (ligne pleine). On a moyenné sur 100 estimations indépendantes (intervalle de confiance en traits pointillés).

Exemple 5.8 (Détection de la métastabilité.). Considérons le potentiel double-puits $V(q) = (q^2 - 1)^2$, et une trajectoire de la dynamique (5.11) avec $\beta = 5$ et $\Delta t = 0,05$. On s'attend à une métastabilité du potentiel, et donc à ce que la particule reste longtemps confiné dans un puits avant de passer à un autre. Deux observables ont un comportement très différent : l'énergie potentielle, qui

est symétrique ($V(-q) = V(q)$) et la position q . La position moyenne devrait être égale à 0, mais il est possible que la lente convergence numérique empêche une estimation précise de cette moyenne d'ensemble; alors que la convergence de l'observable énergie potentielle devrait être nettement plus rapide puisque la métastabilité ne l'affecte pas. La figure 5.7 présente les estimations de la variance obtenues par les méthodes décrites ci-dessus. Comme prévu, la variance de l'observable position est bien plus grande que celle de l'observable énergie potentielle.

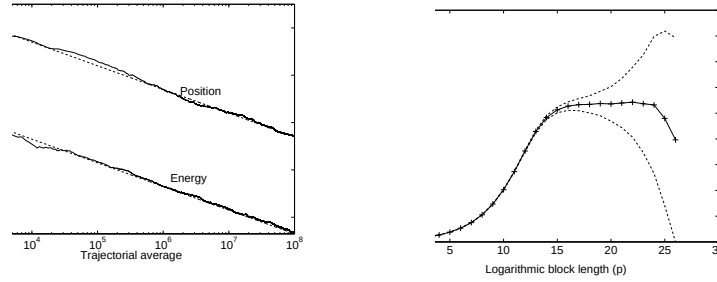


Fig. 5.7. Gauche : Variance $\sigma_{N,M}^2$ en fonction de la longueur N de la trajectoire, avec $M = 100$ trajectoires indépendantes. L'approximation linéaire (ligne en traits interrompus) donne $\sigma^2 \simeq 2500$ pour l'observable position, et $\sigma^2 \simeq 0.20$ pour l'observable énergie potentielle. Droite : Estimation de la déviation standard par moyenne par blocs, en fonction de la taille $n = 2^p$ des blocs (ligne pleine), pour l'observable position. On a moyenné sur 100 estimations indépendantes (intervalle de confiance en traits pointillés). La valeur estimée de la variance est $\sigma^2 \simeq 2800$.

Conclusion et mise en garde

On retiendra donc que

- seule la *non-convergence* est décidable : tout ce qu'on peut faire en pratique, c'est considérer plusieurs observables, aussi différentes que possibles (énergie, température cinétique, pression,...), et vérifier que chacune semble avoir convergé, au sens où l'intervalle de confiance est raisonnablement petit à la fin de la simulation, et où plusieurs simulations partant de conditions initiales différentes donnent les mêmes résultats (avec une variance qui n'augmente pas lorsqu'on rassemble les configurations microscopiques correspondantes);
- il faut, dans tous les cas, rester humble face à la caractérisation de la convergence, et plutôt faire la liste des indicateurs qui sont en accord avec l'hypothèse selon laquelle la convergence a eu lieu, que de dire qu'on a convergé.

On trouvera plus de références et de précisions sur la mesure de la convergence d'un échantillonnage par chaîne de Markov dans [17, chapitres 3 et 8].

5.4 Application numérique : fluide de Lennard-Jones

On considère un fluide d'argon à une densité $\rho = 1400 \text{ kg/m}^3$, simulé avec une dynamique de Langevin à la température cible $T = 400 \text{ K}$. On utilise des pas de temps $\Delta t = 5 \times 10^{-15} \text{ s}$, avec $\xi \Delta t / m = 0,005$ (m étant la masse d'un atome d'argon, de masse molaire $M = 39,95 \text{ g/mol}$).

Notons qu'il s'agit d'un fluide dense (la densité est supérieure à celle de l'eau liquide, en unités réduites on a $\rho \sigma^3 \simeq 0,833$), supercritique (la température critique de l'argon est environ 151 K, sa densité critique 535 kg/m^3). Nous nous situons dans la région marquée **F** du diagramme de phases de la figure 1.3.

Thermalisation des quantités de mouvement

On peut vérifier que la distribution des quantités de mouvement est conforme à la prédiction théorique (distribution de Maxwell) à la température simulée. En particulier, la température cinétique, définie comme la moyenne de l'observable

$$T_c(q, p) = \frac{1}{3Nk_B} \sum_{i=1}^N \frac{p_i^2}{m_i},$$

doit être égale à la température visée. La figure 5.8 donne les valeurs de T_c observées en fonction de l'indice d'itération, pour des systèmes de tailles différentes. On observe bien la bonne moyenne, et la variance est égale à $2T^2/N$, conformément à la prédiction (exercice).

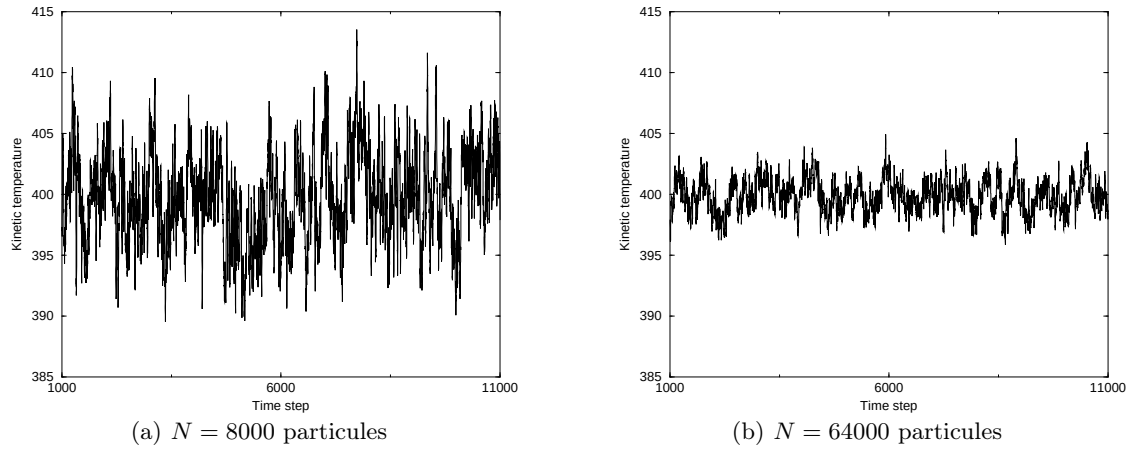


Fig. 5.8. Evolution typique de la température cinétique au cours du temps, pour un fluide d'argon de densité $\rho = 1400 \text{ kg/m}^3$ à la température cible $T = 400 \text{ K}$. On note que les fluctuations de T_c diminuent lorsque N augmente.

Capacité calorifique

On calcule à présent la capacité calorifique molaire à volume constant.

$$C_v = \frac{\mathcal{N}_A}{Nk_B T^2} (\langle H^2 \rangle - \langle H \rangle^2)$$

pour $T = 400$ ou 600 K . On utilise un pas de temps $\Delta t = 10^{-14} \text{ s}$ à $T = 400 \text{ K}$ avec $\xi \Delta t/m = 0,01$, et $\Delta t = 5 \times 10^{-15} \text{ s}$, $\xi \Delta t/m = 0,005$ à $T = 600 \text{ K}$. Les valeurs de référence sont extraites de tables thermodynamiques⁶, et données dans la légende des différents tracés de C_v en fonction du nombre de pas de temps de simulation (figure 5.9).

On observe, en particulier, la convergence des résultats, avec une précision augmentant avec le nombre de pas de temps. En moyenne, 500 000 pas de temps semblent suffisants pour assurer une convergence de l'ordre de $0,2 \text{ J/K/mol}$ (soit une erreur relative de l'ordre de 1 %). On ne l'a pas fait ici, mais une manière plus rigoureuse de proposer une barre d'erreur est de calculer la variance de C_v , ce qui demande le calcul du moment d'ordre 4 de l'énergie ou plusieurs simulations indépendantes.

Bien entendu, la partie « idéale » de C_v se calcule aisément, puisqu'elle vaut simplement ($R = \mathcal{N}_A k_B \simeq 8,314 \text{ J/K}$ étant la constante des gaz parfaits)

6. Toutes les références expérimentales que nous considérons pour l'argon sont tirées du *NIST Chemistry Webbook*, <http://webbook.nist.gov/chemistry/fluid/>

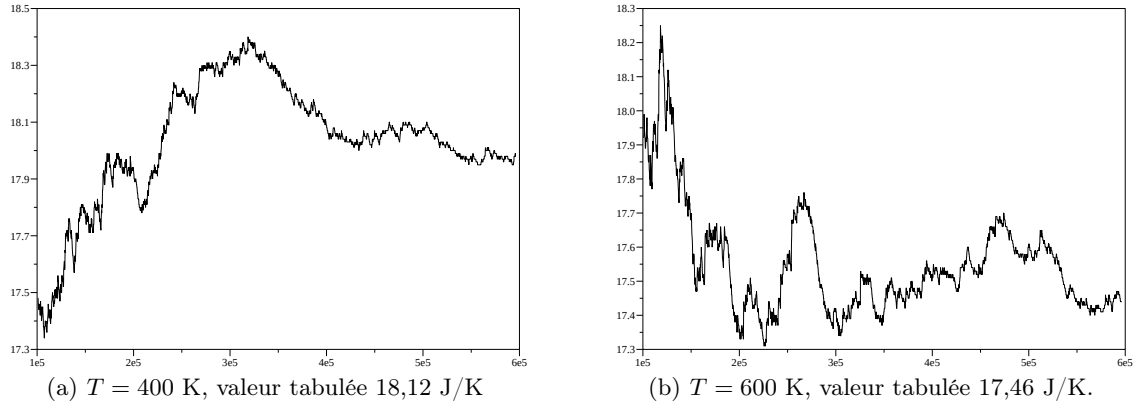


Fig. 5.9. Estimation de la capacité calorifique molaire à volume constant C_v en fonction du nombre de pas de temps.

$$C_v^{\text{id}} = \frac{3}{2}R \simeq 12,471 \text{ J/K.}$$

Équation d'état de l'argon

Nous montrons pour finir le calcul de la relation entre densité et pression dans l'argon à la température de 300 K, depuis la pression atmosphérique ordinaire jusqu'à des pressions 10000 fois plus élevées (pour lesquelles la densité réduite $\rho\sigma^3$ atteint environ 1,07). La température est largement supérieure à la température critique (de 151 K). Dans cette gamme l'argon passe continûment de l'état de vapeur ordinaire, qui est bien décrit par l'équation d'état des gaz parfaits, à celui de fluide très dense supercritique.

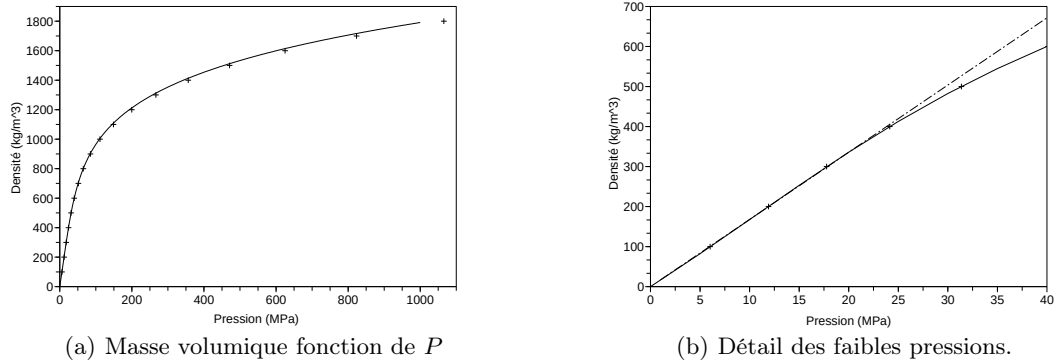


Fig. 5.10. Loi d'état de l'argon à $T = 300$ K ('+') comparée à des mesures expérimentales (courbe pleine).

La figure 5.10a compare les résultats numériques aux données expérimentales, ce qui donne un très bon accord. On notera que l'on peut atteindre par la simulation des régimes inaccessibles expérimentalement. Sur la figure 5.10b apparaît la tangente à l'origine, dont la pente est donnée par la loi des gaz parfaits, soit, avec P en MPa et la densité de masse ρ en kg/m^3 ,

$$P = 6,2433 \times 10^{-2} \rho.$$

Exercices

Exercice 5.1. L'algorithme Hybrid Monte-Carlo a été proposé dans [12] pour échantillonner la mesure canonique. C'est une chaîne de Markov, définie par la procédure suivante.

HYBRID MONTE CARLO

Algorithme 5.3. Pour une configuration initiale $q^0 \in \mathcal{M}$ et $\tau > 0$,

- (1) générer des impulsions p^n selon la maxwellienne à la température inverse β et calculer l'énergie $E^n = H(q^n, p^n)$ de la configuration (q^n, p^n) ;
- (2) calculer $\Phi_\tau(q^n, p^n) = (p^{n,\tau}, q^{n,\tau})$, en intégrant la dynamique hamiltonienne sur l'intervalle de temps $[0, \tau]$ en partant de la condition initiale (q^n, p^n) ;
- (3) calculer l'énergie $E^{n,\tau} = H(q^{n,\tau}, p^{n,\tau})$ de la nouvelle configuration et accepter la nouvelle configuration $q^{n,\tau}$ avec une probabilité

$$\alpha^n = \min(1, e^{-\beta(E^{n,\tau} - E^n)}).$$

Dans ce cas, poser $q^{n+1} = q^{n,\tau}$, sinon $q^{n+1} = q^n$.

- (1) Montrer que les étapes (1) et (2) de l'algorithme (dans le cas où Φ est le flot de la dynamique hamiltonienne) préservent la mesure canonique.
- (2) Montrer que si Φ est le flot de la dynamique hamiltonienne, alors l'étape d'acceptation/rejet n'est pas nécessaire. Qu'en est-il dans le cas où Φ est le flot défini par un schéma numérique ?
- (3) Montrer que si le schéma numérique est réversible et préserve le volume dans l'espace des phases, alors la mesure canonique est invariante par la chaîne de Markov.
- (4) Montrer que cet algorithme n'est pas ergodique pour un potentiel harmonique si τ est égal à un multiple période, mais qu'il est ergodique pour les autres choix de τ . Commenter.

Exercice 5.2. Il est possible de remplacer la règle d'acceptation dans le schéma de Metropolis-Hastings (Algorithme 5.2) par d'autres règles. Montrer par exemple que la règle de Baker

$$r_{\text{Baker}}(q^n, \tilde{q}^{n+1}) = \frac{\pi_{\text{can}}(\tilde{q}^{n+1})T(\tilde{q}^{n+1}, q^n)}{\pi_{\text{can}}(\tilde{q}^{n+1})T(\tilde{q}^{n+1}, q^n) + \pi_{\text{can}}(q^n)T(q^n, \tilde{q}^{n+1})}$$

permet également d'assurer la préservation de la mesure canonique

Exercice 5.3. Le *parallel tempering* consiste à simuler deux répliques d'un même système en parallèle, chacun selon une dynamique de Monte-Carlo par exemple, et à proposer de temps en temps un échange des configurations. Notant β_1, β_2 les températures inverses des systèmes, et q_1, q_2 les configurations microscopiques avant l'échange proposé entre les configurations q_1 et q_2 , montrer que la probabilité d'accepter l'échange des configurations est donnée par

$$r(q_1, q_2) = e^{(\beta_1 - \beta_2)(V(q_1) - V(q_2))}.$$

Un exemple de propriété dynamique : l'autodiffusion dans les fluides

La physique statistique ne se borne pas à étudier les propriétés des systèmes constitués d'un grand nombre de particules en interaction dans leurs états d'équilibre thermodynamique, elle vise également à prédire comment ils vont évoluer au cours du temps sous l'action de sollicitations extérieures. On souhaite ainsi retrouver, par exemple, le comportement visqueux des fluides en écoulement, ou bien les propriétés de conductivité thermique des matériaux, à partir de l'échelle moléculaire. La simulation numérique – par dynamique moléculaire, puisqu'il s'agit de modéliser la dynamique physique du système et non pas seulement d'échantillonner une certaine mesure dans l'espace des phases – a été utilisée pour cela dès son origine.

Nous ne donnons ici qu'une illustration dans un cas simple d'un domaine vaste et complexe de la physique statistique, dont l'étude est plus difficile que celle des propriétés d'équilibre. L'exemple de l'autodiffusion des molécules d'un fluide nous permet d'illustrer le rôle des fonctions de corrélation dépendant du temps et la relation entre réponse linéaire à des sollicitations extérieures et corrélations du système non perturbé.

Tout au long de ce chapitre il est admis que les différentes moyennes sont évaluées selon la distribution de Boltzmann-Gibbs, étant entendu que l'usage de l'ensemble canonique est conforme aux indications du paragraphe 4.4.2 et n'affecte que le tirage aléatoire d'une configuration initiale, la dynamique hamiltonienne mise en œuvre par la suite ignorant tout couplage avec le thermostat. Les résultats concernant la thermodynamique et les coefficients de transport macroscopiques sont indépendants de l'ensemble statistique utilisé pour les obtenir dans la limite thermodynamique (*i.e.* dans la limite $N \rightarrow +\infty$).

6.1	La diffusion de particules marquées dans un fluide	98
6.1.1	Échelles de temps et nature du mouvement	98
6.1.2	Mouvement brownien et diffusion	99
6.2	Propriétés des fonctions de corrélation	101
6.2.1	Propriétés générales des fonctions de corrélation dépendant du temps	101
6.2.2	Propriétés particulières de la fonction d'autocorrélation des vitesses	103
6.3	Relation entre corrélation et réponse	105
6.3.1	Importance des propriétés de réponse linéaire	105
6.3.2	Réponse linéaire à une perturbation du Hamiltonien	105
6.3.3	Mobilité d'une particule	106
6.4	Remarques finales	107

6.1 La diffusion de particules marquées dans un fluide

6.1.1 Échelles de temps et nature du mouvement

Dans un fluide, imaginons que nous marquons une ou plusieurs molécules – s'il y en a plusieurs, elles seront suffisamment éloignées pour ne pas s'influencer mutuellement – pour en suivre les mouvements. C'est bien sûr possible en dynamique moléculaire, et des techniques de marquage isotopique autorisent également de telles expériences en laboratoire. Les interactions de la particule avec ses voisins modifient peu à peu sa vitesse, jusqu'à la décorréler complètement de sa valeur initiale. Mesurée sur une durée assez longue, la vitesse d'une particule échantillonne alors la distribution de Maxwell, conformément à (3.5). Sa moyenne est nulle, et sa variance, soit $\langle \mathbf{v}^2 \rangle$, égale à $3k_B T/m$.

La fonction d'autocorrélation de la vitesse, définie par :

$$C(t) = \langle \mathbf{v}(0) \cdot \mathbf{v}(t) \rangle, \quad (6.1)$$

caractérise la perte d'influence de la condition initiale pour la trajectoire d'une particule (dans un fluide, du fait de l'isotropie, $C(t)$ peut être définie à partir de la corrélation d'une coordonnée quelconque x de la vitesse : on a alors $C(t) = 3 \langle v_x(0)v_x(t) \rangle$).

La fonction d'autocorrélation de la vitesse est représentée sur la figure 6.1¹ dans le liquide de Lennard-Jones, normalisée par sa valeur initiale $C(0) = 3k_B T/m$. Elle présente un début

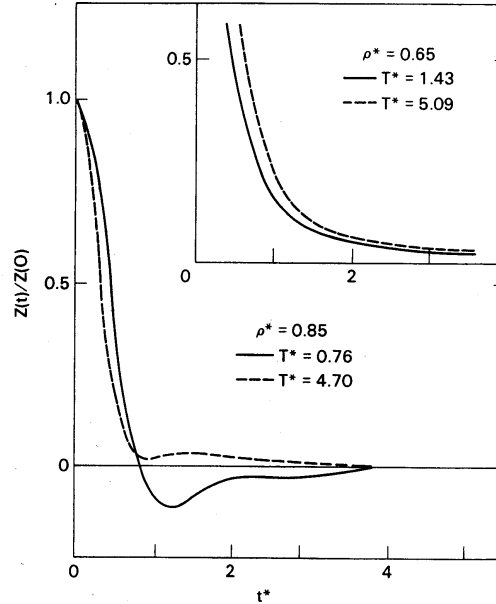


Fig. 6.1. Fonction d'autocorrélation de la vitesse (6.1) dans le liquide de Lennard-Jones. On utilise des unités réduites : $\rho^* = \rho\sigma^3$, $T^* = k_B T/\epsilon$, $t^* = t/\tau$ et $\tau = \sqrt{\frac{48\epsilon}{m\sigma^2}}$.

d'oscillation pour des temps intermédiaires, et ne se distingue plus de zéro pour des temps de l'ordre de quelques unités (réduites). De fait, pendant des intervalles de temps δt successifs grands devant l'unité utilisée sur cette figure, les déplacements $\delta \mathbf{r}$ successifs d'une particule dont on suit le

1. Résultats dus à D. Lévesque et L. Verlet, cités dans [20].

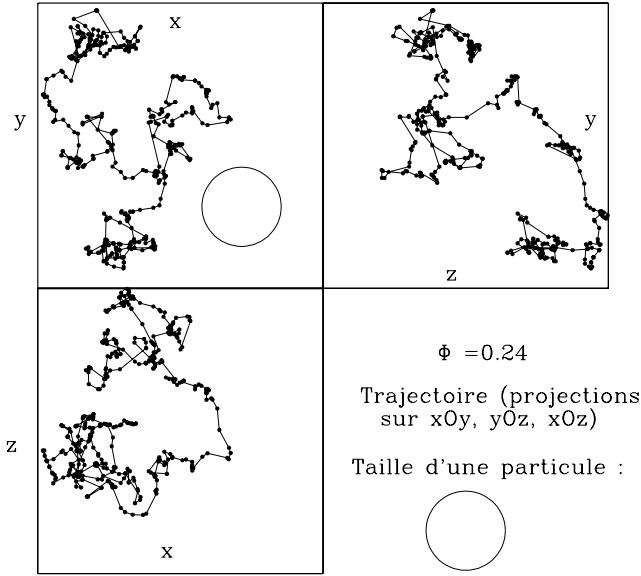


Fig. 6.2. Trajectoire d'une particule dans le fluide de sphères dures pour la compacité $\Phi = 0,24$, vue en projection dans les 3 plans de coordonnées. Les points noirs représentent les positions aux moments des collisions, qui correspondent à un changement de direction de la vitesse.

mouvement sont décorrélés. On a alors une marche aléatoire dont les trajectoires erratiques, comme le montre la figure 6.2, rappellent celles d'un mouvement brownien (comparer à la figure 5.3).

6.1.2 Mouvement brownien et diffusion

Hypothèse d'indépendance et conséquences

Précisons cette analogie, en admettant (c'est une hypothèse très forte, qui n'est pas exacte en pratique) que les déplacements $\delta \mathbf{r}_i$ de la particule pendant des intervalles de temps successifs δt sont *statistiquement indépendants*. Ils sont alors également équidistribués (le mouvement d'une particule est un processus stationnaire à l'équilibre), de moyenne nulle (par isotropie) et admettent une matrice de covariance que nous prenons égale à :

$$\langle \delta r_\alpha \delta r_\beta \rangle = 2D\delta t \delta_{\alpha\beta},$$

définissant ainsi le coefficient D . Pour simplifier, étudions le mouvement selon la coordonnée x – les autres coordonnées variant de façon indépendante. Agrégeons $M \gg 1$ pas δx , définissant des pas plus grands Δx , correspondant aux mouvements pendant des intervalles de temps successifs de longueur $\Delta t = M\delta t$. Nous obtenons alors une chaîne de Markov $(x_n)_{n \in \mathbb{N}}$, avec, en vertu du théorème de la limite centrale,

$$x^{n+1} = x^n + \sqrt{2D\Delta t} G^n,$$

avec des variables aléatoires $G^n \sim \mathcal{N}(0, 1)$ indépendantes, c'est-à-dire exactement (5.8) : on retrouve la discrétisation du mouvement brownien introduite au chapitre 5. À temps long devant Δt , on sait que chacune des coordonnées du vecteur « bout-à-bout » $\mathbf{r}(t) - \mathbf{r}(0)$ de la chaîne est distribuée comme $\mathcal{N}(0, 2Dt)$. Il y a étalement, ou *diffusion* de la densité de probabilité de présence du marcheur aléatoire, et D , dont la dimension est $[L]^2[T]^{-1}$, est par définition le *coefficient de diffusion*. L'étalement de cette densité correspond, physiquement, au mélange progressif des

molécules marquées avec le reste du fluide, jusqu'à ce que leur concentration soit uniforme dans l'échantillon. À tout instant t , la concentration c de particules marquées obéit à l'équation de diffusion :

$$\frac{\partial c}{\partial t} = D \Delta c. \quad (6.2)$$

Il en est de même pour la densité de probabilité de la position d'une seule particule.

Test numérique

On peut vérifier qu'en dépit de l'usage de l'approximation d'indépendance, qui est une simplification, les observations sont bien en accord avec les conclusions qu'on en a tirées concernant la description statistique du mouvement à long terme. Ainsi, on a bien $\Delta \mathbf{r}^2(t) \propto t$ pour le carré de la distance parcourue, comme le montre la figure 6.3, après quelques collisions dans le cas des sphères dures, ou plus généralement une fois passée la durée caractéristique d'oscillation d'une particule dans la « cage » formée par ses proches voisins. Dans les simulations, on peut donc obtenir le

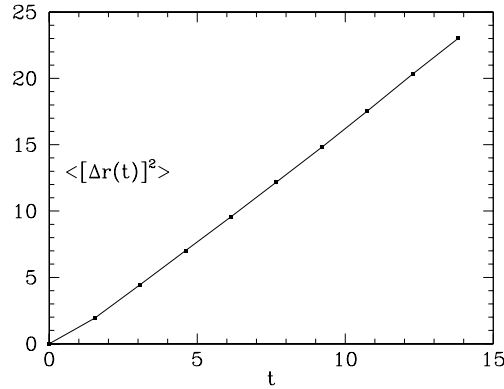


Fig. 6.3. Carré moyen de la distance parcourue, $\langle |\mathbf{r}(t) - \mathbf{r}(0)|^2 \rangle$ fonction de t dans le fluide de sphères dures à $\Phi = 0,24$ (mêmes unités que celles de la figure 2.2).

coefficient de diffusion selon :

$$D = \lim_{t \rightarrow +\infty} \frac{\langle |\mathbf{r}(t) - \mathbf{r}(0)|^2 \rangle}{6t} \quad (6.3)$$

Autodiffusion et diffusion d'une particule différente en solution ou en suspension

Nous avons jusqu'ici considéré que les particules marquées dont on suit le mouvement étaient des molécules du fluide identiques à toutes les autres. C'est pourquoi on parle alors plus précisément d'*autodiffusion* – surtout en présence d'autres phénomènes de type diffusif, pour éviter de les confondre.

Rien ne change qualitativement dans les considérations qui précèdent si la particule suivie est d'une nature différente. Il peut s'agir d'une molécule d'un soluté, ou bien encore d'une particule colloïdale, c'est-à-dire, typiquement, d'une taille comprise entre le nanomètre et le micron. De tels objets peuvent être décrits comme macroscopiques (comme des échantillons de matière continue), mais sont toujours sensibles à l'agitation moléculaire, ils constituent les particules browniennes observables au microscope. Un changement quantitatif notable par rapport au cas des molécules est la brièveté du temps de corrélation de leur vitesse, au regard du temps typique nécessaire pour que leur déplacement soit de l'ordre de leur diamètre. De ce fait, le mouvement de ces particules

peut être décrit comme brownien à toutes les échelles de temps physiquement pertinentes pour décrire leur trajectoire.

Le coefficient de diffusion des particules colloïdales est nettement plus faible que celui des molécules. On verra plus loin pourquoi il est donné par la formule (1.5).

6.2 Propriétés des fonctions de corrélation

6.2.1 Propriétés générales des fonctions de corrélation dépendant du temps

Définition

Rappelons la définition des fonctions de corrélation introduite au Chapitre 2, par (2.15). En général, on appelle fonction de corrélation (on sous-entend alors : dépendant du temps) des observables A et B la covariance de A et de $B(t)$. La fonction $B(t)$ est obtenue à partir de B en remplaçant les coordonnées de $(\mathbf{Q}, \mathbf{P}) \in \mathcal{E}$ par celles de $\phi_t(\mathbf{Q}, \mathbf{P})$, qui résultent d'une évolution par la dynamique hamiltonienne pendant la durée t à partir de $\mathbf{Q}, \mathbf{P}$, pris comme microétat initial. Ainsi, $B(t)$ est par définition l'observable $B \circ \phi_t$. La fonction de corrélation, notée C_{AB} a donc pour définition :

$$C_{AB}(t) = \langle A(0)B(t) \rangle - \langle A \rangle \langle B \rangle = \text{Cov}(A, B \circ \phi_t). \quad (6.4)$$

Sa valeur à $t = 0$ est la covariance $\text{Cov}(A, B)$.

Stationnarité

Nous avons déjà rencontré la propriété de stationnarité, qui justifie la définition sous la forme (6.4), avec un choix arbitraire de l'origine des temps. Elle résulte de l'équilibre, la mesure de probabilité ρ_{can} dans l'espace des phases $\mathcal{E}$ qui définit les moyennes – c'est-à-dire la distribution de Boltzmann-Gibbs de l'ensemble canonique – étant indépendante du temps. La stationnarité signifie que les moyennes sont indépendantes du temps : on a

$$\forall t \in \mathbb{R}, \quad \langle A \circ \phi_t \rangle = \langle A \rangle,$$

et, en considérant le produit de deux observables, fonctions de $(\mathbf{Q}, \mathbf{P})$ que l'on peut composer avec le flot hamiltonien ϕ_t pour différentes valeurs de t , on trouve :

$$\forall t, t' \in \mathbb{R}, \quad \langle A(t')B(t+t') \rangle - \langle A \rangle \langle B \rangle = \langle A(0)B(t) \rangle - \langle A \rangle \langle B \rangle = C_{AB}(t).$$

Pour les observables qui sont des fonctions différentiables des coordonnées des microétats, on peut définir leur dérivée par rapport au temps par :

$$\dot{A} = \sum_{i=1}^N \frac{\partial A}{\partial \mathbf{q}_i} \cdot \dot{\mathbf{q}}_i + \frac{\partial A}{\partial \mathbf{p}_i} \cdot \dot{\mathbf{p}}_i = \sum_{i=1}^N \frac{\partial A}{\partial \mathbf{q}_i} \cdot \frac{\partial H}{\partial \mathbf{p}_i} - \frac{\partial A}{\partial \mathbf{p}_i} \cdot \frac{\partial H}{\partial \mathbf{q}_i}. \quad (6.5)$$

La stationnarité entraîne des relations utiles pour des fonctions de corrélation impliquant de telles dérivées temporelles des observables. En effet, notons d'abord que l'on peut permuter moyenne et dérivée par rapport au temps, on a ainsi

$$\begin{aligned} \frac{d}{dt} \langle A(0)B(t) \rangle &= \frac{d}{dt} \int_{\mathcal{E}} A(\mathbf{Q}, \mathbf{P}) B \circ \phi_t(\mathbf{Q}, \mathbf{P}) d\rho_{\text{can}}(\mathbf{Q}, \mathbf{P}) \\ &= \int_{\mathcal{E}} A(\mathbf{Q}, \mathbf{P}) \dot{B} \circ \phi_t(\mathbf{Q}, \mathbf{P}) d\rho_{\text{can}}(\mathbf{Q}, \mathbf{P}) = \langle A(0) \dot{B}(t) \rangle. \end{aligned}$$

En écrivant ensuite que la dérivée de $\langle A(t')B(t+t') \rangle$ par rapport à t' s'annule (stationnarité), on obtient pour des observables quelconques, fonctions régulières des $\mathbf{q}_i$ et des $\mathbf{p}_i$, l'égalité :

$$\frac{d}{dt} C_{AB}(t) = C_{A\dot{B}}(t) = \langle A(0) \dot{B}(t) \rangle = - \langle \dot{A}(0) B(t) \rangle = -C_{\dot{A}B}(t). \quad (6.6)$$

Effets d'un renversement des vitesses

Exploitions maintenant la propriété de réversibilité de la dynamique hamiltonienne, exprimée par (2.4) à l'aide de l'opérateur de renversement des vitesses $\mathcal{S}$. Notons d'abord que beaucoup d'observables utiles ont une parité définie vis-à-vis du renversement des vitesses, c'est-à-dire qu'elles vérifient :

$$A \circ \mathcal{S} = \epsilon_A A, \quad (6.7)$$

avec $\epsilon_A = 1$ pour certaines, dites paires, comme l'observable énergie, c'est-à-dire le Hamiltonien H , ou l'observable pression (dont la moyenne apparaît dans la formule 2.23), et $\epsilon_A = -1$ pour d'autres, dites impaires, comme la vitesse d'une particule marquée que nous étudions dans ce chapitre. On vérifie aisément que si l'observable A , fonction différentiable de $(\mathbf{Q}, \mathbf{P})$ possède la parité ϵ_A , alors sa dérivée temporelle, donnée par (6.5), est dotée de la parité opposée.

Les quantités de mouvement n'apparaissant que sous la forme de leurs carrés (distribution de Maxwell) dans la densité de probabilité canonique, on aura, pour toute observable A ,

$$\langle A \rangle = \langle A \circ \mathcal{S} \rangle, \quad (6.8)$$

de sorte que *la moyenne d'une observable impaire est nulle, ainsi que la covariance de deux observables de parités opposées.*

Considérons maintenant deux observables A et B ayant des parités définies ϵ_A et ϵ_B , et cherchons à relier $C_{AB}(t)$ à $C_{BA}(t)$. Pour simplifier l'écriture, nous supposons – quitte à remplacer A par $A - \langle A \rangle$ et B par $B - \langle B \rangle$ – que leurs moyennes sont nulles. Exploitant d'abord (2.4), sous la forme $\phi_t = \mathcal{S} \circ \phi_{-t} \circ \mathcal{S}$, on a :

$$C_{AB}(t) = \langle A(B \circ \phi_t) \rangle = \langle A(B \circ \mathcal{S} \circ \phi_{-t} \circ \mathcal{S}) \rangle,$$

qui entraîne d'après (6.8) appliquée au produit des deux observables, sachant que $\mathcal{S} \circ \mathcal{S} = I$,

$$C_{AB}(t) = \langle (A \circ \mathcal{S})(B \circ \mathcal{S} \circ \phi_{-t}) \rangle,$$

que l'on transforme alors par (6.7) appliquée à A et à B en

$$C_{AB}(t) = \epsilon_A \epsilon_B \langle A(0)B(-t) \rangle.$$

Invokant alors la stationnarité qui entraîne que

$$C_{AB}(-t) = C_{BA}(t), \quad (6.9)$$

il vient :

$$C_{AB}(t) = \epsilon_A \epsilon_B C_{BA}(t). \quad (6.10)$$

Nous verrons plus loin que les fonctions de corrélation donnent directement la réponse des systèmes étudié en physique statistique à des sollicitations extérieures. Dans ce contexte la propriété (6.10), issue de la réversibilité en temps de la dynamique hamiltonienne au niveau microscopique, a des conséquences remarquables au niveau de la thermodynamique hors d'équilibre macroscopique.

Décorrélation à temps long

Une propriété générale des fonctions de corrélation que nous admettons est qu'elles possèdent la limite nulle pour $t \rightarrow +\infty$: pour deux observables quelconques A et B , on a

$$\lim_{t \rightarrow +\infty} C_{AB}(t) = 0, \quad \text{c'est-à-dire} \quad \lim_{t \rightarrow +\infty} \langle A(0)B(t) \rangle = \langle A \rangle \langle B \rangle.$$

Il s'agit d'une propriété, liée à l'ergodicité, appelée *propriété de mélange* de la dynamique hamiltonienne, dont nous admettons que les systèmes d'intérêt physique constitués d'un grand nombre de particules en interaction sont dotés en général (cf. la discussion de l'ergodicité au Chapitre 3, Remarques 3.1 et 3.2). La propriété de mélange est plus forte que la propriété d'ergodicité au sens de l'égalité des moyennes sur le temps et des moyennes d'ensemble.

Une majoration

Une application directe de l'inégalité de Schwarz, en utilisant la stationnarité, donne, pour toutes observables A et B à tout instant t :

$$|C_{AB}(t)| \leq [\text{Var}(A)]^{\frac{1}{2}} [\text{Var}(B)]^{\frac{1}{2}}.$$

Ainsi, dans le cas $B = A$, la fonction d'autocorrélation d'un observable est majorée, en valeur absolue, par sa variance, c'est-à-dire par sa valeur initiale.

6.2.2 Propriétés particulières de la fonction d'autocorrélation des vitesses

Fonction d'autocorrélation de la vitesse et coefficient de diffusion

On peut, en appliquant les propriétés générales des fonctions de corrélation à la fonction d'autocorrélation de la vitesse transformer le résultat (6.3) et faire apparaître la fonction d'autocorrélation des vitesses $C(t)$:

$$\langle |\mathbf{r}(t) - \mathbf{r}(0)|^2 \rangle = \int_0^t \int_0^t C(t_2 - t_1) dt_1 dt_2,$$

soit, $C(t)$ étant paire d'après (6.9),

$$\langle |\mathbf{r}(t) - \mathbf{r}(0)|^2 \rangle = 2 \int_0^t \int_{t_1}^t C(t_2 - t_1) dt_1 dt_2.$$

Avec le changement de variable (de jacobien 1)

$$\begin{cases} u = t_2 - t_1, & 0 \leq u \leq t, \\ v = \frac{1}{2}(t_1 + t_2), & u/2 \leq v \leq t - u/2, \end{cases}$$

on trouve

$$\langle |\mathbf{r}(t) - \mathbf{r}(0)|^2 \rangle = 2 \int_0^t (t - u) C(u) du,$$

de sorte que l'on a :

$$\lim_{t \rightarrow +\infty} \frac{1}{t} \langle |\mathbf{r}(t) - \mathbf{r}(0)|^2 \rangle = 2 \lim_{t \rightarrow +\infty} \int_0^t \left(1 - \frac{u}{t}\right) C(u) du = 2 \int_0^{+\infty} C(u) du$$

en admettant que $C \in L^1([0, +\infty[)$, par application du théorème de convergence dominée de Lebesgue. Ceci demande que l'autocorrélation de la vitesse décroît suffisamment rapidement. D'où la formule :

$$D = \frac{1}{3} \int_0^{+\infty} C(t) dt. \quad (6.11)$$

Ce résultat (6.11) est un exemple d'une famille de formules auxquelles on donne le nom de *formules de Green-Kubo*, qui expriment un coefficient de transport macroscopique comme l'intégrale d'une fonction de corrélation d'observables microscopiques.

Développement à temps court et fréquence d'Einstein

Du fait de la parité de $C(t)$ et de la régularité des trajectoires (avec un potentiel d'interaction indéfiniment dérivable), on a le développement limité :

$$C(t) = C(0)[1 - \Omega_0^2 t^2 + O(t^4)],$$

avec, $\mathbf{F}$ étant la force subie par la particule suivie,

$$\Omega_0^2 = \frac{m}{3k_B T} \left\langle \left(\frac{d\mathbf{v}}{dt} \right)^2 \right\rangle = \frac{1}{3mk_B T} \langle |\mathbf{F}|^2 \rangle.$$

En convenant que les gradients sont pris par rapport à la position de la particule marquée, et en notant V (fonction de toutes les positions) l'énergie potentielle totale, on a :

$$\langle |\mathbf{F}|^2 \rangle = -\frac{k_B T}{Z_q} \int_{\mathcal{D}^N} \nabla V \cdot \nabla (\exp[-\beta V]) d\mathbf{Q},$$

d'où, par une intégration par parties,² on tire

$$\langle |\mathbf{F}|^2 \rangle = \frac{k_B T}{Z_q} \int_{\mathcal{D}^N} \nabla^2 V \exp[-\beta V] d\mathbf{Q}.$$

Avec le potentiel de paire $\mathcal{V}(r)$, fonction de la distance seule, dont le laplacien est $\mathcal{V}''(r) + 2\frac{\mathcal{V}'(r)}{r}$, cela s'écrit :

$$\langle |\mathbf{F}|^2 \rangle = 4\pi\rho k_B T \int_0^{+\infty} r^2 g(r) \left[\mathcal{V}''(r) + 2\frac{\mathcal{V}'(r)}{r} \right] dr.$$

On a donc :

$$\Omega_0^2 = \frac{4\pi\rho}{3m} \int_0^{+\infty} r^2 g(r) \left[\mathcal{V}''(r) + 2\frac{\mathcal{V}'(r)}{r} \right] dr. \quad (6.12)$$

La quantité Ω_0 est appelée *fréquence d'Einstein*. Autour de la particule marquée et à la distance r le champ de densité moyen est $\rho g(r)$. Il en résulte une force nulle à l'origine, qui correspond à un minimum local du potentiel associé à ce champ de densité. La fréquence Ω_0 est exactement égale à la pulsation propre des oscillations d'une particule de masse m au voisinage de ce minimum.

La fréquence d'Einstein fixe le temps caractéristique des variations rapides de la vitesse d'une particule. La connaissance de Ω_0 permet un choix approprié du pas de temps dans les simulations. Le pas de temps Δt se comparera alors à $1/\Omega_0$ (typiquement on prendra $\Omega_0 \Delta t$ d'ordre $1/10$).

Comportement à temps long

Le retour à zéro de la fonction d'autocorrélation de la vitesse $C(t)$ détermine l'oubli des conditions initiales et la possibilité de passer, pour les temps plus longs, à une description markovienne du mouvement aléatoire d'une particule. Il est donc important de connaître le comportement asymptotique de la fonction d'autocorrélation à temps long. Les modèles les plus simples, fondés sur l'approximation d'une suite de collisions décorréelées, par exemple, dans le cas des sphères dures (hypothèse simplificatrice remontant à Boltzmann, dite du « chaos moléculaire »), prédisent une décroissance exponentielle de $C(t)$. Or, on a³ :

2. le terme de flux au travers de $\partial\mathcal{D}^N$ s'annule dans le cas d'un système confiné par un potentiel « de boîte » ; et en fait la nature locale de la moyenne recherchée entraîne que le résultat doit être indépendant de la nature des conditions aux limites.

3. Pour ce comportement, on peut fournir l'explication suivante [20], en comparant la vitesse de la particule suivie à celle d'une petite région dans un fluide, qui serait animée d'une vitesse supérieure, dans une certaine direction, à celle du fluide ambiant. La décorrélation correspond alors à la perte progressive de cette différence de vitesse. Celle-ci se produit, dans un écoulement, par diffusion : la quantité de mouvement est diffusée dans une région dont le diamètre croît comme $\sqrt{\nu t}$, ν étant la viscosité cinématique. Cette quantité de mouvement est mv , avec une masse m qui augmente comme $t^{3/2}$. La vitesse v doit donc décroître comme $t^{-3/2}$.

$$C(t) = O(t^{-3/2}) \text{ pour } t \rightarrow +\infty. \quad (6.13)$$

Cette décroissance algébrique de $C(t)$ témoigne d'effets collectifs complexes et persistants, difficiles à modéliser. En fait, la fonction d'autocorrélation de la vitesse d'une particule colloïdale présente également ce comportement à temps long, ce qui montre qu'elle n'est pas correctement décrite par l'équation de Langevin (c'est-à-dire (5.13) avec un potentiel V nul), censée modéliser de telles particules mésoscopiques ! En effet (exercice) la dynamique de Langevin pour une particule isolée prédit une décroissance exponentielle de la fonction d'autocorrélation de sa vitesse.

Ce comportement à temps long ne remet pas en cause, du moins dans un espace à trois dimensions, l'existence d'un coefficient de diffusion. Dans un espace à deux dimensions, en revanche, on ne parvient pas à définir un coefficient de diffusion indépendant de la taille du système.

6.3 Relation entre corrélation et réponse

6.3.1 Importance des propriétés de réponse linéaire

Les sollicitations extérieures sur le système sont d'ordinaire de très faibles perturbations du mouvement spontané des particules qui le constituent. Pour le montrer sur un exemple, discutons des perturbations appliquées à un fluide dans un écoulement avec un fort taux de déformation : considérons un écoulement dans l'air avec un taux de cisaillement de $\dot{\gamma} = 10^4 \text{ s}^{-1}$. Cela entraîne une différence de vitesse de 10^3 m/s en des points distants de 10 cm , alors que la vitesse du son est d'environ 330 m/s . Au niveau microscopique, on a affaire à des molécules dont les vitesses sont distribuées selon une loi gaussienne dont la moyenne est la vitesse d'ensemble locale (vitesse de l'écoulement), et dont l'écart-type u_0 est donné par la loi de Maxwell : $u_0 \sqrt{\frac{3k_B T}{m}}$, ou quelques centaines de m/s (cf. § 3.1.3). Le gaz ne sera pas beaucoup perturbé tant que les différences des vitesses des molécules lors des collisions ne sont pas sensiblement affectées par l'écoulement d'ensemble, c'est-à-dire tant que la condition

$$\dot{\gamma} l \ll u_0,$$

dans laquelle l est le libre parcours moyen, reste satisfaite. Or l est de l'ordre de 10^{-7} m dans l'air à température et pression ordinaires (comme on pourra le vérifier en exercice), d'où $\dot{\gamma} l \sim 10^{-3} \text{ m/s}$ avec la valeur de $\dot{\gamma}$ choisie, ce qui est bien négligeable devant les vitesses thermiques. Même un très fort taux de cisaillement ne perturbe qu'à peine les mécanismes moléculaires, et représente donc une faible modification du macroétat d'équilibre. La viscosité de cisaillement, dans un fluide, exprime une variation des contraintes au premier ordre dans la petite perturbation introduite par l'application d'un taux de cisaillement.

D'une façon générale, la recherche d'une *réponse linéaire*, c'est-à-dire de variations des moyennes des observables entraînées par une perturbation extérieure, est justifiée par la faiblesse, dans les cas pratiques, des actions extérieures, comparées aux interactions des molécules du système. Outre le comportement visqueux des fluides, on peut citer, comme lois phénoménologiques exprimant une réponse linéaire la loi de Fourier pour la conduction de la chaleur, la loi d'Ohm pour la conduction électrique, ou la loi de Fick pour l'interdiffusion dans un mélange. La physique statistique permet de fonder ces lois sur une approche microscopique, et de donner des expressions analogues à (6.11) pour les coefficients qu'elles contiennent. Ces relations font intervenir des fonctions de corrélation entre observables du système étudié, caractéristiques de sa dynamique *en l'absence de perturbation*.

6.3.2 Réponse linéaire à une perturbation du Hamiltonien

Considérons à nouveau le Hamiltonien perturbé $H_0 - \lambda B$, où B est une observable et λ un petit paramètre scalaire, tel que nous l'avons introduit en (4.13) pour en étudier les effets en statique. Nous supposons pour simplifier qu'en l'absence de perturbation on a $\langle B \rangle = 0$.

Le paramètre λ peut dépendre du temps. Admettre une *réponse linéaire* d'une observable quelconque A aux perturbations passées, c'est postuler l'existence d'une *fonction de réponse* que nous noterons $\chi_{AB}(t)$, telle que l'on ait à tout instant t :

$$\Delta \langle A(t) \rangle \equiv \langle A(t) \rangle - \langle A \rangle_0 = \int_{-\infty}^t \chi_{AB}(t-u) \lambda(u) du. \quad (6.14)$$

Dans cette formule, $\langle A(t) \rangle$ est la moyenne selon un macroétat (mesure de probabilité) qui est influencé par l'histoire de la perturbation $\lambda(t)$, mais proche du macroétat d'équilibre canonique du Hamiltonien H_0 du système non perturbé. Pour écrire (6.14), nous avons supposé que, la perturbation étant faible, son application n'entraînait aucune évolution importante du système, qui ne « vieillit » pas, c'est pourquoi l'effet à l'instant t d'une perturbation à t' ne dépend que de la différence $t - t'$. De plus, on doit bien sûr imposer la condition que la réponse ne dépende que du passé : $\chi_{AB}(t)$ est nulle pour tout $t < 0$.

Supposons maintenant que nous avons appliqué une perturbation constante (ou très lentement croissante) $\lambda = \lambda_0$ à tout instant $t < 0$, de sorte qu'à $t = 0$ on a atteint le macroétat d'équilibre (canonique) avec le Hamiltonien (4.13) pour la valeur λ_0 . Alors les moyennes des observables peuvent être calculées, au premier ordre en λ , selon (4.14). À l'instant $t = 0$, supprimons la perturbation : on aura $\lambda(t) = 0$ pour tout $t > 0$. Le système évolue alors avec la dynamique hamiltonienne en l'absence de perturbation, caractérisée par le flot ϕ_t . Nous pouvons alors évaluer $\Delta \langle A(t) \rangle$, définie en (6.14), de deux façons différentes. Selon (6.14), nous avons :

$$\Delta \langle A(t) \rangle = \lambda_0 \int_{-\infty}^0 \chi_{AB}(t-u) du = \lambda_0 \int_t^{+\infty} \chi_{AB}(u) du. \quad (6.15)$$

Par ailleurs, la relation (4.14), appliquée à l'observable $A(t) = A \circ \phi_t$ fournit :

$$\Delta \langle A(t) \rangle = \lambda_0 \beta \text{Cov}(B, A(t)) = \lambda_0 \beta C_{BA}(t). \quad (6.16)$$

En notant $H(t)$ la fonction de Heaviside (qui vaut 1 si $t \geq 0$ et 0 si $t < 0$), l'égalité pour tout t des seconds membres de (6.15) et de (6.16) donne le résultat :

$$\chi_{AB}(t) = -\beta \frac{d}{dt} C_{BA}(t) H(t) = -\beta C_{BA}(t) H(t). \quad (6.17)$$

La formule (6.17) relie donc simplement la *fonction de réponse* χ_{AB} à la *fonction de corrélation* des observables A et B en l'absence de perturbation.

6.3.3 Mobilité d'une particule

Revenant à l'étude particulière du mouvement d'une particule marquée dans le fluide, perturbons le Hamiltonien de façon à exercer une force constante $\mathbf{F}$ sur cette particule, que nous prenons conventionnellement parallèle à l'axe de la coordonnée x . Dans (4.13), on prendra donc $\lambda = F$ et $B = x$, l'abscisse de la particule.

Nous nous intéressons à la vitesse moyenne v que la particule acquière dans la direction x de la force appliquée, et nous notons α la fonction de réponse correspondante. On a donc, au premier ordre en F ,

$$\langle v(t) \rangle = \int_{-\infty}^t \alpha(t-u) F(u) du,$$

avec d'après (6.17), et en utilisant (6.6),

$$\alpha(t) = -\beta C_{xv}(t) H(t) = \beta C_{vv}(t) H(t) = \frac{\beta}{3} C(t) H(t),$$

où nous retrouvons la fonction d'autocorrélation de la vitesse $C(t)$, qui exprime donc la *réponse de la vitesse à une histoire de la force appliquée*.

On appelle *mobilité* d'une particule dans un fluide le coefficient de proportionnalité μ entre la vitesse moyenne qu'elle acquière en régime stationnaire et la force constante qu'on lui applique. On a par définition

$$\mu = \int_0^{+\infty} \alpha(t) dt = \frac{\beta}{3} \int_0^{+\infty} C(t) dt.$$

Par comparaison avec (6.11), on a donc

$$D = \mu k_B T. \quad (6.18)$$

Lorsque la particule suivie est de taille mésoscopique, c'est-à-dire grande à l'échelle moléculaire (mais elle doit rester assez petite toutefois si on veut pouvoir déceler son mouvement brownien), sa mobilité se calcule simplement par l'hydrodynamique macroscopique. Le mouvement d'une telle particule est généralement en régime de Stokes (très faible nombre de Reynolds), et on a pour une sphère de rayon R animée de la vitesse v la force visqueuse

$$F_v = -6\pi\eta Rv,$$

η étant la viscosité de cisaillement du fluide. Sous l'action d'une force appliquée F , la vitesse en régime stationnaire est telle que $F + F_v = 0$, ce qui fournit la mobilité :

$$\mu = \frac{1}{6\pi\eta R}.$$

En combinant cette relation avec (6.18), on en tire la valeur du coefficient de diffusion, lié $= R$, T et η par la relation de Stokes-Einstein (1.5) :

$$D = \frac{k_B T}{6\pi\eta R},$$

qui est donc établie dans le cas d'une particule colloïdale dont la mobilité est donnée par l'hydrodynamique de Stokes macroscopique.

Avec $T = 293$ K, $\eta = 10^{-3}$ Pa.s (viscosité de l'eau à 293 K), $R = 10^{-7}$ m, on a ainsi $D \simeq 2 \times 10^{-12}$ m²s⁻¹. Comme on l'a signalé au Chapitre 1, la formule de Stokes-Einstein donne un très bon ordre de grandeur des coefficients d'autodiffusion dans les liquides constitués de petites molécules, qui sont donc de l'ordre de 10^{-9} m²s⁻¹.

On remarquera que la diffusion moléculaire dans les liquides est un phénomène très lent à notre échelle. Le temps caractéristique τ_D pour un transport diffusif sur une distance L est $\tau_D = L^2/D$, il est inversement proportionnel à la taille des particules diffusantes et atteint déjà 10^5 secondes (une journée) pour $L \sim 1$ cm – d'où l'utilité de la petite cuiller pour homogénéiser la concentration du sucre dans une tasse de café !

6.4 Remarques finales

- L'approche de la théorie de la réponse linéaire que nous avons esquissée et appliquée au cas simple de la diffusion de particules marquées permet de donner des expressions microscopiques, à partir des fonctions de corrélation, des coefficients macroscopiques qui apparaissent dans les lois linéaires de la thermodynamique hors d'équilibre. On peut consulter les ouvrages de référence sur la simulation moléculaire [15, 1] pour en savoir plus. Il convient au préalable d'identifier des observables appropriées pour les variables ou les champs macroscopiques qui interviennent dans ces lois, comme par exemple le flux de chaleur.

Citons la formule de Green-Kubo pour la viscosité de cisaillement des fluides, qui fait intervenir le volume V et un terme non diagonal quelconque σ_{xy} de l'observable tenseur des contraintes définie par (2.22) :

$$\eta = \frac{V}{k_B T} \int_0^{+\infty} \langle \sigma_{xy}(0) \sigma_{xy}(t) \rangle dt.$$

- La propriété de symétrie (6.10) des fonctions de corrélation est en fait à l'origine des *relations de réciprocité d'Onsager* [23], important résultat de la thermodynamique des phénomènes hors d'équilibre couplés. Un tel couplage se rencontre par exemple dans les conducteurs, pour lesquels un gradient de potentiel électrique entraîne un courant électrique (loi d'Ohm) et aussi un flux de chaleur (effet Peltier), tandis qu'un gradient de température entraîne un flux de chaleur (loi de Fourier) et aussi un courant électrique (effet Seebeck). Les réponses des flux (chaleur, courant) aux « forces » conjuguées (gradients de température ou de potentiel) font intervenir une matrice de coefficients de réponse dont les relations de réciprocité, issues de (6.10), établissent la *symétrie*.

Introduction à la physique quantique

Introduction historique : la révolution quantique

La mécanique quantique est une théorie physique apparue au début du vingtième siècle. Venant au secours de la mécanique Newtonienne et de l'électromagnétisme pour décrire le monde de l'infiniment petit, elle proposait une vision pour le moins révolutionnaire du fonctionnement des constituants intimes de la matière et de la lumière. Incroyable pour certains (même pour le grand Albert Einstein !), choquante, étrange, elle a pourtant donné le coup d'envoi d'un siècle de progrès technologiques sans précédents.

Le monde physique en 1900

Comme pour toute théorie, ses débuts furent difficiles. A la fin du 19^{ème} siècle, le temps des conservatismes religieux de l'Europe, qui avaient fait abjurer Galilée en 1633, est certes révolu. Pour autant, le monde scientifique connaît un moment d'autosatisfaction, qui le rend peu réceptif aux idées nouvelles. Les physiciens ont en effet à leur disposition trois théories, dites de nos jours théories 'classiques', très abouties :

- la *mécanique Newtonienne*, développée à partir du 17^{ème} siècle, qui met en équations le mouvement des corps, et dont les concepts se retrouvent à la base des disciplines que sont la mécanique des fluides et la résistance des matériaux ;
- la *thermodynamique*, née au début du 19^{ème} siècle, que Boltzmann vient vers la fin du 19^{ème} siècle de faire évoluer jusqu'à sa version statistique ;
- l'*électromagnétisme*, élaboré dans la deuxième moitié du 19^{ème} siècle, qui appréhende dans une théorie unifiée le caractère ondulatoire de la lumière et les phénomènes d'optique, les phénomènes d'électro et de magnétostatique, ainsi que ceux de conduction et d'induction électrique.

Dans ces théories, les constituants de ce monde se partagent en deux entités clairement distinctes :

- la *matière*, d'une part, c'est à dire les corps solides, liquides ou gazeux, mais surtout 'palpables', que la mécanique classique appréhende au travers de la notion de particule, ou plus précisément, de *point matériel*. Cette entité a un caractère ponctuel et se meut dans l'espace le long d'une trajectoire. Soulignons qu'il n'est pas encore question d'atome à l'époque. L'idée de ces grains élémentaires de matière, empruntée à la philosophie grecque antique, fait encore débat. Dès 1869, Mendeleïev a certes déjà élaboré sa fameuse classification des éléments. Pour autant, il faudra attendre la preuve expérimentale de l'existence de l'électron en 1897, puis de celle du noyau atomique en 1912, pour convaincre les derniers incrédules ;
- les *champs* diffus (champ de gravitation, champ électrique, champ magnétique), qui occupent tout l'espace, vide et matière compris, et sont capables d'interaction avec la matière. Leurs 'mouvements' (au moins pour les champs électromagnétiques) prennent la forme de déformations, appelées *ondes*, plus ou moins étendues à travers l'espace, et qui ont la propriété de se propager.

Force est de constater que ces théories expliquent remarquablement grand nombre d'aspects du monde tel que nous le percevons avec nos cinq sens. Elles sont par ailleurs entièrement suffisantes aux développements technologiques de l'époque. Il n'est ainsi pas très étonnant qu'en 1892,

William Thomson, alias Lord Kelvin, co-fondateur de l'électromagnétisme et de la thermodynamique, déclare :

La physique est définitivement constituée dans ses concepts fondamentaux : tout ce qu'elle peut désormais apporter, c'est la détermination précise de quelques décimales supplémentaires.

Cependant, cette citation se poursuit :

Il y a bien deux petits problèmes : celui du résultat de l'expérience de Michelson et celui du corps noir, mais ils seront rapidement résolus et n'altèrent en rien notre confiance. . .

Deux petits problèmes...

Deux petits problèmes en effet. . . En mesurant avec précision la vitesse de la lumière, Michelson vient de montrer en 1887 que la théorie de l'électromagnétisme s'accommode mal des changements de repère, et qu'il n'existe surtout pas de repère absolu pouvant servir à écrire les équations de Maxwell. Quant au fameux *corps noir*, bloc de matière chauffé dont on mesure le spectre d'émission radiatif (il est connu qu'un fer chaud prend une couleur rouge, mais des radiations sont en fait émises à toutes les longueurs d'onde), nul n'est à l'époque capable de comprendre la forme des spectres obtenus.

Et ces deux petits problèmes ne sont bientôt pas les seuls. Les progrès techniques aidant, les physiciens se dotent en cette fin de siècle de nouveaux instruments de mesure et d'observation. Ils imaginent de nouvelles expériences, et auscultent la nature avec une précision et une minutie sans précédent. L'*effet photoélectrique*, mis en évidence en 1887 par Hertz, soulève le problème de la compréhension du rôle joué par la longueur d'onde de la lumière dans sa capacité à extraire les électrons d'un solide. Viendra rapidement l'énigme de la stabilité de l'atome : en 1912, les électrons et les noyaux atomiques voient leur existence prouvée expérimentalement. Pour autant, les théories physiques classiques prédisent qu'un assemblage noyau-électrons ne peut que s'effondrer sur lui-même. Comment comprendre le fait que l'atome 'tient' ?

Autant de questions auxquelles, malgré l'optimisme affiché de Kelvin, la science 'classique' s'avère finalement incapable de répondre. Les théories de l'époque, que l'on croyait parfaitement abouties, et surtout, de portée universelle, montrent leurs limites. On découvre peu à peu qu'il existe, au delà des phénomènes classiques, tout un monde de nouveaux phénomènes dont on ignorait jusqu'alors l'existence. Les théories classiques ne semblent finalement s'appliquer qu'à une portion restreinte d'entre eux.

Tracer les frontières du domaine d'applications des théories classiques, construire une nouvelle théorie adaptée à la description et à la prédiction de ces nouveaux phénomènes, telle fut la tâche des physiciens au début du 20^{ème} siècle. Au final, ce ne sont pas une, mais plusieurs théories qui verront le jour, parfois d'ailleurs développées par les mêmes personnes. De l'expérience de Michelson naîtra la *relativité*. Le rayonnement du corps noir, l'effet photo-électrique, et le problème de la stabilité de l'atome conduiront quant à eux à la construction d'une nouvelle physique : la *physique quantique*.

Une théorie choquante

La physique quantique s'attache essentiellement (mais, nous le verrons, pas exclusivement) à rendre compte du mode de fonctionnement de la matière et de la lumière aux échelles microscopiques. Elle choque, intrigue parfois. Tout d'abord, par les principes physiques qu'elle défend. Citons pêle-mêle :

- la *dualité onde-corpuscule* : la lumière, par essence une onde en physique classique, peut parfois se 'matérialiser' en un point précis de l'espace, à la manière d'une particule. Réciproquement, des particules telles que des électrons et des neutrons peuvent se comporter aussi comme des ondes, pouvant se superposer et interférer entre elles ;
- le *principe d'incertitude* : Il est par exemple, impossible physiquement de connaître simultanément la position et la vitesse d'une particule ;

- le *don d'ubiquité* des particules : une particule peut occuper simultanément plusieurs points de l'espace ;
- l'impossibilité de faire une mesure sur un système sans modifier l'état de ce dernier ;
- le fait que le résultat de cette mesure est quasi-systématiquement aléatoire, tout en ne pouvant prendre que des valeurs bien précises.

Et la liste n'est pas exhaustive. « *Quiconque n'est pas choqué par la théorie quantique ne l'a pas comprise* », disait Niels Bohr, un des ses fondateurs. De nos jours, près de 100 ans après ses premiers balbutiements, et après pratiquement autant d'années de mise en oeuvre sans faille, la mécanique quantique heurte toujours le sens commun. Le fait est que loin de se limiter à nos fournir de nouveaux outils mathématiques pour décrire la nature, la mécanique quantique nous met face à un monde que nous connaissons mal, et dont nous n'avons en général aucune intuition.

Au delà de ses principes physiques, la théorie quantique choque aussi par son formalisme mathématique. Les équations peuvent paraître, d'une certaine manière, plus compliquées, et la réalité physique plus lointaine que dans les théories classiques, à tel point qu'au début du 20^{ème} siècle, cette théorie des quantas fut parfois appelée *mécanique des matrices*. Soulignons qu'à l'époque, les concepts mathématiques sur lesquels elle s'appuie, notamment les espaces de Hilbert et dans une moindre mesure, les distributions, sont eux-même loin d'être aboutis. C'est ainsi à un développement conjoint de la physique et des mathématiques que l'on a assisté. La théorie quantique doit beaucoup pour sa mise en forme à des mathématiciens comme Hilbert ou Von Neumann. En retour, elle a été source d'inspiration pour lancer de nouveaux sujets de recherche en mathématique. Il n'est d'ailleurs pas rare que certains résultats de la théorie quantique, initialement admis par les physiciens en tant que faits expérimentaux, mais sans que l'on sache les démontrer, n'aient vu leurs preuves mathématiques apportées que longtemps après.

La théorie quantique étonne donc, et pourtant, après un siècle de confrontation continue avec l'expérience, force est de constater que... ça marche ! Jamais mise en défaut à ce jour, elle parvient, en ne s'appuyant que sur un petit nombre de principes, à rendre compte avec une extrême finesse de toujours plus d'observations expérimentales. Construite à l'origine sur la base d'expériences impliquant des électrons, sa validité a pu depuis être vérifiée sur les noyaux atomiques puis les constituants les plus intimes de la matière. La compréhension qu'elle a donné du monde microscopique a permis d'imaginer, puis de réaliser, nombre de développements technologiques et scientifiques du 20^{ème} siècle :

- les *machines électroniques à semi-conducteurs*, que l'on retrouve aujourd'hui partout, de la simple montre à l'ordinateur le plus perfectionné, en passant par le lave linge, les systèmes de communication, et les panneaux solaires ;
- les *nouvelles sources lumineuses* : la lumière LASER, bien sur, et ses innombrables applications scientifiques, industrielles et médicales, mais aussi les diodes électroluminescentes (LED), qui, simples voyants d'appareils ou éléments de télécommandes infra-rouges à l'origine, deviennent aujourd'hui une véritable source économique d'éclairage ;
- les techniques de *spectroscopie analytique* (spectroscopie d'absorption UV-visible-Infra rouge, spectroscopie de fluorescence, Résonance Magnétique Nucléaire, ...) qui permettent en peu de temps de retrouver la structure chimique de composés, que ce soit dans un tube à essai ou aux confins de l'espace. Citons également un appareil de notre quotidien dérivé de ces techniques : le four à micro-ondes ;
- une nouvelle discipline scientifique : la *chimie quantique*. La physique quantique donne les clefs de compréhension du mécanisme de la liaison chimique, et permet maintenant de prédire, à l'avance, par le biais de la simulation numérique et -à nouveau- grâce à de considérables développements en mathématiques appliquées, parfois bien mieux que l'expérience, quelles seront les propriétés d'une molécule donnée ;
- loin de l'infiniment petit, on a découvert qu'il existait aussi des *systèmes de grande taille* régis par la mécanique quantique : des corps célestes tels que les étoiles à neutrons, mais également des corps à notre échelle, tels que les aimants, dont seule la physique quantique explique l'aimantation à température ambiante. N'oublions pas enfin les supraconducteurs, riches d'acquis autant que d'espoirs au plan technologique.

Une fois que l'on a, sinon compris, du moins assimilé et admis les fondements de la physique quantique, il est frappant de réaliser que loin d'être abstraite et inaccessible, elle est à la base de nombre de petites choses de notre quotidien. La couleur des objets, par exemple, ou le fait, tout simplement, que tous les matériaux denses qui nous entourent ont des masses volumiques comparables, de l'ordre de 1 à 10 kg par litre.

Les théories quantiques : Mécanique quantique et Théorie des champs

Parlons maintenant de *mécanique* quantique. Car peut-être l'avez vous remarqué, nous n'avons jusqu'ici employé que le terme de *physique* quantique. La physique quantique désigne l'ensemble des adaptations qu'ont connues au début du vingtième siècle la mécanique Newtonienne et l'électromagnétisme. La distinction entre les deux disciplines, clairement séparées à l'origine, est devenue un peu plus floue sur le fond, mais s'est néanmoins maintenue en grande partie dans la forme, si bien que la physique quantique se retrouve divisée en deux branches :

- la *mécanique quantique*, qui se préoccupe, comme la mécanique classique, du mouvement et de la dynamique de la matière, et qui peut être vue comme une extension de la mécanique classique vers les échelles microscopiques ;
- la *théorie quantique des champs* : elle englobe en premier lieu l'*optique quantique*, attachée à la description des champs électromagnétiques, et qui manipule entre autres la notion de *photons*. Mais elle s'intéresse de manière plus générale à tous les phénomènes vibratoires non localisés (ondes acoustiques dans les solides, avec la notion de *phonons*). Beaucoup plus difficile d'accès que la mécanique quantique, nous nous contenterons d'y faire ponctuellement allusion.

Physique quantique vs physique classique

L'histoire nous enseigne donc que la théorie quantique a été pensée comme complémentaire de la théorie classique. Ce serait un contresens que de vouloir les opposer. La théorie quantique est un élargissement et un raffinement de la théorie classique, et permet de décrire de nouveaux phénomènes physiques. Réciproquement, la théorie classique peut être vue comme un moyen de simplifier la théorie quantique dans certaines situations.

Dans ce cours, nous verrons que, bien que le formalisme de la mécanique quantique soit clairement distinct de celui de la mécanique classique, il contient en lui nombre de résultats dits classiques. Un de nos soucis dans les pages qui viennent sera de faire ressortir, chaque fois que cela sera possible, ces similarités.

La particule quantique isolée

Le premier système que nous étudions dans le cadre de la théorie quantique est une particule isolée qui se déplace dans le vide, sans la moindre interaction avec l'extérieur. Ce système simple permet d'introduire le formalisme des *fonctions d'onde*. Alors que la mécanique classique enseigne que la particule, pour peu qu'on l'observe dans un repère Galiléen, suit un mouvement rectiligne et uniforme, la théorie quantique prédit déjà un comportement beaucoup plus riche.

8.1	Description quantique de la particule en translation	115
8.1.1	La fonction d'onde	115
8.1.2	L'équation de Schrödinger	116
8.1.3	La constante de Planck	117
8.1.4	Remarques	117
8.2	Conservation des probabilités - dynamique de la particule	117
8.2.1	Equation de Schrödinger et normalisation	117
8.2.2	Forme locale de la conservation des probabilités : <i>Courant et Vitesse</i> de probabilité	118
8.3	La phase de la fonction d'onde	119
8.4	Solutions de l'équation de Schrödinger dans le vide	119
8.4.1	Etude du Hamiltonien quantique de la particule libre	119
8.4.2	Solution générale : Le paquet d'ondes	121
8.5	Quelques solutions particulières	122
8.5.1	Les ondes de De Broglie	122
8.5.2	Le paquet d'onde quasi-monochromatique	124
8.5.3	Le paquet d'ondes gaussien	125
8.5.4	Superposition d'états	127
Exercices		129

8.1 Description quantique de la particule en translation

8.1.1 La fonction d'onde

En mécanique quantique, la particule perd son aspect localisé. Elle peut s'étaler sur un domaine plus ou moins étendu de l'espace, et marquer de son empreinte simultanément plusieurs positions. Son état ne se décrit plus à l'aide d'un vecteur position $\mathbf{r}(t)$ et d'un vecteur impulsion $\mathbf{p}(t)$. Ces deux notions sont remplacées par celle de *fonction d'onde*.

Définition 8.1. *L'état de la particule quantique au temps t est entièrement décrit par une fonction d'onde notée $\psi(\mathbf{r}, t)$. C'est un champ scalaire à valeurs complexes, défini en tout point $\mathbf{r}$ de l'espace.*

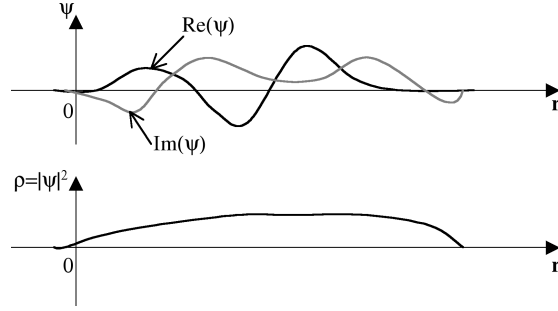


Fig. 8.1. La fonction d'onde complexe $\psi(\mathbf{r}, t)$ et la probabilité de présence $\rho(\mathbf{r}, t) = |\psi(\mathbf{r}, t)|^2$.

Une interprétation physique de $\psi(\mathbf{r}, t)$ est donnée par le phénomène expérimental suivant : lorsqu'on cherche, au moyen d'un instrument de mesure, à localiser la position de la particule dans l'espace, cette dernière se 'matérialise' en un point, choisi aléatoirement parmi les positions où $\psi(\mathbf{r}, t) \neq 0$. Au cours de ce processus, tous les points possibles d'apparition ne sont pas équiprobables. La probabilité $p(B)$ de trouver la particule dans un domaine B de l'espace est

$$p(B) = \int_B |\psi(\mathbf{r}, t)|^2 d\mathbf{r}.$$

Comme on ne considère qu'une seule particule¹, l'espace ambiant est ici $\mathbb{R}^3$.

Définition 8.2. La densité de probabilité de faire apparaître la particule au point $\mathbf{r}$ lors d'une mesure est $\rho(\mathbf{r}, t) = |\psi(\mathbf{r}, t)|^2$. Elle vérifie la condition de normalisation :

$$\forall t \in \mathbb{R}, \quad \int_{\mathbb{R}^3} \rho(\mathbf{r}, t) d\mathbf{r} = \int_{\mathbb{R}^3} |\psi(\mathbf{r}, t)|^2 d\mathbf{r} = 1. \quad (8.1)$$

Par abus de langage, $\rho(\mathbf{r}, t)$ est souvent simplement appelée *probabilité de présence*. Au vu de ces relations, la probabilité de présence ρ s'exprime en m^{-3} , et l'amplitude ψ en $\text{m}^{-3/2}$.

D'un point de vue mathématique, l'espace des fonctions d'onde est un espace de Hilbert², que l'on notera $\mathcal{H}$ par la suite. Dans le cas d'une particule sans spin³, l'espace $\mathcal{H}$ est identifié à l'espace $L^2(\mathbb{R}^3, \mathbb{C})$.

8.1.2 L'équation de Schrödinger

Postulat 8.1. L'évolution temporelle en repère Galiléen de la fonction d'onde d'une particule isolée de masse m est régie par l'équation de Schrödinger :

$$i\hbar \frac{\partial \psi}{\partial t} = -\frac{\hbar^2}{2m} \Delta \psi. \quad (8.2)$$

Dans cette équation, i est le nombre complexe tel que $i^2 = -1$, et Δ désigne l'opérateur Laplacien par rapport à la variable d'espace $\mathbf{r}$. L'équation de Schrödinger est ainsi une équation d'évolution. Le fait qu'elle est bien posée mathématiquement sera vu dans un cadre plus général au §9.1.2.

1. Le formalisme des fonctions d'onde permet également de décrire simultanément plusieurs particules (voir le chapitre 13).

2. On rappelle qu'un espace de Hilbert est un espace vectoriel normé complet muni d'un produit scalaire.

3. Vous avez certainement entendu parler de cette petite flèche orientée vers le haut ou le bas dont on affuble l'électron dans certaines situations. Nous verrons au § 12.6 à quoi elle correspond précisément, et surtout, dans l'immédiat, nous l'oublions !

8.1.3 La constante de Planck

L'équation de Schrödinger fait apparaître une nouvelle constante physique : la *constante de Planck*. Elle est notée $\hbar$, et vaut dans le système d'unités internationales :

$$\hbar = 1.055 \times 10^{-34} \text{ Js.}$$

On prononce 'hache barre'. Elle est au coeur des théories quantiques et permet aussi de définir les ordres de grandeurs typiques des caractéristiques physiques des systèmes quantiques (taille de l'atome, énergie des photons, ... - voir Appendice A.2).

Exprimée dans les unités plus conventionnelles que sont le *mètre*, le *kilogramme*, et la *seconde*, l'unité de la constante de Planck est le $\text{kg m}^2 \text{s}^{-1}$. Elle a donc même unité qu'un *moment cinétique*.

8.1.4 Remarques

- (1) On retrouve ici les analogues aux ingrédients de la théorie classique :
 - on a un *objet mathématique* pour décrire l'état du système à un instant donné : c'est la fonction d'onde, qui remplace le couple position-impulsion ;
 - on a également une *loi* qui détermine la manière dont le système évolue au cours du temps : c'est l'équation de Schrödinger, que l'on peut voir comme une version quantique de l'équation fondamentale de la dynamique.
- (2) La particule dont il est question peut être une particule élémentaire au sens physique du terme, telle qu'un électron, un proton, ou un neutron. Mais il peut s'agir plus généralement du centre de gravité d'un corps solide que l'on désirerait décrire de manière quantique. De même qu'en mécanique classique, la théorie quantique ne fait aucune distinction, dans son formalisme, entre une particule physique et le centre de masse d'un objet volumineux.
- (3) *Probabilité de présence* ou *densité de probabilité d'apparition*? Nous avons dit que $|\psi(\mathbf{r}, t)|^2$ est la densité de probabilité du lieu de matérialisation de la particule lorsqu'on cherche à en mesurer la position. On pourrait être tenté d'en déduire que la particule ne perd jamais son caractère ponctuel, que pour une raison obscure, on ne pourrait jamais exactement connaître sa position, et que la théorie quantique serait chargée de gérer l'information manquante de manière statistique.

Gardons nous bien de cette erreur d'interprétation. Car à chaque instant, la particule est bel et bien présente *simultanément* en divers points de l'espace. Malheureusement, la mécanique quantique affuble la quantité $|\psi(\mathbf{r}, t)|^2$ du nom de *probabilité de présence*... Tradition dangereuse qui pousse à la confusion. D'ailleurs, si on cherche à être précis sur les termes, $|\psi(\mathbf{r}, t)|^2$ n'est même pas au sens strict une probabilité, mais une *densité* de probabilité.

Ce qui pousse la particule à se montrer en un seul point est un phénomène bien particulier de la mécanique quantique, nommé *réduction du paquet d'onde*. Nous y reviendrons au § 11.4.

8.2 Conservation des probabilités - dynamique de la particule

Nous examinons un certain nombre de propriétés mathématiques remarquables de l'équation de Schrödinger. Certaines sont nécessaires pour assurer le caractère cohérent de la théorie quantique. Elles permettent également de saisir le contenu *dynamique* de la fonction d'onde.

8.2.1 Equation de Schrödinger et normalisation

Proposition 8.1. *L'équation de Schrödinger est compatible avec la condition de normalisation de la fonction d'onde, au sens où, si $\|\psi(0, \cdot)\|_{L^2(\mathbb{R}^3)} = 1$, alors, pour tout $t \in \mathbb{R}$, $\|\psi(t, \cdot)\|_{L^2(\mathbb{R}^3)} = 1$.*

Preuve. La démonstration sera développée dans un cadre plus général au § 9.1.4. Elle consiste formellement à considérer la quantité :

$$N(t) = \int_{\mathbb{R}^3} |\psi(\mathbf{r}, t)|^2 d\mathbf{r}.$$

On établit, sur la base de l'équation de Schrödinger, que $\frac{dN}{dt} = 0$. On en déduit que s'il existe t_0 tel que $N(t_0) = 1$ alors, $N(t) = 1$ pour tout $t \in \mathbb{R}$. $\square$

Ce résultat permet de ne pas faire de la condition de normalisation une préoccupation permanente : il suffit en particulier qu'elle soit vérifiée dans les conditions initiales d'un problème pour qu'elle soit vérifiée tout le temps.

8.2.2 Forme locale de la conservation des probabilités : Courant et Vitesse de probabilité

La conservation globale de la probabilité de présence $\rho(\mathbf{r}, t) = |\psi(\mathbf{r}, t)|^2$ peut s'écrire de manière locale, à la manière d'une conservation de la masse en mécanique des fluides⁴. Dans les définitions qui suivent, $\theta(\mathbf{r}, t)$ désigne la *phase* de la fonction d'onde.

Définition 8.3 (Courant et vitesse de probabilité). *Etant donnée une fonction d'onde $\psi(\mathbf{r}, t) = \sqrt{\rho(\mathbf{r}, t)}e^{i\theta(\mathbf{r}, t)}$ qui décrit l'état quantique d'une particule de masse m , on appelle courant de probabilité le champ de vecteurs à valeurs réelles :*

$$\mathbf{J}(\mathbf{r}, t) = -i\frac{\hbar}{2m} (\bar{\psi}\nabla\psi - \psi\nabla\bar{\psi}). \quad (8.3)$$

On appelle *vitesse de probabilité* le champ de vecteur à valeurs réelles :

$$\mathbf{v}(\mathbf{r}, t) = \frac{\hbar}{m} \nabla\theta. \quad (8.4)$$

Le caractère réel du courant de probabilité vient du fait que le terme de droite de (8.3) est la différence de deux quantités complexes conjuguées. Ces deux nouvelles grandeurs prennent pleinement leur sens physique grâce à la propriété (formelle) suivante :

Propriété 8.1 (Conservation locale). *La probabilité de présence $\rho(\mathbf{r}, t)$, le courant de probabilité $\mathbf{J}(\mathbf{r}, t)$, et la vitesse de probabilité $\mathbf{v}(\mathbf{r}, t)$ vérifient les relations suivantes de conservation locale de la probabilité de présence :*

$$\mathbf{J}(\mathbf{r}, t) = \rho(\mathbf{r}, t)\mathbf{v}(\mathbf{r}, t), \quad (8.5)$$

$$\frac{\partial\rho}{\partial t} + \nabla \cdot \mathbf{J} = \frac{\partial\rho}{\partial t} + \nabla \cdot (\rho\mathbf{v}) = 0. \quad (8.6)$$

Ces équations forment un système hyperbolique de lois de conservation.

La première des deux relations se montre simplement en considérant des fonctions d'onde de la forme $\psi(\mathbf{r}, t) = \sqrt{\rho(\mathbf{r}, t)}e^{i\theta(\mathbf{r}, t)}$ dans l'expression de $\mathbf{J}$. On laisse en exercice le fait que

$$\mathbf{J} = \rho \frac{\hbar}{m} \nabla\theta.$$

Pour la seconde relation, il suffit de vérifier l'équation concernant le vecteur $\mathbf{J}$. Par l'équation de Schrödinger (8.2),

$$\frac{\partial\psi}{\partial t} = \frac{i\hbar}{2m} \Delta\psi, \quad \text{et} \quad \frac{\partial\bar{\psi}}{\partial t} = -\frac{i\hbar}{2m} \Delta\bar{\psi}.$$

Il vient ainsi :

4. D'ailleurs, ce champ scientifique est celui de l'hydrodynamique quantique (*quantum hydrodynamics*).

$$\frac{\partial \rho}{\partial t} = \bar{\psi} \frac{\partial \psi}{\partial t} + \frac{\partial \bar{\psi}}{\partial t} \psi = \frac{i\hbar}{2m} (\bar{\psi} \Delta \psi - \psi \Delta \bar{\psi}).$$

Par ailleurs,

$$\nabla \cdot \mathbf{J} = -i \frac{\hbar}{2m} (\nabla \bar{\psi} \cdot \nabla \psi + \bar{\psi} \cdot \Delta \psi - \nabla \psi \cdot \nabla \bar{\psi} - \psi \cdot \Delta \bar{\psi}) = -i \frac{\hbar}{2m} (\bar{\psi} \cdot \Delta \psi - \psi \cdot \Delta \bar{\psi}).$$

On retrouve ainsi l'expression précédente.

L'équation (8.6) a la forme d'une équation de conservation de matière. Elle permet d'interpréter $\mathbf{J}(\mathbf{r}, t)$ comme une densité de flux de probabilité de présence. La densité $\mathbf{J}(\mathbf{r}, t)$ indique comment la particule se déplace localement pour réarranger la manière dont elle occupe l'espace.

De même qu'en mécanique des fluides, l'origine du courant peut être associée, au moins de manière formelle, à l'existence d'un champ $\mathbf{v}(\mathbf{r}, t)$ de 'vitesses d'écoulement des probabilités'. Il est remarquable que ce champ ne dépende que du gradient de la phase de la fonction d'onde. C'est ce *gradient de phase* $\theta(\mathbf{r}, t)$, que nous avons ignoré jusque là, qui constitue l'*équivalent quantique du champ de vitesse* classique.

8.3 La phase de la fonction d'onde

Les quantités $|\psi(\mathbf{r}, t)|$ et $\nabla \theta(\mathbf{r}, t)$ sont les analogues des vecteurs position et impulsion (à la masse près) qui en mécanique classique décrivent complètement l'état d'une particule ponctuelle. Il est cependant intrigant que la phase $\theta(\mathbf{r}, t)$ n'intervienne que par son gradient. En fait, sa valeur précise n'a aucune signification physique en elle-même. Seules les *différences* de phases au sein d'une même fonction d'onde sont porteuses d'information.

En effet, si on déphase globalement d'une constante θ_0 la fonction d'onde $\psi(\mathbf{r}, t)$ en la multipliant par $e^{i\theta_0}$, on obtient une nouvelle fonction d'onde $\tilde{\psi}(\mathbf{r}, t) = \psi(\mathbf{r}, t) e^{i\theta_0}$ de phase $\tilde{\theta}(\mathbf{r}, t) = \theta(\mathbf{r}, t) + \theta_0$ et qui vérifie :

$$\begin{cases} |\tilde{\psi}(\mathbf{r}, t)|^2 = |\psi(\mathbf{r}, t)|^2, \\ \nabla \tilde{\theta}(\mathbf{r}, t) = \nabla \theta(\mathbf{r}, t). \end{cases}$$

Les fonctions d'onde $\tilde{\psi}$ et ψ contiennent donc la même information physique de localisation et de cinétique sur la particule. Il en ressort que :

Propriété 8.2. *La fonction d'onde d'une particule n'est définie qu'à une phase près sans importance. Deux fonctions d'onde qui ne diffèrent l'une de l'autre qu'à un déphasage global près décrivent le même état quantique.*⁵

8.4 Solutions de l'équation de Schrödinger dans le vide

On étudie dans cette section les solutions de (8.2). Avant de donner la forme générale des solutions, nous étudions au préalable l'opérateur $-\Delta$ dans le cadre de la théorie spectrale. Cet exercice très mathématique s'avèrera particulièrement utile par la suite.

8.4.1 Etude du Hamiltonien quantique de la particule libre

On définit l'opérateur $\hat{H}_0 = -\frac{\hbar^2}{2m} \Delta$ sur $\mathcal{H} = L^2(\mathbb{R}^3)$, dit *opérateur Hamiltonien*. Pour plus de clarté, on notera par la suite les opérateurs avec un $\hat{\cdot}$. Cette notation n'a rien à voir avec la notation $\hat{u}$ qui désigne usuellement la transformée de Fourier

5. Il existe d'ailleurs pour la mécanique quantique un formalisme particulier, dit des *opérateurs densité* (hors programme pour nous !) qui ne travaille qu'en terme de différences de phases, et élimine naturellement ce problème de redondance dans la transcription des états.

$$\forall \mathbf{k} \in \mathbb{R}^3, \quad \hat{u}(\mathbf{k}) = \mathcal{F}(u)(\mathbf{k}) = (2\pi)^{-3/2} \int_{\mathbb{R}^3} u(\mathbf{r}) e^{-i\mathbf{k} \cdot \mathbf{r}} d\mathbf{r}$$

de u . On notera par défaut

$$\langle f, g \rangle = \int_{\mathbb{R}^3} \overline{f(\mathbf{r})} g(\mathbf{r}) d\mathbf{r}$$

le produit scalaire sur $\mathcal{H}$. L'opérateur $\hat{H}_0$ est évidemment symétrique sur $C_0^\infty(\mathbb{R}^3)$: pour $\psi, \phi \in C_0^\infty(\mathbb{R}^3)$,

$$\int_{\mathbb{R}^3} -\frac{\hbar^2}{2m} \Delta \psi \cdot \phi = \frac{\hbar^2}{2m} \int \nabla \bar{\psi} \cdot \nabla \phi = \int_{\mathbb{R}^3} \bar{\psi} \cdot \left(-\frac{\hbar^2}{2m} \Delta \phi \right).$$

La propriété fondamentale qui suit permet d'appliquer beaucoup de résultats importants (rappelés en Appendice C).

Proposition 8.2 (Caractère auto-adjoint de $\hat{H}_0$). *L'opérateur*

$$\hat{H}_0 = -\frac{\hbar^2}{2m} \Delta$$

défini sur le domaine

$$D(\hat{H}_0) = H^2(\mathbb{R}^3) = \{\psi \in L^2(\mathbb{R}^3), \Delta \psi \in L^2(\mathbb{R}^3)\}$$

dense dans $\mathcal{H} = L^2(\mathbb{R}^3)$, est un opérateur auto-adjoint non-borné sur $L^2(\mathbb{R}^3)$.

Preuve. Les propriétés de $\hat{H}_0$ sont directement héritées de celles de l'opérateur Laplacien Δ . On vérifie facilement, en utilisant la transformée de Fourier, que $\psi, \Delta \psi \in L^2(\mathbb{R}^3)$, si et seulement si $\psi \in H^2(\mathbb{R}^3)$. Comme $C_0^\infty(\mathbb{R}^3) \subset D(\hat{H}_0)$, $\hat{H}_0$ est bien à domaine dense dans $\mathcal{H}$. Par définition,

$$D(\hat{H}_0^*) = \left\{ u \in L^2(\mathbb{R}^3), \exists v \in L^2(\mathbb{R}^3), \forall w \in H^2(\mathbb{R}^3), \left\langle -\frac{\hbar^2}{2m} \Delta w, u \right\rangle = \langle w, v \rangle \right\}.$$

Or,

$$\left\langle -\frac{\hbar^2}{2m} \Delta w, u \right\rangle = \frac{\hbar^2}{2m} \langle |\mathbf{k}|^2 \hat{w}, \hat{u} \rangle = \frac{\hbar^2}{2m} \int_{\mathbb{R}^3} \hat{w}(\mathbf{k}) (|\mathbf{k}|^2 \hat{u}(\mathbf{k})) d\mathbf{k}.$$

Ceci montre que $u \in D(\hat{H}_0^*)$ si et seulement si $u \in L^2(\mathbb{R}^3)$ et $|\mathbf{k}|^2 \hat{u} \in L^2(\mathbb{R}^3)$, i.e. $u \in H^2(\mathbb{R}^3) = D(\hat{H}_0)$. On a en outre

$$\mathcal{F}(\hat{H}_0^* u)(\mathbf{k}) = \frac{\hbar^2}{2m} |\mathbf{k}|^2 \hat{u}(\mathbf{k}),$$

soit $\hat{H}_0^* u = -\frac{\hbar^2}{2m} \Delta u$. Ceci montre finalement que $\hat{H}_0$ est auto-adjoint.

Il est clair que $\hat{H}_0$ est non-borné. On peut toutefois s'en convaincre en considérant la fonction $u(x) = \chi(x) |x|^{-1}$ avec χ régulière à support dans $B(0, 1)$, qui est bien $L^2(\mathbb{R}^3)$. En revanche, Δu n'est pas dans $L^2(\mathbb{R}^3)$. On peut alors trouver une suite de fonctions régulières u_n telles que $\|u_n\|_{L^2(\mathbb{R}^3)} < +\infty$ alors que $\|\Delta u_n\|_{L^2(\mathbb{R}^3)} \rightarrow +\infty$ et donc $\|\hat{H}_0 u_n\|_{L^2(\mathbb{R}^3)} \rightarrow +\infty$ $\square$

On peut à présent caractériser le spectre de l'opérateur $\hat{H}_0$:

Proposition 8.3 (Spectre du hamiltonien de la particule libre). *Le spectre de $\hat{H}_0 = -\frac{\hbar^2}{2m} \Delta$ est purement continu :*

$$\sigma(\hat{H}_0) = \sigma_{\text{ess}}(\hat{H}_0) = \sigma_c(\hat{H}_0) = [0, +\infty[.$$

Preuve. Comme $\hat{H}_0$ est auto-adjoint, son spectre est contenu dans $\mathbb{R}$ et son spectre résiduel est vide. On remarque ensuite que s'il existait $\lambda \in \mathbb{R}$ et $u \in D(\hat{H}_0)$ tel que $\hat{H}_0 u = \lambda u$, alors $(\lambda - \frac{\hbar^2}{2m} |\mathbf{k}|^2) \hat{u}(\mathbf{k}) = 0$ pour tout $\mathbf{k} \in \mathbb{R}^3$, ce qui donne $\hat{u} = 0$. Ceci montre que $\hat{H}_0$ n'a pas de valeur propre, et le spectre ponctuel est vide. On remarque ensuite que $-\frac{\hbar^2}{2m} \Delta - \lambda \geq -\lambda$, et donc que $-\frac{\hbar^2}{2m} \Delta - \lambda$ est inversible si $\lambda < 0$. En effet,

$$\left(-\frac{\hbar^2}{2m}\Delta - \lambda\right)u = v \in L^2(\mathbb{R}^3)$$

est équivalent à

$$u = \mathcal{F}^{-1} \left(\frac{\hat{v}}{|\lambda|^2 + \frac{\hbar^2|\xi|^2}{2m}} \right) \in H^2(\mathbb{R}^3).$$

L'inverse $(-\frac{\hbar^2}{2m}\Delta - \lambda)^{-1}$ est donc de domaine dense et borné.

On a donc à ce stade $\sigma(\hat{H}_0) = \sigma_{\text{ess}}(\hat{H}_0) = \sigma_c(\hat{H}_0) \subset [0, +\infty[$. Montrons à présent $\sigma_{\text{ess}}(\hat{H}_0) = [0, +\infty[$. On considère pour ce faire $\lambda \geq 0$, $\chi \in C^\infty(\mathbb{R}^3)$ telle que $\hat{\chi}$ soit à support dans $B(0, 2)$, et de norme $\|\chi\|_{L^2(\mathbb{R}^3)} = 1$. On définit alors la suite de fonctions

$$u_n(\mathbf{r}) = n^{-3/2} \chi\left(\frac{\mathbf{r}}{n}\right) \quad u(\mathbf{r}) = n^{-3/2} \chi\left(\frac{\mathbf{r}}{n}\right) e^{i\mathbf{k}' \cdot \mathbf{r}},$$

où $\mathbf{k}' \in \mathbb{R}^3$ est tel que $\frac{\hbar^2|\mathbf{k}'|^2}{2m} = \lambda$. On a, au sens des distributions, $-\frac{\hbar^2}{2m}\Delta u = \lambda u$, mais $u \notin L^2(\mathbb{R}^3)$. En revanche, $u_n \in L^2(\mathbb{R}^3)$ pour tout n et

$$\|u_n\|_{L^2(\mathbb{R}^3)} = 1.$$

Par ailleurs, comme

$$\left| -\frac{\hbar^2}{2m}\Delta u_n - \lambda u_n \right| \leq \frac{\hbar^2}{m} \frac{1}{n^{5/2}} |\mathbf{k}'| \left| \nabla \chi\left(\frac{\mathbf{r}}{n}\right) \right| + \frac{\hbar^2}{2m} \frac{1}{n^{7/2}} \left| \Delta \chi\left(\frac{\mathbf{r}}{n}\right) \right|,$$

on a bien $\| -\frac{\hbar^2}{2m}\Delta u_n - \lambda u_n \|_{L^2(\mathbb{R}^3)} \rightarrow 0$. Enfin, pour $v \in L^2(\mathbb{R}^3)$,

$$\left| \int_{\mathbb{R}^3} \overline{u_n} v \right| = \left| \int_{\mathbb{R}^3} \overline{\mathcal{F}(u_n)} \hat{v} \right| \rightarrow 0$$

car $\mathcal{F}(u_n)(\mathbf{k}) = \sqrt{n} \hat{\chi}(n(\mathbf{k} + \mathbf{k}')) \neq 0$ seulement si $\mathbf{k} \in B(-\mathbf{k}', 2n^{-1})$. Ceci montre donc $u_n \rightharpoonup 0$. Finalement, le lemme de Weyl (Appendice C, Lemme C.1) montre que $\lambda \in \sigma_{\text{ess}}(-\frac{\hbar^2}{2m}\Delta)$. $\square$

Remarque 8.1. La preuve de la Proposition 8.3 montre en particulier que les ondes planes $e^{i\mathbf{k} \cdot \mathbf{r}}$ sont des quasi-fonctions propres, ou des fonctions propres généralisées, du Hamiltonien $\hat{H}_0$ (et donc de l'opérateur Laplacien), au sens où ce sont des limites de suites de Weyl.

8.4.2 Solution générale : Le paquet d'ondes

Nous allons voir à présent comment utiliser la caractérisation ci-dessus du spectre de $\hat{H}_0 = -\frac{\hbar^2}{2m}\Delta$ pour construire la forme générale des solutions de l'équation de Schrödinger (8.2). On vérifie facilement que, pour $\mathbf{k} \in \mathbb{R}^3$, si

$$E_{\mathbf{k}} = \frac{\hbar^2|\mathbf{k}|^2}{2m},$$

alors les fonctions

$$\nu_{\mathbf{k}}(\mathbf{r}, t) = e^{i(\mathbf{k} \cdot \mathbf{r} - \frac{E_{\mathbf{k}}}{\hbar} t)}$$

sont des solutions au sens des distributions de (8.2). Cette forme de fonctions, appelées *ondes planes*, ne correspond toutefois pas à des fonctions d'états physiquement admissibles car elles ne sont pas de norme $L^2(\mathbb{R}^3)$ finie. Ce sont des états propres généralisés (mathématiquement, des limites de suites de Weyl).

La solution générale *normalisable* de l'équation de Schrödinger s'obtient en considérant un *paquet d'ondes*, c'est à dire une superposition linéaire de plusieurs ondes planes (en utilisant une fonction de pondération $\phi(\mathbf{k})$) :

$$\psi(\mathbf{r}, t) = \frac{1}{(2\pi)^{3/2}} \int_{\mathbb{R}^3} \phi(\mathbf{k}) e^{i(\mathbf{k} \cdot \mathbf{r} - \frac{E_{\mathbf{k}}}{\hbar} t)} d\mathbf{k}, \quad (8.7)$$

avec $\phi \in L^2(\mathbb{R}^3)$. Notons que $\psi(\mathbf{r}, t)$ est la transformée de Fourier inverse de $\phi(\mathbf{k}) e^{-\frac{i}{\hbar} E_{\mathbf{k}} t}$. On vérifie que la fonction ψ ainsi construite vérifie bien l'équation de Schrödinger dans le vide (8.2) dès que

$$\int_{\mathbb{R}^3} |\phi(\mathbf{k})| |\mathbf{k}|^2 d\mathbf{k} < +\infty.$$

On a par ailleurs, par la relation de Parseval,

$$\int_{\mathbb{R}^3} |\psi(\mathbf{r}, t)|^2 d\mathbf{r} = \int_{\mathbb{R}^3} |\phi(\mathbf{k})|^2 d\mathbf{k}.$$

La condition de normalisation (8.1) aboutit au résultat :

Proposition 8.4. *La forme générale des solutions $\psi(\mathbf{r}, t)$ de l'équation de Schrödinger dans le vide est donnée par :*

$$\psi(\mathbf{r}, t) = \frac{1}{(2\pi)^{3/2}} \int_{\mathbb{R}^3} \phi(\mathbf{k}) e^{i(\mathbf{k} \cdot \mathbf{r} - \frac{E_{\mathbf{k}}}{\hbar} t)} d\mathbf{k}, \quad (8.8)$$

avec :

$$\int_{\mathbb{R}^3} |\phi(\mathbf{k})|^2 d\mathbf{k} = 1.$$

8.5 Quelques solutions particulières

En dépit de la forme très souple de la solution générale, la fonction d'onde d'une particule se déplaçant dans le vide prend en pratique quelques formes types que l'on examine à présent. Ce petit catalogue de solutions particulières de l'équation de Schrödinger va aussi permettre de nous fixer les idées.

8.5.1 Les ondes de De Broglie

Les *ondes de De Broglie* sont les ondes planes :

$$\nu_{\mathbf{k}}(\mathbf{r}, t) = e^{i(\mathbf{k} \cdot \mathbf{r} - \frac{E_{\mathbf{k}}}{\hbar} t)}. \quad (8.9)$$

Une représentation schématique de ces ondes est donnée par la figure 8.2. Bien qu'ils ne soient pas normalisables, ces états sont beaucoup utilisés en pratique pour décrire la situation limite dans laquelle la fonction de pondération ϕ devient extrêmement piquée autour d'un vecteur d'onde $\mathbf{k}$ donné. C'est le cas bien souvent des particules émises depuis une source lointaine⁶. En tant qu'état limite, la norme de l'onde de Broglie perd son sens physique de probabilité de présence. En revanche, sa phase garde toute sa pertinence.

Les ondes de De Broglie sont caractérisées par leur vecteur d'onde $\mathbf{k}$, et leur pulsation ω qui vaut ici $\omega = \frac{E_{\mathbf{k}}}{\hbar} = \frac{\hbar k^2}{2m}$. Ce résultat confirme, s'il était encore besoin de s'en convaincre, que l'équation de Schrödinger se comporte comme une équation des ondes. La particule peut donc prendre des aspects très différents :

6. Nous nous garderons bien de préciser ce que signifie 'lointain' dans le cas présent, mais pour nous fixer les idées, cela s'applique aux électrons émis par le canon à électrons d'un tube cathodique, après quelques dizaines de centimètres de parcours. La fonction d'onde tend naturellement à s'étaler sur une distance certes modeste pour nous, mais considérable pour cette particule. Les conditions particulières d'accélération initiale de la particule lui garantissant son caractère monochromatique, on peut alors, en première approximation, dire que l'état de l'électron en question est une onde de De Broglie vis à vis d'un obstacle que l'électron viendrait percuter. Cette propriété est à la base du fonctionnement des microscopes électroniques à transmission.

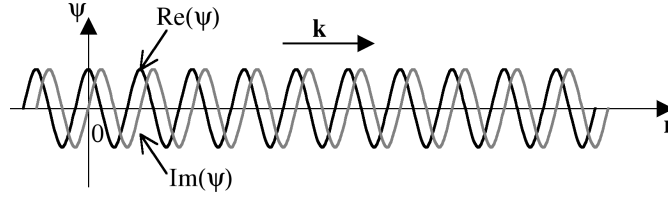


Fig. 8.2. L'onde de De Broglie $\psi(\mathbf{r}, t) = \nu_{\mathbf{k}}(\mathbf{r}, t)$.

- d'une part, elle est capable de se matérialiser en un seul point de l'espace lorsqu'on cherche à mesurer sa position ;
 - d'autre part, elle peut, avec des états s'approchant des ondes de De Broglie, également exhiber un comportement purement ondulatoire totalement délocalisé dans l'espace.
- Ces deux situations extrêmes fixent le cadre de ce qu'on a coutume d'appeler la *dualité onde-corpuscule*.

En tant que limite d'états normalisables (au sens des suites de Weyl), il est utile de donner un sens physique aux ondes de De Broglie, par exemple en les reliant à des concepts de physique classique.

- (1) Quelle est la vitesse de propagation de l'onde ? Les résultats de la mécanique classique donnent

$$\mathbf{v}_{\text{onde}} = \frac{\omega \mathbf{k}}{|\mathbf{k}|^2} = \frac{\hbar \mathbf{k}}{2m}.$$

Une première remarque vient du fait que contrairement, par exemple, à l'électromagnétisme et aux ondes acoustiques, les ondes de de Broglie ne se propagent pas à une vitesse unique. Leur vitesse de propagation est proportionnelle au vecteur d'onde $\mathbf{k}$ considéré.

- (2) Quelle est la vitesse de la particule ? Il ne s'agit pas directement de la vitesse de l'onde. Car l'onde *n'est pas* la particule, mais uniquement un objet mathématique qui la décrit. Nuance de taille ! La bonne vitesse à considérer ici est celle du déplacement des probabilités de présence. Elle ne dépend de la fonction d'onde que par sa phase, et le problème de normalisation n'empêche donc pas de la calculer :

$$\mathbf{v}_{\text{probabilité}}(\mathbf{r}, t) = \frac{\hbar}{m} \nabla \theta = \frac{\hbar \mathbf{k}}{m}.$$

La vitesse $\mathbf{v}_{\text{probabilité}}$ ne dépend ni du temps, ni de $\mathbf{r}$. L'onde de De Broglie décrit une particule qui se déplace d'un bloc, avec la même vitesse locale de déplacement en tous les points de l'espace.

Ce résultat offre une surprise de taille : l'onde et la particule ne se déplacent pas à la même vitesse ! En effet,

$$\mathbf{v}_{\text{probabilité}} = 2\mathbf{v}_{\text{onde}}.$$

Nous verrons au paragraphe suivant que ces deux vitesses n'ont en fait pas à être égales. Voilà qui ajoute à la richesse de la dualité onde-corpuscule. En physique, ces deux vitesses portent des noms différents $\mathbf{v}_{\text{probabilité}}$ est dite *vitesse de groupe*, alors que $\mathbf{v}_{\text{onde}}$ est la *vitesse de phase*.

- (3) Quelles sont la quantité de mouvement et l'énergie cinétique de la particule ? La vitesse de la particule ayant été identifiée, nous pouvons en déduire, en nous appuyant sur la mécanique classique, sa quantité de mouvement

$$\mathbf{p} = m\mathbf{v}_{\text{probabilité}} = \hbar \mathbf{k},$$

et son énergie cinétique

$$E_c = \frac{1}{2} m \mathbf{v}_{\text{probabilité}}^2 = \frac{\hbar^2 \mathbf{k}^2}{2m} = E_{\mathbf{k}} = \hbar \omega.$$

La quantité $E_{\mathbf{k}}$ intervenant dans la définition (8.9) des ondes de De Broglie s'identifie donc à l'énergie cinétique⁷.

Les deux dernières relations ci-dessus se remettent en forme sous la forme condensée :

$$\mathbf{p} = \hbar \mathbf{k}, \quad E = \hbar \omega. \quad (8.10)$$

Elles ont été établies par *Louis De Broglie* en 1923. Elles dictent une règle assez générale qui permet, lorsqu'on ne connaît pas le détail de la fonction d'onde d'une particule, de définir les caractéristiques essentielles de son comportement ondulatoire à partir de la seule connaissance de sa vitesse. Elles s'appliquent dans toutes les situations où, de par sa préparation physique, la particule acquiert un comportement monochromatique et délocalisé. Elles relient quantité de mouvement et énergie cinétique respectivement au vecteur d'onde et à la pulsation. Remarquons le rôle central joué par la constante de Planck.

Historiquement, en 1923, l'expérience dite d'*effet Compton* venait de mettre clairement en évidence le caractère corpusculaire de la lumière, et les relations (8.10) venaient d'être identifiées comme celles donnant l'interprétation corpusculaire des ondes lumineuses (dans le sens où les photons associés à une onde de pulsation ω et de vecteur d'onde $\mathbf{k}$ ont une impulsion $\mathbf{p}$ et une énergie cinétique E définies par (8.10)). De Broglie, cherchant à transposer ce concept de dualité à la matière, reprit ces relations sur des bases purement pragmatiques.

8.5.2 Le paquet d'onde quasi-monochromatique

Ce modèle de fonction d'onde correspond à la situation dans laquelle la distribution ϕ utilisée dans la construction du paquet d'onde (8.7) est une approximation d'une distribution de Dirac (centrée en un vecteur d'onde $\mathbf{k}_0$). Il offre un moyen de résoudre la question de normalisation de l'onde de De Broglie. La fonction d'onde proposée est de la forme :

$$\psi(\mathbf{r}, t = 0) \simeq e^{i\mathbf{k}_0 \cdot \mathbf{r}} \times F(\mathbf{r}).$$

L'onde de De Broglie de vecteur $\mathbf{k}_0$ est multipliée par une fonction enveloppe $F(\mathbf{r})$, que l'on suppose très régulière, au sens où ses dérivées d'espace première et seconde sont négligeables devant les quantités $|\mathbf{k}_0|$ et $|\mathbf{k}_0|^2$ respectivement. La fonction $F(\mathbf{r})$ permet de satisfaire à la normalisation, et fait apparaître explicitement le fait que la particule ne remplit jamais en pratique tout l'espace (voir la figure 8.3).

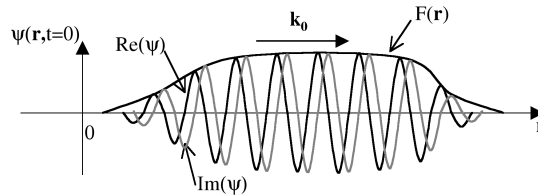


Fig. 8.3. Le paquet d'onde quasi-monochromatique d'enveloppe F et de vecteur d'onde $\mathbf{k}_0$.

Si l'on regarde $\psi(\mathbf{r}, t = 0)$ comme un paquet d'onde (8.7), la pondération $\phi(\mathbf{k})$ est, à une translation près de vecteur $\mathbf{k}_0$, la transformée de Fourier de $F(\mathbf{r})$, au moins au temps $t = 0$. On peut alors donner une expression approchée de $\psi(\mathbf{r}, t)$, en effectuant un développement limité à l'ordre 2 de l'exponentielle dans l'intégrale (8.8),

7. Nous verrons en fait dans les chapitres prochains qu'il faut interpréter $E_{\mathbf{k}}$ comme l'énergie totale de la particule

$$\begin{aligned}\psi(\mathbf{r}, t) &\simeq e^{i\left(\mathbf{k}_0 \cdot \mathbf{r} - \frac{E_{\mathbf{k}_0}}{\hbar} t\right)} \frac{1}{(2\pi)^{3/2}} \int_{\mathbb{R}^3} \phi(\mathbf{k}) e^{i(\mathbf{k} - \mathbf{k}_0) \cdot \left(\mathbf{r} - \frac{1}{\hbar} \nabla E_{\mathbf{k}}(\mathbf{k}_0) t\right)} d\mathbf{k} \\ &= \nu_{\mathbf{k}_0}(\mathbf{r}, t) \quad \times \quad F\left(\mathbf{r} - \frac{1}{\hbar} \nabla E_{\mathbf{k}}(\mathbf{k}_0) t\right).\end{aligned}$$

On observe que les composantes 'onde' $\nu_{\mathbf{k}_0}$ et 'enveloppe' F ont chacune leur comportement propre (voir la figure 8.4) :

- l'onde de De Broglie $\nu_{\mathbf{k}_0}$ se déplace à la vitesse constante,

$$\mathbf{v}_{\text{phase}} \equiv \frac{1}{\hbar} \frac{E_{\mathbf{k}_0} \mathbf{k}_0}{|\mathbf{k}_0|^2} = \frac{1}{2m} \hbar \mathbf{k}_0,$$

comme si l'enveloppe n'était pas là. Nous retrouvons ici la vitesse que nous avons appelée *vitesse de phase*;

- la fonction enveloppe F se translate dans l'espace, apparemment sans se déformer, à une vitesse définie par :

$$\mathbf{v}_{\text{groupe}} \equiv \frac{1}{\hbar} \nabla E_{\mathbf{k}}(\mathbf{k}_0) = \frac{1}{m} \hbar \mathbf{k}_0.$$

Nous retrouvons la valeur de ce que nous avons appelé *vitesse de groupe*.

Précisons d'emblée que l'absence de déformation de l'enveloppe n'est qu'apparente. Elle résulte des approximations faites dans le calcul de l'intégrale. Des calculs plus précis montreraient qu'une déformation, même lente, a toujours lieu (voir § 8.5.3 pour l'étude détaillée d'un cas particulier). En négligeant toutefois les effets de déformation de F , la probabilité de présence de la particule est

$$\rho(\mathbf{r}, t) = |F(\mathbf{r} - \mathbf{v}_{\text{groupe}} t)|^2.$$

La translation subie par la fonction F rappelle le mouvement rectiligne et uniforme de la particule classique. La vitesse de la particule est donc bien $\mathbf{v}_{\text{groupe}}$. Sous cette enveloppe, la particule exhibe néanmoins le caractère ondulatoire d'une onde de De Broglie qui se déplace à une autre vitesse. Cette caractéristique ne disparaît que lorsque la particule est immobile, et donc que $\mathbf{k}_0 = 0$. Par rapport à la mécanique classique, la nécessaire légère délocalisation de la particule et le comportement oscillant sous-jacent sont des faits expérimentaux dont seule la mécanique quantique est capable de rendre compte.

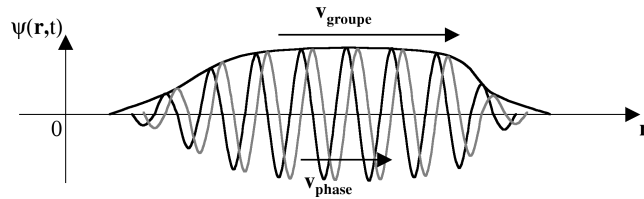


Fig. 8.4. Vitesse de phase et vitesse de groupe pour le paquet d'onde monochromatique étalé.

8.5.3 Le paquet d'ondes gaussien.

Afin d'observer de manière plus précise le phénomène de déformation de l'enveloppe d'une particule quantique (à peu près) localisée, nous traitons une situation dans laquelle tous les calculs ont le mérite de pouvoir être menés de manière exacte. Prenons

$$\psi(\mathbf{r}, t = 0) = \left(\frac{1}{2\pi\sigma_0^2}\right)^{3/4} \exp\left(-\frac{|\mathbf{r}|^2}{4\sigma_0^2}\right) \exp(i\mathbf{k}_0 \cdot \mathbf{r}).$$

Il s'agit d'une onde de De Broglie d'impulsion $\mathbf{p}_0 = \hbar \mathbf{k}_0$ contenue dans une enveloppe gaussienne. L'écriture de ψ sous sa forme 'paquet d'onde' (8.8) utilise la fonction de pondération

$$\phi(\mathbf{k}) = \left(\frac{\sigma^2}{\pi}\right)^{3/4} e^{-\sigma_0^2(\mathbf{k}-\mathbf{k}_0)^2}.$$

L'évolution temporelle s'obtient sans difficulté de principe en calculant de manière exacte l'intégrale (8.8). L'expression de $\psi(\mathbf{r}, t)$ à tous temps est cependant assez compliquée, et nous ne nous intéressons ci-dessous qu'à la probabilité de présence, pour laquelle on obtient (voir Exercice 8.5) :

$$\rho(\mathbf{r}, t) = |\psi(t, \mathbf{r})|^2 = \left(\frac{1}{2\pi\sigma^2(t)}\right)^{3/2} \exp\left(-\frac{(\mathbf{r} - \frac{\mathbf{p}_0}{m}t)^2}{2\sigma^2(t)}\right), \quad \sigma(t) = \sqrt{\sigma_0^2 + \left(\frac{\hbar}{2m\sigma_0}\right)^2 t^2}. \quad (8.11)$$

Cette expression montre que l'enveloppe de la fonction d'onde reste de forme gaussienne, et que son centre se déplace de manière rectiligne et uniforme à la vitesse de groupe $\mathbf{p}_0/m$. Nous retrouvons en cela (de manière rigoureuse !) les résultats obtenus pour le paquet d'onde quasi-monochromatique.

Pour autant, maintenant que nous disposons de résultats exacts, nous voyons que la largeur de cette gaussienne, que l'on peut typiquement prendre égale à $\sigma(t)$, augmente progressivement au cours du temps (voir la figure 8.5). Cette évolution, est lente jusqu'à des temps de l'ordre de $t_{\text{transition}} = 2m\sigma_0^2/\hbar$, puis s'accélère ensuite pour atteindre sur les temps longs un accroissement linéaire en $\sigma(t) = \hbar t/2m\sigma_0$ (voir la figure 8.6). Cette vitesse d'étalement est d'autant plus rapide que la particule est légère et initialement localisée.

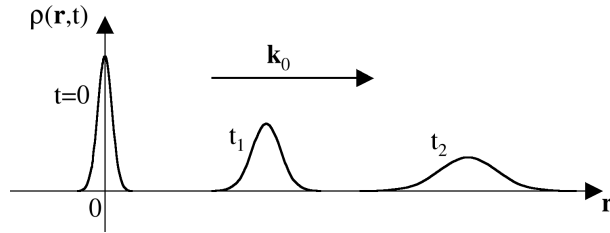


Fig. 8.5. Etalement du paquet d'onde gaussien au cours de sa propagation. La probabilité de présence $\rho(\mathbf{r}, t)$ est représentée à trois instants successifs $t = 0 < t_1 < t_2$.

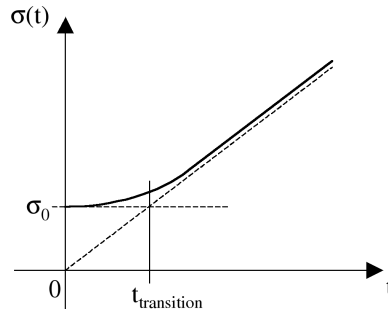


Fig. 8.6. Evolution temporelle de la largeur $\sigma(t)$ du paquet d'onde gaussien.

Ce résultat a une portée très générale en mécanique quantique :

Propriété 8.3 (Etalement du paquet d'onde). *Quelques soient les conditions initiales, les paquets d'onde des particules évoluant dans le vide ont vocation à se délocaliser de plus en plus aux temps longs. On dit qu'ils s'évalent.*

Fixons nous les idées dans l'exemple du paquet gaussien :

- considérons un électron fraîchement arraché à son atome, et que l'on suppose initialement confiné sur un domaine de l'ordre de : $\sigma_0 \sim 10^{-10}$ m. L'étalement entre dans son régime linéaire au bout de $t_{\text{transition}} = 1.8 \cdot 10^{-16}$ s. Si cet électron est accéléré dans un canon à électrons sous une tension de 1 kV, on obtient au bout de 50 cm parcourus (c'est à dire à peine $t = 20$ ns de parcours), un étalement déjà de l'ordre de 1 cm. Cela souligne la caractère extrêmement éphémère de la localisation des particules élémentaires livrées à elles-mêmes. Le caractère insaisissable de l'*électron libre* n'est pas qu'une expression de langage! Nous pouvons juger de la rapidité avec laquelle son état se rapproche d'une onde de De Broglie très délocalisée;
- faisons de même avec une petite masse de 1 g, initialement localisée sur 1 cm. Le temps $t_{\text{transition}}$ passe alors à 6×10^{19} années. L'étalement du paquet d'onde est en fait imperceptible pour les objets suffisamment massifs et/ou volumineux. L'indéformabilité du paquet d'onde quasi-monochromatique devient en particulier vraie dans la limite des très grands étalements.

La mécanique classique, par ses concepts de coordonnées d'espace et de trajectoire, fait l'hypothèse que les particules (ou centres de masse d'objets), quelques soient leur nature, conservent indéfiniment leur caractère localisé. Nous comprenons ici pourquoi cette mécanique ne peut s'appliquer de manière satisfaisante qu'à des objets suffisamment lourds, et pourquoi on doit changer de théorie physique pour décrire les protons, électrons, neutrons et autres particules élémentaires qui constituent la matière.

8.5.4 Superposition d'états

Nous abordons là une classe d'états quantiques impossibles à concevoir en mécanique classique. Ils découlent du principe suivant :

Propriété 8.4 (Principe de superposition). *Si deux fonctions d'onde $\psi_1(\mathbf{r}, t)$ et $\psi_2(\mathbf{r}, t)$ sont solutions de l'équation de Schrödinger, alors toute fonction d'onde obtenue par combinaison linéaire est encore solution :*

$$\psi(\mathbf{r}, t) = \alpha\psi_1(\mathbf{r}, t) + \beta\psi_2(\mathbf{r}, t)$$

avec $\alpha, \beta \in \mathbb{C}$ assurant la condition de normalisation $\|\psi(t, \cdot)\|_{L^2(\mathbb{R}^3)} = 1$.

Cette propriété est une conséquence directe de la linéarité de l'équation de Schrödinger. La possibilité pratique de trouver des coefficients α et β constants qui permettent de satisfaire en permanence à la condition de normalisation découle de la compatibilité entre l'équation de Schrödinger et la normalisation de la fonction d'onde : il suffit que α et β soient choisis de sorte que $\psi(\mathbf{r}, t)$ soit normalisée à $t = 0$ pour qu'elle le soit tout le temps.

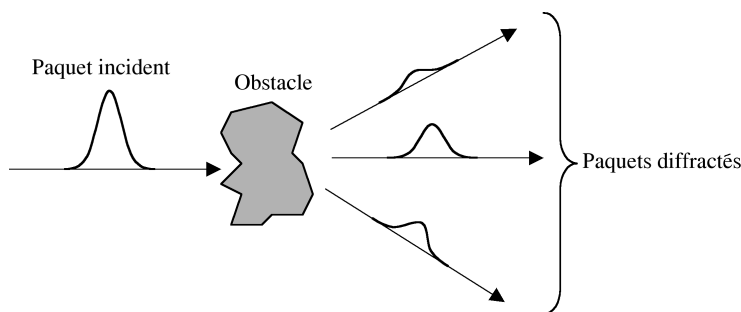


Fig. 8.7. 'Eclatement' d'un paquet d'onde en une superposition de trois paquets indépendants lors de sa diffraction sur un obstacle. Après l'impact, la particule se trouve *simultanément* répartie dans les trois paquets diffractés.

On peut donc avoir des états quantiques qui sont des superpositions de plusieurs paquets d'onde indépendants, par exemple deux paquets quasi-monochromatiques :

$$\psi(\mathbf{r}, t) = \alpha e^{i(\mathbf{k}_1 \cdot \mathbf{r} - \frac{E_{\mathbf{k}_1}}{\hbar} t)} F_1\left(\mathbf{r} - \frac{\hbar \mathbf{k}_1}{m} t\right) + \beta e^{i(\mathbf{k}_2 \cdot \mathbf{r} - \frac{E_{\mathbf{k}_2}}{\hbar} t)} F_2\left(\mathbf{r} - \frac{\hbar \mathbf{k}_2}{m} t\right).$$

La particule est alors partagée entre deux paquets qui ont chacun leur propre fonction enveloppe et leur propre vitesse de déplacement. D'une certaine manière, la particule est comme scindée en deux portions indépendantes.

De telles situations sont courantes en pratique, lorsque qu'une particule vient à heurter un obstacle (voir la figure 8.7). Le caractère ondulatoire de la fonction d'onde donne alors souvent lieu à un phénomène de diffraction : la particule se partage en plusieurs morceaux, qui repartent chacun dans des directions différentes.

Dans un tel état superposé, certaines questions perdent leur sens :

- quelle est l'impulsion de la particule ? On sent que la réponse doit faire intervenir, d'une certaine manière, les impulsions $\hbar \mathbf{k}_1$ et $\hbar \mathbf{k}_2$, mais impossible d'en dire plus⁸, car la particule possède ces deux impulsions *simultanément*;
- quelle est son énergie ? Là aussi, nous sommes face à un subtil mélange entre $E_{\mathbf{k}_1} = \frac{\hbar^2 \mathbf{k}_1^2}{2m}$ et $E_{\mathbf{k}_2} = \frac{\hbar^2 \mathbf{k}_2^2}{2m}$.

La suite du cours permettra, peu à peu, de mieux appréhender ce type de situation.

Il peut enfin arriver que deux fractions indépendantes de la fonction d'onde d'une même particule viennent au cours de leur dynamique à se rencontrer. Elles donnent lieu, comme en optique, à des phénomènes d'*interférences*, dans lesquels la probabilité de présence de la particule acquiert des formes très singulières.

8. Au moins à ce stade du cours !

Exercices

Exercice 8.1 (Longueur d'onde de de Broglie).

Calculer la longueur d'onde de de Broglie dans les cas suivants :

- (1) Un électron d'énergie 1 eV.
- (2) Un proton d'énergie 1 MeV.
- (3) Le centre de gravité d'une masse de 1 kg se déplaçant à 1 m/s.

Commenter ces résultats, en considérant la longueur d'onde comme la précision avec laquelle on peut 'raisonnablement' localiser ces particules.

Exercice 8.2 (Exploration du noyau atomique).

Une expérience d'exploration du noyau consiste à bombarder une feuille d'or par des électrons ayant une énergie d'une centaine de MeV. Montrer que c'est bien dans cette gamme d'énergie que les électrons peuvent effectivement être diffractés par un noyau d'or. Pour l'or, le nombre de charge vaut $Z = 79$ et le nombre de masse vaut $A = 197$.

On indique que la taille des noyaux est $r = A^{1/3}r_0$ avec $r_0 = 1.2$ fm.

Exercice 8.3 (Sensibilité de l'oeil humain).

En 1917, Fabry et Buisson mesurent la limite de sensibilité de l'oeil et montrent que le plus faible point lumineux perceptible dans l'obscurité totale est une bougie située à 27 km.

- (1) La puissance de l'énergie lumineuse visible émise par une bougie est de l'ordre de 0.01 W. Quelle est la puissance collectée par la pupille de l'oeil d'un observateur situé à 27 km de celle-ci ? (*On assimilera la pupille à un trou de 5 mm de diamètre*)
- (2) La longueur d'onde typique de la lumière visible est de l'ordre de $\lambda = 500$ nm. Quelle est, en fonction de λ , de la constante de Planck $h = 2\pi\hbar$ et la vitesse de la lumière c l'énergie E_λ portée par un photon visible ? Faire l'application numérique.
- (3) Pour décider ou non d'envoyer au cerveau l'influx nerveux annonçant la détection d'un point lumineux, l'oeil 'intègre' le nombre de photons reçus en 100 ms. Dédurre des résultats qui précèdent l'ordre de grandeur du nombre de photons nécessaires pour déclencher l'influx.

Exercice 8.4 (Interférence quantique).

On considère, dans une zone Z de l'espace, une particule dont l'état est localement donné par la superposition de deux ondes de De Broglie :

$$\psi(\mathbf{r}, t) = \alpha \exp \left(i \left(\mathbf{k}_1 \cdot \mathbf{r} - \frac{E_{\mathbf{k}_1}}{\hbar} t \right) \right) + \beta \exp \left(i \left(\mathbf{k}_2 \cdot \mathbf{r} - \frac{E_{\mathbf{k}_2}}{\hbar} t \right) \right).$$

En dépit du problème de normalisation des ondes planes, nous considérons que la norme de $\psi(\mathbf{r}, t)$ ainsi définie reste physiquement interprétable dans la zone Z .

- (1) Calculer, pour les points $\mathbf{r}$ intérieurs à la zone Z , la densité de probabilité de présence $\rho(\mathbf{r}, t)$ de la particule.
- (2) Montrer que $\rho(\mathbf{r}, t)$ possède un caractère ondulatoire dont on précisera le vecteur d'onde et la pulsation.

Exercice 8.5 (Paquet d'ondes gaussien).

L'objectif de cet exercice est de montrer la formule (8.11) dans le cas où l'espace ambiant est $\mathbb{R}$. Initialement, on considère la fonction d'onde

$$\psi(0, x) = \psi_0(x) = \left(\frac{1}{2\pi\sigma_0^2} \right)^{1/4} e^{-x^2/4\sigma_0^2} e^{ik_0x}.$$

- (1) Montrer que $\|\psi(0, \cdot)\|_{L^2(\mathbb{R})} = 1$.

(2) Donner une équation différentielle ordinaire satisfaite par

$$\phi_0(k) = \int_{\mathbb{R}} \psi_0(x) e^{-ikx} dx,$$

et en déduire la transformée de Fourier de ψ_0 .

(3) Donner, en adaptant (8.7), la forme générale de $\psi(t, x)$.

(4) Montrer que $|\rho(t, x)|$ est de forme gaussienne et donner sa variance $\sigma^2(t)$.

Exercice 8.6 (Étalement du paquet d'onde).

Un atome est mis au contact d'une masse métallique de taille caractéristique $a = 1$ cm. On suppose qu'un des électrons de l'atome, initialement maintenu dans un volume d'une longueur caractéristique de 10^{-10} m, peut alors s'échapper et voir sa fonction d'onde envahir l'ensemble du métal. En utilisant les résultats sur le paquet d'onde gaussien, estimer le temps caractéristique nécessaire à cet étalement.

La particule quantique en interaction avec un potentiel

Nous exposons ici comment la mécanique quantique rend compte des interactions d'une particule avec des forces dérivant d'un potentiel. La théorie quantique permet, dans certains cas particuliers, de retrouver la relation fondamentale de la dynamique de la mécanique classique. Nous décrivons à l'inverse des situations dans lesquelles la particule quantique défie les lois classiques, et joue aux passes-murailles.

9.1	Dynamique de la particule en interaction	131
9.1.1	L'équation de Schrödinger avec potentiel	131
9.1.2	Caractère bien posé de l'équation de Schrödinger	132
9.1.3	Régularité de la fonction d'onde	133
9.1.4	Le terme potentiel et la normalisation	134
9.2	Le paquet d'onde en interaction : la limite classique	135
9.3	Etats stationnaires de la particule quantique	137
9.3.1	Que sont les états stationnaires ?	137
9.3.2	L'équation de Schrödinger indépendante du temps	137
9.3.3	Energie de la particule	138
9.3.4	Régularité des états stationnaires unidimensionnels	138
9.3.5	Minoration de l'énergie des états stationnaires	139
9.3.6	Importance des états stationnaires en mécanique quantique	139
9.4	Une situation purement quantique : la barrière de potentiel	139
9.4.1	La barrière de potentiel	140
9.4.2	Etats stationnaires de la barrière de potentiel	140
9.4.3	Effet tunnel et transmission résonnante	142
9.4.4	Et la limite classique ?	145
9.5	Implémentation numérique	146
9.5.1	Schéma de Crank-Nicholson	146
9.5.2	Décomposition d'opérateurs	148
	Exercices	149

9.1 Dynamique de la particule en interaction

9.1.1 L'équation de Schrödinger avec potentiel

En mécanique quantique, la particule soumise à des forces est encore décrite par une fonction d'onde $\psi(\mathbf{r}, t)$. En particulier, l'interprétation de $\rho(\mathbf{r}, t) = |\psi(\mathbf{r}, t)|^2$ en terme de densité de probabilité de présence reste inchangée. L'équation de Schrödinger est cependant modifiée.

Définition 9.1. *L'interaction d'une particule avec des forces extérieures dérivant d'un potentiel $V(\mathbf{r}, t)$ est décrite par l'équation de Schrödinger avec potentiel :*

$$i\hbar \frac{\partial \psi}{\partial t} = -\frac{\hbar^2}{2m} \Delta \psi(\mathbf{r}, t) + V(\mathbf{r}, t) \psi(\mathbf{r}, t). \quad (9.1)$$

Rappelons qu'une force $\mathbf{F}$ est dite *dérivée d'un potentiel* si $\mathbf{F} \equiv \mathbf{F}(\mathbf{r}, t)$ et s'il existe un champ scalaire à valeur réelles $V(\mathbf{r}, t)$ tel que

$$\mathbf{F}(\mathbf{r}, t) = -\nabla V(\mathbf{r}, t).$$

On peut citer par exemple :

- (1) la *force de gravitation* : pour une particule attirée par un corps de masse M dont le centre de gravité est pris comme origine de l'espace, la force et le potentiel sont respectivement :

$$\mathbf{F}_{\text{gravitation}}(\mathbf{r}) = -\mathcal{G} \frac{Mm}{r^2} \mathbf{e}_r, \quad V_{\text{gravitation}}(\mathbf{r}) = -G \frac{Mm}{r},$$

où $\mathcal{G} = 6.67 \times 10^{-11} \text{ m}^3 \text{ kg}^{-1} \text{ s}^{-2}$ est la constante de gravitation ;

- (2) la *force de Coulomb* (en électrostatique) : dans le cas particulier d'une particule de charge électrique q se déplaçant dans un champ électrique statique $\mathbf{E}(\mathbf{r})$,

$$\mathbf{F}_{\text{Coulomb}}(\mathbf{r}) = q\mathbf{E}(\mathbf{r}), \quad V_{\text{Coulomb}}(\mathbf{r}) = qU(\mathbf{r}),$$

où $U(\mathbf{r})$ est le potentiel électrique, défini par la relation $\mathbf{E}(\mathbf{r}) = -\nabla U(\mathbf{r})$.

Pour toutes ces forces, le champ de potentiel qui intervient dans l'équation de Schrödinger s'identifie au champ de potentiel classique.

Citons toutefois un contre-exemple : la *force de Lorentz*

$$\mathbf{F}_{\text{Lorentz}} = q\mathbf{v}(t) \wedge \mathbf{B}(\mathbf{r}, t).$$

Cette force est ressentie par une particule classique chargée qui se déplace à la vitesse $\mathbf{v}$ dans un champ magnétique $\mathbf{B}$. Elle ne peut pas se mettre explicitement sous la forme d'une fonction de la seule position $\mathbf{r}$ de la particule. Elle peut donc encore moins s'écrire comme le gradient d'un champ de potentiel, et on ne peut pas utiliser l'équation de Schrödinger (9.1) dans ce cas.¹

9.1.2 Caractère bien posé de l'équation de Schrödinger

On se place dans le cas où le potentiel extérieur V ne dépend pas du temps. On introduit l'opérateur suivant, noté $\hat{H}$, que nous nommons *opérateur Hamiltonien* du système

$$\hat{H} = -\frac{\hbar^2}{2m} \Delta + V(\mathbf{r}). \quad (9.2)$$

L'opérateur $\hat{H}$ est la somme de deux opérateurs symétriques, l'opérateur de multiplication par V et $\hat{H}_0 = -\frac{\hbar^2}{2m} \Delta$ (voir le § 8.4.1 pour le caractère symétrique de ce dernier opérateur). En général, $\hat{H}$ est également auto-adjoint², mais cette propriété est à vérifier au préalable dans tous les cas. Des indications sur la manière de mener cette vérification seront données au chapitre 10.

L'équation de Schrödinger est bien posée au sens mathématique du terme dès que l'opérateur Hamiltonien est autoadjoint :

1. Il existe en fait une troisième forme de l'équation de Schrödinger, qui permet de traiter l'interaction d'une particule chargée avec un champ électromagnétique quelconque. Elle fait intervenir, en plus du potentiel Coulombien $U(\mathbf{r}, t)$, le potentiel vecteur $\mathbf{A}(\mathbf{r}, t)$. Le lecteur intéressé pourra trouver ces développements dans tout ouvrage de référence en mécanique quantique.

2. Rappelons qu'un opérateur auto-adjoint est symétrique, mais pour qu'un opérateur symétrique soit auto-adjoint, il faut vérifier des conditions sur le domaine de l'adjoint (voir § C.1).

Théorème 9.1 (Stone). Soit $\hat{H}$ un opérateur auto-adjoint non-borné sur un espace de Hilbert $\mathcal{H}$. Alors $-i\hat{H}$ est le générateur infinitésimal d'un groupe fortement continu $(\hat{U}(t))_{t \in \mathbb{R}}$ d'opérateurs sur $\mathcal{H}$, c'est-à-dire

- $\hat{U}(0) = \text{Id}$, $\hat{U}(t_1 + t_2) = \hat{U}(t_1)\hat{U}(t_2)$ pour tout $t_1, t_2 \in \mathbb{R}$ (propriété de groupe);
- pour tout $u_0 \in \mathcal{H}$ et $t_0 \in \mathbb{R}$, $\lim_{t \rightarrow t_0} \|\hat{U}(t)u_0 - \hat{U}(t_0)u_0\|_{\mathcal{H}} = 0$ (continuité forte);
- pour tout $t \in \mathbb{R}$ et $u_0 \in \mathcal{H}$, $\|\hat{U}(t)u_0\|_{\mathcal{H}} = \|u_0\|_{\mathcal{H}}$ (unitarité).

En conséquence, (9.1) admet une unique solution $\psi \in C^1(\mathbb{R}, \mathcal{H}) \cap C^0(\mathbb{R}, D(\hat{H}))$ pour $\psi_0 \in D(\hat{H})$.

On rappelle que le générateur infinitésimal d'un groupe³ $(\hat{S}(t))_{t \in \mathbb{R}}$ est l'opérateur (non-borné) $\hat{A}$ de domaine $D(\hat{A}) \subset \mathcal{H}$, les éléments $u \in D(\hat{A})$ étant caractérisés par le fait que l'application $t \mapsto \hat{S}(t)u$ soit dérivable en 0, avec

$$\hat{A}u = \left. \frac{d}{dt} (\hat{S}(t)u) \right|_{t=0}.$$

On utilise la notation formelle $\hat{S}(t) = e^{t\hat{A}}$ dans ce cas.

Ainsi, on peut définir la solution de (9.1) pour des données initiales suffisamment régulières. L'évolution est donnée par l'application d'un propagateur $\hat{U}(t) = e^{-it\hat{H}}$. En effet, si la condition initiale est suffisamment régulière ($\psi_0 \in D(\hat{H})$), alors $\psi(t) = \hat{U}(t)\psi_0$ vérifie l'équation de Schrödinger (9.1). De plus, grâce au propagateur $\hat{U}(t)$, on va pouvoir définir formellement des évolutions physiques pour des données initiales moins régulières : on définit, par densité, la quantité $\hat{U}(t)\psi_0$ pour $\psi_0 \in \mathcal{H}$. En effet, pour $\psi_0 \in \mathcal{H}$, il existe une suite $(\psi_0^n)_{n \in \mathbb{N}}$ telle que $\|\psi_0 - \psi_0^n\|_{\mathcal{H}} \rightarrow 0$ lorsque $n \rightarrow +\infty$. On définit alors $\psi(t, \cdot) = \lim_{n \rightarrow +\infty} \hat{U}(t)\psi_0^n$ en utilisant la propriété de continuité forte du groupe $(\hat{U}(t))_{t \in \mathbb{R}}$.

9.1.3 Régularité de la fonction d'onde

Connaître *a priori* la régularité de la fonction d'onde et de ses dérivées se révèle très utile dans l'étude explicite de solutions. On se place dans le cas de potentiels $V(\mathbf{r})$ ne dépendant pas du temps. On note

$$\Omega = \{\mathbf{r} \in \mathbb{R}^3 \mid V(\mathbf{r}) < +\infty\}.$$

On suppose que Ω est ouvert, et que $V \in L^2(\Omega) + L^\infty(\Omega)$ par exemple (afin que $\hat{H}$ soit auto-adjoint, voir la Remarque 10.2). On considère la quantité suivante, que nous verrons plus tard être l'énergie moyenne du système :

$$E = \int_{\mathbb{R}^3} \hat{H}\psi \cdot \bar{\psi} \, d\mathbf{r} = \frac{\hbar^2}{2m} \int_{\mathbb{R}^3} |\nabla \psi(t, \mathbf{r})|^2 \, d\mathbf{r} + \int_{\mathbb{R}^3} V(\mathbf{r}) |\psi(\mathbf{r}, t)|^2 \, d\mathbf{r}. \quad (9.3)$$

On admet que le premier terme représente l'énergie cinétique, et le second l'énergie potentielle de la particule. Pour des raisons physiques, ces deux quantités ne peuvent être infinies. Ceci montre alors que

(1) le support de ψ est contenu dans Ω . En effet

$$\int_{\mathbb{R}^3} V(\mathbf{r}) |\psi(\mathbf{r}, t)|^2 \, d\mathbf{r} < +\infty.$$

Ceci montre que $\psi = 0$ presque partout sur le complémentaire de Ω ;

(2) de plus,

$$\int_{\mathbb{R}^3} |\nabla \psi(t, \mathbf{r})|^2 \, d\mathbf{r} < +\infty.$$

montre que $\psi \in H^1(\omega)$ pour tout domaine $\omega \subset \Omega$ borné;

3. En fait, cette définition peut être un peu modifiée et avoir un sens dès qu'on a affaire à un semi-groupe.

On résume ces propriétés en disant que

Propriété 9.1. *Si l'énergie du système est finie (i.e. (9.3) est vérifiée), alors $\psi \in H^1(\Omega)$, avec $\Omega = \{\mathbf{r} \in \mathbb{R}^3 \mid V(\mathbf{r}) < +\infty\}$.*

Remarque 9.1. Dans des cas particuliers, on peut montrer une régularité plus grande (voir § 9.3.2).

Remarque 9.2 (Continuité de ψ en dimension 1). Dans le cas particulier de la dimension 1 d'espace (que nous utiliserons souvent dans les exercices), on a une injection (compacte) $H^1(\omega) \hookrightarrow C^0(\omega)$ pour des domaines ω bornés. On en déduit que ψ est $C^0(\Omega)$. Ces arguments ne peuvent être transposés directement en dimensions d'espace supérieures, mais doivent être adaptés au cas par cas en fonction des potentiels considérés (en utilisant typiquement des injections de Sobolev).

9.1.4 Le terme potentiel et la normalisation

La conservation de la normalisation de ψ au cours du temps est encore assurée pour l'équation de Schrödinger générale (9.1). On a les résultats suivants.

Proposition 9.1. *Pour tout potentiel $V \equiv V(\mathbf{r})$ tel que $\hat{H}$ soit auto-adjoint, l'équation de Schrödinger (9.1) est compatible avec la relation de normalisation, au sens où si $\|\psi(0, \cdot)\|_{L^2(\mathbb{R}^3)} = 1$, alors, pour tout $t \in \mathbb{R}$, $\|\psi(t, \cdot)\|_{L^2(\mathbb{R}^3)} = 1$.*

On énonce ici cette propriété de manière explicite, et on en fait une démonstration ; cependant, elle était déjà contenue dans l'énoncé du Théorème 9.1 (unitarité). Formellement, cette propriété (et le calcul qui suit) se généralise au cas où V dépend du temps (cependant, dans ce cas, le caractère bien posé de l'équation de Schrödinger (9.1) n'est pas clair).

Preuve. On peut faire une preuve explicite pour une donnée initiale suffisamment régulière. Pour un état initial $\psi_0 \in D(\hat{H})$, ψ solution de (9.1) est à valeurs dans $D(\hat{H})$ par le Théorème 9.1. Or, on a

$$\begin{aligned} \frac{d}{dt} \int_{\mathbb{R}^3} |\psi(x, t)|^2 dx &= \langle \partial_t \psi, \psi \rangle + \langle \psi, \partial_t \psi \rangle = \left\langle \frac{1}{i\hbar} \hat{H} \psi, \psi \right\rangle + \left\langle \psi, \frac{1}{i\hbar} \hat{H} \psi \right\rangle \\ &= \frac{1}{i\hbar} \left(\langle \hat{H} \psi, \psi \rangle - \langle \psi, \hat{H} \psi \rangle \right) = 0 \end{aligned}$$

si $\hat{H}$ est symétrique. Ceci montre que

$$\forall t \in \mathbb{R}, \quad \int_{\mathbb{R}^3} |\psi(x, t)|^2 dx = 1.$$

Dans le cas général d'une donnée simplement $L^2(\mathbb{R}^3)$, ce résultat découle directement de l'unitarité du propagateur (voir le Théorème 9.1). $\square$

Comme au chapitre précédent, la conservation de la normalisation des fonctions d'onde peut aussi se voir par le biais des formules de l'hydrodynamique quantique.

Propriété 9.2. *On considère une particule quantique décrite par la fonction d'onde $\psi(\mathbf{r}, t) = \sqrt{\rho(\mathbf{r}, t)} e^{i\theta(\mathbf{r}, t)}$. Si ψ satisfait à l'équation de Schrödinger avec potentiel (9.1), alors la probabilité de présence $\rho(\mathbf{r}, t)$, le courant de probabilité $\mathbf{J}(\mathbf{r}, t)$, et la vitesse de probabilité $\mathbf{v}(\mathbf{r}, t)$ donnés respectivement par*

$$\rho(\mathbf{r}, t) = |\psi(\mathbf{r}, t)|^2, \quad \mathbf{J}(\mathbf{r}, t) = -i \frac{\hbar}{2m} (\bar{\psi} \nabla \psi - \psi \nabla \bar{\psi}), \quad \mathbf{v}(\mathbf{r}, t) = \frac{\hbar}{m} \nabla \theta,$$

vérifient encore les relations de conservation locale de la probabilité de présence :

$$\mathbf{J}(\mathbf{r}, t) = \rho(\mathbf{r}, t) \mathbf{v}(\mathbf{r}, t)$$

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \mathbf{J} = \frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{v}) = 0$$

Ces équations forment un système hyperbolique de lois de conservation.

Le résultat de conservation locale se démontre de la même manière qu'au chapitre précédent, en tenant compte du terme potentiel supplémentaire dans $\frac{\partial \rho}{\partial t}$. On a formellement :

$$\frac{\partial \psi}{\partial t} = \frac{i\hbar}{2m} \Delta \psi - \frac{i}{\hbar} V \psi, \quad \frac{\partial \bar{\psi}}{\partial t} = -\frac{i\hbar}{2m} \Delta \bar{\psi} + \frac{i}{\hbar} V \bar{\psi},$$

d'où

$$\frac{\partial \rho}{\partial t} = \bar{\psi} \frac{\partial \psi}{\partial t} + \frac{\partial \bar{\psi}}{\partial t} \psi = \bar{\psi} \left(\frac{i\hbar}{2m} \Delta \psi - \frac{i}{\hbar} V \psi \right) + \psi \left(-\frac{i\hbar}{2m} \Delta \bar{\psi} + \frac{i}{\hbar} V \bar{\psi} \right) = \frac{i\hbar}{2m} (\bar{\psi} \Delta \psi - \psi \Delta \bar{\psi}).$$

Les termes potentiels se simplifient, et ne modifient finalement pas le résultat de ce calcul. Les propriétés démontrées en sections 8.2.2 et 8.3 restent donc encore vraies.

La dérivée temporelle de la probabilité de présence ρ dépend donc uniquement de la fonction d'onde, c'est à dire de l'état du système à l'instant considéré. En particulier, elle ne dépend pas du terme potentiel de l'équation de Schrödinger (9.1), c'est à dire des sollicitations extérieures appliquées au système. Ce résultat est à rapprocher de la mécanique classique, pour laquelle la dynamique de Newton s'écrit

$$\begin{cases} \frac{d\mathbf{r}}{dt} = \frac{\mathbf{p}}{m}, \\ \frac{d\mathbf{p}}{dt} = \sum_i \mathbf{F}_i. \end{cases}$$

On constate, ici aussi, que la présence des forces n'affecte pas l'expression de la dérivée temporelle du vecteur position. Cette dérivée dépend exclusivement du vecteur vitesse $\mathbf{v} = \frac{\mathbf{p}}{m}$, qui est, avec $\mathbf{r}$, une des informations décrivant l'état du système.

9.2 Le paquet d'onde en interaction : la limite classique

Dans le cas de l'équation de Schrödinger (8.2) pour la particule libre, les paquets d'onde quasi-monochromatiques se déplacent de manière rectiligne et uniforme (voir aux § 8.5.2 et 8.5.3). Qu'en est-il dans le cas de l'équation de Schrödinger avec potentiel (9.1) ?

Considérons une particule quantique de masse m , de fonction d'onde $\psi(\mathbf{r}, t)$, soumise à un potentiel $V(\mathbf{r}, t)$ pouvant dépendre du temps. Etant donné l'interprétation de $\rho(\mathbf{r}, t) d\mathbf{r} = |\psi(\mathbf{r}, t)|^2 d\mathbf{r}$ en tant que probabilité de présence de la particule, on peut définir pour cette particule une position moyenne

$$\mathbf{R}(t) = \int_{\mathbb{R}^3} \mathbf{r} |\psi(\mathbf{r}, t)|^2 d\mathbf{r}. \quad (9.4)$$

On montre alors que :

Proposition 9.2. *La vitesse de déplacement et l'accélération du point moyen $\mathbf{R}(t)$ du paquet d'onde sont respectivement données par :*

$$\frac{d\mathbf{R}}{dt} = \int_{\mathbb{R}^3} \mathbf{v}(\mathbf{r}, t) |\psi(\mathbf{r}, t)|^2 d\mathbf{r} = \frac{1}{m} \frac{\hbar}{i} \int_{\mathbb{R}^3} \bar{\psi}(\mathbf{r}, t) \cdot \nabla \psi(\mathbf{r}, t) d\mathbf{r}, \quad (9.5)$$

$$m \frac{d^2 \mathbf{R}}{dt^2} = - \int_{\mathbb{R}^3} \nabla V(\mathbf{r}, t) \cdot |\psi(\mathbf{r}, t)|^2 d\mathbf{r}. \quad (9.6)$$

Preuve. On présente ici une démonstration formelle, se limitant à des fonctions régulières pour lesquelles les intégrations par parties ne posent pas problème, par exemple $C_0^\infty(\mathbb{R}^3)$. On en déduit par densité que (9.5) et (9.6) sont valables lorsque l'énergie totale est finie.

Pour obtenir (9.5), on utilise la relation de conservation de la masse $\partial_t \rho + \nabla(\rho \mathbf{v}) = 0$. Il vient alors :

$$\frac{d\mathbf{R}}{dt} = \int_{\mathbb{R}^3} \mathbf{r} \frac{\partial \rho}{\partial t} d\mathbf{r} = - \int_{\mathbb{R}^3} \nabla(\rho \mathbf{v}) \mathbf{r} d\mathbf{r} = \int_{\mathbb{R}^3} \mathbf{v}(\mathbf{r}, t) \rho(\mathbf{r}, t) d\mathbf{r}.$$

La deuxième expression proposée pour $\frac{d\mathbf{R}}{dt}$ se démontre en réintroduisant le vecteur courant de probabilité $\mathbf{J}(\mathbf{r}, t)$:

$$\frac{d\mathbf{R}}{dt} = \int_{\mathbb{R}^3} \mathbf{J}(\mathbf{r}, t) d\mathbf{r} = -i \frac{\hbar}{2m} \int_{\mathbb{R}^3} (\bar{\psi} \nabla \psi - \psi \nabla \bar{\psi}) d\mathbf{r} = \frac{1}{m} \frac{\hbar}{i} \int_{\mathbb{R}^3} \psi \nabla \bar{\psi} d\mathbf{r}.$$

La dérivée seconde de $\mathbf{R}$ s'obtient ensuite selon :

$$m \frac{d^2 \mathbf{R}}{dt^2} = m \frac{d}{dt} \left(\frac{d\mathbf{R}}{dt} \right) = \frac{\hbar}{i} \int_{\mathbb{R}^3} \left(\frac{\partial \bar{\psi}}{\partial t} \nabla \psi + \bar{\psi} \frac{\partial (\nabla \psi)}{\partial t} \right) d\mathbf{r}.$$

Or, l'équation de Schrödinger (9.1) donne

$$\frac{\partial \bar{\psi}}{\partial t} = -\frac{i\hbar}{2m} \Delta \bar{\psi} + \frac{i}{\hbar} V \bar{\psi}, \quad \frac{\partial (\nabla \psi)}{\partial t} = \frac{i\hbar}{2m} \nabla (\Delta \psi) - \frac{i}{\hbar} V \cdot \nabla \psi - \frac{i}{\hbar} \nabla V \cdot \psi.$$

Il vient finalement :

$$m \frac{d^2 \mathbf{R}}{dt^2} = -\frac{\hbar^2}{2m} \int_{\mathbb{R}^3} (\Delta \bar{\psi} \cdot \nabla \psi - \bar{\psi} \cdot \nabla (\Delta \psi)) d\mathbf{r} - \int_{\mathbb{R}^3} \nabla V \bar{\psi} \psi d\mathbf{r}.$$

Une dernière intégration par parties montre que la première intégrale est nulle. $\square$

Ces résultats ont une interprétation particulièrement intuitive. L'impulsion moyenne $m \frac{d\mathbf{R}}{dt}$ n'est autre que la moyenne des impulsions locales $m\mathbf{v}(\mathbf{r}, t)$, pondérées au prorata de la probabilité de présence $\rho(\mathbf{r}, t)$. De même, la force responsable de l'accélération $\frac{d^2 \mathbf{R}}{dt^2}$ s'obtient en moyennant le champ de forces classique $\mathbf{F}(\mathbf{r}, t) = -\nabla V(\mathbf{r}, t)$. L'expression de $m \frac{d\mathbf{R}}{dt}$ directement en fonction de ψ et de son gradient peut quant à elle ne paraître à ce stade du cours qu'un tour de passe-passe mathématique. Son utilité se révélera par la suite (voir le chapitre 11).

Limite classique

Si la fonction d'onde est très piquée autour du point $\mathbf{R}$ (au sens où le paquet d'onde a une extension très limitée devant l'échelle caractéristique d'évolution du champ de force, voir figure 9.1), l'accélération de la position moyenne $\mathbf{R}(t)$ du paquet d'onde est donnée par :

$$m \frac{d^2 \mathbf{R}}{dt^2} = -\nabla V(\mathbf{R}(t), t). \quad (9.7)$$

Ce résultat s'obtient rigoureusement dans la limite où $|\psi(\mathbf{r}, t)|^2 \rightarrow \delta_{R_0(t)}$ au sens des distributions, auquel cas $R(t) \rightarrow R_0(t)$ et $\int_{\mathbb{R}^3} \nabla V(\mathbf{r}, t) \cdot |\psi(\mathbf{r}, t)|^2 d\mathbf{r} \rightarrow \nabla V(R_0(t), t)$.

Si la particule quantique possède à un moment donné l'aspect ponctuel de la particule classique, alors son vecteur position obéit à ce même instant à la *relation fondamentale de la dynamique*. C'est là un résultat très important, qui prouve que la mécanique classique est contenue dans la mécanique quantique. La principale différence de fond entre les deux théories est que la mécanique quantique se montre capable de décrire un plus grand nombre de situations physiques, celles notamment où le caractère 'étalé' ou 'ondulatoire' de la particule prend de l'importance. Réciproquement, la mécanique classique doit être vue comme une approximation, un comportement asymptotique de la mécanique quantique dans les quelques situations simples où la particule est suffisamment localisée pour pouvoir être considérée comme ponctuelle.

En pratique, le phénomène d'étalement du paquet d'onde fait qu'une particule perd toujours à terme son aspect localisé, et ce d'autant plus vite qu'elle est légère. En conséquence, les corps lourds sont plus enclins à se laisser décrire par la mécanique classique⁴ sur des périodes longues (typiquement plusieurs milliards de milliards d'années en ce qui concerne les objets de notre quotidien), que les électrons (quelques 10^{-16} secondes tout au plus).

4. Mais attention, pas de manière systématique ! Voir la suite de ce chapitre...

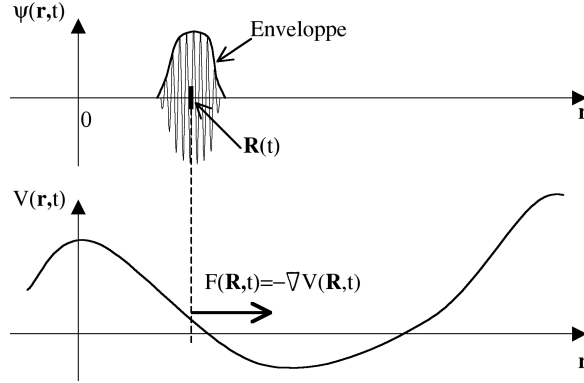


Fig. 9.1. La limite classique : lorsque le paquet d'onde est très piqué, l'accélération de sa position moyenne $\mathbf{R}(t)$ s'identifie, au terme de masse près, à la dérivée locale du champ de potentiel.

9.3 Etats stationnaires de la particule quantique

Afin de nous armer pour l'examen de situations dans lesquelles la particule exhibe pleinement son comportement quantique, nous étudions en préliminaire une classe particulière de solutions de l'équation de Schrödinger : les *solutions stationnaires*. Elles sont définies pour des champs de potentiel ne dépendant pas du temps.

9.3.1 Que sont les états stationnaires ?

Définition 9.2. Pour une particule évoluant dans un champ de potentiel $V(\mathbf{r})$ indépendant du temps, on nomme 'états stationnaires' les solutions de l'équation de Schrödinger qui se factorisent sous la forme :

$$\psi(\mathbf{r}, t) = \phi(\mathbf{r})\varphi(t),$$

avec les conditions de normalisation

$$\int_{\mathbb{R}^3} |\phi(\mathbf{r})|^2 d\mathbf{r} = 1, \quad \forall t \in \mathbb{R} \quad |\varphi(t)| = 1.$$

Les fonctions $\phi(\mathbf{r})$ et $\varphi(t) = e^{-i\alpha(t)}$ sont des fonctions à valeurs complexes, nommées respectivement parties spatiale et temporelle de l'état stationnaire.

Du point de vue physique, ces états stationnaires décrivent des situations dans lesquelles la probabilité de présence $\rho(\mathbf{r}, t) = |\phi(\mathbf{r})|^2$ reste constante au cours du temps. Cela ne signifie cependant pas que la particule reste immobile. Par exemple, dans le cas des ondes de De Broglie, la fonction d'onde $e^{i(\mathbf{k}\cdot\mathbf{r} - \frac{E_{\mathbf{k}}}{\hbar}t)}$ est associée à un champ uniforme de vitesse de probabilité $\mathbf{v} = \frac{\hbar\mathbf{k}}{m}$ non nul.

9.3.2 L'équation de Schrödinger indépendante du temps

Le résultat qui suit est capital dans la recherche et le calcul des états stationnaires.

Proposition 9.3 (Equation de Schrödinger indépendante du temps). Pour une particule de masse m évoluant dans un champ de potentiel $V(\mathbf{r})$ indépendant du temps, $\phi(\mathbf{r})$ et $\varphi(t)$ sont les parties spatiale et temporelle d'un état stationnaire, si et seulement si $\phi(\mathbf{r})$ vérifie la condition de normalisation $\|\phi\|_{L^2(\mathbb{R}^3)} = 1$, et s'il existe une constante $E \in \mathbb{R}$ homogène à une énergie telle que

$$-\frac{\hbar^2}{2m}\Delta\phi(\mathbf{r}) + V(\mathbf{r})\phi(\mathbf{r}) = E\phi(\mathbf{r}), \quad (9.8)$$

et

$$\varphi(t) = e^{-i(\frac{E}{\hbar}t + \theta)}, \quad (9.9)$$

avec une phase $\theta \in \mathbb{R}$ quelconque. L'équation (9.8) est l'équation de Schrödinger indépendante du temps. La fonction $\phi(\mathbf{r})$ est appelée état d'énergie d'énergie E .

Remarque 9.3. Une autre manière de voir la chose est de dire que les états stationnaires $\phi(\mathbf{r})$ et leurs énergies E sont respectivement les vecteurs propres et les valeurs propres de l'opérateur Hamiltonien :

$$\hat{H} = -\frac{\hbar^2}{2m}\Delta + V.$$

Preuve. Supposons que l'on puisse écrire $\psi(\mathbf{r}, t) = \phi(\mathbf{r})\varphi(t)$ avec les conditions de normalisation données dans la Définition 9.2. Introduisant cette expression dans l'équation de Schrödinger dépendante du temps, il vient :

$$i\hbar\phi(\mathbf{r})\frac{d\varphi}{dt}(t) = \alpha'(t)\hbar\phi(\mathbf{r})\varphi(t) = -\frac{\hbar^2}{2m}\Delta\phi(\mathbf{r})\varphi(t) + V(\mathbf{r})\phi(\mathbf{r})\varphi(t).$$

Simplifiant les deux membres par $\varphi(t) \neq 0$,

$$\alpha'(t)\hbar\phi(\mathbf{r}) = -\frac{\hbar^2}{2m}\Delta\phi(\mathbf{r}) + V(\mathbf{r})\phi(\mathbf{r}).$$

Comme le second membre de cette égalité ne dépend pas du temps, $\alpha'(t)$ est constant et $\alpha(t) = E/\hbar \cdot t + \theta$ pour des constantes $E, \theta \in \mathbb{R}$. Finalement,

$$-\frac{\hbar^2}{2m}\Delta\phi(\mathbf{r}) + V(\mathbf{r})\phi(\mathbf{r}) = E\phi(\mathbf{r}).$$

La réciproque, c'est à dire, le caractère suffisant des relations (9.8) et (9.9) est quant à lui immédiat. $\square$

9.3.3 Energie de la particule

La constante E qui apparaît dans l'équation de Schrödinger indépendante du temps a la dimension d'une énergie. Nous admettons dès à présent que E correspond sans ambiguïté à l'énergie totale de la particule placée dans l'état $\phi(\mathbf{r})$. Les états stationnaires sont d'ailleurs les seuls états quantiques dont l'énergie est parfaitement connue. Nous en verrons les raisons plus tard (voir le chapitre 11). Cela justifie que les états stationnaires soient appelés *états d'énergie*.

9.3.4 Régularité des états stationnaires unidimensionnels

On peut montrer, pour des solutions stationnaires, une régularité plus grande pour la fonction d'onde. En dimension 1 d'espace par exemple, une solution stationnaire $\phi(x)$ est de classe C^1 sur le domaine $\Omega = \{V < +\infty\}$ avec $V \in L^2(\mathbb{R}) + L^\infty(\mathbb{R})$ (on suppose encore Ω ouvert). En effet, pour une solution stationnaire,

$$-\frac{\hbar^2}{2m}\frac{d^2\phi}{dx^2} + (V(x) - E)\phi(x) = 0.$$

Or, $(V - E)\phi$ est dans $L^2_{\text{loc}}(\Omega)$, donc ϕ est dans $H^2_{\text{loc}}(\Omega)$. Comme on a l'injection (compacte) $H^1(\omega) \hookrightarrow C^0(\omega)$ pour des ensembles ouverts et bornés $\omega \subset \Omega$, on en déduit que $\partial_x\phi$ est $C^0(\omega)$, et donc que $\phi \in C^1(\omega)$. Ceci signifie donc que

- ϕ est continue partout sur $\overline{\Omega}$, s'annule sur $\partial\Omega$, et est nulle sur Ω^c ;
- $\partial_x\phi$ est continue sur l'ouvert Ω , mais peut présenter des discontinuités au niveau de $\partial\Omega$.

9.3.5 Minoration de l'énergie des états stationnaires

Le résultat qui suit compare, pour un état stationnaire, les valeurs possibles de E aux valeurs prises par le potentiel $V(\mathbf{r})$.

Proposition 9.4. *L'énergie E associée à un état stationnaire $\phi(\mathbf{r})$ d'une particule quantique évoluant dans un champ de potentiel $V(\mathbf{r})$ satisfait la minoration*

$$E \geq \inf_{\mathbf{r} \in \mathbb{R}^3} V(\mathbf{r}). \quad (9.10)$$

Preuve. On considère l'équation de Schrödinger indépendante du temps (9.8). On prend ϕ régulière à support compact, et on utilise un argument de densité. Multipliant les deux membres de (9.8) par $\bar{\phi}(\mathbf{r})$ et intégrant sur $\mathbb{R}^3$:

$$-\frac{\hbar^2}{2m} \int_{\mathbb{R}^3} \bar{\phi}(\mathbf{r}) \Delta \phi(\mathbf{r}) d\mathbf{r} + \int_{\mathbb{R}^3} V(\mathbf{r}) |\phi(\mathbf{r})|^2 d\mathbf{r} = E.$$

Or,

$$\int_{\mathbb{R}^3} \bar{\phi}(\mathbf{r}) \Delta \phi(\mathbf{r}) d\mathbf{r} = - \int_{\mathbb{R}^3} |\nabla \phi|^2 d\mathbf{r} \leq 0,$$

et comme $|\phi(\mathbf{r})|^2 \geq 0$, on a finalement, grâce à la condition de normalisation,

$$E \geq \int_{\mathbb{R}^3} V(\mathbf{r}) |\phi(\mathbf{r})|^2 d\mathbf{r} \geq \inf_{\mathbb{R}^3} V \int_{\mathbb{R}^3} |\phi(\mathbf{r})|^2 d\mathbf{r} = \inf_{\mathbb{R}^3} V.$$

□

Etant donné que nous savons déjà, par anticipation, que E est l'énergie totale de la particule, ce résultat se comprend simplement comme le fait que cette énergie ne peut être inférieure au minimum d'énergie potentielle accessible. Résultat intuitif? Peut être d'un point de vue classique. Pour ce qui est de l'intuition *quantique* attendons cependant un peu...

9.3.6 Importance des états stationnaires en mécanique quantique

Lors de l'étude d'une situation donnée, la recherche des états stationnaires constitue en général le premier travail à effectuer. Ils fournissent une base d'exemples de fonctions d'onde acceptables pour la particule, et représentent en cela une précieuse source d'information pour mener des raisonnements physiques. Nous le verrons plus tard, ils constituent également une base au sens mathématique du terme, sur laquelle peuvent se décomposer toutes les solutions de l'équation de Schrödinger (voir § 10.4.3). Ces états sont enfin des états dans lesquels l'énergie de la particule est connue sans ambiguïté (à l'opposé, par exemple, de fonctions d'ondes obtenues par superposition d'états très différents l'un de l'autre). Cela en fait des états très particuliers, qui se retrouvent à la base de nombre de développements théoriques.

9.4 Une situation purement quantique : la barrière de potentiel

Le résultat sur la limite classique de l'équation de Schrödinger (§ 9.2) donne une clef quant aux situations dans lesquelles le comportement de la particule est susceptible d'échapper à la mécanique classique. Il faut pour cela que la fonction d'onde soit au moins aussi étalée que la distance caractéristique sur laquelle évolue le champ de potentiel. C'est une situation de ce type, traitée dans un cadre unidimensionnel, que l'on étudie dans ce paragraphe.

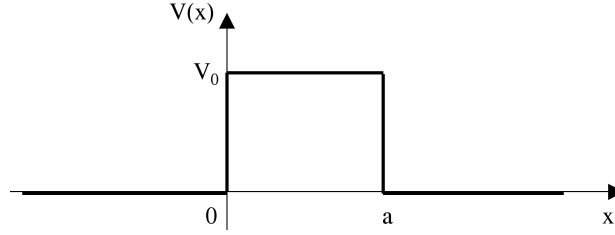


Fig. 9.2. Le modèle unidimensionnel de la *barrière* de potentiel.

9.4.1 La barrière de potentiel

On se place en dimension 1 d'espace et on considère le potentiel (voir figure 9.2)

$$V(x) = \begin{cases} V = 0 & \text{si } x \leq 0, \\ V = V_0 \geq 0 & \text{si } 0 < x < a, \\ V = 0 & \text{si } x \geq a. \end{cases}$$

On cherche à savoir comment une particule quantique, arrivant de la gauche avec une énergie cinétique E_c , interagit avec la barrière. Rappelons ce que nous enseignerait la mécanique classique en pareille situation :

- si $E_c < V_0$: la particule n'a pas l'énergie suffisante pour 'grimper' sur la marche. Elle rebondit et repart vers la gauche. La vitesse ne fait que changer de signe, et l'énergie cinétique est conservée ;
- si $E_c \geq V_0$: la particule passe par dessus la marche et continue sa route vers la droite. Une fois la marche franchie, elle retrouve sa vitesse et son énergie cinétique initiales.

9.4.2 Etats stationnaires de la barrière de potentiel

Calculons les états stationnaires $\psi(x, t) = \phi(x)\varphi(t)$ de la particule quantique et les énergies E associées, en supposant que la particule arrive sur la barrière par la gauche (ceci sera utile par la suite). La Proposition 9.4 montre que dans tous les cas, $E \geq 0$. On cherche ainsi des solutions de

$$-\frac{\hbar^2}{2m} \frac{d^2\phi}{dx^2}(x) + V(x)\phi(x) = E\phi(x).$$

Des résultats généraux de théorie spectrale (voir le chapitre 10) montrent que le spectre de l'opérateur Hamiltonien est purement continu. On ne s'attend donc à trouver que des pseudo-états propres (*i.e.* des limites de suites de Weyl, du type ondes planes). On peut préciser ce résultat par une construction analytique des solutions. Pour ce faire, il faut distinguer deux cas, selon que $E < V_0$ ou $E > V_0$.

Lorsque $E < V_0$.

La forme générale de la solution de l'équation de Schrödinger indépendante du temps, résolue par morceaux sur les intervalles $] -\infty, 0]$, $] 0, a[$, et $[a, +\infty[$ est :

$$\phi(x) = \begin{cases} \alpha_g e^{ikx} + \beta_g e^{-ikx} & \text{si } x \leq 0, \\ \gamma e^{Kx} + \delta e^{-Kx} & \text{si } 0 < x < a, \\ \alpha_d e^{ikx} + \beta_d e^{-ikx} & \text{si } x \geq a, \end{cases} \quad (9.11)$$

avec

$$k = \frac{\sqrt{2mE}}{\hbar}, \quad K = \frac{\sqrt{2m(V_0 - E)}}{\hbar}.$$

On a donc des solutions de type ondes planes à l'infini, qui ne sont pas normalisables (sauf dans le cas trivial $\phi = 0$). Du point de vue physique, on les considère cependant comme des états limites de paquets d'onde très étalés qui viendraient à interagir avec la barrière⁵.

La forme de la solution (9.11) est la forme la plus générale possible mathématiquement, mais n'est pas pertinente en tant que telle du point de vue physique. Rétablissons tout d'abord la dépendance temporelle de la fonction d'onde en la multipliant par la quantité $e^{-i\frac{E}{\hbar}t}$. De part et d'autre de la barrière, là où le potentiel est nul, les termes en e^{ikx} et e^{-ikx} correspondent aux ondes de De Broglie $e^{i(kx - \frac{E}{\hbar}t)}$ et $e^{i(-kx - \frac{E}{\hbar}t)}$, et représentent des états quantiques dans lesquels la particule se déplace respectivement aux vitesses $+\frac{\hbar k}{m}$ et $-\frac{\hbar k}{m}$. Ainsi,

- le terme $\alpha_g e^{ikx}$ décrit une particule arrivant par la gauche. Cette composante correspond à la situation initiale du problème que nous nous posons. Pour imposer la présence de cette composante dans les états recherchés, on prend arbitrairement $\alpha_g = 1$. La valeur précise de cette constante n'a pas grande importance, du fait que nos états ne sont de toute manière pas normalisables ;
- les termes $\beta_g e^{-ikx}$ et $\alpha_d e^{ikx}$ décrivent respectivement une particule qui repart vers la gauche après avoir rebondi sur la barrière, et une particule qui s'éloigne vers la droite après l'avoir franchie. Il s'agit des deux destins envisageables pour la particule en mécanique classique, et nous considérons, pour l'instant sur la base de notre seul bon sens, que ces termes sont en accord avec le problème posé ;
- le terme $\beta_d e^{-ikx}$ décrit en revanche une particule arrivant de la droite. Comme nous considérons une particule arrivant par la gauche, on prend $\beta_d = 0$.

Au final, on se restreint à la recherche de solutions de la forme (voir figure 9.3) :

$$\phi(x) = \begin{cases} \phi = e^{ikx} + \beta_g e^{-ikx} & \text{si } x \leq 0, \\ \phi = \gamma e^{Kx} + \delta e^{-Kx} & \text{si } 0 < x < a, \\ \phi = \alpha_d e^{ikx} & \text{si } x \geq a. \end{cases}$$

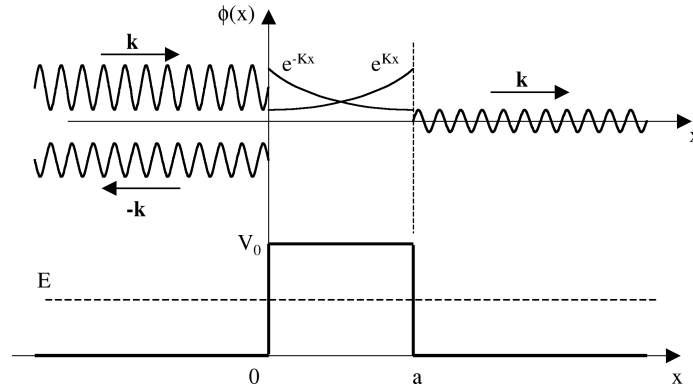


Fig. 9.3. Les cinq composantes des états stationnaires d'énergie $E < V_0$ correspondant à une particule arrivant sur la barrière par la gauche.

Le tri qui a été fait parmi les termes proposés doit être vu comme une tentative de donner aux états stationnaires une signification physique qui répond à l'esprit du problème. Ce sens physique est ici celui de la mécanique classique. Il reste cependant à vérifier que ces choix sont bien autorisés par la théorie quantique. Pour cela, il faut s'assurer qu'il est possible de satisfaire aux conditions de raccordement des différents morceaux de solution. En vertu de leurs propriétés de régularité

5. On verra au § 10.3 un autre cas particulier, pour lequel il faudra explicitement exclure certaines solutions non normalisables. Lesdites solutions ne seront pas de type ondes planes dans ce cas.

(voir § 9.3.4), il faut que ϕ et $\partial_x \phi$ soient continues en $x = 0$ et $x = a$. Les coefficients β_g , γ , δ , et α_d doivent donc vérifier le système linéaire de quatre équations à quatre inconnues

$$\begin{cases} 1 + \beta_g = \gamma + \delta, \\ \gamma e^{Ka} + \delta e^{-Ka} = \alpha_d e^{ika}, \\ ik - ik\beta_g = K\gamma - K\delta, \\ K\gamma e^{Ka} - K\delta e^{-Ka} = ik\alpha_d e^{ika}, \end{cases}$$

dont on peut vérifier qu'il possède effectivement une unique solution :

$$\begin{aligned} \beta_g &= \frac{(K^2 + k^2) \sinh(Ka)}{(k^2 - K^2) \sinh(Ka) + 2ikK \cosh(Ka)}, & \gamma &= \frac{i(K + ik)ke^{-Ka}}{(k^2 - K^2) \sinh(Ka) + 2ikK \cosh(Ka)}, \\ \delta &= \frac{i(K - ik)ke^{Ka}}{(k^2 - K^2) \sinh(Ka) + 2ikK \cosh(Ka)}, & \alpha_d &= \frac{2iKke^{-ika}}{(k^2 - K^2) \sinh(Ka) + 2ikK \cosh(Ka)}. \end{aligned}$$

Lorsque $E > V_0$.

La forme générale de la solution de l'équation de Schrödinger dans le cas où on se restreint à des ondes arrivant par la gauche est (voir figure 9.4)

$$\phi(x) = \begin{cases} \phi = e^{ikx} + \beta_g e^{-ikx} & \text{si } x \leq 0, \\ \phi = \gamma e^{ik'x} + \delta e^{-ik'x} & \text{si } 0 < x < a, \\ \phi = \alpha_d e^{ikx} & \text{si } x \geq a, \end{cases}$$

avec

$$k = \frac{\sqrt{2mE}}{\hbar}, \quad k' = \frac{\sqrt{2m(E - V_0)}}{\hbar}.$$

De la même manière que précédemment, la continuité de ϕ et de sa dérivée $\frac{d\phi}{dx}$ en $x = 0$ et $x = a$ demande que les coefficients β_g , γ , δ , et α_d satisfassent un système linéaire de 4 équations à 4 inconnues dont l'unique solution est

$$\begin{aligned} \beta_g &= \frac{(k^2 - k'^2) \sin(k'a)}{(k^2 + k'^2) \sin(k'a) + 2ikk' \cos(k'a)}, & \gamma &= \frac{i(k + k')ke^{-ik'a}}{(k^2 + k'^2) \sin(k'a) + 2ikk' \cos(k'a)}, \\ \delta &= \frac{-i(k - k')ke^{ik'a}}{(k^2 + k'^2) \sin(k'a) + 2ikk' \cos(k'a)}, & \alpha_d &= \frac{-2kk'e^{-ika}}{(k^2 + k'^2) \sin(k'a) + 2ikk' \cos(k'a)}. \end{aligned}$$

9.4.3 Effet tunnel et transmission résonnante

Rupture avec la mécanique classique

Dans les deux situations $E > V_0$ et $E < V_0$, il est possible de construire des états stationnaires avec $\beta_d = 0$. Cela prouve *a posteriori* la cohérence et la légitimité des choix qui nous ont fait abandonner l'onde de De Broglie venant de la droite. La théorie quantique autorise donc bien que la particule soit réfléchie ou transmise, à l'exclusion d'autres possibilités.

Un fait surprenant apparaît cependant : en dehors de quelques cas particuliers, aucun des coefficients β_g , γ , δ , et α_d obtenus n'est nul. Cela signifie qu'il est *impossible* dans le cas général, de construire d'état quantique dans lequel la particule est entièrement réfléchie vers la gauche, ou entièrement transmise vers la droite. La théorie quantique impose que la particule soit *simultanément* transmise *et* réfléchie par l'obstacle.

Cette réflexion-transmission simultanée constitue un comportement qui tranche avec la mécanique classique :

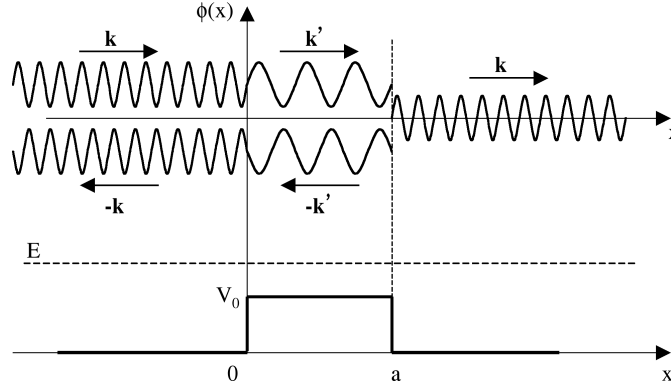


Fig. 9.4. Les cinq composantes des états stationnaires d'énergie $E > V_0$ correspondant à une particule arrivant sur la barrière par la gauche.

- il faut tout d'abord comprendre ce phénomène comme le fait qu'en interagissant avec la barrière, la fonction d'onde se sépare en deux composantes : une qui continue sa course vers la droite, et une qui revient vers la gauche. La *particule quantique perd son unité*, et son état après l'interaction est une *superposition* de deux états transmis et réfléchi ;
- ensuite, même avec une énergie E inférieure à la hauteur V_0 de la barrière, une fraction de la particule est capable, au moins en partie, de poursuivre sa route vers la droite. Plus fort encore, la fonction d'onde, et donc, la probabilité de présence, ne sont pas nulles dans la zone $0 < x < a$ où l'énergie est pourtant inférieure au potentiel $V(x)$. Cela serait parfaitement inconcevable en mécanique classique. Pour autant, la théorie quantique autorise la particule à faire des excursions modérées dans ces 'zones interdites'. Ces phénomènes sont désignés sous le terme d'*effet tunnel* ;
- les coefficients β_g et α_d sont enfin des nombres complexes. La reflexion ou la transmission de la particule s'accompagne donc d'un *déphasage*. Cette notion, propre à la fonction d'onde, échappe complètement à la mécanique classique, pour qui seules importent les informations de position et de vitesse.

Conservation de l'énergie

Analysons les ondes transmises et réfléchies. Toutes deux sont des ondes de De Broglie qui se déplacent dans une zone où le potentiel $V(x)$ est nul, et qui ont, au signe près, le même vecteur d'onde k que la particule incidente. Elles décrivent donc toutes deux une particule qui se déplace à la vitesse $v = \frac{\hbar k}{m}$, et avec une énergie cinétique $E_c = \frac{1}{2}mv^2 = \frac{\hbar^2 k^2}{2m}$. Nous retrouvons le résultat classique qui dit que la vitesse et l'énergie cinétique de la particule se conservent au cours du rebond ou du franchissement de la barrière.

Probabilités de rebond et de franchissement

Considérons enfin les courants de probabilité associés aux ondes incidente, réfléchie, et transmise. Ces ondes de De Broglie ne sont pas normalisables, mais nous pouvons écrire, en valeurs absolues, et à une constante multiplicative C près (la même pour les 3 ondes) :

$$J_i = C \frac{\hbar k}{m}, \quad J_r = C |\beta_g|^2 \frac{\hbar k}{m}, \quad J_t = C |\alpha_d|^2 \frac{\hbar k}{m}.$$

On peut alors constater que, quelque soit la valeur de l'énergie E , $J_r + J_t = J_i$. Le courant de probabilité, et donc d'une certaine manière la particule, se conservent au franchissement de la barrière. La particule incidente se retrouve en entier dans les ondes réfléchies et transmises,

et surtout, ne part pas ailleurs. Ce résultat souligne à nouveau la cohérence physique des états stationnaires qui viennent d'être construits.

On peut définir, à partir des flux J_i , J_r et J_t des *coefficients de réflexion et de transmission* selon :

$$R = \frac{J_r}{J_i}, \quad T = \frac{J_t}{J_i},$$

avec $R + T = 1$. Une interprétation de ces coefficients est la suivante : si l'on effectue une mesure de la position de la particule juste après qu'elle a fini d'interagir avec la barrière, la particule aura une probabilité R d'apparaître à gauche, et une probabilité T d'apparaître à droite (voir figure 9.5). Les coefficients R et T prennent ainsi valeur de *probabilité de rebond* et de *probabilité de franchissement* de la barrière.⁶

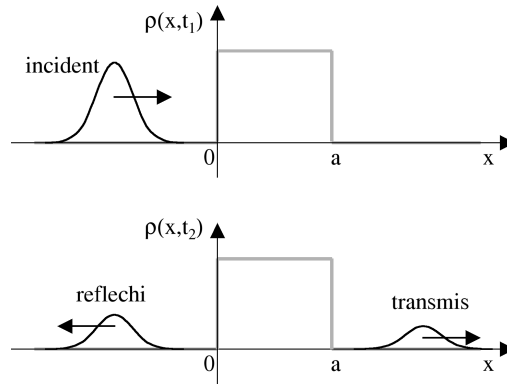


Fig. 9.5. La diffraction d'un paquet d'onde arrivant sur la barrière. Instant t_1 : avant l'impact. Instant t_2 : après l'impact. Dans le traitement mathématique de cette section, ces paquets d'onde sont approximatés par des ondes de De Broglie.

La manière dont R et T évoluent en fonction des conditions du problème est également remarquable. Considérons par exemple T . On déduit des relations qui précèdent que :

$$T = |\alpha_d|^2 = \begin{cases} \left(1 + \frac{V_0^2}{E(V_0 - E)} \sinh^2 \left(\frac{\sqrt{2m(V_0 - E)}}{\hbar} a \right) \right)^{-1} & \text{si } E < V_0, \\ \left(1 + \frac{V_0^2}{E(E - V_0)} \sin^2 \left(\frac{\sqrt{2m(E - V_0)}}{\hbar} a \right) \right)^{-1} & \text{si } E > V_0. \end{cases} \quad (9.12)$$

Effet Tunnel

Ce cas concerne les énergies inférieures à V_0 . Fixons quelques ordres de grandeur. Pour un électron de masse $m = 9.1 \times 10^{-31}$ kg, devant traverser une barrière d'épaisseur $a = 1$ Å, de hauteur $V_0 = 1$ eV, avec une énergie incidente $E = 0.5$ eV, on trouve $T \approx 0.62$, soit 62% de chances de traverser la barrière! Loin d'être anecdotique, l'effet tunnel est donc un phénomène prépondérant à l'échelle atomique.

Dès que la masse et les énergies mises en jeu sont ne serait-ce que légèrement supérieures, le terme en sinus hyperbolique devient rapidement très grand, et on peut faire l'approximation

6. Le cas de la barrière de potentiel présente cependant un certain niveau de subtilité. En toute rigueur, il faudrait, pour bien définir ces coefficients de transmission et de réflexion, intégrer la probabilité de présence associée à chaque composante de l'état quantique, respectivement à gauche ou à droite de la barrière selon la composante, et comparer les valeurs des intégrales obtenues. Du fait que l'état manipulé ici n'est pas normalisable, nous avons utilisé une approche alternative qui consiste à comparer les flux de probabilité.

$$T \approx \frac{16E(V_0 - E)}{V_0^2} \exp\left(-2\frac{\sqrt{2m(V_0 - E)}}{\hbar}a\right).$$

La probabilité de traverser la barrière décroît alors exponentiellement avec l'épaisseur de celle-ci. Lorsqu'on travaille avec des électrons, l'effet tunnel est extrêmement sensible à des fluctuations de l'épaisseur a de l'ordre de la fraction d'Angström. La technique de *microscopie par effet tunnel* met cela à profit pour sonder le relief de surface de matériaux avec une résolution sub-atomique. Cependant, dès que la particule est trop massive, la probabilité de traverser la barrière devient vite négligeable. Dans l'exemple ci-dessus, on trouve ainsi $T \approx 10^{-19}$ pour un simple proton de 1.6×10^{-27} kg, et cette quantité n'a plus grande signification à notre échelle (voir figure 9.6).

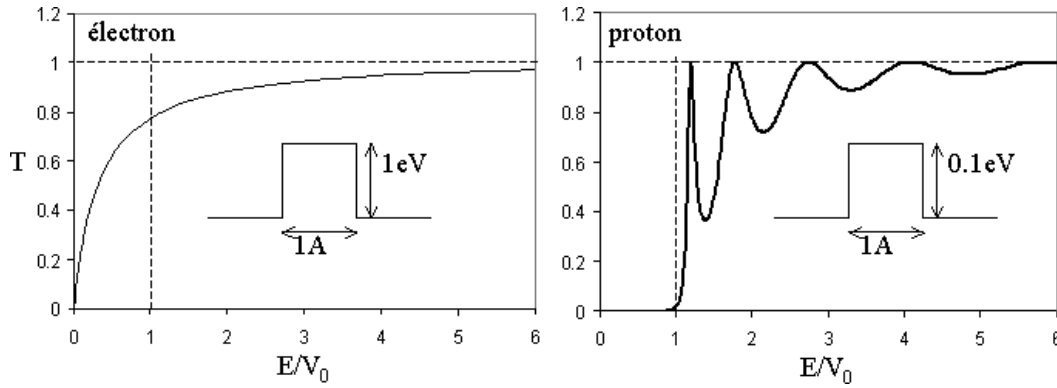


Fig. 9.6. Probabilité T de traverser une barrière de $1\text{\AA}$ de large en fonction du rapport E/V_0 . *A gauche* : cas d'un électron sur une barrière de 1eV . T garde des valeurs très significatives pour $E < V_0$. C'est l'effet tunnel. *A droite* : cas d'un proton sur une barrière de 0.1eV . T présente d'importantes oscillations pour $E > V_0$. C'est la transmission résonante.

Transmission résonante

Pour les énergies E supérieures à V_0 , la probabilité de traverser la barrière T oscille entre une situation de transfert partiel de la particule, et des situations de transfert total où $T = 1$. Ces maxima de taux de transmission sont atteints quand $k'a$ est un multiple de π , c'est à dire, quand a est égal à un nombre entier de fois la demi longueur d'onde $\frac{\pi}{k'}$ qui décrit la particule à l'intérieur de la barrière. Ces oscillations sont ainsi le résultat d'interférences quantiques de la particule avec elle-même au niveau de l'obstacle. Cette propriété peut servir en pratique à concevoir, dans les machines électroniques, des filtres qui ne se laissent franchir que par des électrons possédant une énergie donnée.

9.4.4 Et la limite classique ?

La limite classique est obtenue dans les situations où la taille du paquet d'onde devient très petite devant l'échelle caractéristique de variation du potentiel $V(x)$. Une astuce efficace pour trouver cette limite sans construire explicitement ces paquets d'onde est de faire tendre dans les formules de la mécanique quantique la masse m de la particule étudiée vers l'infini, tout en maintenant l'énergie constante. On observe en effet que les particules massives sont en général suffisamment bien décrites par la mécanique classique.

Faisons donc tendre m vers l'infini dans les résultats (9.12). Pour $E < V_0$, nous avons vu que l'effet tunnel devient insignifiant, et nous retrouvons le fait classique de la particule rebondit entièrement en arrière. Pour $E > V_0$, en revanche, augmenter la masse m n'empêche pas les oscillations

de T de se produire. Elles deviennent plus rapides, mais leur amplitude reste invariablement bornée inférieurement par $\left(1 + \frac{V_0^2}{4E(E-V_0)}\right)^{-1}$ et supérieurement par 1. On ne parvient donc pas à tendre vers la limite classique d'une transmission totale.

Cette bizarrerie provient des discontinuités du champ de potentiel. La distance caractéristique sur laquelle ce potentiel évolue est en effet la distance nulle ! Ainsi, aussi localisé qu'il soit, un paquet d'onde présentera toujours une largeur supérieure à l'échelle d'évolution de V . En toute rigueur, le résultat concernant la limite classique de l'équation de Schrödinger ne s'applique donc pas. Cela ne signifie pas qu'il est impossible de retrouver le comportement classique : nous y parvenons dans le cas de l'effet tunnel. En revanche, cela n'est plus une obligation, comme nous le constatons pour la transmission résonante.

Pour forcer le comportement classique à réapparaître, il faudrait adoucir la forme de la barrière, en remplaçant, par exemple, les discontinuités en $x = 0$ et $x = a$ par des rampes. Même abruptes, ces rampes possèderaient alors une largeur finie. On verrait dans ce cas que les fonctions d'onde plus localisées que cette largeur retrouvent un comportement classique avec réflexion ou transmission totale.

9.5 Implémentation numérique

On présente dans cette section des schémas numériques d'intégration de (9.1). Pour simplifier, nous nous restreindrons au cas d'une particule dans un potentiel extérieur indépendant du temps dans un espace à une dimension.

On adimensionnalise le problème en effectuant un changement de variables tel que $\hbar = 2m = 1$ (dans l'esprit des unités atomiques, voir § A.2). Dans ce cas, l'évolution de la fonction d'onde est régie par

$$i\partial_t\psi = -\partial_x^2\psi + V(x)\psi. \quad (9.13)$$

On considère par ailleurs que la solution reste dans un domaine⁷ $[0, R]$. On impose ainsi les conditions de bord $\psi(0, t) = \psi(1, t) = 0$. D'autres types de conditions de bord sont possibles, en particulier des conditions de bord périodiques.

9.5.1 Schéma de Cranck-Nicholson.

Ce schéma repose sur une approximation de type différences finies. Pour un maillage uniforme $(x_0, x_1, \dots, x_{N+1})$ de $[0, R]$ (ainsi, $x_i = i\Delta x$ avec $(N+1)\Delta x = R$), on considère une approximation du Laplacien par la différence finie centrée

$$\partial_x^2\psi(x, t) = \frac{\psi(x - \Delta x, t) - 2\psi(x, t) + \psi(x + \Delta x, t)}{\Delta x^2} + O(\Delta x).$$

Introduisant le vecteur Ψ des approximations de ψ aux points x_i , ainsi que les matrices L du Laplacien discret et $V_{\Delta x}$ du potentiel, soit

$$\Psi = \begin{pmatrix} \psi_0 \\ \psi_1 \\ \vdots \\ \psi_{N+1} \end{pmatrix}, \quad L_{\Delta x} = \frac{1}{\Delta x^2} \begin{pmatrix} 2 & -1 & & \\ -1 & 2 & -1 & \\ & \ddots & \ddots & \ddots \\ & & -1 & 2 & -1 \\ & & & -1 & 2 \end{pmatrix}, \quad V_{\Delta x} = \begin{pmatrix} V(x_0) & & \\ & \ddots & \\ & & V(x_N) \end{pmatrix},$$

on a une approximation $H_{\Delta x} = -L_{\Delta x} + V_{\Delta x}$ du Hamiltonien sur la grille $\{x_i\}$. Ainsi, l'équation de Schrödinger s'écrit, dans sa version discrétisée en espace,

7. On peut en effet, par un changement de variables bijectif $\mathbb{R} \rightarrow [0, 1]$, se ramener à ce cas. On peut également considérer un domaine suffisamment grand, un paquet d'ondes suffisamment localisé et un temps d'évolution suffisamment court.

$$i\partial_t \Psi = H_{\Delta x} \Psi.$$

On en déduit que

$$\Psi(t) = e^{-itH_{\Delta x}} \Psi(0).$$

Cette formule d'intégration exacte en temps peut être utilisée dans des cas simples où le calcul de l'exponentielle de matrice (qui demande une diagonalisation préalable de $H_{\Delta x}$) peut être effectué. On obtient dans ce cas le schéma numérique

$$\Psi^{n+1} = e^{-i\Delta t H_{\Delta x}} \Psi^n. \quad (9.14)$$

Pour calculer en pratique cet opérateur, on peut diagonaliser $H_{\Delta x}$ selon $H_{\Delta x} = P^T D P$. Dans ce cas, $e^{-i\Delta t H_{\Delta x}} = P^T e^{-i\Delta t D} P$. Comme la diagonalisation d'une matrice demande de l'ordre de $O(N^3)$ opérations (où N est le nombre de points de grille), cette opération est rapidement prohibitive. Dans ce cas, il faut approximer $e^{-i\Delta t H_{\Delta x}}$ par une formule moins coûteuse⁸. On peut proposer par exemple

$$e^{-i\Delta t H_{\Delta x}} \Psi = \left(1 - \frac{i\Delta t}{2} H_{\Delta x}\right) \left(1 + \frac{i\Delta t}{2} H_{\Delta x}\right)^{-1} \Psi + O(\Delta t^2),$$

ce qui conduit au schéma numérique

$$\Psi^{n+1} = A(\Delta t, \Delta x) \Psi^n, \quad A(\Delta t, \Delta x) = \left(1 - \frac{i\Delta t}{2} H_{\Delta x}\right) \left(1 + \frac{i\Delta t}{2} H_{\Delta x}\right)^{-1}. \quad (9.15)$$

En pratique, l'itération (9.15) est réécrite sous la forme d'un système tridiagonal, dont l'inversion ne demande que $O(N)$ itérations :

$$\begin{cases} \Psi_1^{n+1} & = 0 \\ \Psi_1^{n+1} - \left(2 + V(x_2)\Delta x^2 + \frac{2i\Delta x^2}{\Delta t}\right) \Psi_2^{n+1} + \Psi_3^{n+1} & = R_2^n, \\ & \vdots \\ \Psi_{N-1}^{n+1} - \left(2 + V(x_2)\Delta x^2 + \frac{2i\Delta x^2}{\Delta t}\right) \Psi_N^{n+1} + \Psi_{N+1}^{n+1} & = R_N^n, \\ \Psi_{N+1}^{n+1} & = 0, \end{cases} \quad (9.16)$$

avec

$$R_k^n = \Psi_{k-1}^n - \left(2 + V(x_2)\Delta x^2 - \frac{2i\Delta x^2}{\Delta t}\right) \Psi_k^n + \Psi_{k+1}^n, \quad 1 \leq k \leq N.$$

Des généralisations au cas des dimensions supérieures existent (on a une forme tridiagonale généralisée).

Ce schéma a de bonnes propriétés numériques. On vérifie en particulier que la probabilité numérique

$$p^n = \sum_{k=1}^N |\Psi_k^n|^2 \Delta x$$

est conservée exactement. Ceci provient du fait que les valeurs propres de $A(\Delta t, \Delta x)$ sont de module 1. La vérification de cette dernière propriété est immédiate au vu de la définition donnée par (9.15).

⁸. Il faut également recourir à d'autres méthodes si le potentiel dépend du temps ! Le schéma de Crank-Nicholson est précisément utilisable dans ce cadre.

9.5.2 Décomposition d'opérateurs.

Comme on l'a vu pour certains intégrateurs utilisés en dynamique moléculaire, il est possible de décomposer un opérateur en opérateurs élémentaires dont l'évolution est facile à calculer, afin d'obtenir une approximation de l'évolution globale en composant lesdits opérateurs élémentaires. Cette technique est ainsi une alternative à une itération en temps à la Cranck-Nicholson dans le cas où le calcul de l'exponentielle de matrice pour le schéma (9.14) n'est pas faisable.

Plus précisément, on décompose l'opérateur Hamiltonien en ses parties cinétique et potentielle, et on considère les flots élémentaires $\phi_{\Delta t}^{\text{pot}}$, $\phi_{\Delta t}^{\text{cin}}$ correspondant respectivement aux équations

$$i\partial_t \psi = V\psi, \quad i\partial_t \psi = -\Delta\psi.$$

La première équation peut s'intégrer analytiquement selon

$$\phi_{\Delta t}^{\text{pot}}(\psi)(x) = e^{itV(x)} \psi(x).$$

La seconde peut également être intégrée analytiquement en variables de Fourier. Considérant la transformée de Fourier $\hat{\psi}(k) = \mathcal{F}(\psi)(k)$ de ψ , la seconde équation peut être réécrite

$$i\partial_t \hat{\psi}(k) = |k|^2 \hat{\psi}(k).$$

On a alors

$$\phi_{\Delta t}^{\text{cin}}(\psi)(x) = \mathcal{F}^{-1}(e^{it|k|^2} \hat{\psi}).$$

Au final, un schéma numérique fondé sur cette décomposition et les approximations correspondantes des opérateurs élémentaires est donné par

$$\Psi^{n+1} = \Phi_{\Delta t/2}^{\text{pot}} \circ \Phi_{\Delta t}^{\text{cin}} \circ \Phi_{\Delta t/2}^{\text{pot}}(\Psi^n), \quad (9.17)$$

où $\psi(t_n)$ est approximé par Ψ^n comme pour le schéma de Cranck-Nicholson, et les applications $\phi_{\Delta t}^{\text{pot}}$, $\phi_{\Delta t}^{\text{cin}}$ sont approximées en considérant une grille uniforme $(x_0, \dots, x_{N+1})$ de l'espace réel $[0, R]$. Dans ce cas,

$$\Phi_{\Delta t}^{\text{pot}}\Psi = \text{Diag}(e^{-i\Delta t V(x_0)}, \dots, e^{-i\Delta t V(x_{N+1})})\Psi.$$

Pour la partie cinétique, on utilise une transformée de Fourier discrète.

Exercices

Exercice 9.1 (Particule uniformément accélérée).

On considère une particule quantique de masse m décrite à $t = 0$ par l'onde de De Broglie

$$\psi(\mathbf{r}, t = 0) = e^{i\mathbf{k}_0 \cdot \mathbf{r}}.$$

On la soumet, à partir de $t = 0$ à l'action d'une force $\mathbf{F}$.

- (1) Ecrire l'équation de Schrödinger pour la fonction d'onde $\psi(\mathbf{r}, t)$ pour les temps positifs.
- (2) Résoudre cette équation. Indication : on montrera qu'il existe une solution de la forme générale

$$\psi(\mathbf{r}, t) = e^{i(\mathbf{k}(t) \cdot \mathbf{r} - \theta(t))}.$$

et on se limitera à expliciter le vecteur $\mathbf{k}(t)$.

- (3) Exprimer en fonction de $\mathbf{k}(t)$ l'impulsion $\mathbf{p}(t)$ de la particule. Montrer que $\mathbf{p}(t)$ ainsi calculée obéit à la relation fondamentale de la dynamique.

Exercice 9.2 (Marche de potentiel).

On considère, dans un contexte unidimensionnel, le paysage de potentiel suivant, dit *marche de potentiel* :

$$V(x) = \begin{cases} 0 & \text{si } x < 0, \\ V_0 \geq 0 & \text{si } x \geq 0. \end{cases}$$

On souhaite étudier ce qu'il advient d'une particule quantique de masse m arrivant sur la marche par la gauche.

- (1) Représenter graphiquement ce champ de potentiel.
- (2) Donner, sans se préoccuper pour l'instant des conditions de raccordement de la fonction d'onde en $x = 0$, la forme générale de la partie spatiale $\phi(x)$ d'un état stationnaire d'énergie E . On distinguera les cas $E > V_0$ et $E < V_0$.

On pourra poser :

$$k = \frac{\sqrt{2mE}}{\hbar}, \quad k' = \frac{\sqrt{2m(E - V_0)}}{\hbar} \quad \text{si } E > V_0, \quad K = \frac{\sqrt{2m(V_0 - E)}}{\hbar} \quad \text{si } E < V_0.$$

- (3) Quelle est, parmi ces états, la forme des états physiquement acceptables décrivant une particule arrivant par la gauche ?
- (4) En utilisant les conditions de raccordement, préciser la forme des états stationnaires pour $E < V_0$. Montrer que la particule est entièrement réfléchi par la barrière. Montrer également qu'elle pénètre légèrement dans la zone $x > 0$, sur une distance typique l_{tunnel} :

$$l_{\text{tunnel}} = \frac{\hbar}{\sqrt{2m(V_0 - E)}}.$$

- (5) Application numérique : on considère un électron arrivant avec une énergie de 1 eV sur une marche de 2 eV de haut. Calculer l_{tunnel} . Même question pour un cycliste de 70 kg arrivant à une vitesse de 10 m/s sur une rampe très raide de 2 m de haut.

La particule quantique confinée

Jusqu'ici, nous avons considéré des particules libres de se déplacer dans tout l'espace. L'exemple des ondes de De Broglie a montré qu'elles pouvaient adopter n'importe quelle impulsion, n'importe quelle énergie cinétique et n'importe quelle position. Cette liberté est cependant loin d'être la règle en mécanique quantique. Nous regardons ici la situation dans laquelle un champ de potentiel indépendant du temps piège la particule dans une région de taille limitée. L'énergie, et bien d'autres grandeurs, ne peuvent alors prendre que certaines valeurs bien précises : c'est la *quantification*.

Ceci relève mathématiquement de propriétés spectrales particulières des Hamiltoniens quantiques, qui sont étudiées au § 10.2 au travers d'exemples fondamentaux : puits infini, oscillateur harmonique quantique, potentiel coulombien (atome de type hydrogénoïde). Au-delà des propriétés abstraites données par la théorie spectrale, des études plus fines peuvent être menées sur des cas particuliers. Un tel exemple est le cas du trou de potentiel, traité analytiquement au § 10.3. Enfin, une discussion des implications physiques profondes du phénomène de quantification clôt ce chapitre (§ 10.4).

10.1 Cas des potentiels confinants	152
10.1.1 Puits infini	152
10.1.2 Oscillateur harmonique	153
10.2 Théorie spectrale des Hamiltoniens quantiques - exemple du potentiel coulombien	155
10.2.1 Caractère auto-adjoint	155
10.2.2 Spectre essentiel	157
10.2.3 Spectre ponctuel	157
10.3 Un exemple quasi-analytique : le trou de potentiel	159
10.3.1 Définition	159
10.3.2 Etats d'énergie liés	160
10.3.3 Etats d'énergie de diffusion	162
10.4 Universalité et implications du phénomène de quantification	164
10.4.1 La quantification	164
10.4.2 Caractère général du phénomène de piégeage	164
10.4.3 Quantification et dynamique de la particule	164
10.4.4 Propriétés spectrales des atomes, des molécules, et des corps denses	165
10.4.5 La quantification comme élément stabilisateur de la matière	167
10.4.6 Et la limite classique ?	168
10.4.7 Et ce n'est pas tout...	169
Exercices	170

10.1 Cas des potentiels confinants

Le puits de potentiel et le potentiel harmonique constituent deux cas très particuliers de champs de potentiels. Dans les deux cas, en effet, V est égal à ou tend vers $+\infty$ lorsque $|\mathbf{r}|$ tend vers $+\infty$. La particule, quelque soit son énergie, se retrouve donc toujours confinée dans une petite région d'espace. C'est ce confinement systématique qui est à l'origine du caractère purement discret du spectre de $\hat{H}$. D'un point de vue mathématique, cela signifie en particulier que le spectre de l'opérateur Hamiltonien est purement discret, et d'un point de vue physique, les systèmes sont complètement quantifiés.

10.1.1 Puits infini

Le puits infini correspond au cas d'un confinement strict dans un domaine Ω ouvert borné ; dans ce cas, le potentiel V est de la forme

$$V(x) = \begin{cases} 0 & \text{si } x \in \Omega, \\ +\infty & \text{si } x \notin \Omega. \end{cases}$$

Nous recherchons les parties spatiales $\phi(\mathbf{r})$ des états stationnaires, et leurs énergies E associées. Comme on ne considère que des fonctions d'onde d'énergie finie, on se restreint à des fonctions à support dans Ω , par exemple $\phi \in L^2(\Omega)$. On peut montrer en fait (voir § 9.1.3) que

$$\phi \in H_0^1(\Omega), \quad \int_{\Omega} |\phi(t, x)|^2 dx = 1.$$

Chercher les états stationnaires de l'équation de Schrödinger (9.1) revient donc à chercher les $\phi \in H_0^1(\Omega)$ tels que $-\frac{\hbar^2}{2m}\Delta\phi = E\phi$.

Caractère discret du spectre

Or, par le théorème de Lax-Milgram, pour $f \in L^2(\Omega)$, l'équation $-\Delta u = f$ admet une et une seule solution, et l'application $f \mapsto u$ est linéaire bornée de $L^2(\Omega)$ dans $H_0^1(\Omega)$. Alors, l'opérateur

$$\begin{array}{ccc} \hat{T} = -\Delta^{-1} & : & L^2(\Omega) \rightarrow H_0^1(\Omega) \hookrightarrow L^2(\Omega) \\ f & \mapsto & u \mapsto u \end{array}$$

est linéaire borné auto-adjoint compact (car l'injection $H_0^1(\Omega) \hookrightarrow L^2(\Omega)$ est compacte¹). Par ailleurs, cet opérateur est positif car

$$\langle \hat{T}f, f \rangle = \langle u, -\Delta u \rangle = \|\nabla u\|_{L^2(\Omega)}^2 \geq 0.$$

Ceci montre, en utilisant le théorème C.2, qu'il existe une suite $(\mu_n)_{n \geq 1}$ de réels strictement positifs convergeant vers 0, et une base hilbertienne (ϕ_n) de $L^2(\Omega)$ tels que

$$\forall n \geq 1, \quad \hat{T}\phi_n = \mu_n \phi_n.$$

Finalement, il existe une suite $(E_n)_{n \geq 1}$ de réels strictement positifs

$$E_n = \frac{\hbar^2}{2m\mu_n},$$

convergeant vers $+\infty$, et une base hilbertienne $(\phi_n)_{n \geq 1}$ de $L^2(\Omega)$ tels que

$$\forall n \geq 1, \quad \phi_n \in H_0^1(\Omega), \quad -\frac{\hbar^2}{2m}\Delta\phi_n = E_n\phi_n.$$

De manière générale, on ne peut pas plus préciser les suites $(E_n)_{n \geq 1}$ et $(\phi_n)_{n \geq 1}$. Pour certaines géométries du domaine Ω , on a toutefois des expressions analytiques.

1. Rappelons qu'un opérateur $T : E \rightarrow F$ (E, F Banach) est compact si, de toute suite bornée (u_n) donnée, on peut extraire une sous-suite $(\tilde{u}_k) = (Tu_{n_k})$ qui converge fortement.

Le puits infini en dimension un d'espace

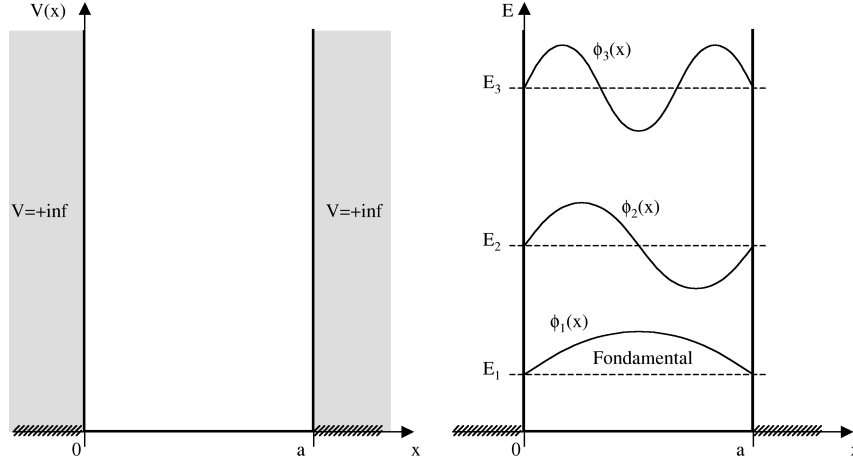


Fig. 10.1. Le puits de potentiel infini en dimension 1. *À gauche* : Représentation du champ de potentiel. Les parties grisées correspondent à $V = +\infty$. *À droite* : Les trois premiers niveaux d'énergie.

En dimension un d'espace, pour le domaine $\Omega =]0, a[$, les fonctions ϕ_n sont

$$\phi_n(x) = \sqrt{\frac{2}{a}} \sin\left(\frac{n\pi}{a}x\right),$$

les valeurs propres associées étant

$$E_n = \frac{\hbar^2 \pi^2 n^2}{2ma^2}.$$

Les valeurs admissibles de l'énergie sont donc à chercher dans un ensemble limité : on dit qu'elles sont quantifiées. On généralise aisément au cas d'un domaine rectangulaire en dimension supérieure. Dans tous les cas, on remarque que seul l'état de plus basse énergie a une fonction d'onde qui ne s'annule pas dans l'intérieur du domaine. Il est dit *état fondamental* (voir figure 10.1).

10.1.2 Oscillateur harmonique

Le Hamiltonien d'un oscillateur harmonique est de la forme

$$\hat{H} = -\frac{\hbar^2}{2m}\Delta + \frac{1}{2}m\omega^2|\mathbf{r}|^2,$$

avec $\omega > 0$. En mécanique classique, cela correspond au cas d'une particule de masse m qui serait reliée au point origine par un ressort de raideur k . Dans ce cas, ω correspond à la pulsation des oscillations qu'effectue la particule après avoir été éloignée de sa position d'équilibre. Le potentiel classique est dans ce cas $V(\mathbf{r}) = k|\mathbf{r}|^2/2$ et $\omega = \sqrt{\frac{k}{m}}$ (voir figure 10.2).

Caractère quantifié du spectre d'énergie

Le domaine de l'opérateur $\hat{H}$ est

$$D(\hat{H}) = \left\{ u \in H^2(\mathbb{R}^3) \mid \int_{\mathbb{R}^3} |\mathbf{r}|^4 |u(\mathbf{r})|^2 d\mathbf{r} < +\infty \right\},$$

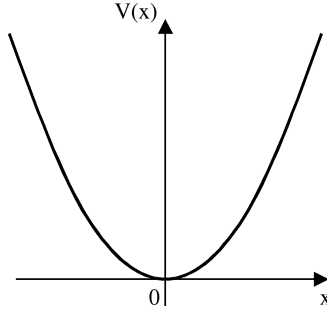


Fig. 10.2. Le potentiel harmonique en dimension 1.

soit

$$D(\hat{H}) = \left\{ u \in L^2(\mathbb{R}^3) \mid \int_{\mathbb{R}^3} |\mathbf{r}|^4 |u(\mathbf{r})|^2 d\mathbf{r} + \int_{\mathbb{R}^3} |\mathbf{k}|^4 |\hat{u}(\mathbf{k})|^2 d\mathbf{k} < +\infty \right\}.$$

On vérifie facilement que $\hat{H}$ est auto-adjoint.

Pour vérifier que le spectre de l'oscillateur harmonique est purement ponctuel, on peut mener une étude similaire au cas puits infini (voir Exercice 10.7). On peut également recourir à un résultat général de théorie spectrale (voir [29, Théorème XIII.67]) qui montre que les Hamiltoniens de la forme $-\Delta + V$ avec $V \in L^1_{\text{loc}}(\mathbb{R}^3)$ borné inférieurement et tel que $V(\mathbf{r}) \rightarrow +\infty$ lorsque $|\mathbf{r}| \rightarrow +\infty$, ont une résolvante compacte. Ainsi, $\hat{H}$ a un spectre purement ponctuel et on peut également montrer qu'il est injectif.

Etats d'énergie de l'oscillateur harmonique 1D

Les états d'énergie de l'oscillateur harmonique peuvent être calculés de manière analytique. Pour un espace de dimension $d = 1$, ce sont les fonctions de Hermite

$$\forall n \geq 0, \quad \phi_n(x) = \frac{\pi^{-1/4}}{\sqrt{2^n n!}} \left(\frac{m\omega}{\hbar} \right)^{\frac{1}{4} - \frac{n}{2}} e^{\frac{m\omega}{2\hbar} x^2} \frac{d^n}{dx^n} \left(e^{-\frac{m\omega}{\hbar} x^2} \right),$$

les niveaux d'énergie correspondants (figure 10.3) étant

$$\forall n \geq 0, \quad E_n = \hbar\omega \left(n + \frac{1}{2} \right).$$

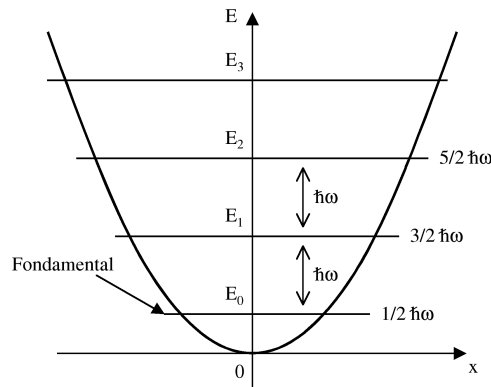


Fig. 10.3. Les 4 premiers niveaux d'énergie de l'oscillateur harmonique en dimension 1.

10.2 Théorie spectrale des Hamiltoniens quantiques - exemple du potentiel coulombien

Cette section a pour objectif d'étudier des cas plus généraux de champ de potentiel, et notamment ceux où la particule a encore la possibilité, en fonction de son énergie, d'explorer librement tout l'espace, et pas uniquement par le truchement de l'effet tunnel. Nous montrons qu'alors, le Hamiltonien possède en général un spectre qui se scinde en deux parties :

- une *partie ponctuelle* : des valeurs propres qui, en somme, correspondent à de “vrais” états physiques (normalisables) ;
- une *partie continue* : des valeurs propres qui correspondent à des états non normalisables, pouvant néanmoins s'exprimer comme des limites de suites de Weyl d'états normalisables.

La propriété fondamentale sous-jacente dont nous avons besoin pour l'étude des spectres des opérateurs est le caractère autoadjoints des Hamiltoniens quantiques. C'est à cette propriété que nous nous intéressons en premier lieu. On applique ensuite divers résultats de l'Appendice C pour caractériser le spectre des différents systèmes considérés ici.

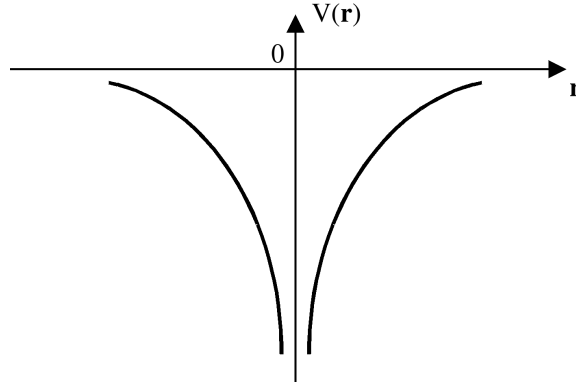


Fig. 10.4. Le potentiel coulombien. Très creusé au voisinage de l'origine, il tend vers 0 à l'infini. Une particule d'énergie positive peut *a priori* explorer librement tout l'espace.

Tout au long de cette partie, les résultats essentiels sont illustrés par le cas du potentiel Coulombien que ressent un électron au voisinage d'un noyau atomique de nombre de charge Z (voir figure 10.4) :

$$V(\mathbf{r}) = -\frac{e^2 Z}{4\pi\epsilon_0 |\mathbf{r}|}.$$

10.2.1 Caractère auto-adjoint

Pour montrer qu'un Hamiltonien quantique avec potentiel est auto-adjoint, on peut soit faire une preuve directe au cas par cas, soit utiliser un théorème général qui garantit la conservation du caractère auto-adjoint pour des Hamiltoniens de la forme $\hat{H} = -\frac{\hbar^2}{2m}\Delta + V$ moyennant certaines conditions sur V (V “perturbation pas trop importante” du Laplacien).

Définition 10.1. Soient $\hat{H}_0$ et $\hat{W}$ deux opérateurs non-bornés sur un espace de Hilbert $\mathcal{H}$. On dit que $\hat{W}$ est $\hat{H}_0$ -borné si $D(\hat{H}_0) \subset D(\hat{W})$ et s'il existe des constantes a et $b > 0$ telles que

$$\forall u \in D(\hat{H}_0), \quad \|\hat{W}u\| \leq a\|\hat{H}_0 u\| + b\|u\|.$$

L'infimum des $a > 0$ tels que cette inégalité soit vraie est appelée la borne relative de $\hat{W}$ par rapport à $\hat{H}_0$.

Remarque 10.1. Un opérateur borné de $\mathcal{H}$ dans lui-même est $\hat{H}_0$ -borné avec une borne relative de 0 pour tout opérateur $\hat{H}_0$.

Connaissant un opérateur auto-adjoint $\hat{H}_0$, on montre alors que tous les opérateurs $\hat{H}_0 + W$ avec $\hat{W}$ $\hat{H}_0$ -bornés de borne suffisamment petite sont auto-adjoints. On a en effet le

Théorème 10.1 (Kato-Rellich). *Soit $\hat{H}_0$ un opérateur auto-adjoint, et $\hat{W}$ un opérateur symétrique $\hat{H}_0$ -borné, de borne relative strictement inférieure à 1. Alors $\hat{H} = \hat{H}_0 + \hat{W}$ défini sur le domaine $D(\hat{H}) = D(\hat{H}_0)$ est auto-adjoint.*

Exemple 10.1 (Potentiel coulombien). Montrons que pour le potentiel

$$V(\mathbf{r}) = -\frac{e^2 Z}{4\pi\epsilon_0 |\mathbf{r}|},$$

le Hamiltonien quantique

$$\hat{H} = -\frac{\hbar^2}{2m} \Delta + V$$

à domaine $D(\hat{H}) = H^2(\mathbb{R}^3)$ est auto-adjoint. L'opérateur $\hat{H}_0$ est ici le Hamiltonien privé de son terme potentiel :

$$\hat{H}_0 = -\frac{\hbar^2}{2m} \Delta.$$

Remarquons tout d'abord que l'opérateur $\hat{V}$ associé à la multiplication par $V(\mathbf{r})$ est symétrique de domaine $D(\hat{V}) \supset H^1(\mathbb{R}^3)$ par l'inégalité de Hardy (voir l'exercice 10.1). L'opérateur $\hat{H}$ est donc symétrique de domaine $D(\hat{H}) = D(\hat{H}_0) = H^2(\mathbb{R}^3)$.

Soit alors $u \in H^2(\mathbb{R}^3)$. L'injection de Sobolev $H^2(\mathbb{R}^3) \hookrightarrow L^\infty(\mathbb{R}^3)$ (voir par exemple [4]) donne l'existence d'une constante $C > 0$ telle que

$$\forall u \in H^2(\mathbb{R}^3), \quad \|u\|_{L^\infty(\mathbb{R}^3)} \leq C \|u\|_{H^2(\mathbb{R}^3)},$$

où on utilise la norme sur $H^2(\mathbb{R}^3)$ définie par (voir Exercice 10.2)

$$\|u\|_{H^2(\mathbb{R}^3)} = (\|u\|_{L^2(\mathbb{R}^3)}^2 + \|\Delta u\|_{L^2(\mathbb{R}^3)}^2)^{1/2}.$$

On a alors (utilisant pour la fin de cette preuve les unités atomiques, voir Appendice A.2)

$$\begin{aligned} \|Vu\|_{L^2(\mathbb{R}^3)} &= Z \left\| \frac{u}{|\mathbf{r}|} \right\|_{L^2(\mathbb{R}^3)} \\ &= Z \left(\int_{B_r} \frac{|u(\mathbf{r})|^2}{|\mathbf{r}|^2} d\mathbf{r} + \int_{\mathbb{R}^3 \setminus B_r} \frac{|u(\mathbf{r})|^2}{|\mathbf{r}|^2} d\mathbf{r} \right)^{1/2} \\ &\leq Z \left(4\pi r \|u\|_{L^\infty(\mathbb{R}^3)}^2 + \frac{1}{r^2} \|u\|_{L^2(\mathbb{R}^3)}^2 \right)^{1/2} \\ &\leq Z \left(4\pi C r \|u\|_{H^2(\mathbb{R}^3)}^2 + \frac{1}{r^2} \|u\|_{L^2(\mathbb{R}^3)}^2 \right)^{1/2} \\ &\leq Z \left(16\pi C r \|H_0 u\|_{L^2(\mathbb{R}^3)}^2 + \left(4\pi C r + \frac{1}{r^2} \right) \|u\|_{L^2(\mathbb{R}^3)}^2 \right)^{1/2} \\ &\leq \epsilon \|H_0 u\|_{L^2(\mathbb{R}^3)} + Z \left(4\pi C r + \frac{1}{r^2} \right)^{1/2} \|u\|_{L^2(\mathbb{R}^3)} \end{aligned}$$

avec le choix $r = \epsilon^2 / (16\pi C Z^2)$. Ceci montre que V est H_0 -borné de borne relative égale à 0, et H est auto-adjoint par le théorème de Kato-Rellich (Théorème 10.1).

On aurait également pu utiliser l'inégalité de Hardy pour voir que

$$\|Vu\|_{L^2(\mathbb{R}^3)} \leq \|\nabla u\|_{L^2(\mathbb{R}^3)} = \int_{\mathbb{R}^3} |\xi|^2 |\hat{u}|^2 d\xi \leq \frac{\epsilon}{2} \int_{\mathbb{R}^3} |\xi|^4 |\hat{u}|^2 d\xi + \frac{1}{2\epsilon} \int_{\mathbb{R}^3} |\hat{u}|^2 d\xi$$

pour tout $\epsilon > 0$.

Remarque 10.2. On peut en fait généraliser directement la preuve précédente au cas d'un potentiel $V \in L^2(\mathbb{R}^3) + L^\infty(\mathbb{R}^3)$ (voir Exercice 10.3).

10.2.2 Spectre essentiel

On a vu que le caractère auto-adjoint d'un Hamiltonien quantique peut se déduire du caractère auto-adjoint du Hamiltonien quantique libre $\hat{H}_0$, moyennant certaines conditions sur le potentiel $\hat{V}$. Sous des conditions plus fortes, on peut également obtenir des renseignements sur le spectre essentiel de l'opérateur $\hat{H} = \hat{H}_0 + \hat{V}$.

Théorème 10.2 (Stabilité du spectre essentiel par perturbation compacte). *Soit $\hat{H}_0$ un opérateur auto-adjoint de domaine $D(\hat{H}_0)$ et $\hat{V}$ un opérateur symétrique. On suppose que $\hat{V}$ est une perturbation $\hat{H}_0$ -compacte (i.e. $\hat{V}$ est un opérateur compact de $D(\hat{H}_0)$ muni de la norme du graphe $|u|^2 = \|\hat{H}_0 u\|^2 + \|u\|^2$ dans $\mathcal{H}$). Alors l'opérateur $\hat{H} = \hat{H}_0 + \hat{V}$ défini sur $D(\hat{H}) = D(\hat{H}_0)$ est auto-adjoint et vérifie $\sigma_{\text{ess}}(\hat{H}) = \sigma_{\text{ess}}(\hat{H}_0)$.*

On peut appliquer ce résultat par exemple lorsque le potentiel $V(\mathbf{r})$ est tel que pour tout $\epsilon > 0$, on puisse écrire $V(\mathbf{r}) = V_2(\mathbf{r}) + V_\infty(\mathbf{r})$ avec $V_2 \in L^2(\mathbb{R}^3)$, $V_\infty \in L^\infty(\mathbb{R}^3)$, et $\|V_\infty\|_{L^\infty(\mathbb{R}^3)} \leq \epsilon$. Un exemple de tel potentiel est le potentiel coulombien.

On vérifie en premier lieu que $\hat{V}$ définit un opérateur multiplicatif symétrique, auto-adjoint selon les résultats du § 10.2.1. Montrons ensuite que $\hat{V}$ est $\hat{H}_0$ -compact. On montre en premier lieu que $D(\hat{H}_0) \subset D(\hat{V})$. En effet, soit $u \in D(\hat{H}_0) = H^2(\mathbb{R}^3)$. Alors $u \in L^2(\mathbb{R}^3) \cap L^\infty(\mathbb{R}^3)$ par les injections de Sobolev, et

$$\|\hat{V}u\|_{L^2(\mathbb{R}^3)} \leq \|\hat{V}_2\|_{L^2(\mathbb{R}^3)}\|u\|_{L^\infty(\mathbb{R}^3)} + \|\hat{V}_\infty\|_{L^\infty(\mathbb{R}^3)}\|u\|_{L^2(\mathbb{R}^3)} < +\infty.$$

On montre ensuite que si (u_n) est une suite bornée dans $H^2(\mathbb{R}^3)$ (par une constante M), alors on peut extraire de $(\hat{V}u_n)$ une sous-suite convergente au sens de $L^2(\mathbb{R}^3)$. En effet, comme (u_n) est une suite bornée dans $H^2(\mathbb{R}^3)$, on peut extraire une sous-suite $(u_{n_k})_{k \geq 0}$ telle que

$$u_{n_k} \rightharpoonup u^* \quad H^2(\mathbb{R}^3) \text{ -- faible}, \quad u_{n_k} \rightarrow u^* \quad \text{presque partout.}$$

Soit $\epsilon > 0$. Une décomposition $V = V_2 + V_\infty$ avec $\|V_\infty\|_{L^\infty(\mathbb{R}^3)} \leq \epsilon/4M$ donne :

$$\begin{aligned} \|\hat{V}u_{n_k} - \hat{V}u^*\|_{L^2(\mathbb{R}^3)} &\leq \|\hat{V}_2(u_{n_k} - u^*)\|_{L^2(\mathbb{R}^3)} + \|\hat{V}_\infty\|_{L^\infty(\mathbb{R}^3)}\|u_{n_k} - u^*\|_{L^2(\mathbb{R}^3)} \\ &\leq \|\hat{V}_2(u_{n_k} - u^*)\|_{L^2(\mathbb{R}^3)} + \frac{\epsilon}{2}. \end{aligned}$$

Il suffit donc de montrer que $\|\hat{V}_2(u_{n_k} - u^*)\|_{L^2(\mathbb{R}^3)} \rightarrow 0$ lorsque $k \rightarrow +\infty$. Pour ce faire, on applique un théorème de Lebesgue en remarquant que $|V_2|^2|u_{n_k} - u^*|^2 \rightarrow 0$ presque partout, et que cette fonction est dominée par la fonction intégrable $|V_2(\mathbf{r})|^2 \sup \|u_{n_k}\|_{L^\infty(\mathbb{R}^3)} \leq C|V_2(\mathbf{r})|^2 M$.

10.2.3 Spectre ponctuel

Une manière de préciser le spectre ponctuel est de recourir au théorème suivant :

Théorème 10.3 (Principe du Min-Max). *Soit $\hat{H}$ un opérateur auto-adjoint, que l'on suppose borné inférieurement, i.e. il existe $c \in \mathbb{R}$ tel que $\hat{H} - c \geq 0$. On définit, pour $n \geq 1$,*

$$\lambda_n(\hat{H}) = \inf_{E \in \mathcal{E}_n} \sup_{\phi \in E, \|\phi\|=1} \langle \hat{H}\phi, \phi \rangle,$$

où $\mathcal{E}_n$ désigne l'ensemble des sous-espaces de $D(\hat{H})$ de dimension n . Alors, pour tout $n \geq 1$,

- soit il y a n valeurs propres (comptées avec leur multiplicité) en dessous du spectre essentiel de $\hat{H}$, et $\lambda_n(\hat{H})$ est la n -ième valeur propre de $\hat{H}$;
- soit $\lambda_n(\hat{H})$ est le bas du spectre essentiel, auquel cas il y a au plus $n - 1$ valeurs propres (comptées avec leur multiplicité) en dessous de $\lambda_n(\hat{H})$.

2. Pour un opérateur $\hat{T}$, on utilise la notation $\hat{T} \geq 0$ si $\langle \hat{T}u, u \rangle \geq 0$ pour tout $u \in D(\hat{T})$.

Application au potentiel coulombien

Nous avons montré que le Hamiltonien correspondant au potentiel coulombien a pour spectre essentiel $\sigma_{\text{ess}}(\hat{H}) = [0, +\infty[$ (voir § 10.2.2). On va à présent caractériser le spectre ponctuel. Plus précisément, nous allons montrer qu'il existe une infinité de valeurs propres négatives, bornées inférieurement. On a déjà montré au § 11.5.3 que le Hamiltonien correspondant au potentiel coulombien est borné inférieurement. Par conséquent, ces valeurs propres ont comme point d'accumulation 0 (si elles avaient un point d'accumulation strictement négatif, ce point appartiendrait au spectre essentiel, ce qui ne se peut puisque nous avons montré que $\sigma_{\text{ess}}(\hat{H}) = \mathbb{R}_+$).

Il suffit de trouver une suite de fonctions $(\psi_k)_{k \in \mathbb{N}}$ orthonormales telles que $\langle \hat{H} \psi_k, \psi_k \rangle < 0$. En effet, on aura dans ce cas $\lambda_k(\hat{H}) < 0 < \inf \sigma_{\text{ess}}(\hat{H})$ pour tout k . Une telle suite de fonctions peut être construite en prenant une fonction régulière $\phi \in C_0^\infty(\mathbb{R}^3)$, à support dans la couronne $\{1 < |\mathbf{r}| < 2\}$, et en posant

$$\psi_k(\mathbf{r}) = \frac{1}{k^{3/2}} \phi\left(\frac{\mathbf{r}}{k}\right).$$

pour $k \in \mathbb{K} = \{2^n, n \in \mathbb{N}^*\}$. On vérifie facilement que

$$\forall (k, l) \in \mathbb{K}^2, \quad \langle \psi_k, \psi_l \rangle = \delta_{kl}.$$

On calcule ensuite

$$\langle \hat{H} \psi_k, \psi_k \rangle = \frac{\hbar^2}{2mk^2} \int_{\mathbb{R}^3} |\nabla \phi(\mathbf{r})|^2 d\mathbf{r} - \frac{e^2 Z}{4\pi\epsilon_0 k} \int_{\mathbb{R}^3} \frac{|\phi(\mathbf{r})|^2}{|\mathbf{r}|} d\mathbf{r} < 0$$

si $k \geq k_0$ est assez grand. La suite $(\psi_{k_0+k})_{k \geq 0}$ convient.

La figure 10.5 résume la structure du spectre de l'atome hydrogénoïde. La forme précise des fonctions propres sera donnée au chapitre 12. On peut toutefois d'ores et déjà préciser une borne supérieure pour l'énergie de l'état fondamental. On considère pour ce faire une famille de fonctions tests $(\psi_\mu)_{\mu > 0}$, donnée par

$$\psi_\mu(\mathbf{r}) = \sqrt{\frac{\mu^3}{\pi}} e^{-\mu|\mathbf{r}|}, \quad \|\psi_\mu\|_{L^2(\mathbb{R}^3)} = 1,$$

et l'Hamiltonien de l'atome d'hydrogène ($Z = 1$)

$$\hat{H} = -\frac{\hbar^2}{2m} \Delta - \frac{e^2}{4\pi\epsilon_0 |\mathbf{r}|}.$$

Alors, passant en coordonnées sphériques, on montre que

$$\langle \psi_\mu, \hat{H} \psi_\mu \rangle = \frac{\hbar^2}{2m} \int_{\mathbb{R}^3} |\nabla \psi_\mu|^2 d\mathbf{r} - \frac{e^2}{4\pi\epsilon_0} \int_{\mathbb{R}^3} \frac{|\psi_\mu(\mathbf{r})|^2}{|\mathbf{r}|} d\mathbf{r} = \frac{\hbar^2}{2m} \mu^2 - \frac{e^2}{4\pi\epsilon_0} \mu.$$

L'expression du membre de droite atteint son minimum en $\mu = me^2/(4\pi\epsilon_0 \hbar^2)$, et ce minimum vaut $-me^4/2\hbar^2(4\pi\epsilon_0)^2$. On en déduit, par le principe du min-max (Théorème 10.3), que la première valeur propre E_0 vérifie

$$E_0 \leq -\frac{m}{2\hbar^2} \left(\frac{e^2}{4\pi\epsilon_0} \right)^2.$$

L'application numérique donne $E_0 \leq -13.6 \text{ eV}$. Cette valeur est en fait l'énergie trouvée expérimentalement pour l'état fondamental de l'électron : -13.6 eV (il se trouve en effet que l'état fondamental est de la forme des fonctions tests utilisées ici, voir § 12.5). On verra également au § 11.5.4 une borne inférieure de cette première valeur propre.

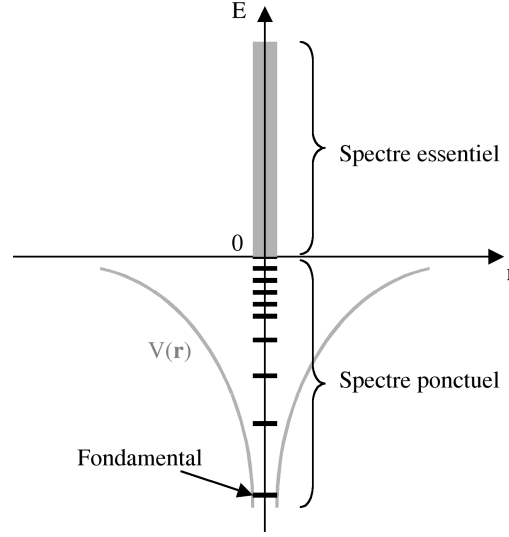


Fig. 10.5. Allure général du spectre d'énergie dans le potentiel coulombien. Les énergies négatives sont quantifiées, et correspondent au spectre ponctuel. Les énergies positives forment un continuum correspondant au spectre essentiel.

10.3 Un exemple quasi-analytique : le trou de potentiel

La théorie spectrale offre des résultats généraux permettant de préciser sans effort l'allure du spectre d'énergie des systèmes quantiques. Elle ne donne en revanche aucune information quantitative sur ceux-ci, ni sur la forme spatiale des états propres associés. Nous étudions ici un système pour lequel une analyse quantitative est possible. Nous constaterons que cette analyse demande, par rapport à l'application des théorèmes spectraux, un peu plus de travail !

Le trou de potentiel est un modèle unidimensionnel dans lequel le potentiel est nul partout, sauf sur un petit intervalle où il est négatif. Il rend compte de manière grossière de la situation courante dans laquelle une particule quantique peut se faire piéger par un centre attracteur, mais peut également par un apport suffisant d'énergie, retrouver la liberté de se déplacer partout dans l'espace. Il est adapté à la description qualitative de la plupart des situations de piégeage rencontrées dans la nature.

10.3.1 Définition

Nous nous restreignons à la dimension 1 d'espace. Le trou de potentiel est défini par (figure 10.6) :

$$V(x) = \begin{cases} 0 & \text{si } x < -a, \\ -V_0 \geq 0 & \text{si } -a \leq x \leq a, \\ 0 & \text{si } x > a. \end{cases}$$

Nous recherchons les états stationnaires d'une particule de masse m placée dans ce potentiel. La partie spatiale $\phi(x)$ de la fonction d'onde et son énergie E doivent vérifier l'équation de Schrödinger indépendante du temps :

$$-\frac{\hbar^2}{2m} \frac{d^2 \phi}{dx^2}(x) + V(x)\phi(x) = E\phi(x).$$

Le résultat de minimisation de l'énergie par le potentiel (Proposition 9.4) impose ici $E \geq -V_0$. On distingue deux situations, selon que $E \geq 0$ ou $-V_0 \leq E \leq 0$.

On vérifie facilement que le trou de potentiel est une perturbation compacte du Hamiltonien libre. Son spectre essentiel est donc $[0, +\infty[$. Nous allons le vérifier par la recherche explicite des

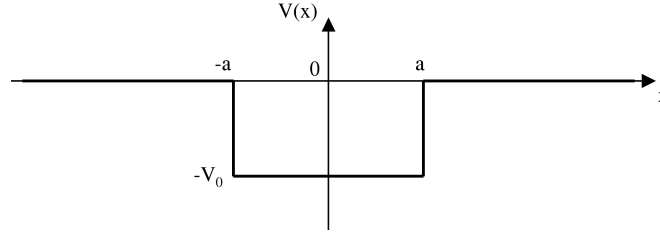


Fig. 10.6. Le trou de potentiel en dimension 1. Il décrit de manière simpliste le potentiel ressenti par une particule au voisinage d'un centre attracteur placé à l'origine.

fonctions d'onde. Ce calcul permettra également de caractériser les valeurs de l'énergie des états d'énergie $E < 0$.

10.3.2 Etats d'énergie liés

On considère $-V_0 \leq E < 0$. Dans ce cas, les éléments du spectre ne peuvent être que dans le spectre discret.

Solutions physiquement admissibles de l'équation de Schrödinger

Mathématiquement, la solution générale par morceaux de l'équations de Schrödinger est donnée par :

$$\phi(x) = \begin{cases} \alpha e^{Kx} + \beta e^{-Kx} & \text{si } x < -a, \\ \gamma e^{ikx} + \delta e^{-ikx} & \text{si } -a \leq x \leq a, \\ \sigma e^{Kx} + \mu e^{-Kx} & \text{si } x > a. \end{cases}$$

avec

$$K = \frac{\sqrt{-2mE}}{\hbar}, \quad k = \frac{\sqrt{2m(E + V_0)}}{\hbar}. \quad (10.1)$$

D'un point de vue physique, les composantes βe^{-Kx} et σe^{Kx} ne sont cependant pas acceptables, car elles conduisent à une fonction d'onde non normalisable, et qui ne peut pas non plus être une limite de suite de Weyl.³ Les fonctions $\phi(x)$ physiquement admissibles sont donc de la forme

$$\phi(x) = \begin{cases} \alpha e^{Kx} & \text{si } x < -a, \\ \gamma e^{ikx} + \delta e^{-ikx} & \text{si } -a \leq x \leq a, \\ \mu e^{-Kx} & \text{si } x > a. \end{cases}$$

Quantification des énergies

Les conditions de continuité sur $\phi(x)$ et sa dérivée en $x = \pm a$ conduisent au système de quatre équations à quatre inconnues $\alpha, \gamma, \delta, \mu$:

$$\begin{cases} \gamma e^{-ika}(ik - K) = \delta e^{ika}(ik + K), \\ \delta e^{-ika}(ik - K) = \gamma e^{ika}(ik + K), \\ \alpha = e^{Ka}(\gamma e^{-ika} + \delta e^{ika}), \\ \mu = e^{Ka}(\gamma e^{ika} + \delta e^{-ika}). \end{cases}$$

Ce système possède la solution évidente $(0, 0, 0, 0)$. Il existe donc une solution non nulle si et seulement si ce système linéaire est singulier. Les deux dernières lignes montrent que α et μ sont déterminés de manière univoque par γ et δ . C'est donc la singularité du système formé par les deux premières lignes qui doit être examinée. Une rapide étude de son déterminant montre que cette singularité est obtenue dans deux situations :

3. Comparer avec le calcul du § 9.4.

- (i) $K = k \tan(ka)$: les solutions du système global sont alors obtenues avec $\gamma = \delta \neq 0$;
- (ii) $K = -k \cot(ka)$: les solutions du système global sont alors obtenues avec $\gamma = -\delta \neq 0$.

Ces deux conditions sont en fait deux équations de *quantification de l'énergie* : elles fixent des conditions que E doit respecter pour pouvoir être associée à un état stationnaire (rappelons en effet que k et K sont données par (10.1)).

Il est commode, au lieu de résoudre ces équations (par une méthode numérique ici car il n'y a pas de solution analytique), de les discuter graphiquement, en se plaçant dans le plan (ka, Ka) (voir figure 10.7). On considère alors

$$Ka = ka \tan(ka), \quad Ka = -ka \cot(ka).$$

Par ailleurs,

$$(Ka)^2 + (ka)^2 = -\frac{2mE}{\hbar^2}a^2 + \frac{2m(E + V_0)}{\hbar^2}a^2 = \frac{2mV_0}{\hbar^2}a^2.$$

Quand E décrit $[-V_0, 0]$, le couple (ka, Ka) décrit donc, dans le quart de plan des coordonnées positives, un quart de cercle de rayon $\frac{\sqrt{2mV_0}}{\hbar}a$. Les énergies E autorisées correspondent aux intersections de cet arc de cercle avec les courbes définies par chacune des deux équations de quantification.

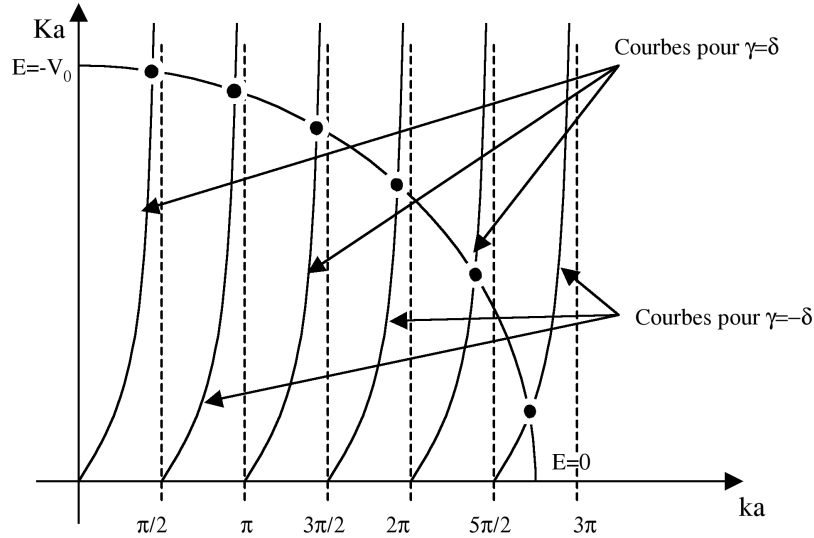


Fig. 10.7. Résolution graphique des équations de quantification dans le plan (ka, Ka) . L'arc de cercle peut être vu comme une courbe paramétrée par l'énergie E quand elle varie dans l'intervalle $[-V_0, 0]$. L'extrémité de l'arc en haut à gauche correspond à l'énergie $E = -V_0$. L'extrémité en bas à droite correspond à l'énergie $E = 0$. Les intersections de cet arc avec les autres courbes fixent les valeurs d'énergie quantifiées.

Dans l'exemple ci-dessus, 6 points d'intersection apparaissent, qui signifient l'existence de 6 états stationnaires d'énergies différentes, chacun d'ordre de multiplicité 1. Dans le cas général, on déduit de la représentation graphique que le nombre d'états stationnaires est donné par :

$$N = 1 + E \left(\frac{2a\sqrt{2mV_0}}{\pi\hbar} \right),$$

où $E(x)$ désigne la partie entière de x . En particulier, il existe toujours au moins un état stationnaire dans le trou.

Forme des états d'énergie liés

Le détail précis des calculs serait sans intérêt ici. Contentons nous d'en donner les grandes lignes. Quelle que soit la valeur de V_0 , on trouve que les états stationnaires d'énergie négative peuvent être décrits, à un déphasage global près, par des fonctions d'onde $\phi(x)$ à valeurs réelles. Nous en donnons en figure 10.8 l'allure dans une situation à 4 états stationnaires.

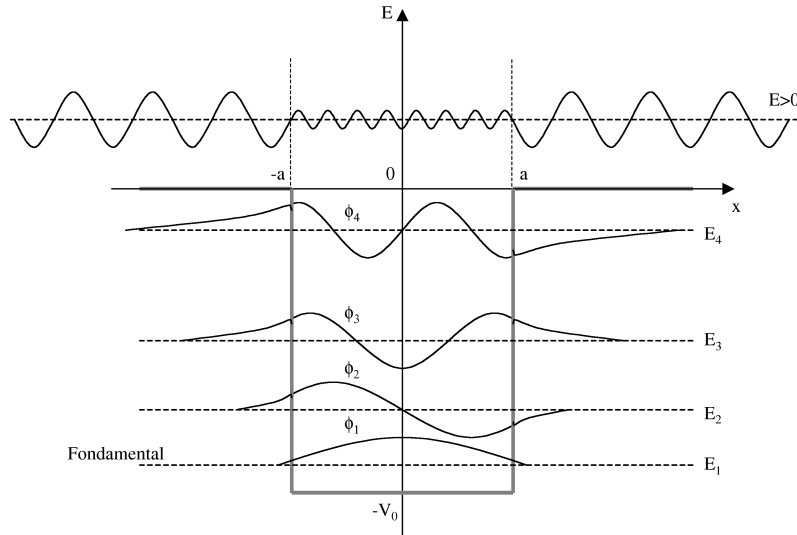


Fig. 10.8. Etats d'énergie d'un trou de potentiel possédant 4 états liés. Tous les états liés sont représentés. Noter la caractéristique alternativement symétrique et antisymétrique des fonctions d'onde associées. En haut du schéma, la fonction d'onde d'énergie positive donne l'allure typique d'un état de diffusion.

A l'intérieur du trou, les états stationnaires ont une allure similaire à celle trouvée dans le cas du puits infini. Pour autant, ces fonctions ne sont pas ici strictement contenues dans le puits. Elles peuvent s'étendre à gauche et à droite, montrant des ailes exponentielles décroissantes, d'autant plus larges que l'énergie de l'état considéré est importante. Ainsi, l'état fondamental reste relativement bien confiné, alors que l'état ϕ_4 s'étend de manière plus généreuse à l'extérieur. Ceci est une conséquence de l'effet tunnel, qui autorise la particule quantique à faire des excursions modérées là où son énergie est inférieure à l'énergie potentielle.

L'alternance de fonctions d'onde symétriques et antisymétriques n'est pas un hasard. Cette curiosité est une conséquence directe, bien que technique à démontrer, du fait que le 'paysage' de potentiel utilisé est symétrique autour du point origine de l'espace.

Tous les états obtenus décrivent une particule qui reste essentiellement localisée au voisinage du trou, avec une probabilité de présence qui tend rapidement vers 0 à l'infini. On les appelle états *liés*. En tant qu'états propres d'un opérateur auto-adjoint, ils forment une famille d'états orthonormés, et constituent une base sur laquelle peuvent se décomposer toutes les fonctions d'onde décrivant une particule piégée dans le trou. La fonction d'onde de la particule quantique placée dans un des ces états, ou dans une combinaison linéaire de ces états, a la propriété de *ne pas s'étaler* dans le temps.

10.3.3 Etats d'énergie de diffusion

On examine le cas $E > 0$. La théorie spectrale enseigne que ces valeurs de E correspondent aux éléments du spectre essentiel. La forme générale des solutions de l'équation de Schrödinger est :

$$\phi(x) = \begin{cases} \alpha e^{ik'x} + \beta e^{-ik'x} & \text{si } x < -a, \\ \gamma e^{ikx} + \delta e^{-ikx} & \text{si } -a \leq x \leq a, \\ \sigma e^{ik'x} + \mu e^{-ik'x} & \text{si } x > a. \end{cases}$$

avec

$$k' = \frac{\sqrt{2mE}}{\hbar}, \quad k = \frac{\sqrt{2m(E + V_0)}}{\hbar}.$$

On trouve de part et d'autre du trou de potentiel une superposition d'ondes de De Broglie. Elles signifient, par leur caractère complètement délocalisé, que la particule a la possibilité d'explorer tout l'espace, sans restriction. Cela correspond bien à l'intuition de la physique classique, selon laquelle une particule peut accéder à toutes les zones où son énergie E est supérieure à l'énergie potentielle. De même que les ondes de De Broglie, cette forme de solution conduit à des fonctions d'ondes qui ne sont pas normalisables. Néanmoins, comme les ondes de De Broglie, elle reste physiquement pertinente en tant que limite de suite de Weyl. Les conditions de raccordement en $x = \pm a$ s'écrivent :

$$\begin{cases} \alpha e^{-ik'a} + \beta e^{ik'a} = \gamma e^{-ika} + \delta e^{ika}, \\ k'(\alpha e^{-ik'a} - \beta e^{ik'a}) = k(\gamma e^{-ika} - \delta e^{ika}), \\ \gamma e^{ika} + \delta e^{-ika} = \sigma e^{ik'a} + \mu e^{-ik'a}, \\ k(\gamma e^{ika} - \delta e^{-ika}) = k'(\sigma e^{ik'a} - \mu e^{-ik'a}). \end{cases}$$

Il s'agit d'un système linéaire de 4 équations à 6 inconnues. L'espace des solutions est donc au moins de dimension 2. Ainsi, *toute énergie E positive possède plusieurs états stationnaires associés*. A l'inverse des états liés, pour qui apparaissait le phénomène de quantification, les états stationnaires que nous trouvons ici peuvent prendre n'importe quelle valeur d'énergie positive. Cela peut s'associer au fait physique qu'ils ne correspondent plus à des états confinés, mais à des états dans lesquels la particule explore librement tout l'espace. Ils sont appelés *états de diffusion*. L'ensemble des énergies qui leurs sont associées forme un *continuum d'énergies*. Notons les paquets d'onde construits à partir des états de diffusion s'étalent, au contraire des états liés.

On représente schématiquement en figure 10.9 le diagramme énergétique complet du trou quantique. On retrouve la forme très caractéristique prédite par la théorie spectrale : on a un continuum d'états de diffusion, associés au spectre essentiel, et qui occupe tout le domaine où la particule a une énergie suffisante pour se déplacer librement partout dans l'espace. En dessous, se trouvent quelques états liés, associés au spectre ponctuel, dont l'énergie est donc quantifiée, et qui décrivent physiquement une particule piégée dans un minimum du potentiel. Quelque soit la taille du trou (largeur a et profondeur V_0) il existe toujours au moins un niveau d'énergie quantifié.

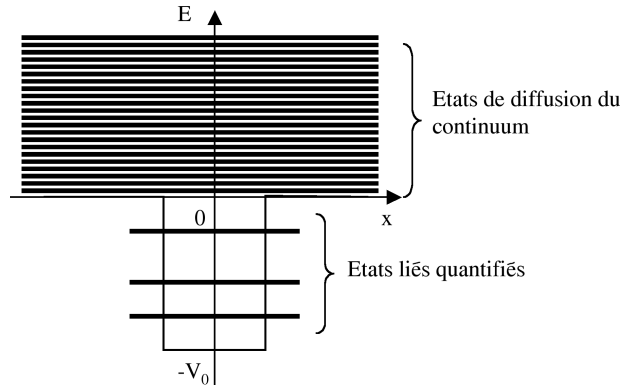


Fig. 10.9. Structure du spectre énergétique complet du trou de potentiel à une dimension. Les énergies négatives sont quantifiées. Les énergies positives sont toutes accessibles, et forment un continuum d'états de diffusion. Cette structure est à rapprocher de celle du spectre énergétique du potentiel Coulombien.

Cette forme de diagramme est de portée assez universelle. Elle se retrouve qualitativement quelque soit la forme de trou considérée (voir par exemple le spectre du potentiel Coulombien, figure 10.5).

10.4 Universalité et implications du phénomène de quantification

10.4.1 La quantification

Les états stationnaires sont les seuls états pour lesquels l'énergie de la particule est connue sans ambiguïté. Nous comprendrons au chapitre prochain que même lorsque, par exemple, une particule est dans une superposition à poids égaux de deux états stationnaires ϕ_1 et ϕ_2 d'énergies E_1 et E_2 , elle n'acquiert pas l'énergie moyenne $\frac{E_1+E_2}{2}$, mais elle possède simultanément, à poids égaux, les énergies E_1 et E_2 . Les énergies du spectre de $\hat{H}$ sont donc les *seules* autorisées.

Il ressort en particulier que lorsqu'une portion du spectre du Hamiltonien est ponctuelle, la particule n'a accès qu'aux quelques valeurs du spectre. Les autres valeurs lui sont interdites. Ce phénomène est nommé *quantification de l'énergie*.

Il a une portée très générale qui dépasse largement la seule grandeur énergie. Nous verrons qu'il existe, pour toutes les grandeurs physiques, un ensemble restreint de valeurs 'autorisées'. Comme pour l'énergie, ces valeurs sont uniquement dictées par la masse m de la particule et la forme du potentiel $V(\mathbf{r})$ dans lequel elle évolue.

10.4.2 Caractère général du phénomène de piégeage

La quantification est une conséquence du piégeage d'une particule dans un espace restreint. Dans la nature, les phénomènes de piégeage sont très courants, car les trous de potentiels sont légion :

- potentiel Coulombien des noyaux atomiques, qui retient les électrons dans les atomes et les molécules ;
- électrons piégés dans un corps solide. Les parois du solide agissent alors comme des frontières que l'électron ne peut franchir sans apport important d'énergie ;
- en astrophysique, neutrons confinés dans une étoile à neutrons.

Dans toutes ces situations, des phénomènes de quantification apparaissent, plus ou moins marqués. Souvent, le comportement des particules échappe alors complètement à la description donnée par la mécanique classique, donnant toute leur richesse aux propriétés des matériaux (électriques, thermiques, mécaniques, optiques, ...)

10.4.3 Quantification et dynamique de la particule

Solution générale de l'équation de Schrödinger dans le cas discret

Lorsque le spectre de $\hat{H}$ est purement ponctuel (cela correspond notamment au cas d'un Hamiltonien diagonalisable, par exemple auto-adjoint d'inverse compact, comme c'est le cas si le potentiel est confinant, voir sections 10.1.1 et 10.1.2), on peut résoudre l'équation de Schrödinger par une décomposition sur une base orthonormale d'états stationnaires⁴.

Considérons donc $\hat{H}$ auto-adjoint diagonalisable sur $L^2(\mathbb{R}^3)$, et une base hilbertienne d'états stationnaires $(\phi_n)_{n \geq 0} \in L^2(\mathbb{R}^3)$ d'énergies $(E_n)_{n \geq 0}$ réelles tels que

$$\hat{H}\phi_n = E_n\phi_n.$$

Pour un état initial $\psi_0 \in L^2(\mathbb{R}^3)$, on a la décomposition

4. Cette base n'étant pas unique si certaines valeurs propres ont une multiplicité supérieure ou égale à 2.

$$\psi_0(\mathbf{r}) = \sum_{n=0}^{+\infty} a_n \phi_n(\mathbf{r}), \quad (10.2)$$

où les coefficients complexes a_n sont donnés par

$$a_n = \int_{\mathbb{R}^3} \overline{\phi_n(\mathbf{r})} \psi_0(\mathbf{r}) d\mathbf{r} \in \mathbb{C}.$$

L'unique solution de l'équation de Schrödinger avec condition initiale ψ_0 est alors donnée par

$$\psi(\mathbf{r}, t) = \sum_{n=0}^{+\infty} a_n \phi_n(\mathbf{r}) e^{-i \frac{E_n}{\hbar} t}. \quad (10.3)$$

Cette formule peut être vue comme une simple conséquence du principe de superposition.

Remarque 10.3 (Schéma numérique). La décomposition de la fonction d'onde sur la base d'états stationnaires donne un schéma numérique de résolution de l'équation de Schrödinger : on peut en effet, pour une condition initiale donnée, calculer les N premiers modes propres de la condition initiale selon (10.2), et en déduire une approximation de ψ en tronquant la somme (10.3) à l'ordre N .

Remarque 10.4 (Formulation spectrale générale). Lorsque $\hat{H}$ est un opérateur auto-adjoint quelconque, et que son spectre comporte une composante continue, il est possible de généraliser la décomposition en modes propres (10.3). Partant de $\psi(\mathbf{r}, t = 0) = \psi_0(\mathbf{r})$, la théorie spectrale donne

$$\psi(\mathbf{r}, t) = \int_{\lambda \in \mathbb{R}} e^{-i \frac{\lambda}{\hbar} t} dP_\lambda \psi_0(\mathbf{r}),$$

où $(P_\lambda)_{\lambda \in \mathbb{R}}$ est la famille spectrale associée au Hamiltonien $\hat{H}$ (voir Appendice C.5).

Fréquences propres d'évolution

Calculons l'évolution temporelle de la probabilité de présence de la particule. Il vient :

$$\begin{aligned} \rho(\mathbf{r}, t) &= \left(\sum_{k=0}^{+\infty} a_k \phi_k(\mathbf{r}) e^{-i \frac{E_k}{\hbar} t} \right) \left(\sum_{l=0}^{+\infty} \overline{a_l} \overline{\phi_l(\mathbf{r})} e^{i \frac{E_l}{\hbar} t} \right) \\ &= \sum_{k,l=0}^{+\infty} a_k \overline{a_l} \phi_k(\mathbf{r}) \overline{\phi_l(\mathbf{r})} e^{i \frac{E_l - E_k}{\hbar} t}. \end{aligned}$$

La densité $\rho(\mathbf{r}, t)$ apparaît ainsi comme une superposition de composantes oscillantes de pulsations

$$\omega_{kl} = \frac{E_l - E_k}{\hbar}.$$

La probabilité de présence n'est capable, dans son évolution, de n'exhiber qu'un nombre restreint de fréquences. Il en irait en fait de même pour toutes les autres grandeurs physiques (position, accélération, courant de probabilité, ...) Les pulsations ω_{kl} sont les *pulsations propres* d'évolution du système, et sont soumises au phénomène de quantification.

10.4.4 Propriétés spectrales des atomes, des molécules, et des corps denses

Corrolaire de la quantification des fréquences propres d'évolution, la quantification de l'énergie a une conséquence directe sur la capacité d'un système à absorber ou émettre des rayonnements électromagnétiques.

Le spectre d'absorption-émission

Même si ce cours ne traite pas d'optique quantique, nous savons tous, de manière plus ou moins précise, qu'une manière quantique de regarder une onde électromagnétique de pulsation ω est de dire que celle-ci est divisée en fines particules, appelées *photons*, chacun portant individuellement une énergie

$$E_{\text{photon}} = \hbar\omega.$$

Vis à vis de l'électromagnétisme classique, l'énergie du photon correspond donc à la *fréquence* de l'onde. C'est le *nombre* de photons transportés qui détermine quant à lui le *flux d'énergie* du faisceau lumineux.

L'absorption ou l'émission d'un rayonnement par un système quantique ne peut se faire que photon par photon (voir figure 10.10). Du fait de la quantification des énergies, pour qu'un photon puisse être absorbé ou émis, il faut que son énergie corresponde à la différence entre deux niveaux d'énergie du système, c'est à dire qu'il existe 2 indices n et m tels que

$$E_{\text{photon}} = E_m - E_n = \hbar\omega_{nm}.$$

Si les énergies accessibles forment un continuum, toutes les fréquences sont ainsi susceptibles d'être absorbées ou ré-émises. En revanche, en présence du phénomène de quantification, seules les fréquences propres ω_{nm} d'évolution du système sont autorisées. Le terme courant pour les désigner est celui de *spectre d'absorption-émission* du système.

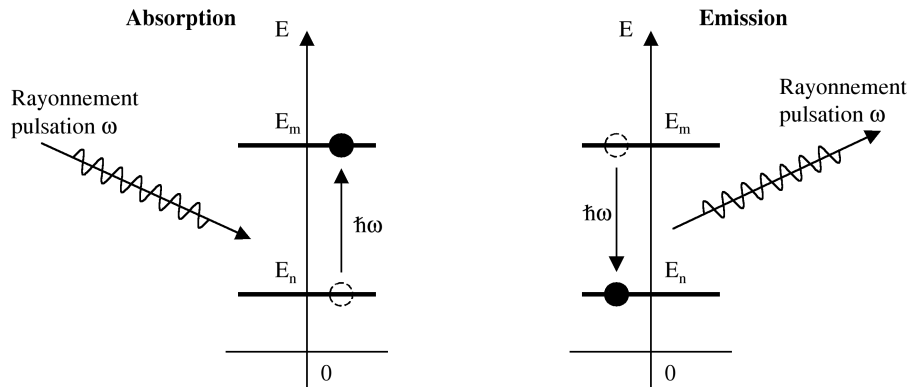


Fig. 10.10. Les mécanismes d'absorption et d'émission de rayonnement de pulsation ω par transition d'un système quantifié entre deux niveaux d'énergie E_n et E_m .

Exemple du puits quantique infini

Dans le cas du puits quantique infini unidimensionnel, de largeur a , les pulsations permises sont données par

$$\omega_{nm} = \frac{\hbar\pi^2}{2ma^2} (m^2 - n^2), \quad n, m \in \mathbb{N}^*.$$

On peut se limiter, pour n'avoir que des pulsations positives, à prendre $n \leq m$. On démontre alors (petit exercice intéressant d'arithmétique), que les pulsations obtenues sont de la forme :

$$\omega_N = \frac{\hbar\pi^2}{2ma^2} N,$$

avec comme valeurs possibles pour N :

- 0 ;
- les nombres impairs supérieurs ou égaux à 3 ;
- les multiples de 4 supérieurs ou égaux à 8.

L'allure du spectre correspondant est donnée à la figure 10.11.

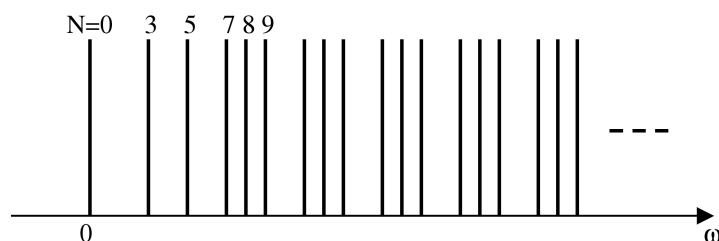


Fig. 10.11. Structure du spectre d'émission/absorption du puits de potentiel infini à une dimension.

Importance du spectre en physique et en chimie

En présence d'énergies quantifiées, le spectre prend donc, en fonction du système quantique, une structure très particulière. Tout une gamme de techniques expérimentales d'analyse de la matière, dites *techniques de spectroscopie*, retiennent ces spectres comme des *signatures* des espèces chimiques mises en jeu. On peut citer notamment :

- la spectroscopie d'absorption de rayonnement infrarouge, visible, et ultraviolet ;
- la spectroscopie d'émission par fluorescence ;
- la résonance magnétique nucléaire (voir figure 10.12).

Historiquement, les spectres d'émission-absorption de certains atomes étaient connus bien avant l'avènement de la mécanique quantique. On comprend aisément qu'ils aient beaucoup intrigué leurs observateurs à l'époque.

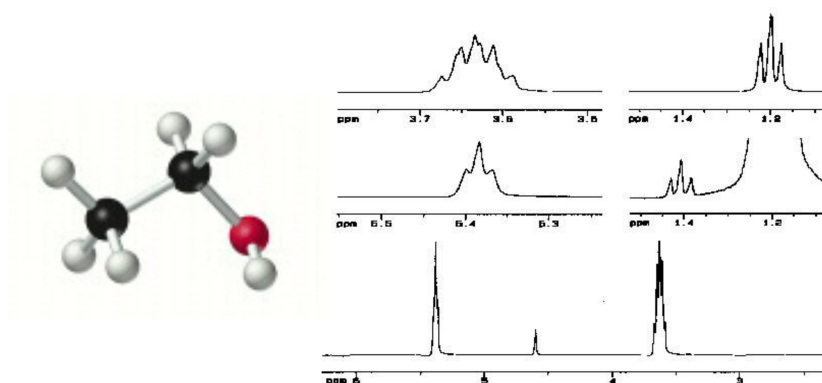


Fig. 10.12. Spectres de résonance magnétique nucléaire (RMN) de la molécule d'éthanol $\text{CH}_3\text{-CH}_2\text{-OH}$. En RMN, le système quantique étudié évolue sur un nombre fini d'états quantifiés. Le spectre ne possède ainsi qu'un nombre fini de raies.

10.4.5 La quantification comme élément stabilisateur de la matière

Nous remarquons, dans tous les exemples abordés dans ce chapitre (puits de potentiel, oscillateur harmonique, potentiel Coulombien, trou de potentiel) que l'énergie la plus basse accessible à la particule reste toujours strictement supérieure à l'énergie potentielle régnant au fond de la zone de piégeage. D'une certaine manière, le phénomène de quantification empêche la particule de tomber tout au fond du trou. Par ailleurs, nous avons vu que dans les trous de potentiel, l'existence d'états liés permet de contrecarrer le phénomène d'étalement des paquets d'onde. Ce double phénomène est lourd de conséquence, et en même temps salutaire, pour la stabilité de la matière :

- dans les atomes, il est ainsi interdit aux électrons de tomber au fond du potentiel attractif du noyau. Les électrons sont 'quantiquement' maintenus à bonne distance de ce dernier. Par

ailleurs, le potentiel retient les électrons près de lui dans des états liés. C'est cela qui explique que la structure de l'atome est stable, avec toujours une taille de l'ordre de l'Angström (voir aussi le § 11.5.4 pour une démonstration mathématique de ces considérations physiques) ;

- en astrophysique, l'étoile à neutrons est un assemblage compact de neutrons qui contient dans une sphère de 10 km de diamètre une masse équivalente à celle du soleil. Une telle masse devrait en temps normal s'effondrer encore plus sous son propre poids. Le niveau fondamental de ce corps possède en fait une énergie énorme qui oblige les neutrons à maintenir une agitation minimale même dans la limite des très basses températures, et entretient dans la sphère une pression qui compense les forces de gravitation.

10.4.6 Et la limite classique ?

La quantification des énergies prend à contre pied la mécanique classique, qui considère que cette quantité peut prendre toutes les valeurs, du moment qu'elle reste supérieure au potentiel. Pour examiner ce résultat classique avec un oeil quantique, nous pouvons utiliser la règle heuristique consistant à regarder ce que devient le spectre des énergies lorsque la masse m de la particule tend vers $+\infty$.

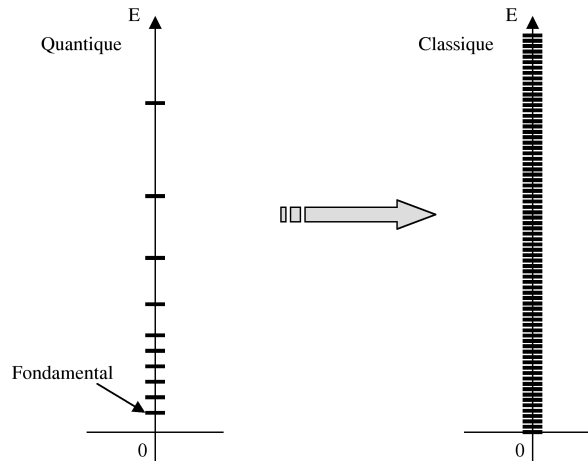


Fig. 10.13. Densification du spectre quantifié du Hamiltonien lors du passage à la limite classique (cas du puits infini et de l'oscillateur harmonique unidimensionnel).

Reprenons par exemple les énergies propres du puits quantique unidimensionnel de largeur a :

$$E_n = \frac{\hbar^2 \pi^2}{2ma^2} n^2, \quad n \geq 1.$$

Ces valeurs sont toutes espacées entre elles d'un nombre entier de fois la quantité $\frac{\hbar^2 \pi^2}{2ma^2}$. Lorsque m croît, cette quantité tend vers 0. Les énergies E_n se rapprochent donc les unes des autres (voir figure 10.13), pendant que l'énergie du niveau fondamental tend vers 0. Au final, le spectre énergétique tout entier se densifie ainsi sur le domaine $[0, +\infty[$. Cela s'observe de même dans le cas de l'oscillateur harmonique, avec les énergies

$$E_n = \hbar \sqrt{\frac{k}{m}} \left(n + \frac{1}{2} \right).$$

Dans la limite $m \rightarrow +\infty$, on a ainsi l'impression que toutes les énergies sont à nouveau accessibles au système, comme en physique classique. Pour autant, la quantification ne disparaît pas : elle n'est simplement plus mesurable.

10.4.7 Et ce n'est pas tout...

Là ne s'arrêtent pas les conséquences de la quantification. Nous aurons loisir dans la suite de ce cours de montrer qu'elle est à l'origine des propriétés chimiques des atomes et des molécules, expliquant par exemple le phénomène de liaison chimique. Elle déterminera aussi le caractère conducteur ou isolant électrique de la matière, ainsi que ses éventuelles propriétés de transparence optique. Ce phénomène est enfin au coeur des technologies de l'électronique à semi-conducteurs et ... du four à micro-ondes !

Exercices

Exercice 10.1 (Inégalité de Hardy).

On va montrer ici l'inégalité de Hardy

$$\forall u \in H^1(\mathbb{R}^3), \quad \left\| \frac{u}{|x|} \right\|_{L^2(\mathbb{R}^3)} \leq 2 \|\nabla u\|_{L^2(\mathbb{R}^3)}.$$

(1) Montrer que, dans $\mathbb{R}^d$,

$$\nabla \cdot \left(\frac{x}{|x|^2} \right) = \frac{d-1}{|x|^2}.$$

(2) En utilisant la formule ci-dessus et en effectuant une intégration par parties, montrer alors l'inégalité de Hardy (on utilisera également une inégalité de Cauchy-Schwarz).

Exercice 10.2. Montrer que la norme

$$\|u\|_{H^2(\mathbb{R}^3)} = (\|u\|_{L^2(\mathbb{R}^3)} + \|\Delta u\|_{L^2(\mathbb{R}^3)})^{1/2}$$

est bien équivalente à la norme usuelle (somme L^2 de toutes les dérivées).

Exercice 10.3. Montrer que si $u \in H^2(\mathbb{R}^3)$, alors pour tout $\epsilon > 0$, il existe $M_\epsilon > 0$ tel que

$$\|u\|_{L^\infty(\mathbb{R}^3)} \leq \epsilon \|\Delta u\|_{L^2(\mathbb{R}^3)} + M_\epsilon \|u\|_{L^2(\mathbb{R}^3)}.$$

On pourra passer en variables de Fourier, utiliser une inégalité de Cauchy-Schwarz avec $(1 + |\mathbf{k}|^2)\hat{u}$ et $(1 + |\mathbf{k}|^2)^{-1}$ pour montrer

$$\|\hat{u}\|_{L^1(\mathbb{R}^3)} \leq c (\| |\mathbf{k}|^2 u \|_{L^2(\mathbb{R}^3)} + \|u\|_{L^2(\mathbb{R}^3)}),$$

et conclure en considérant $\hat{u}_n(\mathbf{k}) = n^3 \hat{u}(n\mathbf{k})$. Montrer alors que l'affirmation de la Remarque 10.2 en s'inspirant de la preuve donnée dans l'Exemple 10.1.

Exercice 10.4 (Particule dans la boîte).

On considère dans l'espace $\mathbb{R}^3$ une particule de masse m confinée dans le domaine parallélépipédique $[0, a] \times [0, b] \times [0, c]$ avec a , b et c les longueurs des côtés de la boîte. A l'intérieur de la boîte, le potentiel est nul. Il est infini à l'extérieur.

(1) Calculer les états stationnaires qui peuvent se mettre sous la forme factorisée :

$$\phi(\mathbf{r}) = \phi_x(x)\phi_y(y)\phi_z(z)$$

ainsi que les énergies associées.

(2) Montrer que l'on peut extraire de cette famille d'états factorisés une base complète d'états propres du Hamiltonien $\hat{H}$.

Exercice 10.5 (Superposition d'états dans le puits quantique).

On considère, dans le fond d'un puits quantique unidimensionnel infini de largeur a , une particule de masse m dont l'état à $t = 0$ est donné par :

$$\psi(x, t = 0) = \frac{1}{\sqrt{2}} (\phi_1(x) + \phi_2(x)),$$

avec ϕ_1 et ϕ_2 les deux premiers états d'énergie du puits.

(1) Rappeler l'expression analytique de ϕ_1 et ϕ_2 .

(2) Donner, en explicitant ϕ_1 et ϕ_2 , l'expression de $\psi(x, t)$ pour $t > 0$.

(3) Calculer, en fonction de t , la probabilité de présence $\rho(x, t)$, le courant de probabilité $J(x, t)$ et la vitesse de probabilité $v(x, t)$.

- (4) Montrer que la particule est animée d'un mouvement périodique dont on donnera la période T .
 (5) Représenter graphiquement l'allure de $\rho(x, t)$, $J(x, t)$ et $v(x, t)$ pour $t = 0$, $t = \frac{T}{4}$, $t = \frac{T}{2}$ et $t = \frac{3T}{4}$.

Exercice 10.6 (Le puits de Dirac).

On considère une particule de masse m placée dans le champ de potentiel unidimensionnel suivant :

$$V(x) = -\alpha\delta(x), \quad \alpha \geq 0,$$

c'est à dire un potentiel nul partout, sauf au point $x = 0$, où il prend la forme d'un pic de Dirac négatif. Nous recherchons les états stationnaires d'énergie E négative.

- (1) Montrer que $\hat{H} = -\frac{\hbar^2}{2m}\Delta + V$ est autoadjoint de domaine $D(\hat{H}) = H_0^1(\mathbb{R})$. En déduire que les parties spatiales des états stationnaires de la particule restent des fonctions continues de x
 (2) Ecrire l'équation de Schrödinger indépendante du temps pour la partie spatiale $\phi(x)$ d'un état d'énergie E .
 (3) Pour $E < 0$, donner, sans se préoccuper des conditions de raccordement en $x = 0$, la forme générale des états stationnaires dans les domaines $x < 0$ et $x > 0$. Lesquels sont physiquement acceptables vis à vis de la condition de normalisation ? (*On posera $K = \frac{\sqrt{-2mE}}{\hbar}$*)
 (4) En intégrant l'équation de Schrödinger sur un petit intervalle $[-\epsilon, \epsilon]$, puis en faisant tendre ϵ vers 0, montrer que :

$$\frac{d\phi}{dx}(0^-) - \frac{d\phi}{dx}(0^+) = \frac{2m\alpha}{\hbar^2}\phi(0).$$

- (5) En tenant compte, par ailleurs, de la continuité de $\phi(x)$ en $x = 0$, en déduire une condition de quantification sur l'énergie E .
 (6) Combien existe-t-il d'états stationnaires d'énergie négative ?
 (7) Vérifier qu'aucune condition de quantification ne s'applique pour les énergies positives.

Exercice 10.7 (Le spectre de l'oscillateur harmonique est purement ponctuel). Considérons, pour $f \in L^2(\mathbb{R}^3)$, le problème

$$\text{Trouver } u \in \mathcal{D}'(\mathbb{R}^3) \text{ tel que } \hat{H}u = f,$$

posé sur le sous-espace vectoriel

$$\mathcal{V} = \{u \in L^2(\mathbb{R}^3) \mid |u|_{\mathcal{V}} < +\infty\},$$

où la semi-norme $|\cdot|_{\mathcal{V}}$ est donnée par

$$|u|_{\mathcal{V}}^2 = \int_{\mathbb{R}^3} |\mathbf{r}|^2 |u(\mathbf{r})|^2 d\mathbf{r} + \int_{\mathbb{R}^3} |\mathbf{k}|^2 |\hat{u}(\mathbf{k})|^2 d\mathbf{k}.$$

On voudrait appliquer un théorème de Lax-Milgram pour chercher $u \in \mathcal{V}$ tel que $a(u, v) = l(v)$ pour tout $v \in \mathcal{V}$, avec la forme bilinéaire $a(u, v) = \langle \hat{H}u, v \rangle$ et la forme linéaire $l(v) = \langle f, v \rangle$, et définir ainsi une application $\hat{T} : L^2(\mathbb{R}^3) \rightarrow \mathcal{V} \rightarrow L^2(\mathbb{R}^3)$, qui à f associe u .

- (1) Vérifier que $|\cdot|_{\mathcal{V}}$ est une norme.
 (2) Vérifier que a est bilinéaire continue et coercive.
 (3) Montrer que l est linéaire et continue pour la norme $|\cdot|_{\mathcal{V}}$. On pourra montrer que $\|u\|_{L^2(\mathbb{R}^3)} \leq |u|_{\mathcal{V}}$ en remarquant que

$$d \int_{\mathbb{R}^3} u^2 = -2 \int_{\mathbb{R}^3} u \mathbf{r} \cdot \nabla u ;$$

- (4) Montrer que l'injection $\mathcal{V} \rightarrow L^2(\mathbb{R}^3)$ est compacte.
 (5) Montrer finalement que $\hat{T}$ est auto-adjoint et positif et conclure.

Exercice 10.8 (Exercice de révision sur la théorie spectrale).

On considère l'espace de Hilbert $\mathcal{H} = L^2(\mathbb{R}^3)$. On définit sur $\mathcal{H}$ le Hamiltonien

$$H_{a,R} = -\frac{1}{2}\Delta + V_{a,R},$$

où le potentiel $V_{a,R}$ est défini, pour $a \in \mathbb{R}$ et $R \geq 0$, par

$$V_{a,R}(x) = \begin{cases} a & \text{si } |x| \leq R, \\ 0 & \text{si } |x| > R. \end{cases}$$

- (1) Préciser le domaine de $H_{a,R}$. Montrer que $H_{a,R}$ est autoadjoint sur $D(H_{a,R})$ (on pourra utiliser le théorème de Kato-Rellich). Que peut-on en déduire pour $\sigma(H_{a,R})$ et $\sigma_{\text{res}}(H_{a,R})$?
- (2) Quel est le spectre essentiel de $H_{a,R}$? (on pourra montrer que $V_{a,R}$ est une perturbation Δ -compacte de $H_{a,R}$).
- (3) Montrer que $\sigma_d(H_{a,R}) = \emptyset$ si $a \geq 0$ (on pourra utiliser le principe du min-max).
- (4) Montrer que $\sigma_d(H_{a,R})$ n'est pas vide lorsque $a < 0$ si R est choisi assez grand. Comment peut-on interpréter physiquement ce résultat en comparaison de la situation évoquée à la Question 3?
- (5) Comment ces résultats sont-ils modifiés si l'espace ambiant est de dimension 1 ou 2?

Le formalisme de Dirac

Les fonctions d'onde que nous avons utilisées jusqu'à présent ne sont aptes à décrire qu'une particule en translation. On montre dans ce chapitre qu'il est possible, au moyen d'une reformulation plus abstraite, d'obtenir un formalisme capable de plus de généralité. Le formalisme vectoriel introduit ici, dit encore *formalisme de Dirac*, permet, au travers d'une série de postulats, d'asseoir de manière rigoureuse l'ensemble des concepts que nous avons entrevus en mécanique ondulatoire. Il donne une description précise des phénomènes quantiques survenant lors de la mesure d'une grandeur physique, et explique, au travers des *relations d'incertitude de Heisenberg*, pourquoi certaines de ces grandeurs ne peuvent être connues simultanément. Ce formalisme permet d'étendre la théorie quantique à des situations autres que la particule en translation, et autorise par exemple la description des systèmes à plusieurs particules ou la rotation des particules sur elles-mêmes. Au plan technique, on verra que les nouveaux outils introduits permettent de mener des calculs plus efficacement.

11.1 Formalisme vectoriel	173
11.1.1 Grandeurs physiques et observables	173
11.1.2 Notations vectorielles : bra-ket	174
11.1.3 Généralité du formalisme vectoriel	175
11.2 Construction des observables : principe de correspondance	175
11.3 Equation de Schrödinger	176
11.4 La mesure	176
11.4.1 Observable à spectre discret	177
11.4.2 Observable générale	178
11.5 Commutation des observables	179
11.5.1 Crochets de commutation	179
11.5.2 Evolution d'une observable	180
11.5.3 Relations d'incertitude de Heisenberg	181
11.5.4 Commutation et stabilité de l'atome d'hydrogène	183
Exercices	185

11.1 Formalisme vectoriel

11.1.1 Grandeurs physiques et observables

On a vu quelques exemples d'expressions de moyennes d'une quantité physique dans les chapitres précédents. Rappelons par exemple, les expressions respectives des moyennes des composantes selon x de la position, l'impulsion, et l'accélération moyenne du paquet d'onde d'une particule de masse m décrite par une fonction d'onde $\psi(\mathbf{r}, t)$:

$$\begin{aligned}
X(t) &= \int_{\mathbb{R}^3} \overline{\psi(\mathbf{r}, t)} x \psi(\mathbf{r}, t) d\mathbf{r}, \\
m \frac{dX}{dt} &= \frac{\hbar}{i} \int_{\mathbb{R}^3} \overline{\psi(\mathbf{r}, t)} \cdot \frac{\partial \psi}{\partial x}(\mathbf{r}, t) d\mathbf{r}, \\
\frac{d^2 X}{dt^2} &= -\frac{1}{m} \int_{\mathbb{R}^3} \overline{\psi(\mathbf{r}, t)} \frac{\partial V}{\partial x}(\mathbf{r}, t) \psi(\mathbf{r}, t) d\mathbf{r}.
\end{aligned}$$

On peut remarquer que toutes ces quantités sont de la forme :

$$\text{moyenne}(t) = \int_{\mathbb{R}^3} \overline{\psi(\mathbf{r}, t)} \cdot \hat{A} \psi(\mathbf{r}, t) d\mathbf{r}.$$

avec $\hat{A}$ un opérateur agissant sur la fonction $\psi(\mathbf{r}, t)$. Nous avons dans le cas présent :

- pour la position, $\hat{A}\psi(\mathbf{r}, t) = x\psi(\mathbf{r}, t)$;
- pour l'impulsion, $\hat{A}\psi(\mathbf{r}, t) = \frac{\hbar}{i} \frac{\partial \psi}{\partial x}(\mathbf{r}, t)$;
- pour l'accélération, $\hat{A}\psi(\mathbf{r}, t) = -\frac{1}{m} \frac{\partial V}{\partial x}(\mathbf{r}, t) \psi(\mathbf{r}, t)$.

De tels opérateurs sont omniprésents dans le formalisme quantique. Ils sont appelés *observables*, et sont le pendant quantique des quantités scalaires usuelles de la mécanique classique.

Définition 11.1 (Observable en mécanique ondulatoire). *A chaque grandeur scalaire A de la mécanique classique est associé un opérateur auto-adjoint sur l'espace d'états $L^2(\mathbb{R}^3)$, appelé 'observable'. La valeur moyenne au temps t d'une observable $\hat{A}(t)$ pour un système physique placé dans l'état $\psi(\mathbf{r}, t)$ est notée :*

$$\langle A(t) \rangle_\psi = \int_{\mathbb{R}^3} \overline{\psi(\mathbf{r}, t)} \cdot \widehat{A(t)} \psi(\mathbf{r}, t) d\mathbf{r}.$$

Dans les exemples qui précèdent, le caractère symétrique et autoadjoint des observables position, impulsion et accélération est une propriété qui peut effectivement être démontrée (voir Exercice 11.1).

11.1.2 Notations vectorielles : bra-ket

Dans l'exemple des fonctions d'onde, nous voyons que l'état d'un système quantique est décrit par un vecteur de norme 1 dans un espace de Hilbert $\mathcal{H}$ (typiquement ici, $\mathcal{H} = L^2(\mathbb{R}^3)$). Pour désigner ce vecteur, sans faire référence à la complexité du détail de la fonction $\psi(\mathbf{r}, t)$, le formalisme vectoriel utilise la notation $|\psi\rangle$. On le prononce "ket". Le ket $|\psi(t)\rangle$ désigne donc l'état quantique $\psi(\mathbf{r}, t)$ à l'instant t . Dans ce formalisme, le vecteur conjugué de $|\psi\rangle$ est noté $\langle\psi|$. On le prononce "bra"¹. Le bra $\langle\psi(t)|$ correspond au vecteur d'état conjugué $\overline{\psi(\mathbf{r}, t)}$.

Le produit scalaire sur cet espace entre un bra $\langle\psi|$ et un ket $|\phi\rangle$ est noté

$$(\langle\psi|, |\phi\rangle) = \langle\psi|\phi\rangle \equiv \int_{\mathbb{R}^3} \overline{\psi(\mathbf{r}, t)} \cdot \phi(\mathbf{r}, t) d\mathbf{r}.$$

Le bra $\langle\psi|$ est ainsi identifié à une forme linéaire $\mathcal{H} \rightarrow \mathbb{R}$

$$\langle\psi| : |\phi\rangle \rightarrow \langle\psi|\phi\rangle.$$

La valeur moyenne d'une observable pour un système physique placé dans l'état $|\psi\rangle$ est alors donnée par

$$\langle A(t) \rangle = \left(\langle\psi|, \widehat{A(t)}|\psi\rangle \right) = \left\langle \psi | \widehat{A(t)} | \psi \right\rangle = \int_{\mathbb{R}^3} \overline{\psi(\mathbf{r}, t)} \cdot \widehat{A(t)} \psi(\mathbf{r}, t) d\mathbf{r}.$$

Le formalisme bra/ket est une manière (un peu ludique) de réécrire les formules précédentes. Mais on verra que ce formalisme se montre extrêmement puissant pour nombre de manipulations mathématiques nécessitées par la théorie quantique.

1. Le jeu de mot fondé sur l'anglais *bracket* n'aura échappé à personne. Ces appellations sont devenues l'usage en mécanique quantique.

11.1.3 Généralité du formalisme vectoriel

Dans le contexte de la mécanique ondulatoire, les notations fonctionnelle $\psi(\mathbf{r}, t)$ et vectorielle $|\psi(t)\rangle$ sont évidemment redondantes. L'utilisation de l'une ou de l'autre relève du contexte et/ou d'un choix tout à fait subjectif. Pour autant, alors que le formalisme des fonctions d'onde est spécifique de la seule description des particules en translation, le formalisme vectoriel se généralise à la description de systèmes quantiques plus généraux.

Postulat 11.1 (Description vectorielle des systèmes quantiques).

- (1) *Tout système physique voit ses états décrits par les vecteurs $|\psi\rangle$ de la sphère unité d'un espace de Hilbert $\mathcal{H}$. La nature précise de $\mathcal{H}$, ainsi que sa dimension, varient selon la nature du système considéré.*
- (2) *A chaque grandeur physique scalaire classique A se rapportant au système est associé un opérateur auto-adjoint $\hat{A}$ de $\mathcal{H}$, appelé 'observable associée'. Cet opérateur est tel que la valeur moyenne de la grandeur physique pour un système dans l'état $|\psi\rangle$ est :*

$$\langle A(t) \rangle = \langle \psi | \hat{A}(t) | \psi \rangle.$$

Ainsi, la rotation d'une particule physique sur elle-même (le *spin*), que nous abordons au chapitre 12, se décrit dans un espace de Hilbert judicieusement choisi de dimension finie. Ce sont les vecteurs $|\psi\rangle$ de norme unité de cet espace qui deviennent les représentants des états physiques accessibles au système.

11.2 Construction des observables : principe de correspondance

Dans le cas de la mécanique ondulatoire, les calculs que nous avons menés avec les fonctions d'onde ont permis d'identifier l'expression de quelques observables. Notamment, celles associées aux composantes des vecteurs position et impulsion, respectivement notées $\hat{x}$, $\hat{y}$, $\hat{z}$ et $\hat{p}_x$, $\hat{p}_y$, $\hat{p}_z$, s'expriment selon

$$\begin{cases} \hat{x} : \psi(\mathbf{r}, t) \rightarrow x \psi(\mathbf{r}, t), \\ \hat{y} : \psi(\mathbf{r}, t) \rightarrow y \psi(\mathbf{r}, t), \\ \hat{z} : \psi(\mathbf{r}, t) \rightarrow z \psi(\mathbf{r}, t), \end{cases} \quad \begin{cases} \hat{p}_x : \psi(\mathbf{r}, t) \rightarrow \frac{\hbar}{i} \frac{\partial \psi}{\partial x}(\mathbf{r}, t), \\ \hat{p}_y : \psi(\mathbf{r}, t) \rightarrow \frac{\hbar}{i} \frac{\partial \psi}{\partial y}(\mathbf{r}, t), \\ \hat{p}_z : \psi(\mathbf{r}, t) \rightarrow \frac{\hbar}{i} \frac{\partial \psi}{\partial z}(\mathbf{r}, t). \end{cases}$$

Il n'existe malheureusement pas de principe ni de méthode générale pour fixer la forme des observables relevant d'un système quantique quelconque : seul l'empirisme est de rigueur. . . Toutefois, lorsque pour un système donné, quelques observables sont déjà connues, une règle de construction se dégage. Nous en donnons un énoncé dans le contexte de la mécanique ondulatoire.

Postulat 11.2 (Principe de correspondance en mécanique ondulatoire). *Pour toute grandeur classique qui s'exprime comme une fonction des vecteurs position $\mathbf{r}$ et impulsion $\mathbf{p}$, l'observable associée s'obtient en remplaçant dans l'expression classique les composantes x , y , z et p_x , p_y , p_z par les observables $\hat{x}$, $\hat{y}$, $\hat{z}$ et $\hat{p}_x$, $\hat{p}_y$, $\hat{p}_z$.*

Remarque 11.1. On fera bien attention au fait que les observables ne commutent pas en général ! Ainsi par exemple, $\hat{x}\hat{p}_x \neq \hat{p}_x\hat{x}$. L'ordre dans les notations de type $\hat{A}\hat{B}$ est donc important. Pour appliquer le principe de correspondance à la quantité classique $x p_x$, la théorie quantique propose, pour conserver à l'observable son caractère auto-adjoint, de symétriser son expression selon :

$$\widehat{x p_x} = \frac{\hat{x}\hat{p}_x + \hat{p}_x\hat{x}}{2}.$$

Ces considérations ne tiennent que pour des formules relativement simples. Dans le cas d'une fonction générale $f(x, p)$, trouver l'observable associée peut demander beaucoup plus de travail.

Ce principe se généralise à tous les autres espaces vectoriels utilisés en mécanique quantique, à cela près que les 'observables de référence' ne sont alors plus la position et l'impulsion, mais d'autres opérateurs adaptés à chaque situation. Pour un système physique quelconque, seule la construction des quelques premières observables de référence relève donc d'un véritable travail de recherche tant expérimentale que théorique, les observables suivantes s'obtenant, pour la plupart, par déduction.

En mécanique ondulatoire, l'application de ce principe à quelques grandeurs usuelles mène au Tableau 11.1. Notons qu'il n'existe pas en tant que telle d'observable 'vectorielle'. Aux grandeurs classiques vectorielles sont en fait systématiquement associées trois observables (une par composante).

Tableau 11.1. Principe de correspondance entre les grandeurs physiques et les observables quantiques associées.

Grandeur physique	Notation classique $a(\mathbf{r}, \mathbf{p})$	Observable $\hat{A}$
Position	x, y, z	multiplication par x, y, z
Impulsion	p_x, p_y, p_z	$\hat{p}_x = -i\hbar\partial_x, \hat{p}_y = -i\hbar\partial_y, \hat{p}_z = -i\hbar\partial_z$
Energie cinétique	$E_c = \frac{\mathbf{p}^2}{2m}$	$\hat{E}_c = -\frac{\hbar^2}{2m}\Delta$
Energie potentielle	$V(\mathbf{r}, t)$	$\widehat{V(t)} =$ multiplication par $V(\mathbf{r}, t)$
Energie totale	$E = E_c + V(\mathbf{r}, t)$	$\widehat{H(t)} = -\frac{\hbar^2}{2m}\Delta + V(\mathbf{r}, t)$
Force (dérivée du potentiel)	$\mathbf{F} = -\nabla V$	
Composantes F_x, F_y, F_z	$-\frac{\partial V}{\partial x}, -\frac{\partial V}{\partial y}, -\frac{\partial V}{\partial z}$	multiplication par $-\frac{\partial V}{\partial x}, -\frac{\partial V}{\partial y}, -\frac{\partial V}{\partial z}$
Moment cinétique total	$\mathbf{L} = \mathbf{r} \wedge \mathbf{p}$	
Moment cinétique selon x	$L_x = yp_z - zp_y$	$\hat{L}_x = -i\hbar(y\partial_z - z\partial_y)$
Moment cinétique selon y	$L_y = zp_x - xp_z$	$\hat{L}_y = -i\hbar(z\partial_x - x\partial_z)$
Moment cinétique selon z	$L_z = xp_y - yp_x$	$\hat{L}_z = -i\hbar(x\partial_y - y\partial_x)$

11.3 Equation de Schrödinger

Le Tableau 11.1 montre qu'en mécanique ondulatoire, l'observable énergie du système est donnée par l'opérateur

$$\widehat{H(t)} = \frac{\hat{p}^2}{2m} + \hat{V}(t) = -\frac{\hbar^2}{2m}\Delta + V(\mathbf{r}, t).$$

Cet opérateur est précisément celui qui apparaît dans l'équation de Schrödinger. Cette propriété s'érige en principe général, et se transpose à tous les systèmes quantiques, quelque soit leur espace de description.

Postulat 11.3 (Hamiltonien et équation de Schrödinger). *L'évolution du vecteur d'état $|\psi\rangle$ de tout système physique est régie par l'observable $\hat{H}$ associée à l'énergie totale, selon l'équation de Schrödinger :*

$$i\hbar\frac{\partial}{\partial t}|\psi\rangle = \widehat{H(t)}|\psi\rangle. \quad (11.1)$$

L'opérateur $\hat{H}$ est appelé 'opérateur Hamiltonien' ou encore 'Hamiltonien' du système.

11.4 La mesure

La théorie spectrale s'applique aux opérateurs que sont les observables. Ce faisant, elle permet de préciser un postulat fondamental de la mécanique quantique, le *postulat de la mesure*. Curieusement, ce postulat ne possède un énoncé rigoureux que dans le cas où le spectre de l'observable

est complètement discret. C'est à cette situation que nous nous intéressons en premier lieu. Dans un second temps, nous discutons de la difficulté à généraliser cet énoncé au cas d'observables dont le spectre possède des composantes continues.

11.4.1 Observable à spectre discret

On rappelle que la mesure spectrale associée à une observable $\hat{A}(t)$ de spectre purement discret est définie à l'aide de la famille de projecteurs

$$\hat{P}_\lambda = \sum_{\mu \leq \lambda, \mu \in \sigma(\hat{A}(t))} \hat{P}_{E_\mu}, \quad \lambda \in \mathbb{R},$$

où E_μ est l'espace propre associé à la valeur propre μ et $\hat{P}_{E_\mu}$ est le projecteur orthogonal sur E_μ (voir Appendice C.5).

Postulat 11.4 (Mesure d'une grandeur physique : cas d'une observable à spectre discret). *Soit un système physique placé à l'instant t dans un état $|\psi_{\text{avant}}\rangle$, et une observable $\hat{A}(t)$ de spectre discret se rapportant à ce système. Alors :*

- (1) *lorsque l'on mesure à l'instant t la grandeur physique A , on ne peut trouver qu'un point du spectre de $\hat{A}(t)$;*
- (2) *le résultat de la mesure résulte d'un tirage aléatoire parmi les valeurs propres de $\hat{A}(t)$. La probabilité p_λ qu'une valeur propre λ donnée soit tirée vaut :*

$$p_\lambda = \|\hat{P}_{E_\lambda} |\psi_{\text{avant}}\rangle\|^2 ;$$

- (3) *si λ est la valeur propre tirée lors de la mesure, l'état du système immédiatement après la mesure devient :*

$$|\psi_{\text{après}}\rangle = \frac{1}{\sqrt{p_\lambda}} \hat{P}_{E_\lambda} |\psi_{\text{avant}}\rangle .$$

Il apparaît au travers de ce postulat que le résultat d'une mesure est fondamentalement probabiliste. Le choix se fait au prorata du carré des poids de décomposition de l'état $|\psi_{\text{avant}}\rangle$ sur les divers sous espaces propres de l'observable $\hat{A}$. La Définition C.11 et le Théorème C.3 assurent que la somme des probabilités ainsi obtenues pour les diverses valeurs propres est bien égale à 1 :

$$\sum_{\lambda \in \sigma(\hat{A}(t))} p_\lambda = \sum_{\lambda \in \sigma(\hat{A}(t))} \|\hat{P}_{E_\lambda} |\psi_{\text{avant}}\rangle\|^2 = \|\psi_{\text{avant}}\|^2 = 1$$

Ce principe choqua les physiciens du début du 20^{ème} siècle. L'histoire dit qu'il ne fut notamment jamais accepté par Einstein qui aurait déclaré : "*Dieu ne joue pas aux dés !*" Et pourtant, un siècle de physique à la clé, force est de constater que... ça marche !

La seule situation dans laquelle le résultat d'une mesure est complètement prédictible est celui où le ket $|\psi_{\text{avant}}\rangle$ est déjà un vecteur propre de $\hat{A}$. C'est le cas, par exemple, des états stationnaires $|\phi\rangle$ des Hamiltoniens $\hat{H}$ indépendants du temps. Construits par définition comme des états propres selon

$$\hat{H} |\phi\rangle = E |\phi\rangle ,$$

ce sont les seuls états pour lesquels l'énergie totale du système est connue sans ambiguïté.

S'il faut se résoudre au fait que le résultat d'une mesure est probabiliste, on peut en revanche remarquer que si l'on répète un grand nombre de fois la même expérience, en partant des mêmes conditions initiales, le résultat moyen de la mesure est bien $\langle A \rangle$, avec un écart quadratique moyen $\Delta A = \sqrt{\text{Var}(A)}$. La variance $\text{Var}(A)$ sur les résultats de mesure d'une observable $\hat{A}$ est pour un état initial $|\psi\rangle$ est

$$\text{Var}(A) \equiv \langle (A - \langle A \rangle)^2 \rangle = \langle A^2 \rangle - \langle A \rangle^2 = \langle \psi | \hat{A}^2 | \psi \rangle - \langle \psi | \hat{A} | \psi \rangle^2 .$$

Ceci provient du fait que l'opérateur $\hat{A}$ peut s'écrire (voir Appendice C)

$$\hat{A} = \sum_{\lambda \in \sigma(\hat{A})} \lambda \hat{P}_{E_\lambda},$$

et plus généralement,

$$\hat{A}^n = \sum_{\lambda \in \sigma(\hat{A})} \lambda^n \hat{P}_{E_\lambda}.$$

Ainsi, le moment d'ordre n sur le résultat de la mesure s'exprime de manière générale :

$$\sum_{\lambda \in \sigma(\hat{A})} p_\lambda \lambda^n = \langle \psi | \hat{A}^n | \psi \rangle.$$

Une autre nouveauté est le fait que la mesure engendre une *modification de l'état du système*, qui se trouve instantanément projeté et renormalisé selon un vecteur propre associé au résultat trouvé. Ce phénomène porte le nom de *réduction du paquet d'onde*. Aussi choquant qu'il puisse paraître, on peut toutefois se rassurer en se disant que la projection engendrée par la mesure assure que si on remesure le système immédiatement après une première mesure, on retrouve le même résultat avec une probabilité égale à 1. Encore mal comprise d'un point de vue physique, la réduction du paquet d'onde fait toujours l'objet de recherches au plan expérimental, visant à en percer la véritable nature.

A cause du phénomène de réduction du paquet d'onde, la mesure simultanée de deux grandeurs physiques A et B peut prendre des allures insolubles : si on observe d'abord la grandeur A , on va projeter l'état du système dans un sous espace propre de $\hat{A}$. Si on veut ensuite mesurer la quantité B ce nouveau vecteur d'état sera alors projeté une seconde fois selon un sous espace propre de $\hat{B}$. Rien ne garantit alors que ce nouvel état soit encore un état propre de $\hat{A}$. Si on remesure A , on aura ainsi en général une valeur différente de la première valeur mesurée ! L'observation selon B perturbe le résultat de la mesure selon A ...

Une exception importante concerne cependant le cas où les observables $\hat{A}$ et $\hat{B}$ commutent. En effet, les sous espaces de $\hat{A}$ sont alors stables par la projection perpendiculaire sur les sous espaces de $\hat{B}$, et réciproquement. Plusieurs mesures successives de A et B donnent alors toujours le même résultat, quelque soit l'ordre des mesures. On parle, selon que les observables commutent ou ne commutent pas, de mesures *compatibles* ou *incompatibles*.

11.4.2 Observable générale

Certaines parties du postulat de la mesure (Postulat 11.4) se transposent assez directement au cas d'observables auto-adjointes générales, dont le spectre peut comporter des composantes continues. Ces parties sont celles qui traitent des résultats possibles et de leurs probabilités de tirages. Nous faisons appel ci-dessous à la notion de mesure spectrale d'un opérateur auto-adjoint quelconque (voir Appendice C.5).

Postulat 11.5 (Valeurs mesurables d'une grandeur physique : cas général). *Pour une grandeur physique A , et son observable $\hat{A}$ associée sur $\mathcal{H}$, on a les propriétés suivantes :*

- (1) *lorsque l'on mesure la grandeur physique A , on ne peut trouver qu'un point du spectre de $\hat{A}$;*
- (2) *si le système se trouve à l'instant t précédant la mesure dans l'état $|\psi_{\text{avant}}\rangle$, alors le résultat de la mesure est dans l'intervalle $]a, b]$ avec une probabilité $\|(\hat{P}_b^A - \hat{P}_a^A) \cdot |\psi_{\text{avant}}\rangle\|^2$, où $(\hat{P}_\lambda^A)_{\lambda \in \mathbb{R}}$ est la famille spectrale associée à $\hat{A}$.*

On retrouve bien l'idée d'un tirage aléatoire du résultat de la mesure parmi les points du spectre. La véritable adaptation de l'énoncé concerne la généralisation de la notion de poids de décomposition de $|\psi_{\text{avant}}\rangle$ sur les sous-espaces propres de $\hat{A}$.

Concernant la nature de l'état $|\psi_{\text{apres}}\rangle$, tous les énoncés que nous connaissons [28, 13, 6] font intervenir une certaine erreur de mesure de l'appareil physique, contrairement au cas d'observables à spectre purement discret... mais nécessaire pour la normalisation de la fonction d'onde après mesure.

Tentative de postulat (Réduction du paquet d'onde : cas général). Pour une grandeur physique A , et son observable $\hat{A}$ associée sur $\mathcal{H}$, si le résultat de la mesure est a à Δa près, l'état du système après la mesure est

$$|\psi_{\text{apres}}\rangle = \frac{(\hat{P}_{a+\Delta a}^A - \hat{P}_{a-\Delta a}^A) \cdot |\psi_{\text{avant}}\rangle}{\|(\hat{P}_{a+\Delta a}^A - \hat{P}_{a-\Delta a}^A) \cdot |\psi_{\text{avant}}\rangle\|}.$$

L'erreur de mesure est motivée par le fait que tous les appareils donnent en effet en pratique un résultat de mesure avec une précision finie (penser aux fentes d'Young par exemple, qui, si étroites soient-elles, ont toujours une largeur donnée). Cet énoncé soulève cependant un certain nombre de questions :

- faire dépendre $|\psi_{\text{apres}}\rangle$ de l'erreur de mesure Δa est une idée angoissante voire inacceptable pour nombre de physiciens : il est en effet possible, pour la mesure d'observables à spectres purement discret de supposer que les appareils sont de précision infinie, pourquoi devrait-il en être autrement dans le cas général ? Et si cette erreur doit effectivement être prise en compte, pourquoi ne concerne-t-elle que la nature de $|\psi_{\text{apres}}\rangle$, alors qu'elle n'influe pas sur les probabilités de tirage des valeurs mesurées ?
- cependant, sans cette notion de précision finie, on ne pourrait pas normaliser les paquets d'onde lorsqu'on les projette en utilisant la mesure spectrale sur le spectre continu de l'observable. Un cas pratique où se pose cette difficulté est celui où on mesure la position moyenne x_0 d'un paquet d'onde, avec une précision Δx_0 . Dans ce cas,

$$\psi_{\text{apres}}(t, x) = \frac{\mathbf{1}_{|x-x_0| \leq \Delta x_0} \psi_{\text{avant}}(t, x)}{\|\mathbf{1}_{|\cdot-x_0| \leq \Delta x_0} \psi_{\text{avant}}(t, \cdot)\|}.$$

On voit que cette expression n'est pas définie dans la limite $\Delta x_0 \rightarrow 0$. Pour autant que certains puissent la juger suspecte, l'erreur de mesure est nécessaire pour assurer que $|\psi_{\text{apres}}\rangle$ reste physiquement acceptable.

La physique quantique et la physique en général sont riches de tels débats quasi-philosophiques, où seule la démonstration expérimentale peut faire figure de juge de paix. Malheureusement, la difficulté à concevoir et à mettre en place les bonnes expériences fait que l'heure du jugement peut parfois se faire attendre plusieurs décennies.

11.5 Commutation des observables

11.5.1 Crochets de commutation

Nous l'avons déjà mentionné dans le § 11.2, le produit de deux observables n'est pas toujours commutatif. Prenons par exemple le cas du produit $\hat{x}\hat{p}_x$, et appliquons le à la fonction d'onde $\psi(\mathbf{r})$. On a alors

$$\hat{x}\hat{p}_x\psi = x \left(\frac{\hbar}{i} \frac{\partial \psi}{\partial x} \right) = \frac{\hbar}{i} x \frac{\partial \psi}{\partial x} \neq \hat{p}_x \hat{x} \psi = \frac{\hbar}{i} \frac{\partial}{\partial x} (x\psi) = \frac{\hbar}{i} x \frac{\partial \psi}{\partial x} + \frac{\hbar}{i} \psi.$$

Les *crochets de commutation* permettent de caractériser les propriétés de commutation ou de non-commutation des observables.

Définition 11.2 (Crochets de commutation). Etant donnés deux opérateurs $\hat{A}$ et $\hat{B}$, on appelle 'crochet de commutation' de $\hat{A}$ et $\hat{B}$ l'opérateur défini par :

$$[\hat{A}, \hat{B}] = \hat{A}\hat{B} - \hat{B}\hat{A}.$$

Lorsque ce crochet de commutation est l'opérateur nul, cela signifie que les opérateurs $\hat{A}$ et $\hat{B}$ commutent.

Lorsque $\hat{A}$ et $\hat{B}$ sont des observables, c'est à dire, des opérateurs auto-adjoints, on montre que le crochet de commutation est encore un opérateur autoadjoint, et que son spectre est à valeurs imaginaires pures (voir Exercice 11.4). On établit sans peine les relations suivantes :

Propriété 11.1. *Pour trois opérateurs quelconques $\hat{A}$, $\hat{B}$ et $\hat{C}$ opérant sur le même espace de Hilbert,*

$$\begin{aligned} [\hat{A}, \hat{B}] &= -[\hat{B}, \hat{A}], \\ [\hat{A}, \hat{B} + \hat{C}] &= [\hat{A}, \hat{B}] + [\hat{A}, \hat{C}], \\ [\hat{A} + \hat{B}, \hat{C}] &= [\hat{A}, \hat{C}] + [\hat{B}, \hat{C}]. \end{aligned}$$

Commutation position-impulsion

Les relations de commutation entre les observables position et impulsion sont à la base de nombre de calculs en mécanique quantique. Il est utile de les connaître par coeur :

$$\begin{aligned} [\hat{x}, \hat{x}] &= 0, & [\hat{x}, \hat{y}] &= 0, & [\hat{x}, \hat{z}] &= 0, & [\hat{p}_x, \hat{p}_x] &= 0, & [\hat{p}_x, \hat{p}_y] &= 0, & [\hat{p}_x, \hat{p}_z] &= 0, \\ [\hat{y}, \hat{y}] &= 0, & [\hat{y}, \hat{z}] &= 0, & & & [\hat{p}_y, \hat{p}_y] &= 0, & [\hat{p}_y, \hat{p}_z] &= 0, & & \\ [\hat{z}, \hat{z}] &= 0, & & & & & [\hat{p}_z, \hat{p}_z] &= 0, & & & & \\ [\hat{x}, \hat{p}_x] &= i\hbar, & [\hat{x}, \hat{p}_y] &= 0, & [\hat{x}, \hat{p}_z] &= 0, \\ [\hat{y}, \hat{p}_x] &= 0, & [\hat{y}, \hat{p}_y] &= i\hbar, & [\hat{y}, \hat{p}_z] &= 0, \\ [\hat{z}, \hat{p}_x] &= 0, & [\hat{z}, \hat{p}_y] &= 0, & [\hat{z}, \hat{p}_z] &= i\hbar. \end{aligned}$$

Dans ces formules, $i\hbar$ désigne par abus de notation l'opérateur *multiplication par le scalaire $i\hbar$* .

11.5.2 Evolution d'une observable

Le commutateur entre une observable donnée et le Hamiltonien détermine à elle seule l'évolution temporelle de la valeur moyenne de cet observable.

Proposition 11.1 (Théorème d'Ehrenfest). *Soit un système physique dont on note $|\psi(t)\rangle$ l'état au cours du temps, $\hat{H}(t)$ le Hamiltonien autoadjoint sous lequel il évolue, et $\hat{A}(t)$ une observable. On suppose que le problème d'évolution temporelle est bien posé². Alors la valeur moyenne $\langle \hat{A}(t) \rangle_\psi = \langle \psi(t) | \hat{A}(t) | \psi(t) \rangle$ évolue au cours du temps selon*

$$\frac{d\langle \hat{A}(t) \rangle_\psi}{dt} = \left\langle \frac{\partial \hat{A}}{\partial t}(t) \right\rangle_\psi + \frac{i}{\hbar} \langle [\hat{H}(t), \hat{A}(t)] \rangle_\psi. \quad (11.2)$$

Preuve. Cette égalité découle simplement de l'équation de Schrödinger $\partial_t |\psi\rangle = \frac{1}{i\hbar} \hat{H} |\psi\rangle$ et (par conjugaison) $\partial_t \langle \psi| = -\frac{1}{i\hbar} \langle \psi| \hat{H}$. Il vient alors

$$\begin{aligned} \frac{d\langle \hat{A}(t) \rangle_\psi}{dt} &= \frac{d}{dt} \langle \psi(t) | \hat{A}(t) | \psi(t) \rangle = \left(\frac{d\langle \psi|}{dt} \right) \hat{A}(t) | \psi \rangle + \langle \psi | \left(\frac{d\hat{A}}{dt} \right) | \psi \rangle + \langle \psi | \hat{A}(t) \left(\frac{d|\psi\rangle}{dt} \right) \\ &= \langle \psi(t) | \left(\frac{d\hat{A}}{dt} \right) | \psi(t) \rangle + \frac{i}{\hbar} \left(\langle \psi(t) | \hat{H}(t) \hat{A}(t) | \psi(t) \rangle - \langle \psi(t) | \hat{A}(t) \hat{H}(t) | \psi(t) \rangle \right), \end{aligned}$$

ce qui donne le résultat. □

2. Ceci est par exemple le cas si $\hat{H}(t) \equiv \hat{H}$ constant avec $\hat{H}$ autoadjoint.

Remarque 11.2. Il découle de ce théorème que toute grandeur physique dont l'observable est indépendante du temps et commute avec le Hamiltonien voit sa valeur moyenne conservée au cours du temps. Parmi les grandeurs ne dépendant pas du temps, seules celles qui ne commutent pas avec $\hat{H}$ sont susceptibles d'évoluer. Comme $[\hat{H}, \hat{H}] = 0$, on montre que la dynamique de Schrödinger conserve l'énergie moyenne du système pour autant que le Hamiltonien ne dépende pas du temps.

Application : dynamique de la position moyenne d'un paquet d'onde

Pour la particule de masse m se déplaçant dans un champ de potentiel $V(\mathbf{r}, t)$, on montre facilement que

$$\frac{i}{\hbar}[\hat{H}, \hat{x}] = -\frac{i\hbar}{2m}(\Delta(x\cdot) - x\Delta) = -\frac{i\hbar}{m}\nabla_x,$$

On montre de même que :

$$\frac{i}{\hbar}[\hat{H}, \hat{p}_x] = \frac{i}{\hbar}[\hat{V}, \hat{p}_x] = -\nabla_x V.$$

On a bien sûr des relations analogues selon y et z . On en déduit, par le théorème d'Ehrenfest (Proposition 11.2),

$$\begin{cases} \frac{d}{dt}\langle x \rangle_\psi = \frac{\langle p_x \rangle_\psi}{m}, \\ \frac{d}{dt}\langle p_x \rangle_\psi = \left\langle -\frac{\partial V}{\partial x} \right\rangle_\psi. \end{cases} \quad (11.3)$$

Ainsi que des relations similaires selon y et z . Nous laissons le lecteur libre d'apprécier la facilité avec laquelle on retrouve le résultat du § 9.2.

11.5.3 Relations d'incertitude de Heisenberg

Une des conséquences du principe de la mesure d'une observable $\hat{A}$ est que le résultat de la mesure connaît une dispersion statistique d'écart type ΔA :

$$\Delta A = \sqrt{\text{Var}(A)} = \sqrt{\langle A^2 \rangle_\psi - \langle A \rangle_\psi^2}.$$

La quantité ΔA est aussi appelée *incertitude* de la mesure. Elle traduit uniquement le fait que l'état $|\psi\rangle$ peut être réparti sur plusieurs sous espaces propres de $\hat{A}$, et n'est liée d'aucune manière à une éventuelle limite de précision de l'appareil servant à la mesure. Le seul cas où cette incertitude est nulle est celui où $|\psi\rangle$ est état propre de $\hat{A}$. On a alors $\Delta A = 0$, et le résultat de la mesure est certain et égal à $\langle \psi | \hat{A} | \psi \rangle$ dans ce cas.

Il se trouve que pour un état $|\psi\rangle$ donné, les incertitudes ΔA et ΔB obtenus pour deux observables différents $\hat{A}$ et $\hat{B}$ ne sont pas totalement indépendantes³. On a en effet la

Théorème 11.1 (Relation d'incertitude). *Pour deux observables $\hat{A}$, $\hat{B}$ données et $|\psi\rangle \in D(\hat{A}) \cap D(\hat{B})$,*

$$\Delta A \Delta B \geq \frac{1}{2} \left| \left\langle [\hat{A}, \hat{B}] \right\rangle_\psi \right|.$$

3. L'idée ici n'est pas de mesurer $\hat{A}$ et $\hat{B}$ simultanément, mais de mesurer au choix $\hat{A}$ ou $\hat{B}$, le système étant dans un état initial $|\psi\rangle$ donné, et de se poser la question de l'incertitude obtenue selon que l'un ou l'autre des deux observables soit choisi. La théorie quantique rejette d'ailleurs l'idée de mesures simultanées. Dire que deux mesures ont lieu au même moment signifiera au mieux que l'on effectue *l'une, puis l'autre*, de manière très rapprochée dans le temps, et dans un ordre donné qui a toute son importance. Ces mesures *successives* constituent alors un tout autre problème que celui que nous nous posons ici.

Preuve. Le ket $|\psi\rangle$ étant fixé, on considère, pour $\lambda \in \mathbb{R}$, la quantité

$$Q(\lambda) = \left| \left((\hat{A} - \langle A \rangle_\psi) + i\lambda(\hat{B} - \langle B \rangle_\psi) \right) |\psi\rangle \right|^2 = \lambda^2 (\Delta B)^2 + \lambda i \langle \psi | [\hat{A}, \hat{B}] | \psi \rangle + (\Delta A)^2.$$

Le deuxième terme du développement ci-dessus est réel, car

$$i \langle [\hat{A}, \hat{B}] \rangle_\psi = -2\mathfrak{Im} \left(\langle \psi | \hat{A} \hat{B} | \psi \rangle \right). \quad (11.4)$$

Le polynôme à coefficients réels $Q(\lambda)$ étant positif pour tout λ , on en déduit que son discriminant est négatif, et donc

$$4(\Delta A)^2 (\Delta B)^2 - \langle [\hat{A}, \hat{B}] \rangle_\psi^2 \leq 0.$$

Ceci donne le résultat escompté. $\square$

Une conséquence est la fameuse relation d'incertitude de Heisenberg position-impulsion, qui correspond au cas particulier $\hat{A} = \hat{x}$ (ou $\hat{A} = \hat{y}, \hat{z}$) et $\hat{B} = \hat{p}_x$ (ou $\hat{B} = \hat{p}_y, \hat{p}_z$). On rappelle que $[\hat{x}, \hat{p}_x] = i\hbar$.

Corollaire 11.1 (Relation d'incertitude position-impulsion). *Pour tout état $|\psi\rangle \in D(\hat{x}) \cap D(\hat{p}_x)$ décrivant une particule en translation dans l'espace, on a*

$$\Delta x \Delta p_x \geq \frac{\hbar}{2},$$

et des relations analogues sur les axes y, z .

Notons que cette relation d'incertitude ne s'applique qu'à des composantes prises sur le même axe. Si on applique la relation d'incertitude aux opérateurs $\hat{y}$ et $\hat{p}_x$, par exemple, il vient :

$$\Delta y \Delta p_x \geq 0.$$

Rien n'empêche (mais rien n'oblige non plus) de connaître parfaitement l'impulsion d'une particule sur un axe, et sa position sur un axe perpendiculaire.

Les relations de commutation font qu'il est parfois *physiquement* impossible qu'un système ait deux caractéristiques A et B parfaitement déterminées. Cela traduit le fait qu'il est mathématiquement impossible que son état $|\psi\rangle$ soit simultanément état propre de $\hat{A}$ et de $\hat{B}$. Dans le cas de la position et de l'impulsion d'une particule selon un axe donné, cette impossibilité est systématique. Dans le cas limite où la position est connue avec une précision arbitraire, on n'a donc aucune information sur l'impulsion (et réciproquement). Cette situation est extrême, et en pratique, positions et impulsions sont généralement chacune distribuées autour d'une moyenne fixée. Répétons que ces dispersions de valeurs sont des propriétés intrinsèques des systèmes quantiques, et n'ont rien à voir avec une quelconque incertitude du procédé de mesure.

On peut remarquer aussi que la relation d'incertitude dans sa formulation positions/vitesses est

$$\Delta x \Delta v_x \geq \frac{\hbar}{2m}.$$

La relation d'incertitude position-vitesse se montre donc d'autant plus contraignante que la particule considérée est légère. C'est ainsi qu'elle est omniprésente lorsqu'on étudie des électrons, alors qu'elle peut parfaitement être négligée pour des corps plus lourds obéissant à la mécanique classique.

11.5.4 Commutation et stabilité de l'atome d'hydrogène

Un grand nombre de résultats de la physique quantique reposent, au plan théorique, sur les propriétés de commutations entre observables. Nous venons d'en voir quelques exemples. En guise de conclusion à ce chapitre, nous donnons ici un résultat concernant non des théorèmes généraux, mais l'étude d'un système physique particulier : l'atome d'hydrogène. Nous montrons que les relations de commutations permettent de prouver sa stabilité en établissant une borne inférieure pour l'énergie de son électron (électron qui ne peut par conséquent pas tomber sur le noyau). Apprécions une fois de plus la puissance du formalisme de Dirac.

Théorème 11.2. *Sur $L^2(\mathbb{R}^3)$, on a*

$$-\Delta \geq \frac{1}{4|\mathbf{r}|^2},$$

au sens où, pour un sous-espace dense $\mathcal{E}$ de $L^2(\mathbb{R}^3)$,

$$\forall \psi \in \mathcal{E}, \quad \left\langle \psi, \left(-\Delta - \frac{1}{4|\mathbf{r}|^2} \right) \psi \right\rangle \geq 0.$$

Preuve. On prend $\mathcal{E} = C_0^\infty(\mathbb{R}^3)$. Dans ce cas, on peut vérifier que toutes les expressions ci-dessous ont un sens. On calcule en premier lieu, pour $\mathbf{r} = (r_1, \dots, r_d)$ et $\hat{\mathbf{p}} = (\hat{p}_1, \dots, \hat{p}_d)$ dans un espace de dimension $d = 3$,

$$\sum_{j=1}^d i \left[\frac{1}{|\mathbf{r}|} \hat{p}_j \frac{1}{|\mathbf{r}|}, r_j \right] = \frac{\hbar d}{|\mathbf{r}|^2}.$$

On en déduit, par (11.4),

$$\hbar d \| |\mathbf{r}|^{-1} \psi \|^2 = -2 \sum_{j=1}^d \Im \left\langle \frac{1}{|\mathbf{r}|} \hat{p}_j \frac{1}{|\mathbf{r}|} \psi, r_j \psi \right\rangle.$$

Comme $[\hat{p}_j, |\mathbf{r}|^{-1}] = i\hbar |\mathbf{r}|^{-3} r_j$, on a alors

$$\hbar(d-2) \| |\mathbf{r}|^{-1} \psi \|^2 = -2 \sum_{j=1}^d \Im \left\langle \hat{p}_j \psi, \frac{r_j}{|\mathbf{r}|^2} \psi \right\rangle.$$

Une inégalité de Cauchy-Schwarz montrant que

$$\left| \sum_{j=1}^d \left\langle \hat{p}_j \psi, \frac{r_j}{|\mathbf{r}|^2} \psi \right\rangle \right| \leq \langle \psi, -\Delta \psi \rangle^{1/2} \| |\mathbf{r}|^{-1} \psi \|,$$

on a alors, avec l'inégalité précédente,

$$\hbar(d-2) \| |\mathbf{r}|^{-1} \psi \| \leq 2 \langle \psi, -\Delta \psi \rangle^{1/2}.$$

On obtient donc finalement

$$\langle \psi, -\Delta \psi \rangle \geq \frac{|d-2|^2}{4} \left\langle \psi, \frac{1}{|\mathbf{r}|^2} \psi \right\rangle.$$

Le résultat s'en déduit pour $d = 3$. □

L'atome d'hydrogène est instable dans la description classique de la matière. On décrit en effet l'électron comme une particule ponctuelle dans le champ coulombien créé par un noyau ponctuel de charge e , *i.e.* on a une dynamique newtonienne dans un champ de potentiel $V(\mathbf{r}) = -e^2(4\pi\epsilon_0|\mathbf{r}|)^{-1}$. Comme l'énergie potentielle décroît quand la distance au noyau $|\mathbf{r}|$ décroît, on voit que cette

dynamique est instable par rapport aux perturbations normales à une orbite dans la direction du noyau.

Dans le formalisme de la mécanique quantique au contraire, on montre avec le Théorème 11.2 que le Hamiltonien est borné inférieurement, ce qui assure la stabilité de l'atome d'hydrogène. En effet, au sens des opérateurs sur $L^2(\mathbb{R}^3)$,

$$\hat{H} = -\frac{\hbar^2}{2m}\Delta - \frac{e^2}{4\pi\epsilon_0|\mathbf{r}|} \geq \frac{\hbar^2}{8m|\mathbf{r}|^2} - \frac{e^2}{4\pi\epsilon_0|\mathbf{r}|}.$$

Le minorant ci-dessus atteint son minimum en $|\mathbf{r}|^{-1} = \pi\epsilon_0\hbar^2/(me^2)$, et ainsi

$$\hat{H} \geq -\frac{me^4}{8\pi^2\epsilon_0^2\hbar^2}.$$

L'application numérique donne $\hat{H} \geq -54.5 \text{ eV}$. Cette valeur est en accord avec l'énergie trouvée expérimentalement pour l'état fondamental de l'électron : -13.6 eV .

Exercices

Exercice 11.1 (Caractère autoadjoint des observables). Montrer, dans le formalisme fonction d'onde, le caractère autoadjoint des observables $\hat{p}_x$ et $\partial_x V$.

Exercice 11.2 (Crochets de commutation). Calculer pour une particule de masse m en translation dans un champ de potentiel $V(\mathbf{r})$ les commutateurs $[\hat{p}_x, \hat{H}_c]$, $[\hat{x}, \hat{H}_c]$, $[\hat{p}_x, \hat{V}]$ et $[\hat{x}, \hat{V}]$, où $\hat{H}_c$ désigne l'observable énergie cinétique $\hat{H}_c = \frac{\hat{p}^2}{2m}$. Parmi la position, l'impulsion, et les énergies potentielle et cinétique, quelles familles de grandeurs est-il possible de connaître avec précision simultanément ?

Exercice 11.3 (Mesure). On considère un système physique décrit par un espace de Hilbert à 2 dimensions. On se donne une base orthonormée d'états $(|1\rangle, |2\rangle)$, et deux observables $\hat{A}$ et $\hat{B}$ qui s'écrivent dans cette base :

$$\hat{A} = a \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}, \quad \hat{B} = b \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix},$$

avec a et b deux constantes.

- (1) Calculer les vecteurs propres et les valeurs propres de $\hat{A}$ et $\hat{B}$.
On place le système dans l'état $|\psi_{\text{avant}}\rangle = \frac{1}{\sqrt{5}}(|1\rangle + 2|2\rangle)$.
- (2) On mesure sur $|\psi_{\text{avant}}\rangle$ l'observable $\hat{A}$. Quels résultats peut-on mesurer, et avec quelles probabilités ? calculer ΔA .
- (3) Même question si au lieu de mesurer $\hat{A}$, on mesure $\hat{B}$. Calculer ΔB .
- (4) Calculer le commutateur $[\hat{A}, \hat{B}]$, et vérifier l'inégalité de Heisenberg.
- (5) Partant de l'état $|\psi_{\text{avant}}\rangle$, on mesure $\hat{A}$, et on trouve la grandeur $2a$. On décide alors de mesurer $\hat{B}$. Quelles sont les nouvelles probabilités sur les résultats de cette deuxième mesure ?

Exercice 11.4 (Spectre du commutateur de deux observables). Soient $\hat{A}, \hat{B}$ deux observables, de domaines $D(\hat{A}), D(\hat{B})$. En calculant

$$\langle \psi | [\hat{A}, \hat{B}] | \psi \rangle = \int_{\mathbb{R}^3} \bar{\psi} \cdot C \psi \, d\mathbf{r},$$

montrer que C est auto-adjoint et que $\sigma(C) \subset i\mathbb{R}$.

Le moment cinétique en mécanique quantique : Moment orbital et Spin

Le formalisme de la fonction d'onde décrit la dynamique de translation d'une particule dans le vide. Il ignore cependant un aspect essentiel du mouvement des corps : leur rotation sur eux-mêmes. C'est cet aspect, appréhendé en mécanique classique par le concept de *moment cinétique*, que nous abordons ici. Le présent chapitre fait quelques rappels sur la manière dont la mécanique classique traite du moment cinétique. Nous voyons ensuite comment la mécanique quantique distingue le moment cinétique autour d'un point d'une particule en translation, dit *moment orbital*, et le moment cinétique interne des corps, dit *spin*.

12.1 Le moment cinétique en mécanique classique	187
12.1.1 Moment cinétique d'une particule ponctuelle autour d'un point	187
12.1.2 Rotation d'un objet rigide sur lui-même	189
12.2 Le moment cinétique orbital	190
12.2.1 Relations de commutation	190
12.2.2 Liens avec la mécanique classique	192
12.3 La définition étendue du moment cinétique en mécanique quantique	193
12.4 Quantification du moment cinétique	193
12.4.1 Que quantifier ?	194
12.4.2 Le résultat de quantification	194
12.4.3 Eléments de démonstration	195
12.5 Quantification du moment cinétique orbital	199
12.6 Le spin	203
12.6.1 Définition	203
12.6.2 Espace de spin	203
12.6.3 Manifestations physiques du spin	204
12.6.4 Un exemple d'espace de spin : l'espace du spin $\frac{1}{2}$	205
Exercices	207

12.1 Le moment cinétique en mécanique classique

Le moment cinétique est une grandeur vectorielle qui permet de décrire les mouvements de rotation des corps rigides.

12.1.1 Moment cinétique d'une particule ponctuelle autour d'un point

Considérons une particule ponctuelle de masse m en mouvement dans $\mathbb{R}^3$. Son état classique de translation est repéré par sa position $\mathbf{r}(t)$ et son impulsion $\mathbf{p}(t)$. Considérons par ailleurs un

point de l'espace, de position $\mathbf{r}_0$, qui, pour une raison quelconque, revêt pour nous une importance particulière. Si un observateur, posté en $\mathbf{r}_0$, suit des yeux la trajectoire de la particule, il sera certainement, à la manière d'un spectateur de course automobile, forcé au cours du temps de tourner progressivement la tête. D'une certaine manière, même dans le cas où la particule se déplace en ligne droite, la direction dans laquelle l'observateur doit orienter son regard *tourne* au cours du temps. Ce mouvement de rotation apparent autour du point $\mathbf{r}_0$ est décrit par le moment cinétique $\mathbf{L}(t)$ selon :

$$\mathbf{L} = (\mathbf{r} - \mathbf{r}_0) \wedge \mathbf{p}.$$

La figure 12.1 présente graphiquement le moment cinétique.

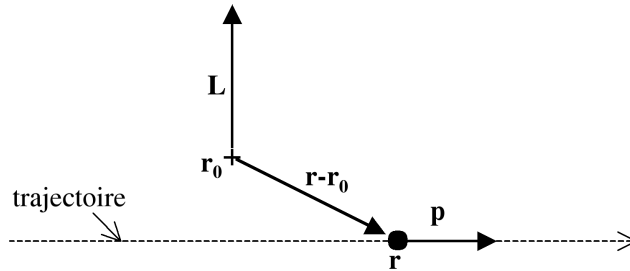


Fig. 12.1. Moment cinétique $\mathbf{L}$ autour d'un point $\mathbf{r}_0$ d'une particule en translation.

La seule situation dans laquelle la particule ne semble pas tourner est celle où, soit elle se dirige droit sur le point $\mathbf{r}_0$, soit elle s'éloigne droit de ce point. Très souvent, pour des raisons pratiques, le point $\mathbf{r}_0$ sert à définir le point origine de l'espace. On a alors simplement :

$$\mathbf{L} = \mathbf{r} \wedge \mathbf{p}. \quad (12.1)$$

Le vecteur $\mathbf{L}(t)$ est une nouvelle grandeur, mais n'est pas à proprement parler un nouveau *paramètre* de description de l'état du système. En effet, il se déduit entièrement de $\mathbf{r}(t)$ et de $\mathbf{p}(t)$, et n'apporte donc pas d'information physique supplémentaire sur la particule. Cependant, il jouit de propriétés remarquables qui en font un outil de raisonnement et de calcul intéressant. La relation fondamentale de la dynamique pour une particule soumise à des forces $\mathbf{F}_i$ implique pour l'évolution temporelle de $\mathbf{L}$:

$$\frac{d\mathbf{L}}{dt} = \frac{d\mathbf{r}}{dt} \wedge \mathbf{p} + \mathbf{r} \wedge \frac{d\mathbf{p}}{dt} = \frac{1}{m} \mathbf{p} \wedge \mathbf{p} + \mathbf{r} \wedge \left(\sum_i \mathbf{F}_i \right),$$

soit

$$\frac{d\mathbf{L}}{dt} = \sum_i \mathbf{r} \wedge \mathbf{F}_i = \sum_i \mathbf{M}_i. \quad (12.2)$$

Le vecteur $\mathbf{L}(t)$ obéit ainsi, à l'instar de $\mathbf{p}(t)$, à une relation conceptuellement proche de la relation fondamentale de la dynamique. Chaque force $\mathbf{F}_i$ est à l'origine d'un *effort en rotation* $\mathbf{M}_i = \mathbf{r} \wedge \mathbf{F}_i$, dit *moment* exercé par la force autour du point de référence (ici, le point origine de l'espace). La quantité $\mathbf{L}$ est particulièrement utile en astronomie ou en électromagnétisme pour étudier la trajectoire d'une particule en interaction avec un centre attractif ou répulsif. Dans cette situation, si ce centre est pris comme référence, les forces appliquées sont toujours colinéaires à $\mathbf{r}$. Le moment appliqué est nul, et $\mathbf{L}$ devient donc une *constante du mouvement*. De même, le moment cinétique d'un système isolé est également une constante du mouvement.

Remarque 12.1. Que l'on considère ou pas le moment cinétique, l'énergie cinétique de la particule ponctuelle continue invariablement à s'écrire :

$$E_c = \frac{p^2}{2m}.$$

Remarque 12.2. Le point matériel considéré ici peut aussi bien être une vraie particule ponctuelle, que le centre de gravité d'un objet, ou encore la particule réduite d'un système de deux corps en interaction.

12.1.2 Rotation d'un objet rigide sur lui-même

Moments cinétiques interne et externe

Pour un objet rigide qui occupe à chaque instant un certain volume $\Omega(t)$ dans l'espace, le moment cinétique est défini comme l'intégrale des moments cinétiques élémentaires

$$\mathbf{L}(t) = \int_{\Omega(t)} \mathbf{r} \wedge \mathbf{v}(\mathbf{r}, t) \rho(\mathbf{r}, t) d\mathbf{r},$$

où $\mathbf{v}(\mathbf{r}, t)$ est le champ de vitesse dans l'objet à l'instant considéré et $\rho(\mathbf{r}, t)$ est la masse volumique locale du corps étudié. Même pour un objet rigide, ρ et Ω évoluent au cours du temps du seul fait des mouvements de l'objet (mais resteraient constants dans un référentiel fixé à l'objet!).

Une particularité de ce moment cinétique est qu'il est possible de le scinder en deux parties externe et interne au corps étudié. Si on note G le centre de gravité de l'objet, $\mathbf{r}_G(t)$ sa position, $\mathbf{v}_G(t)$ sa vitesse, et m la masse totale de l'objet, il vient en effet :

$$\mathbf{L}(t) = m\mathbf{r}_G \wedge \mathbf{v}_G + \int_{\Omega(t)} (\mathbf{r} - \mathbf{r}_G(t)) \wedge (\mathbf{v}(\mathbf{r}, t) - \mathbf{v}_G(t)) \rho(\mathbf{r}, t) d\mathbf{r} = \mathbf{L}_{\text{ext}} + \mathbf{L}_{\text{int}}.$$

Le vecteur $\mathbf{L}_{\text{ext}} = m\mathbf{r}_G \wedge \mathbf{v}_G$ est le moment cinétique du centre de gravité autour du point d'observation. Il est exclusivement dû à la translation de l'objet dans l'espace, et se calcule en fonction de l'impulsion totale $\mathbf{p}_G = m\mathbf{v}_G$. Le moment cinétique $\mathbf{L}_{\text{int}}$ est appelé *moment cinétique interne* de l'objet. Il décrit uniquement le mouvement de rotation autour du centre de gravité (voir la figure 12.2).

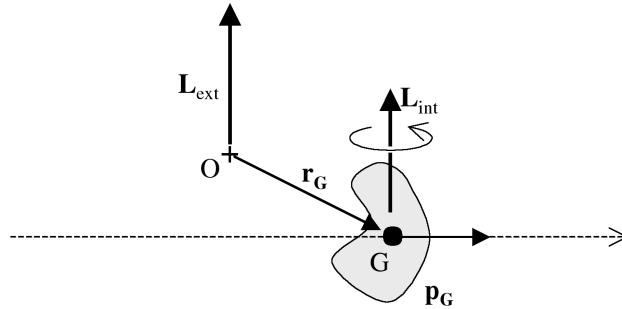


Fig. 12.2. Moments cinétiques externe $\mathbf{L}_{\text{ext}}$ et interne $\mathbf{L}_{\text{int}}$ d'un corps rigide en translation/rotation autour du point origine O .

Les équations d'évolution de $\mathbf{L}_{\text{ext}}$ et $\mathbf{L}_{\text{int}}$ peuvent également être découplées. Pour un objet soumis à une densité volumique de forces $\mathbf{f}(\mathbf{r}, t)$:

$$\frac{d\mathbf{L}_{\text{ext}}}{dt} = \mathbf{r}_G \wedge \mathbf{F}_{\text{totale}} = \mathbf{M}_{\text{ext}}, \quad \text{avec } \mathbf{F}_{\text{totale}} = \int_{\Omega} \mathbf{f}(\mathbf{r}, t) d\mathbf{r},$$

où $\mathbf{F}_{\text{totale}}$ représente donc la somme intégrée de toutes les forces volumiques appliquées au solide. On a également

$$\frac{d\mathbf{L}_{\text{int}}}{dt} = \int_{\Omega} (\mathbf{r} - \mathbf{r}_G(t)) \wedge \mathbf{f}(\mathbf{r}, t) d\mathbf{r} = \mathbf{M}_{\text{int}}.$$

Le moment global appliqué à l'objet peut donc être lui-même scindé en deux composantes $\mathbf{M}_{\text{ext}}$ et $\mathbf{M}_{\text{int}}$ qui agissent respectivement sur le mouvement du centre de gravité et sur le mouvement de rotation du solide autour de ce centre de gravité.

Energie cinétique de rotation

Le moment cinétique $\mathbf{L}_{\text{ext}}$ est une fonction de la position $\mathbf{r}_G$ et de l'impulsion $\mathbf{p}_G$ et n'est donc pas une nouvelle variable de description du système. En revanche, le moment cinétique interne $\mathbf{L}_{\text{int}}$ appréhende un nouvel aspect de la dynamique de l'objet qui échappe à la description par le centre de gravité. C'est un *paramètre supplémentaire* de description de l'état du système. Pour le corps rigide en mouvement, il sous-tend en particulier un terme d'énergie cinétique supplémentaire, selon :

$$E_c = \frac{p_G^2}{2m} + E_{\text{rot}}.$$

Le terme E_{rot} est l'*énergie cinétique de rotation* du solide sur lui-même. Dans le cas d'un corps à symétrie sphérique, elle s'exprime selon :

$$E_{\text{rot}} = \frac{L_{\text{int}}^2}{2I},$$

où I est une grandeur scalaire (de dimension kg m^2), et appelée *moment d'inertie*. Au même titre que la masse, c'est une caractéristique intrinsèque de l'objet. Dans le cas général de corps non symétriques, I prend une forme tensorielle.¹

12.2 Le moment cinétique orbital

Nous revenons à présent dans le monde quantique. Nous nous intéressons ici à une particule en simple translation dans l'espace.

Définition 12.1. *Pour une particule de masse m en mouvement de translation dans l'espace, et décrite par sa fonction d'onde $\psi(\mathbf{r}, t)$, on appelle moment cinétique orbital par rapport à l'origine le triplet d'observables $\hat{L}_x$, $\hat{L}_y$, et $\hat{L}_z$ définies par :*

$$\begin{cases} \hat{L}_x = \hat{y}\hat{p}_z - \hat{z}\hat{p}_y = y\frac{\hbar}{i}\frac{\partial}{\partial z} - z\frac{\hbar}{i}\frac{\partial}{\partial y}, \\ \hat{L}_y = \hat{z}\hat{p}_x - \hat{x}\hat{p}_z = z\frac{\hbar}{i}\frac{\partial}{\partial x} - x\frac{\hbar}{i}\frac{\partial}{\partial z}, \\ \hat{L}_z = \hat{x}\hat{p}_y - \hat{y}\hat{p}_x = x\frac{\hbar}{i}\frac{\partial}{\partial y} - y\frac{\hbar}{i}\frac{\partial}{\partial x}, \end{cases} \quad (12.3)$$

En vertu du principe de correspondance, elle constituent les observables associés aux trois composantes classiques du moment cinétique.

Le moment cinétique orbital décrit le mouvement du centre de gravité de la particule, à l'exclusion de tout mouvement de rotation de la particule sur elle-même (notion qui de toute façon, échappe au formalisme des fonctions d'onde). Il est également utile de définir, toujours à l'aide du principe de correspondance, l'observable *norme au carré* du moment cinétique, selon

$$\hat{L}^2 = \hat{L}_x^2 + \hat{L}_y^2 + \hat{L}_z^2. \quad (12.4)$$

12.2.1 Relations de commutation

La commutation des observables sera au coeur de ce chapitre. Nous donnons en annexe D un formulaire des principales relations impliquant les opérateurs position, impulsion, moment cinétique, énergie, et force appliquée. Soulignons qu'à ce stade du cours, nous en savons suffisamment pour démontrer une à une chacune de ces relations (ce qui constitue par ailleurs un excellent

1. Nous renvoyons le lecteur intéressé aux ouvrages de mécanique de second cycle.

exercice). La plupart se déduisent des 9 relations de commutation position-impulsion désormais connues :

$$\begin{cases} [\hat{x}, \hat{p}_x] = [\hat{y}, \hat{p}_y] = [\hat{z}, \hat{p}_z] = i\hbar, \\ [\hat{x}, \hat{p}_y] = [\hat{x}, \hat{p}_z] = [\hat{y}, \hat{p}_x] = [\hat{y}, \hat{p}_z] = [\hat{z}, \hat{p}_x] = [\hat{z}, \hat{p}_y] = 0. \end{cases}$$

Donnons à titre d'exemple le détail des calculs pour trois de ces relations.

Exemple 12.1. Montrons que $[\hat{L}_x, \hat{p}_x] = 0$. Pour cela, on explicite l'observable $\hat{L}_x$ en fonction des observables position et impulsion :

$$[\hat{L}_x, \hat{p}_x] = [\hat{y}\hat{p}_z - \hat{z}\hat{p}_y, \hat{p}_x] = (\hat{y}\hat{p}_z - \hat{z}\hat{p}_y)\hat{p}_x - \hat{p}_x(\hat{y}\hat{p}_z - \hat{z}\hat{p}_y) = \hat{y}\hat{p}_z\hat{p}_x - \hat{z}\hat{p}_y\hat{p}_x - \hat{p}_x\hat{y}\hat{p}_z + \hat{p}_x\hat{z}\hat{p}_y.$$

Or $\hat{p}_x$ commute avec $\hat{y}$ et $\hat{p}_z$ donc $\hat{y}\hat{p}_z\hat{p}_x = \hat{y}\hat{p}_x\hat{p}_z = \hat{p}_x\hat{y}\hat{p}_z$. De même, $\hat{p}_x$ commute avec $\hat{z}$ et $\hat{p}_y$ donc $\hat{z}\hat{p}_y\hat{p}_x = \hat{z}\hat{p}_x\hat{p}_y = \hat{p}_x\hat{z}\hat{p}_y$. Ceci donne le résultat escompté.

Exemple 12.2. Montrons que $[\hat{L}_x, \hat{p}_y] = i\hbar\hat{p}_z$. On explicite à nouveau l'observable $\hat{L}_x$:

$$[\hat{L}_x, \hat{p}_y] = (\hat{y}\hat{p}_z - \hat{z}\hat{p}_y)\hat{p}_y - \hat{p}_y(\hat{y}\hat{p}_z - \hat{z}\hat{p}_y) = \hat{y}\hat{p}_z\hat{p}_y - \hat{z}\hat{p}_y\hat{p}_y - \hat{p}_y\hat{y}\hat{p}_z + \hat{p}_y\hat{z}\hat{p}_y.$$

Comme $\hat{p}_y$ commute avec $\hat{z}$ et $\hat{p}_z$, on a $\hat{p}_y\hat{z}\hat{p}_y = \hat{z}\hat{p}_y\hat{p}_y$ et $\hat{y}\hat{p}_z\hat{p}_y = \hat{y}\hat{p}_y\hat{p}_z$. Finalement,

$$[\hat{L}_x, \hat{p}_y] = \hat{y}\hat{p}_y\hat{p}_z - \hat{p}_y\hat{y}\hat{p}_z = [\hat{y}, \hat{p}_y]\hat{p}_z = i\hbar\hat{p}_z.$$

Exemple 12.3. Montrons enfin $[\hat{L}_x, \hat{V}] = i\hbar(\hat{y}\hat{F}_z - \hat{z}\hat{F}_y)$. On applique le crochet de commutation à une fonction test quelconque $\psi \in C_0^\infty(\mathbb{R}^3)$:

$$\begin{aligned} [\hat{L}_x, \hat{V}] \psi &= \left((\hat{y}\hat{p}_z - \hat{z}\hat{p}_y)\hat{V} - \hat{V}(\hat{y}\hat{p}_z - \hat{z}\hat{p}_y) \right) \psi \\ &= \frac{\hbar}{i} \left(y \frac{\partial}{\partial z} (V\psi) - z \frac{\partial}{\partial y} (V\psi) \right) - V \frac{\hbar}{i} \left(y \frac{\partial}{\partial z} \psi - z \frac{\partial}{\partial y} \psi \right) \\ &= \frac{\hbar}{i} \left(y \frac{\partial V}{\partial z} - z \frac{\partial V}{\partial y} \right) \psi = i\hbar(\hat{y}\hat{F}_z - \hat{z}\hat{F}_y)\psi, \end{aligned}$$

où $\hat{F}_x$, $\hat{F}_y$ et $\hat{F}_z$ sont les observables associées à la force appliquée au système.

Principales relations utiles

Nous récapitulons ci-dessous les principales relations qui nous serviront dans ce chapitre :

$$\begin{aligned} [\hat{L}_x, \hat{r}^2] &= [\hat{L}_y, \hat{r}^2] = [\hat{L}_z, \hat{r}^2] = 0, \\ [\hat{L}_x, \hat{p}^2] &= [\hat{L}_y, \hat{p}^2] = [\hat{L}_z, \hat{p}^2] = 0, \\ [\hat{L}_x^2, \hat{r}^2] &= [\hat{L}_y^2, \hat{r}^2] = [\hat{L}_z^2, \hat{r}^2] = 0, \\ [\hat{L}_x^2, \hat{p}^2] &= [\hat{L}_y^2, \hat{p}^2] = [\hat{L}_z^2, \hat{p}^2] = 0, \\ [\hat{L}^2, \hat{L}_x] &= [\hat{L}^2, \hat{L}_y] = [\hat{L}^2, \hat{L}_z] = 0, \end{aligned}$$

ainsi que les relations de non-commutation :

$$\begin{aligned} [\hat{L}_x, \hat{V}] &= i\hbar(\hat{y}\hat{F}_z - \hat{z}\hat{F}_y), & [\hat{L}_y, \hat{V}] &= i\hbar(\hat{z}\hat{F}_x - \hat{x}\hat{F}_z), & [\hat{L}_z, \hat{V}] &= i\hbar(\hat{x}\hat{F}_y - \hat{y}\hat{F}_x), \\ [\hat{L}_x, \hat{L}_y] &= i\hbar\hat{L}_z, & [\hat{L}_y, \hat{L}_z] &= i\hbar\hat{L}_x, & [\hat{L}_z, \hat{L}_x] &= i\hbar\hat{L}_y, \end{aligned}$$

avec $\hat{p}^2 = \hat{p}_x^2 + \hat{p}_y^2 + \hat{p}_z^2$ et $\hat{r}^2 = \hat{x}^2 + \hat{y}^2 + \hat{z}^2$. D'autres formules sont données en Appendice D.

12.2.2 Liens avec la mécanique classique

Equations d'évolution du moment cinétique

Le théorème d'Ehrenfest montre que

$$\begin{cases} \frac{d}{dt} \langle \hat{L}_x \rangle = \langle \hat{y} \hat{F}_z - \hat{z} \hat{F}_y \rangle, \\ \frac{d}{dt} \langle \hat{L}_y \rangle = \langle \hat{z} \hat{F}_x - \hat{x} \hat{F}_z \rangle, \\ \frac{d}{dt} \langle \hat{L}_z \rangle = \langle \hat{x} \hat{F}_y - \hat{y} \hat{F}_x \rangle. \end{cases}$$

En effet, pour l'opérateur $\hat{L}_x$ par exemple,

$$\frac{d}{dt} \langle \hat{L}_x \rangle = \frac{1}{i\hbar} \langle [\hat{L}_x, \hat{H}] \rangle,$$

où le Hamiltonien $\hat{H} = \frac{\hat{p}^2}{2m} + \hat{V}$. En utilisant les relations de commutation, il vient :

$$[\hat{L}_x, \hat{H}] = \frac{1}{2m} [\hat{L}_x, \hat{p}^2] + [\hat{L}_x, \hat{V}] = 0 + i\hbar (\hat{y} \hat{F}_z - \hat{z} \hat{F}_y),$$

d'où le résultat. La démonstration se fait de même pour les opérateurs $\hat{L}_y$ et $\hat{L}_z$.

On observe, en accord avec le principe de correspondance, que les termes de droite ne sont rien d'autre que les valeurs moyennes des observables associées aux composantes du moment classique :

$$\mathbf{M} = \mathbf{r} \wedge \mathbf{F} = \begin{pmatrix} yF_z - zF_y \\ zF_x - xF_z \\ xF_y - yF_x \end{pmatrix}.$$

L'équation d'évolution de la moyenne des grandeurs quantiques apparaît donc comme une transposition directe de l'équation classique $\partial_t \mathbf{L} = \mathbf{M}$.

Stabilité du moment cinétique dans un potentiel central

Définition 12.2. *Un potentiel central est un potentiel $V(\mathbf{r})$ qui n'est fonction que de la norme r du vecteur position $\mathbf{r}$.*

Le potentiel Coulombien et le potentiel de gravitation centrés sur l'origine sont des exemples de potentiels centraux.

Proposition 12.1. *Dans un potentiel central, les observables $\hat{L}_x$, $\hat{L}_y$, $\hat{L}_z$ et $\hat{L}^2$ commutent avec le Hamiltonien $\hat{H} = \frac{\hat{p}^2}{2m} + \hat{V}$.*

Preuve. Comme les observables $\hat{L}_x$, $\hat{L}_y$, $\hat{L}_z$ et $\hat{L}^2$ commutent déjà avec le terme $\hat{p}^2$ du Hamiltonien, il suffit de montrer que $\hat{L}_x$, $\hat{L}_y$, $\hat{L}_z$ et $\hat{L}^2$ commutent avec $\hat{V}$. Traitons par exemple le cas de $\hat{L}_x$. Pour ce faire, on considère $\psi \in D([\hat{L}_x, \hat{V}])$, et on calcule

$$\frac{i}{\hbar} [\hat{L}_x, \hat{V}] \cdot \psi = [y\partial_z - z\partial_y](V(r)\psi) - V(r)[y\partial_z - z\partial_y]\psi = [y\partial_z - z\partial_y](V(r)) \cdot \psi.$$

Or, $[y\partial_z - z\partial_y](V(r)) = V'(r)[y\partial_z r - z\partial_y r]$, et $\partial_y r = y/r$, $\partial_z r = z/r$. Ceci montre finalement que $[\hat{L}_x, \hat{V}] = 0$. $\square$

En conséquence, dans un potentiel central, en mécanique quantique comme en mécanique classique, les composantes et la norme du moment cinétique sont des *constantes du mouvement*.

12.3 La définition étendue du moment cinétique en mécanique quantique

Nous avons pu montrer pour le moment cinétique orbital que $\hat{L}_x$, $\hat{L}_y$ et $\hat{L}_z$ vérifient les trois relations de commutation suivantes :

$$\left[\hat{L}_x, \hat{L}_y\right] = i\hbar\hat{L}_z, \quad \left[\hat{L}_y, \hat{L}_z\right] = i\hbar\hat{L}_x, \quad \left[\hat{L}_z, \hat{L}_x\right] = i\hbar\hat{L}_y,$$

c'est à dire

$$\begin{cases} i\hbar\hat{L}_x = \hat{L}_y\hat{L}_z - \hat{L}_z\hat{L}_y, \\ i\hbar\hat{L}_y = \hat{L}_z\hat{L}_x - \hat{L}_x\hat{L}_z, \\ i\hbar\hat{L}_z = \hat{L}_x\hat{L}_y - \hat{L}_y\hat{L}_x. \end{cases}$$

Ces relations revêtent en fait une importance toute particulière. Nous allons voir en effet qu'à elles seules, elles déterminent pour une large part la forme du spectre de quantification du moment cinétique. On a coutume de les noter sous la forme contractée (cela n'est qu'une notation formelle) :

$$i\hbar\hat{\mathbf{L}} = \hat{\mathbf{L}} \wedge \hat{\mathbf{L}},$$

où $\hat{\mathbf{L}} = (\hat{L}_x, \hat{L}_y, \hat{L}_z)$ désigne l'opérateur vectoriel dont les composantes sont les trois observables $\hat{L}_x$, $\hat{L}_y$ et $\hat{L}_z$.

Pour le moment cinétique orbital, il s'agit de relations démontrées. En revanche, lorsque nous nous intéresserons aux objets en rotation sur eux mêmes, il faudra construire de toutes pièces un formalisme spécifique au traitement quantique du moment cinétique interne, mettant en jeu de nouveaux espaces quantiques et de nouveaux opérateurs de moment cinétique. Dans ce contexte, la théorie quantique considère alors ces relations de commutation comme *postulat de base* sur les propriétés des nouveaux opérateurs introduits. La notion de moment cinétique prend ainsi la définition élargie suivante :

Définition 12.3. *En mécanique quantique, on nomme opérateur de moment cinétique tout triplet d'observables $\hat{\mathbf{J}} = (\hat{J}_x, \hat{J}_y, \hat{J}_z)$ qui vérifient la relation de commutation fondamentale :*

$$i\hbar\hat{\mathbf{J}} = \hat{\mathbf{J}} \wedge \hat{\mathbf{J}}.$$

Du seul fait de ces relations de commutation, d'autres propriétés de commutation du moment orbital sont également conservées. On montre en effet la propriété suivante :

Propriété 12.1. *Tout moment cinétique $\hat{\mathbf{J}} = (\hat{J}_x, \hat{J}_y, \hat{J}_z)$ vérifie les relations de commutation :*

$$\left[\hat{J}^2, \hat{J}_x\right] = \left[\hat{J}^2, \hat{J}_y\right] = \left[\hat{J}^2, \hat{J}_z\right] = 0,$$

avec $\hat{J}^2 = \hat{J}_x^2 + \hat{J}_y^2 + \hat{J}_z^2$.

Remarque 12.3. La relation $i\hbar\mathbf{a} = \mathbf{a} \wedge \mathbf{a}$ pour un vecteur $\mathbf{a}$ peut paraître choquante (Chacun sait en effet que pour des vecteurs $\mathbf{a} \in \mathbb{R}^3$, on a $\mathbf{a} \wedge \mathbf{a} = \mathbf{0}$). Ici, cette relation est vraie dans un espace d'opérateurs où le produit n'est pas commutatif. Cette 'bizarrerie' est donc parfaitement possible...

Remarque 12.4. La notation $\hat{\mathbf{J}}$ est traditionnelle en mécanique quantique pour désigner un opérateur de moment cinétique dans le cas général. La notation $\hat{\mathbf{L}}$ est réservée aux seuls moments cinétiques orbitaux.

12.4 Quantification du moment cinétique

Nous considérons un opérateur moment cinétique dans sa définition la plus large, à savoir un triplet d'observables $\hat{\mathbf{J}} = (\hat{J}_x, \hat{J}_y, \hat{J}_z)$ dont la seule propriété connue est $i\hbar\hat{\mathbf{J}} = \hat{\mathbf{J}} \wedge \hat{\mathbf{J}}$.

12.4.1 Que quantifier ?

Quantifier le moment cinétique revient à chercher la nature de ses états propres et de ses valeurs propres. L'opérateur $\hat{\mathbf{J}}$ n'étant pas directement un observable, c'est plus précisément les opérateurs composantes $\hat{J}_x$, $\hat{J}_y$ et $\hat{J}_z$ qui nous intéressent. Le terme de *quantification* présume déjà (mais il s'agit d'un des résultats que nous montrerons ici) du caractère discret des spectres de ces opérateurs.

Ce n'est pas la première fois que nous avons affaire à un triplet d'observables. En effet, l'impulsion $\mathbf{p}$ de la particule qui se déplace dans un espace à 3 dimensions est elle-même prise en compte en mécanique quantique comme une association de trois opérateurs $\hat{p}_x$, $\hat{p}_y$ et $\hat{p}_z$. Dans le cas de l'impulsion, les ondes de De Broglie fournissent une base d'états $|\varphi_{\mathbf{k}}(\mathbf{r})\rangle = e^{i\mathbf{k}\cdot\mathbf{r}}$ qui sont simultanément états propres (généralisés) des 3 opérateurs, selon :

$$\hat{p}_x|\varphi_{\mathbf{k}}\rangle = \hbar k_x|\varphi_{\mathbf{k}}\rangle, \quad \hat{p}_y|\varphi_{\mathbf{k}}\rangle = \hbar k_y|\varphi_{\mathbf{k}}\rangle, \quad \hat{p}_z|\varphi_{\mathbf{k}}\rangle = \hbar k_z|\varphi_{\mathbf{k}}\rangle.$$

On pourrait de même se demander s'il existe une base de kets orthogonaux $|\varphi_n\rangle$ et de valeurs propres a_n , b_n et c_n telles que

$$\hat{J}_x|\varphi_n\rangle = a_n|\varphi_n\rangle, \quad \hat{J}_y|\varphi_n\rangle = b_n|\varphi_n\rangle, \quad \hat{J}_z|\varphi_n\rangle = c_n|\varphi_n\rangle.$$

Malheureusement, les seules familles d'opérateurs à pouvoir être diagonalisés simultanément dans une même base sont les familles d'opérateurs qui commutent deux à deux. Ceci était le cas de $\hat{p}_x$, $\hat{p}_y$ et $\hat{p}_z$, mais ce n'est en revanche pas celui de $\hat{J}_x$, $\hat{J}_y$ et $\hat{J}_z$.

Le traitement du problème de quantification passe ainsi par la diagonalisation d'un seul de ces opérateurs, traditionnellement $\hat{J}_z$. Le résultat de quantification se transpose ensuite à $\hat{J}_x$ et $\hat{J}_y$ au prix de changements de bases dans l'espace des kets, changements de base qui correspondent en fait à des échanges d'axes dans l'espace 3D. Il existe toutefois une relation de commutation qui permet de simplifier le travail : celle entre $\hat{J}_z$ et $\hat{J}^2$. Comme ces deux opérateurs commutent, il est possible de leur rechercher une base d'états propres commun. Certes, seul $\hat{J}_z$ paraissait nous importer, mais nous avons là un 'plus' dont la théorie quantique trouve dommage de se priver.

C'est ainsi qu'en théorie quantique, on nomme *quantification du moment cinétique* le travail de *diagonalisation simultanée* de $\hat{J}_z$ et $\hat{J}^2$.

12.4.2 Le résultat de quantification

Sur la base des seules relations de commutation $i\hbar\hat{\mathbf{J}} = \hat{\mathbf{J}} \wedge \hat{\mathbf{J}}$, il est possible de montrer le résultat suivant :

Théorème 12.1. *Etant donné un triplet d'opérateurs de moment cinétique $\hat{\mathbf{J}} = (\hat{J}_x, \hat{J}_y, \hat{J}_z)$, les opérateurs $\hat{J}_z$ et $\hat{J}^2 = \hat{J}_x^2 + \hat{J}_y^2 + \hat{J}_z^2$ sont diagonalisables dans une même base, avec les propriétés suivantes :*

- (1) *Leurs valeurs propres sont de la forme générale $m\hbar$ pour $\hat{J}_z$ et $j(j+1)\hbar^2$ pour $\hat{J}^2$, avec j et m deux nombres réels sans dimension.*
- (2) *Les seules valeurs possibles pour j sont les nombres positifs entiers ou demi-entiers, c'est à dire :*

$$j = 0, \frac{1}{2}, 1, \frac{3}{2}, 2, \frac{5}{2}, \dots$$

- (3) *Pour une valeur de j fixée, les seules valeurs possibles pour m sont les $(2j+1)$ nombres relatifs entiers ou demi-entiers :*

$$m = -j, -j+1, -j+2, \dots, j-2, j-1, j$$

- (4) *Pour une valeur de j fixée, les sous espaces propres $V_{j,m}$ communs à $\hat{J}_z$ et $\hat{J}^2$ obtenus pour différentes valeurs de m ont même dimension, nulle, finie ou infinie, et notée d_j .*

Ce résultat signifie qu'il existe, pour l'espace de Hilbert sur lequel les observables $\hat{J}_z$ et $\hat{J}^2$ opèrent, une base orthonormale de vecteurs que nous notons $|j, m, k\rangle$, indicés par les nombres j et m autorisés ci-dessus ainsi que par un troisième nombre entier positif k tel que $1 \leq k \leq d_j$, qui vérifient :

$$\forall j, m, k, \quad \begin{cases} \hat{J}_z |j, m, k\rangle = m\hbar |j, m, k\rangle, \\ \hat{J}^2 |j, m, k\rangle = j(j+1)\hbar^2 |j, m, k\rangle. \end{cases}$$

Les indices entiers ou semi entiers j et m permettent de désigner le sous espace propre $V_{j,m}$ d'appartenance du vecteur considéré. L'indice entier k distingue, en cas de sous espaces de dimension d_j supérieure à 1, les différents vecteurs de base de ce sous espace.

Remarque 12.5. Nous apprenons ici quelles sont les valeurs propres *mathématiquement autorisées* pour un moment cinétique, au vu de la relation fondamentale $i\hbar\hat{\mathbf{J}} = \hat{\mathbf{J}} \wedge \hat{\mathbf{J}}$. Cela ne signifie pas pour autant que face à une situation concrète, toutes ces valeurs apparaissent. C'est situation physique par situation physique que sont déterminées les dimensions d_j des sous espaces, et ainsi quelles sont les valeurs propres qui se manifestent effectivement.

Remarque 12.6. Pour énoncer ce résultat, l'opérateur $\hat{J}_z$ a été choisi de manière arbitraire. Il est aussi possible de diagonaliser simultanément $\hat{J}_x$ et $\hat{J}^2$ d'une part, et $\hat{J}_y$ et $\hat{J}^2$ d'autre part. Les valeurs propres et les dimensions d_j trouvées restent alors les mêmes. Les sous espaces propres de $\hat{J}^2$ formés par la somme d'espaces $V_j = \sum_{m=-j}^j V_{j,m}$ restent également inchangés. Seul le détail des sous espaces $V_{j,m}$ obtenus pour les différentes valeurs de m se trouve modifié.

Remarque 12.7. Ce résultat montre que quelque soit la situation considérée, nous sommes en présence de spectres discrets. Ces spectres sont entièrement dictés par la valeur de la constante de Planck, à l'exclusion de tout paramètre spécifique au système. Ces spectres ont donc la particularité de ne pas pouvoir tendre vers un continuum quand on se déplace vers des situations classiques. Le pas de discrétisation des composantes, qui vaut $\hbar \approx 10^{-34} \text{ kg m}^2\text{s}^{-1}$ n'est cependant pas très contraignant, et passe bien souvent inaperçu !

12.4.3 Eléments de démonstration

Nous donnons ici la trame générale de la démonstration du résultat de quantification. On suppose tout d'abord que le spectre des opérateurs de moment cinétique est purement discret. Tous les résultats utilisés ne sont pas démontrés dans le détail. Le point essentiel est de mesurer la précision des renseignements à laquelle on peut parvenir uniquement grâce à un usage habile des relations de commutation.

Etape 1 : $\sigma(\hat{J}^2) \subset [0, +\infty[$

Pour $\phi \in D(\hat{J}_x^2)$, on a $\langle \phi | \hat{J}_x^2 | \phi \rangle = \|\hat{J}_x \phi\|^2$ car $\hat{J}_x$ est auto-adjoint. Ceci montre que $\hat{J}_x^2$ est positif. Ainsi, $\hat{J}^2 = \hat{J}_x^2 + \hat{J}_y^2 + \hat{J}_z^2$ est également positif. On peut ainsi rechercher les valeurs propres de $\hat{J}^2$ sous la forme $j(j+1)\hbar^2$ avec j un réel positif ou nul. En effet, l'application $j \mapsto j(j+1)$ définit bien une bijection de $\mathbb{R}_+$ sur $\mathbb{R}_+$.

Nous pouvons donc utiliser les notations $|j, m, k\rangle$ et $V_{j,m}$ introduites dans le théorème. En effet, comme $\hat{J}_z$ et $\hat{J}^2$ commutent, il existe une base commune de diagonalisation de ces deux opérateurs. On peut alors chercher une base orthonormale de vecteurs $|j, m, k\rangle$ telle que

$$\begin{cases} \hat{J}_z |j, m, k\rangle = m\hbar |j, m, k\rangle, \\ \hat{J}^2 |j, m, k\rangle = j(j+1)\hbar^2 |j, m, k\rangle. \end{cases}$$

La différence avec le théorème vient cependant qu'à ce stade de la démonstration, j et m ne sont *a priori* que des nombres réels dont nous ne connaissons rien (hormis $j \geq 0$). Nous ne savons rien non plus des dimensions $d_{j,m}$ des sous espaces $V_{j,m}$.

Afin de nous repérer dans l'évolution de la démonstration, nous dessinons dans le plan (j, m) les lieux repérés comme possibles pour les valeurs propres (voir la figure 12.3).

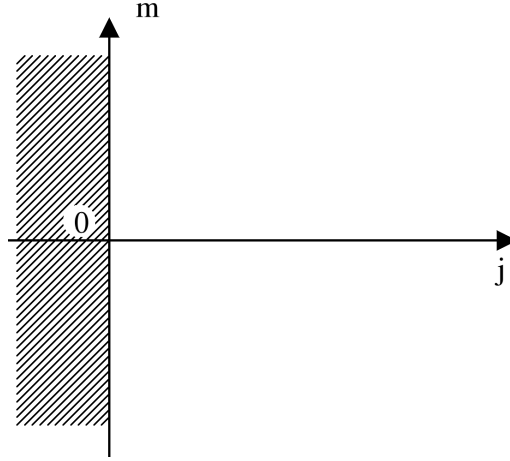


Fig. 12.3. Etape 1 de la démonstration : Plan des nombres quantiques j et m : la zone des couples (m, j) identifiés comme impossibles au vu du résultat sur la positivité de j est hachurée. La zone des couples (m, j) encore potentiellement valables reste en blanc. Cette dernière zone inclue l'axe vertical $j = 0$.

Etape 2 : $-j \leq m \leq j$

On définit les deux opérateurs, conjugués l'un de l'autre

$$\hat{J}_+ = \hat{J}_x + i\hat{J}_y, \quad \hat{J}_- = \hat{J}_x - i\hat{J}_y.$$

Notons qu'ils ne sont pas hermitiens, et ne peuvent donc pas correspondre à des observables. Du fait des relations de commutation entre $\hat{J}_x$, $\hat{J}_y$ et $\hat{J}_z$, on peut montrer que ces deux opérateurs commutent avec $\hat{J}^2$. On peut calculer sur la base des relations de commutation que :

$$\hat{J}_+ \hat{J}_- = \hat{J}^2 - \hat{J}_z^2 + \hbar \hat{J}_z, \quad \hat{J}_- \hat{J}_+ = \hat{J}^2 - \hat{J}_z^2 - \hbar \hat{J}_z.$$

Il vient alors

$$\begin{aligned} \left| \hat{J}_+ |j, m, k\rangle \right|^2 &= \langle j, m, k | \hat{J}_- \hat{J}_+ |j, m, k\rangle = j(j+1)\hbar^2 - m^2\hbar^2 - m\hbar^2 \\ &= \hbar^2 [j(j+1) - m(m+1)] \geq 0. \end{aligned} \quad (12.5)$$

De même,

$$\left| \hat{J}_- |j, m, k\rangle \right|^2 = \hbar^2 [j(j+1) - m(m-1)] \geq 0.$$

On déduit de ces inégalités respectivement que $-(j+1) \leq m \leq j$ et $-j \leq m \leq j+1$, soit $-j \leq m \leq j$.

Etape 3 : $(j - m)$ est un nombre entier positif ou nul

Examinons quelques propriétés du vecteur $\hat{J}_+ |j, m, k\rangle$. Tout d'abord, on déduit des deux résultats précédents, que $\hat{J}_+ |j, m, k\rangle$ est nul à la seule condition que $j = m$. On peut par ailleurs calculer que :

$$\hat{J}_z \hat{J}_+ = \hat{J}_z \hat{J}_x + i\hat{J}_z \hat{J}_y = (i\hbar \hat{J}_y + \hat{J}_x \hat{J}_z) + i(-i\hbar \hat{J}_x + \hat{J}_y \hat{J}_z) = \hbar \hat{J}_+ + \hat{J}_+ \hat{J}_z.$$

Ainsi,

$$\hat{J}_z (\hat{J}_+ |j, m, k\rangle) = (m+1)\hbar \hat{J}_+ |j, m, k\rangle.$$

En outre, comme $\hat{J}_+$ et $\hat{J}^2$ commutent,

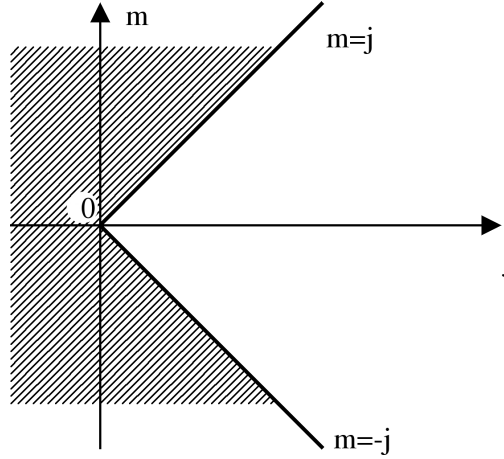


Fig. 12.4. Etape 2 de la démonstration : Représentation graphique de la condition $-j \leq m \leq j$ dans le plan des nombres quantiques j et m . La zone hachurée indique les couples (j, m) identifiés comme impossibles. Les demi-droites $m = j$ et $m = -j$ ($j \geq 0$) restent des lieux possibles.

$$\hat{J}^2 \left(\hat{J}_+ |j, m, k\rangle \right) = \hat{J}_+ \hat{J}^2 |j, m, k\rangle = j(j+1)\hbar^2 \hat{J}_+ |j, m, k\rangle.$$

Dans la mesure où $\hat{J}_+ |j, m, k\rangle$ n'est pas le vecteur nul (c'est à dire $m \neq j$), ceci montre qu'il est vecteur propre commun de $\hat{J}_z$ et $\hat{J}^2$ pour les valeurs propres $(m+1)\hbar$ et $j(j+1)\hbar^2$ respectivement.

Pour une valeur de j fixée, si on dispose d'un vecteur $|j, m, k\rangle$ déjà identifié comme élément de la base recherchée, il est possible de construire des vecteurs de type $\hat{J}_+ |j, m+1, k\rangle, \hat{J}_+ |j, m+2, k\rangle, \dots, \hat{J}_+ |j, m+p, k\rangle, \dots$ (p entier), prouvant par là même l'existence des couples de valeurs propres $(j, m+1), (j, m+2), \dots, (j, m+p), \dots$. Cette construction ne peut être arrêtée que s'il existe un nombre entier positif p tel que $m+p = j$. Dans ce cas en effet $\hat{J}_+ |j, m+p, k\rangle = 0$ (par (12.5)), et lorsqu'on itère par $\hat{J}_+$, on ne génère plus que des vecteurs nuls. Le dernier couple d'indices généré par la récurrence est alors le couple $(j, m+p) = (j, j)$.

Or, les valeurs propres vérifient $m \leq j$. La récurrence ne peut donc jamais se poursuivre indéfiniment. Quel que soit le vecteur $|j, m, k\rangle$ pris dans la base orthonormale recherchée, il existe donc un entier p positif ou nul tel que $j = m+p$ (voir la figure 12.5).

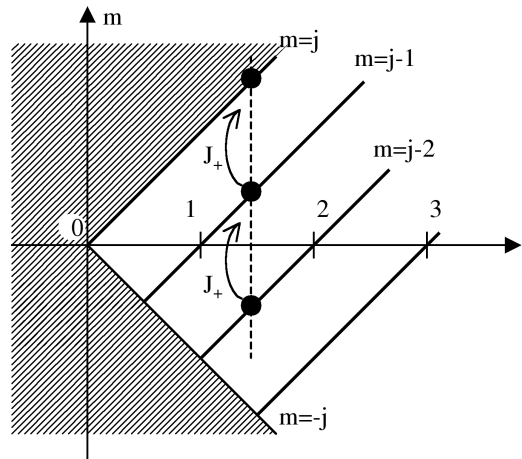


Fig. 12.5. Etape 3 de la démonstration : Action de l'opérateur $\hat{J}_+$ sur les couples de valeurs (j, m) . Les seules zones restant autorisées pour ces valeurs sont les demi-droites $m = j - \text{entier}$ ($m \geq -j$).

Etape 4 : j est un nombre entier ou demi-entier

On montre tout d'abord que $(j + m)$ est un nombre entier positif ou nul, de même manière que le résultat précédent, mais en considérant l'opérateur $\hat{J}_-$ qui induit une récurrence descendante sur les valeurs de m , et ce, toujours à j fixé. Ainsi, $2j$ est entier, et on conclut à l'aide du dessin de la figure 12.6. A ce stade, nous avons déjà prouvé les résultats 1, 2 et 3 du théorème de quantification.

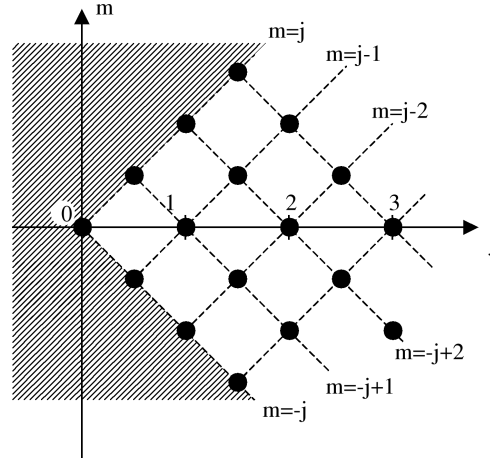


Fig. 12.6. Etape 4 de la démonstration : Seuls subsistent les points d'intersection entre les demi-droites $m = j - \text{entier}$ avec $m \geq -j$ et $m = -j + \text{entier}$ avec $m \leq j$. On observe que l'ensemble de ces points sont de coordonnées entières ou demi-entières.

Etape 5 : Pour $-j \leq m < j$, les espaces $V_{j,m}$ et $V_{j,m+1}$ ont même dimension

Il reste à préciser les dimensions des sous espaces propres. Il peut se produire en premier lieu que les deux sous espaces soient de dimension nulle. Il n'y a alors rien à démontrer. Nous nous intéressons au cas où au moins un des sous-espaces, $V_{j,m}$ par exemple, a une dimension $d_{j,m} \geq 1$. La démonstration qui suit s'adapte sans difficulté au cas où l'information initiale est $d_{j,m+1} \geq 1$.

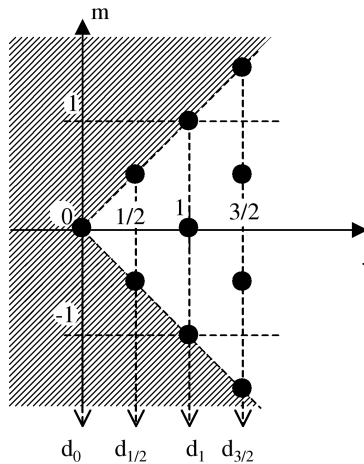


Fig. 12.7. Etape 5 de la démonstration : Egalité des dimensions des sous espaces associés à des mêmes valeurs de j .

Considérons $\hat{J}_+$ comme une application de $V_{j,m}$ vers $V_{j,m+1}$ uniquement. Nous allons montrer que $\hat{J}_+$ est injective. On a déjà établi que

$$\forall |\phi\rangle \in V_{j,m}, \quad \hat{J}_- \hat{J}_+ |\phi\rangle = \hbar^2 [j(j+1) - m(m+1)] |\phi\rangle \in V_{j,m}.$$

Comme ici $-j \leq m \leq j-1$, le terme $[j(j+1) - m(m+1)]$ n'est pas nul. L'opérateur $\hat{J}_- \hat{J}_+$ restreint à $V_{j,m}$ est donc un simple opérateur de multiplication par une constante différente de 0, donc une injection. L'opérateur $\hat{J}_+$ est donc lui même injectif. Ceci montre que

$$d_{j,m+1} \geq d_{j,m}.$$

L'inégalité inverse s'établit en montrant de même que l'application $\hat{J}_-$ est injective de $V_{j,m+1}$ vers $V_{j,m}$ (voir la figure 12.7).

12.5 Quantification du moment cinétique orbital

La recherche des états propres du moment cinétique orbital peut se faire de manière analytique. Les résultats s'énoncent dans le système de coordonnées sphériques (r, θ, φ) (voir la figure 12.8). Dans le formalisme des fonctions d'onde, les opérateurs $\hat{L}_z$, $\hat{L}^2$ et $\hat{L}_\pm$ s'expriment selon :

$$\begin{aligned} \hat{L}^2 &= -\hbar^2 \left(\frac{\partial^2}{\partial \theta^2} + \frac{1}{\tan \theta} \frac{\partial}{\partial \theta} + \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \varphi^2} \right), \\ \hat{L}_z &= \frac{\hbar}{i} \frac{\partial}{\partial \varphi}, \\ \hat{L}_\pm &= \hbar e^{\pm i\varphi} \left(\frac{\partial}{\partial \theta} + i \cot \theta \frac{\partial}{\partial \varphi} \right). \end{aligned}$$

Le résultat de quantification apparaît un cas particulier du résultat général de quantification (Théorème 12.1) :

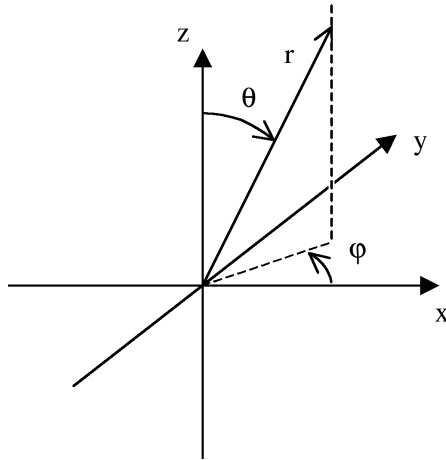


Fig. 12.8. Définition des coordonnées sphériques ρ , θ et φ .

Proposition 12.2. Les observables de moment cinétique orbital $\hat{L}_z$ et $\hat{L}^2$ sont diagonalisables dans une même base orthonormée, avec les propriétés suivantes :

- (1) les valeurs propres de $\hat{L}_z$ et $\hat{L}^2$ sont de la forme $m\hbar$ et $l(l+1)\hbar^2$ respectivement, avec $l \in \mathbb{N}$, et $m \in \mathbb{Z}$ tel que $-l \leq m \leq l$. Les nombres l et m sont appelés nombres quantiques. Ils indiquent les sous espaces propres communs aux deux opérateurs ;

- (2) les sous espaces propres $V_{l,m}$ sont de dimension infinie ;
 (3) La forme générale des états propres du sous-espace $V_{l,m}$ est, dans le formalisme des fonctions d'onde,

$$\phi(r, \theta, \varphi) = Y_l^m(\theta, \varphi) f(r),$$

avec $f(r)$ une fonction quelconque de la variable radiale r , dite 'partie radiale' de la fonction d'onde, et $Y_l^m(\theta, \varphi)$ une famille de fonctions indicées par l et m , dites « harmoniques sphériques » (voir le § 12.5).

On pourra remarquer, c'est une tradition en mécanique quantique, que le nombre quantique que nous avons noté j dans le cas général se note l dans le cas particulier du moment orbital. La liste des valeurs propres et des dimensions de sous espaces qui apparaissent ici est résumée dans le diagramme de la figure 12.9.

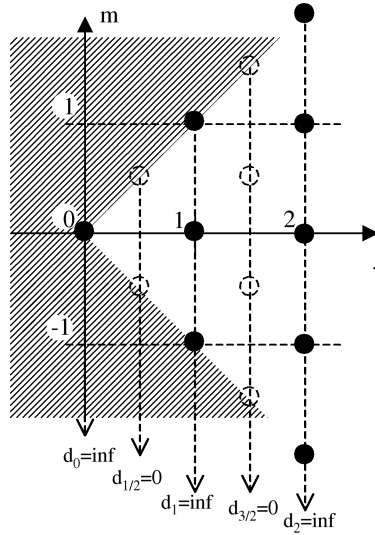


Fig. 12.9. Dimensions des sous espaces $V_{j,m}$ pour le moment cinétique orbital : seules les valeurs propres entières sont autorisées, avec des dimensions de sous-espace infinies.

C'est le fait que, dans la forme générale des états propres, la partie radiale puisse être choisie de manière arbitraire, qui implique que les sous espaces propres sont de dimension infinie. Le fait que les nombres quantiques l et m ne puissent prendre que des valeurs entières est quant à lui une conséquence de ce que nous travaillons dans un espace de fonctions d'onde de l'espace à trois dimensions (variables (r, θ, φ) en coordonnées sphériques), comme le montre la démonstration qui suit. La recherche analytique des états propres du moment cinétique orbital fait plein usage des résultats généraux sur la quantification des moments cinétiques. Nous en donnons ici les grandes lignes.

Etape 1 : Forme factorisée des états du type $|l, l, k\rangle$

Les états $|l, l, k\rangle$, c'est à dire, les états du type $|l, m, k\rangle$ pour lesquels $m = l$, sont tels que

$$\begin{cases} \hat{J}_+ |l, l, k\rangle = 0, \\ \hat{J}_z |l, l, k\rangle = l\hbar |l, l, k\rangle. \end{cases}$$

Notant $\phi_{l,l,k}(r, \theta, \varphi)$ la fonction associée à $|l, l, k\rangle$, on a alors, dans le formalisme des fonctions d'onde,

$$\begin{cases} \hbar e^{i\varphi} \left(\frac{\partial \phi_{l,l,k}}{\partial \theta} + i \cot \theta \frac{\partial \phi_{l,l,k}}{\partial \varphi} \right) = 0 \\ \frac{\hbar}{i} \frac{\partial \phi_{l,l,k}}{\partial \varphi} = l \hbar \phi_{l,l,k}. \end{cases}$$

On constate en premier lieu qu'aucune dérivée en r n'intervient dans ce système. Il est donc possible de factoriser $\phi_{l,l,k}$ selon : $\phi_{l,l,k}(r, \theta, \varphi) = f(r)g(\theta, \varphi)$, avec $f(r)$ arbitraire. On obtient pour la partie angulaire $g(\theta, \varphi)$:

$$\begin{cases} \frac{\partial g}{\partial \theta} + i \cot \theta \frac{\partial g}{\partial \varphi} = \frac{\partial g}{\partial \theta} - l g \cot \theta = 0, \\ \frac{\partial g}{\partial \varphi} = i l g. \end{cases}$$

De la seconde équation, on déduit que $g(\theta, \varphi)$ se factorise lui même en $g(\theta, \varphi) = h(\theta)e^{il\varphi}$, avec $h(\theta)$ une nouvelle fonction qu'il reste à déterminer.

Etape 2 : l est entier

En tant que partie angulaire d'une fonction de l'espace, $g(\theta, \varphi)$ doit vérifier $g(\theta, \varphi) = g(\theta, \varphi + 2\pi)$. Il est donc nécessaire, au vu de la forme factorisée de $g(\theta, \varphi)$, que $e^{i2l\pi} = 1$. On en déduit que l est obligatoirement un nombre entier. Alors même que le théorème général de quantification du moment cinétique montre que les relations de commutation $i\hbar \hat{\mathbf{L}} = \hat{\mathbf{L}} \wedge \hat{\mathbf{L}}$ autorisent *a priori* les valeurs demi-entières de l , ces valeurs se révèlent ici interdites.

Etape 3 : Forme générale des états $|l, l, k\rangle$

Cette forme est obtenue en résolvant, à l entier fixé, l'équation sur $h(\theta)$, *i.e.* $\partial_\theta h = l h \cot \theta$. La solution générale obtenue pour $\phi_{l,l,k}(r, \theta, \varphi)$ est ainsi

$$\phi_{l,l,k}(r, \theta, \varphi) = C_{l,k} (\sin \theta)^l e^{il\varphi} f_k(r),$$

avec $C_{l,k}$ une constante de normalisation. On peut vérifier qu'il s'agit également bien d'états propres de l'opérateur $\hat{L}^2$ avec la valeur propre $l(l+1)\hbar^2$.

Etape 4 : Forme générale des autres états $|l, m, k\rangle$

Elle peut s'obtenir en appliquant la récurrence descendante :

$$\phi_{l,m-1,k}(r, \theta, \varphi) = \frac{1}{\hbar \sqrt{l(l+1) - m(m-1)}} \hat{J}_- \phi_{l,m,k}(r, \theta, \varphi).$$

Dans cette récurrence, la variable r n'est qu'un paramètre. Avec cette manière de procéder, à l et k fixés, toutes les fonctions d'onde $\phi_{l,m,k}(r, \theta, \varphi)$ obtenues pour $m = -l, \dots, l$ ont ainsi la même partie radiale $f_k(r)$, et ne diffèrent que par leur partie angulaire.

Application : les orbitales atomiques pour les atomes hydrogénéoïdes

Nous précisons ici la forme des orbitales atomiques pour des atomes hydrogénéoïdes, c'est-à-dire dans le cas où une charge positive Ze est placée en $x = 0$. Dans ce cas, le Hamiltonien du système est

$$\hat{H} = -\frac{\hbar^2}{2m} \Delta - \frac{Ze^2}{4\pi\epsilon_0|x|}.$$

On se place par la suite en unités atomiques (voir Appendice A, § A.2). Dans ce cas, on a en coordonnées sphériques :

$$\hat{H} = -\frac{1}{2}\Delta - \frac{Z}{|x|} = -\frac{1}{2r}\frac{\partial^2}{\partial r^2}(r\cdot) - \frac{1}{2r^2}\Delta_S - \frac{Z}{r}.$$

L'opérateur Δ_S est l'opérateur de Laplace-Beltrami, qui est l'analogue du Laplacien sur la sphère $\mathbb{S}^2$. On montre que $\Delta_S = -\hat{L}^2$, et cet opérateur a pour vecteurs propres les harmoniques sphériques.

Définition-Proposition 12.1 (Harmoniques sphériques). *Les harmoniques sphériques sont les vecteurs propres de Δ_S , et on les note $(Y_{l,m})_{l \geq 0, -l \leq m \leq l}$. Ces fonctions forment une base hilbertienne de $L^2(\mathbb{S}^2)$:*

- pour tous couples $(l, m), (l', m')$ tels que $l, l' \geq 0, -l \leq m \leq l, -l' \leq m' \leq l'$,

$$\langle Y_l^m, Y_{l'}^{m'} \rangle_{L^2(\mathbb{S}^2)} = \int_0^\pi \int_0^{2\pi} \overline{Y_l^m(\theta, \varphi)} Y_{l'}^{m'}(\theta, \varphi) \sin \theta d\theta d\varphi = \delta_{ll'} \delta_{mm'} ;$$

- toute fonction $f \in L^2(\mathbb{S}^2)$ se décompose de manière unique sur la base des harmoniques sphériques :

$$f = \sum_{l=0}^{+\infty} \sum_{m=-l}^l \langle Y_l^m, f \rangle_{L^2(\mathbb{S}^2)} Y_l^m,$$

la série ci-dessus étant convergente au sens de $L^2(\mathbb{S}^2)$.

- les fonctions Y_l^m sont des vecteurs propres de $\hat{L}^2$ et $\hat{L}_z$:

$$\Delta_S Y_l^m = -l(l+1)Y_l^m, \quad -i\partial_\varphi Y_l^m = mY_l^m.$$

L'expression analytique des harmoniques sphériques est

$$Y_l^m(\theta, \varphi) = \left(\frac{2l+1}{4\pi} \frac{(l-m)!}{(l+m)!} \right)^{1/2} P_l^m(\cos \theta) e^{im\varphi},$$

où les fonctions $P_l^m(\cos \theta)$ sont obtenues à partir des polynômes de Legendre

$$P_l(x) = \frac{(-1)^l}{2^l l!} \frac{d^l}{dx^l} [(1-x^2)^l]$$

par m dérivations :

$$\begin{cases} P_l^m(\cos \theta) = (\sin \theta)^m P_l^{(m)}(\cos \theta) & \text{si } 0 \leq m \leq l, \\ P_l^m(\cos \theta) = (-1)^m \frac{(l+m)!}{(l-m)!} P_l^{-m}(\cos \theta) & \text{si } -l \leq m \leq 0. \end{cases}$$

On vérifie facilement que $\hat{H}$ et $\hat{L}^2, \hat{L}_z$ commutent. On peut donc chercher des vecteurs propres ψ communs à tous ces opérateurs, qui sont donc de la forme :

$$\psi(r, \theta, \varphi) = R(r) Y_l^m(\theta, \varphi).$$

Ainsi, $\hat{H}\psi = E\psi$ donne une équation sur la partie radiale R du vecteur propre ψ :

$$-\frac{1}{2r} \frac{d^2}{dr^2} (rR(r)) + \frac{l(l+1)}{2r^2} R(r) - \frac{Z}{r} R(r) = ER(r).$$

On peut montrer (voir par exemple [6, § 6.1.2]) qu'à l fixé, on a une famille de parties radiales $(R_{n,l})_{n>l}$ de la forme générale

$$R_{n,l}(r) = Q_{n,l}(Zr) e^{-Zr},$$

associées à une énergie $E_{n,l} = -\frac{Z^2}{2n^2}$ qui ne dépend pas de l . Les fonctions $Q_{n,l}$ sont des polynômes de degré $n-l-1 \geq 0$. En résumé,

- les valeurs propres de $\hat{H}$ forment une suite $(E_n)_{n \geq 1}$ avec $E_n = -\frac{Z}{2n^2}$;
- la valeur propre E_n est de multiplicité n^2 , et les n^2 vecteurs propres $\psi_{nlm}(r, \theta, \varphi) = Q_{nl}(r) e^{-Zr} Y_l^m(\theta, \varphi)$ ($0 \leq l \leq n-1$, $-l \leq m \leq l$) forment une base du sous-espace associé ;
- ces vecteurs propres de $\hat{H}$ sont également vecteurs propres de $\hat{L}^2$ et $\hat{L}_z$, associés respectivement aux valeurs propres $l(l+1)$ et m .

On classe les états propres de nombres quantiques n, l, m selon le nombre quantique principal n , suivi d'une lettre correspondant au nombre quantique azimutal l ($l = 0, 1, 2, 3$ correspondent respectivement aux lettres s, p, d, f). Par exemple, le vecteur propre ψ_{100} représente l'orbitale 1s, et les cinq vecteurs propres ψ_{32m} sont les orbitales 3d.

12.6 Le spin

12.6.1 Définition

Les particules du monde physique, telles que les protons, les neutrons, les électrons et plus généralement - et de manière non exhaustive - tous les noyaux d'atomes sont porteurs d'un moment cinétique interne. Au même titre que le moment $\mathbf{L}_{\text{int}}$ des corps rigides en mécanique classique, il constitue pour ces particules un nouveau degré de liberté. Il ne remet pas en cause la description par la fonction d'onde, qui reste le bon outil pour appréhender les états de translation à travers l'espace tridimensionnel, mais vient compléter cette dernière, en apportant une information supplémentaire sur l'état du système. Ce moment cinétique sort donc du cadre de description donné par la fonction d'onde, et ne correspond pas à un moment orbital. En dépit de son analogie avec un moment cinétique interne, il n'a d'ailleurs aucun réel équivalent en mécanique classique. C'est un moment cinétique d'un nouveau type, purement quantique, que l'on nomme *spin*.

Vis à vis du formalisme quantique, qui dit nouveau degré de liberté dit nouvel espace de Hilbert pour le décrire. Par bonheur, alors que maintenant, fonction d'onde *et* spin décrivent conjointement la particule, ces deux aspects peuvent être dissociés en première approche. On peut ainsi construire un nouvel espace quantique propre au spin, nommé *espace de spin*. Ce sont les vecteurs normés de cet espace qui sont représentatifs des différents *états de spin* accessibles au système.

Sur cet espace, le spin prend la forme d'un opérateur de moment cinétique au sens élargi, c'est à dire, un triplet d'observables :

$$\hat{\mathbf{S}} = (\hat{S}_x, \hat{S}_y, \hat{S}_z).$$

Ces *opérateurs de spin* vérifient, comme tout moment cinétique, les relations de commutation² :

$$i\hbar \hat{\mathbf{S}} = \hat{\mathbf{S}} \wedge \hat{\mathbf{S}}.$$

12.6.2 Espace de spin

Nous acceptons les principes suivants, historiquement issus de l'expérience :

Définition 12.4. *L'espace de spin d'une particule est défini par les quatre principes :*

- (1) *à toute particule est associé un nombre I entier ou demi entier, et positif ou nul : le nombre de spin ;*
- (2) *l'espace de spin de la particule est un espace de Hilbert de dimension $2I + 1$;*

2. Voilà qui est un peu flou. Nous ne savons toujours pas exactement sur quel espace nous travaillons, et l'on définit déjà des opérateurs. C'est cependant le maximum que peuvent apporter les principes de la mécanique quantique. Plus choquant, ou plus puissant peut être, est même le fait que nous connaissons déjà pratiquement les spectres de ces opérateurs de spin, avant même de savoir à quel espace ils s'appliquent. Il y a aujourd'hui des théories physiques (hors programme!) qui permettent de prédire la forme de ces espaces. Historiquement, ce n'est pourtant pas la théorie, mais simplement l'expérience, qui a permis d'en deviner la nature.

(3) dans cet espace, les nombres quantiques (j, m) autorisés pour les états propres des opérateurs $\hat{S}_z$ et $\hat{S}^2$ sont uniquement ceux vérifiant :

$$\begin{cases} j = I, \\ m = -I, -I + 1, \dots, I - 1, I ; \end{cases}$$

(4) les sous espaces $V_{I,m}$ obtenus pour les $2I + 1$ valeurs possibles de m sont de dimension 1.

Le nombre de spin n'est pas une propriété qui dépend des circonstances. C'est une caractéristique intrinsèque du type de particule considéré. Par exemple,

- tous les électrons, tous les protons, et tous les neutrons ont un spin $\frac{1}{2}$;
- tous les noyaux de deutérium (isotope de l'hydrogène, dont le noyau est formé d'un proton et d'un neutron) ont un spin 1.

On peut par ailleurs remarquer que pour tout ket $|\sigma\rangle$ de l'espace de spin :

$$\hat{S}^2|\sigma\rangle = I(I+1)\hbar^2|\sigma\rangle.$$

Tous les kets de l'espace de spin sont donc états propres de $\hat{S}^2$ pour la valeur propre $I(I+1)\hbar^2$. On peut interpréter cela 'avec les mains' comme le fait que le moment cinétique interne d'une particule possède une norme parfaitement fixée qu'aucun facteur physique externe ne peut venir modifier. La vitesse de rotation d'une particule sur elle-même ne peut ainsi être ni ralentie, ni accélérée. Tout au plus peuvent avoir lieu des modifications d'orientation de l'axe de rotation.

Base orthonormale de l'espace de spin

Une base orthonormale naturelle de l'espace de spin d'une particule de nombre de spin I est obtenue en prenant les vecteurs propres (tous d'ordre de multiplicité 1) de $\hat{S}_z$ (voir la figure 12.10). Pour le spin $\frac{3}{2}$, on les note par exemple :

$$\left|-\frac{3}{2}\right\rangle, \left|-\frac{1}{2}\right\rangle, \left|\frac{1}{2}\right\rangle, \left|\frac{3}{2}\right\rangle.$$

On ne fait alors référence qu'au nombre quantique $m = -\frac{3}{2}, -\frac{1}{2}, \frac{1}{2}, \frac{3}{2}$, puisque j ici n'a plus grande importance.

Le spin : un moment venu d'ailleurs

Il est possible, et même assez courant, d'avoir des nombres de spin demi-entiers. Ces nombres quantiques sont pourtant strictement interdits pour les moments orbitaux. Il faut accepter l'idée que les particules ne tournent pas dans notre espace à trois dimensions, mais *ailleurs*. Spins et moments orbitaux ont donc bien une nature différente.

12.6.3 Manifestations physiques du spin

Pour autant qu'il vive 'ailleurs', le spin sait se manifester dans notre monde. Il confère en effet à la particule qui le porte un certain nombre de propriétés physiques bien palpables :

- la particule de spin non-nul (c'est à dire $I > 0$) est porteuse d'un *moment magnétique dipolaire*. Ce 'vecteur' moment magnétique est régi par la mécanique quantique. Il est proportionnel au spin selon la relation formelle :

$$\hat{\mu} = \gamma \hat{\mathbf{S}}.$$

On le décrit, au même titre que le moment cinétique ou l'impulsion, par un triplet d'observables $\hat{\mu} = (\hat{\mu}_x, \hat{\mu}_y, \hat{\mu}_z)$. La constante γ qui apparaît est dite *rapport gyromagnétique* de la particule. Elle est, au même titre que I , une caractéristique intrinsèque de cette dernière ;

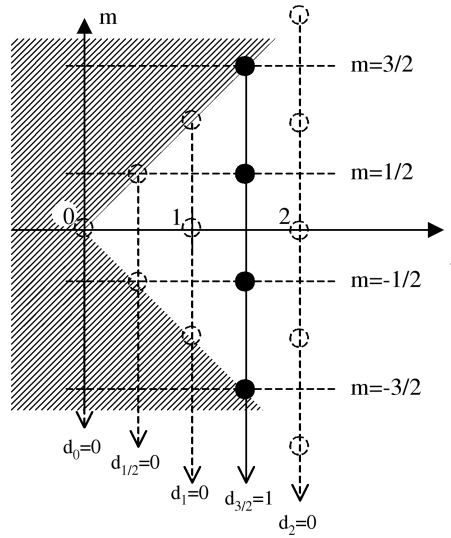


Fig. 12.10. Dimensions des sous espaces $V_{j,m}$ dans l'exemple du spin de nombre $I = \frac{3}{2}$: seuls les 4 espaces pour lesquels $j = \frac{3}{2}$ sont autorisés. Ils sont chacun de dimension 1.

- la particule de spin $I \geq 1$ est également porteuse d'un *quadripôle électrique*.

Le spin des particules est ainsi capable de réagir lorsqu'on lui applique des champs électromagnétiques, et peut en retour lui même en créer. Cette propriété concerne sans distinction les particules chargées électriquement et les particules neutres. Le neutron, de spin $\frac{1}{2}$, a donc droit lui aussi à son dipôle magnétique.

Historiquement, c'est une interaction avec un champ magnétique qui a permis pour la première fois aux spins des noyaux atomiques de trahir leur présence. Depuis, différentes techniques de spectroscopie ont été développées, notamment la résonance magnétique nucléaire (RMN, développée à partir des années 80), qui font usage des ces propriétés pour sonder la matière. Elles permettent d'analyser avec une facilité étonnante la structure chimique des composés, ou encore, depuis peu, de déterminer la conformation spatiale des protéines dans les tissus vivants. Citons également la branche cousine de la RMN qu'est l'Imagerie par Résonance Magnétique (IRM), technique de tomographie 3D, née au même moment, qui a révolutionné le diagnostic médical.

Outre leurs propriétés électromagnétiques, les spins ont des implications profondes sur le comportement des systèmes quantiques à plusieurs particules (voir le chapitre 13, où il sera question du *principe de Pauli*).

12.6.4 Un exemple d'espace de spin : l'espace du spin $\frac{1}{2}$

L'espace de spin $\frac{1}{2}$, premier espace de spin non nul, n'est pas le moindre cependant car il correspond au cas des protons, des électrons, des neutrons, et de plusieurs noyaux atomiques. C'est un espace de dimension 2. La base d'états propres ($|\frac{1}{2}\rangle$, $|\frac{-1}{2}\rangle$) de l'opérateur $\hat{S}_z$ est traditionnellement notée

$$|\uparrow\rangle, |\downarrow\rangle.$$

On prononce : spin *up* et spin *down*. Ces notions de *up* et *down* font référence à l'orientation du spin de long de l'axe z . Attention cependant, car dans l'espace de spin, $|\uparrow\rangle$ et $|\downarrow\rangle$ sont des kets orthogonaux, et ne sont donc en aucun cas les opposés l'un de l'autre. Nous touchons là encore au fait que les kets *rendent compte* de l'état du spin, mais *ne sont pas* le spin. Tout état de spin $|\sigma\rangle$ normalisé s'exprime ainsi selon :

$$|\sigma\rangle = \alpha|\uparrow\rangle + \beta|\downarrow\rangle, \quad |\alpha|^2 + |\beta|^2 = 1.$$

Il est parfois utile, dans cet espace de dimension 2, de travailler avec des notations matricielles. Dans la base $|\uparrow\rangle, |\downarrow\rangle$ on note ainsi

$$|\sigma\rangle = \begin{pmatrix} \alpha \\ \beta \end{pmatrix}.$$

Les observables $\hat{S}_x$, $\hat{S}_y$ et $\hat{S}_z$ sont également représentées sous forme matricielle :

$$\hat{S}_x = \frac{\hbar}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \hat{S}_y = \frac{\hbar}{2} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \hat{S}_z = \frac{\hbar}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Ces formes de matrices, que nous nous contentons ici d'admettre, prennent le nom de *matrices de Pauli*.

On peut également préciser le Hamiltonien dans l'espace de spins. En l'absence de toute sollicitation extérieure, celui-ci est nul.³ L'état de spin est donc stationnaire au cours du temps. En revanche, en présence d'un champ électromagnétique, des termes apparaissent. Par exemple, dans un champ magnétique uniforme $\mathbf{B}$, on a, en vertu du principe de correspondance,

$$\begin{aligned} \hat{H} &= -\widehat{\mathbf{B} \cdot \boldsymbol{\mu}} = -B_x \hat{\mu}_x - B_y \hat{\mu}_y - B_z \hat{\mu}_z \\ &= -\gamma \left(B_x \hat{S}_x + B_y \hat{S}_y + B_z \hat{S}_z \right), \end{aligned}$$

soit, en notations matricielles,

$$\hat{H} = -\frac{\gamma \hbar}{2} \begin{pmatrix} B_z & B_x - iB_y \\ B_x + iB_y & -B_z \end{pmatrix}$$

L'histoire a donné des noms aux divers Hamiltoniens susceptibles d'agir sur les espaces de spins. Le Hamiltonien d'interaction avec $\mathbf{B}$ s'appelle ici *Hamiltonien de Zeeman*.

3. Attention, ceci ne veut pas dire que l'opérateur Hamiltonien est nul dans l'espace des fonctions d'ondes ! On s'intéresse ici seulement à son action sur les spins.

Exercices

Exercice 12.1. Démontrer la Propriété 12.1

Exercice 12.2 (Satellite terrestre et quantification).

Comme n'importe quel système matériel lié, un satellite artificiel de la Terre voit ses mouvements quantifiés. On considère ici un satellite de masse m en orbite circulaire basse ($r \approx R$ = rayon de la Terre).

- (1) Rappeler, en fonction de la constante de gravitation terrestre $g \approx 10 \text{ N m}^{-1}$, la relation entre la vitesse v du satellite, et le rayon r de sa trajectoire.
- (2) Calculer approximativement le nombre quantique l correspondant au moment cinétique d'une telle orbite.
- (3) Quelle est la distance qui sépare deux orbites voisines ?
- (4) Calculer l'énergie nécessaire pour passer d'une orbite basse à l'orbite quantifiée voisine.
- (5) A quelle variation de l correspondrait une modification d'altitude de 1 m ?

Exercice 12.3 (Le four à micro-ondes).

On approxime une molécule d'eau H_2O comme l'assemblage rigide d'un atome d'oxygène (nombre de masse $A = 16$) et de deux atomes d'hydrogènes (nombre de masse $A = 1$). Les noyaux d'hydrogène sont situés à une distance d'environ $1 \text{ \AA}$ de celui d'oxygène, et l'angle $\widehat{HOH}$ est d'environ 105° .

- (1) On assimile le noyau d'oxygène au centre de masse de la molécule. Calculer le moment d'inertie I de la molécule lorsque celle-ci tourne sur elle-même autour d'un axe perpendiculaire à son plan HOH .
- (2) Le moment cinétique de la molécule d'eau est quantifié à la manière d'un moment orbital. Exprimer en fonction de I et de $\hbar$ l'énergie cinétique E_{rot} de rotation de la molécule en fonction de son nombre quantique l .
- (3) Quel est l'ordre de grandeur de l'écart d'énergie entre les niveaux les plus bas ?
- (4) A quelle longueur d'onde λ et quelle fréquence f de photon cette énergie correspond-elle ? (*La gamme des micro-ondes correspond à des fréquences entre 0.4 et 250 GHz.*)

Exercice 12.4 (L'oscillateur harmonique).

Les techniques de calcul utilisées pour extraire les propriétés spectrales des opérateurs de moment cinétique à partir des propriétés de commutations peuvent s'appliquer de manière fructueuse à d'autres systèmes. Nous étudions ici le cas unidimensionnel de l'*oscillateur harmonique*, dans lequel une particule de masse m est confinée dans un potentiel de la forme :

$$V(x) = \frac{1}{2} m \omega^2 x^2.$$

Nous allons montrer que les niveaux d'énergie quantifiés pour cette particule sont

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

- (1) Donner, en fonction des opérateurs position $\hat{x}$ et impulsion $\hat{p}_x$, l'expression du Hamiltonien unidimensionnel $\hat{H}$ qui gouverne la particule.
- (2) On introduit les opérateurs $\hat{a}^+$ et $\hat{a}$, dits respectivement *opérateurs de création* et d'*annihilation* (cette appellation vient de la théorie quantique des champs, dans laquelle ces deux opérateurs sont abondamment utilisés) :

$$\hat{a} = \frac{1}{\sqrt{2}} \left(\sqrt{\frac{m\omega}{\hbar}} \hat{x} + i \frac{\hat{p}_x}{\sqrt{m\hbar\omega}} \right), \quad \hat{a}^+ = \frac{1}{\sqrt{2}} \left(\sqrt{\frac{m\omega}{\hbar}} \hat{x} - i \frac{\hat{p}_x}{\sqrt{m\hbar\omega}} \right).$$

Donner, en fonction de $\hat{x}$ et $\hat{p}_x$, l'expression de l'opérateur $\hat{N} = \hat{a}^+ \hat{a}$. En déduire que

$$\hat{H} = \left(\hat{N} + \frac{1}{2} \right) \hbar \omega.$$

- (3) Calculer les commutateurs $[\hat{a}, \hat{a}^+]$, $[\hat{a}, \hat{N}]$ et $[\hat{a}^+, \hat{N}]$.
- (4) Soit ν une valeur propre de $\hat{N}$, et $|\varphi_\nu\rangle$ un état propre normalisé associé. Montrer que :
- ν est nécessairement positif (*Considérer la quantité $\langle \varphi_\nu | \hat{N} | \varphi_\nu \rangle$*) ;
 - $\hat{a}|\varphi_\nu\rangle$ est un vecteur de norme $\sqrt{\nu}$;
 - si $\nu \neq 0$, $\hat{a}|\varphi_\nu\rangle$ est un vecteur propre de $\hat{N}$ pour la valeur propre $(\nu - 1)$;
 - si $\nu \neq 0$, $\hat{a}^+|\varphi_\nu\rangle$ est un vecteur de norme $\sqrt{\nu + 1}$, et également un vecteur propre de $\hat{N}$ pour la valeur propre $(\nu + 1)$.
- (5) En déduire, à l'aide d'un raisonnement inspiré de la démonstration de la quantification des moments cinétiques, que les valeurs propres de $\hat{N}$ sont entières et que le spectre de $\hat{N}$ est constitué de tous les entiers positifs ou nuls. Conclure sur les valeurs propres de $\hat{H}$.

Le principe de Pauli : fermions et bosons

Le *Spin*, assimilable au moment cinétique interne des particules, est un concept purement quantique. On ne peut qu'imparfaitement l'appréhender dans le monde de la physique classique, et il dérouté notre intuition par le fait que ses nombres quantiques puissent prendre des valeurs demi-entières interdites aux moments orbitaux. Nous voyons dans ce chapitre, qu'outre leurs manifestations électromagnétiques, les spins sont à l'origine de phénomènes d'interactions particulières entre particules de même type. Le *principe de Pauli* explicite ces nouveaux phénomènes, et classe, en fonction des interactions développées, les particules physiques en deux groupes : les *fermions* et les *bosons*. Nous détaillons sur un exemple simple deux conséquences essentielles de ce principe : les *règles de remplissage* des niveaux d'énergie électroniques dans les atomes et les molécules, et le phénomène dit d'*interaction d'échange*, qui conditionne l'existence des aimants à température ambiante.

13.1 Système de plusieurs particules identiques	209
13.1.1 Position du problème	209
13.1.2 Description quantique	210
13.1.3 Indiscernabilité des particules et symétrie de la fonction d'onde	210
13.2 Principe de Pauli : spin, fermions et bosons	211
13.3 Système de deux particules sans spin	211
13.3.1 Deux particules discernables sans spin dans le fond d'un puits quantique	212
13.3.2 Paire de bosons dans le fond d'un puits quantique	213
13.3.3 Paire de fermions dans le fond d'un puits quantique	215
13.3.4 Principe de Pauli et équation de Schrödinger	217
13.4 Paire d'électrons dans un système à deux états	217
13.4.1 Etats quantiques 'à spin' de l'électron unique	218
13.4.2 Etats quantiques accessibles à la paire d'électrons	218
13.5 Règles de remplissage des niveaux électroniques	220
13.6 Interaction d'échange et magnétisme	221
Exercices	222

13.1 Système de plusieurs particules identiques

13.1.1 Position du problème

Les systèmes physiques que nous avons abordés jusqu'à présent n'étaient toujours constitués que d'une seule particule. Ce type de système, certes plaisant dans le cadre d'un cours introductif à la physique quantique, n'est cependant rencontré en pratique que dans quelques rares expériences

de laboratoire. Car même dans l'espace, le vide n'est en général pas assez poussé pour dire que les quelques particules qui s'y trouvent n'interagissent pas entre elles.

Peu nombreux sont dans la nature les systèmes qui ne sont pas constitués d'au moins quelques milliards de milliards de particules. Une seule goutte d'eau, par exemple, contient dans un volume légèrement inférieur à 50 mm^3 environ 3×10^{21} noyaux d'hydrogène (protons seuls), $1,5 \times 10^{21}$ noyaux d'oxygène, et $1,5 \times 10^{22}$ électrons, tous en interactions mutuelles. Sans même aller jusque là, la majorité des atomes et des petites molécules sont déjà en général constitués de quelques dizaines à quelques centaines de particules.

Nous souhaitons ici faire un pas vers l'étude de tels systèmes, en nous limitant cependant dans le cadre de ce cours au cas où toutes les particules sont de même nature - par exemple, uniquement les électrons dans le champ de potentiel Coulombien d'un noyau atomique ou d'une petite molécule.

13.1.2 Description quantique

En mécanique quantique, le système à plusieurs particules est décrit dans toute sa globalité par un ket $|\psi\rangle$ appartenant à un certain espace de Hilbert. Cet espace est de dimension plus grande que ceux que nous avons rencontrés jusqu'ici, du fait du plus grand nombre de degrés de libertés dont il doit rendre compte : position, impulsion, et spin de chaque particule.

Par exemple, omettant la variable de spin, l'état d'un système de N particules de masse m est décrit par une *fonction d'onde à N corps* :

$$\psi(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N, t),$$

avec $\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N$ les vecteurs position des particules $i = 1, \dots, N$. La quantité $|\psi(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N, t)|^2$ s'interprète ici comme la densité de probabilité de trouver au temps t , lors d'une mesure, les N particules aux positions respectives $\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N$. La phase de ψ reste quant à elle porteuse de l'information sur la dynamique instantanée du système. Le principe de correspondance continue à s'appliquer, et le Hamiltonien des N particules sans spin s'écrit en général :

$$\hat{H} = -\frac{\hbar^2}{2m}\Delta_{\mathbf{r}_1} - \frac{\hbar^2}{2m}\Delta_{\mathbf{r}_2} - \dots - \frac{\hbar^2}{2m}\Delta_{\mathbf{r}_N} + V(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N),$$

où $V(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N)$ est l'énergie potentielle d'interaction des N particules. Par exemple, si on considère N électrons dans le vide, ce potentiel d'interaction est la somme des interactions coulombiennes élémentaires

$$V(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N) = \sum_{1 \leq i < j \leq N} \frac{e^2}{4\pi\epsilon_0 |\mathbf{r}_i - \mathbf{r}_j|}.$$

13.1.3 Indiscernabilité des particules et symétrie de la fonction d'onde

Un problème conceptuel pour la construction de la fonction d'onde d'un système à plusieurs particules vient du fait qu'en physique, des particules de même nature sont en principe indiscernables. Cela signifie qu'on ne change pas l'état du système si on *échange* les rôles de deux particules. Or, dans l'exemple qui précède, la fonction d'onde $\psi(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N, t)$, qui affecte à chaque particule un numéro, semble bien loin de cette idée d'indiscernabilité. Comment concilier les deux ?

Si, on remplace dans $\psi(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N, t)$, par exemple, la particule 1 par la particule 2, et la particule 2 par la particule 1 (c'est l'*échange* des particules 1 et 2), on construit pour le système un nouvel état, décrit par la nouvelle fonction d'onde

$$\psi_{12}(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \dots, \mathbf{r}_N, t) = \psi(\mathbf{r}_2, \mathbf{r}_1, \mathbf{r}_3, \dots, \mathbf{r}_N, t).$$

L'échange de particules ne devant pas modifier l'état du système, ψ et ψ_{12} doivent décrire le même état. En vertu de la propriété 8.2, ces deux fonctions ne doivent donc différer au plus que par un déphasage global. Il existe ainsi $\alpha_{12} \in \mathbb{C}$ tel que $|\alpha_{12}| = 1$ et

$$\forall \mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \dots, \mathbf{r}_N \quad \psi(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \dots, \mathbf{r}_N, t) = \alpha_{12} \psi(\mathbf{r}_2, \mathbf{r}_1, \mathbf{r}_3, \dots, \mathbf{r}_N, t).$$

Echangeant les variables $\mathbf{r}_1$ et $\mathbf{r}_2$, on obtient

$$\forall \mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \dots, \mathbf{r}_N \quad \psi(\mathbf{r}_2, \mathbf{r}_1, \mathbf{r}_3, \dots, \mathbf{r}_N, t) = \alpha_{12} \psi(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \dots, \mathbf{r}_N, t),$$

d'où finalement

$$\forall \mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \dots, \mathbf{r}_N \quad \psi(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \dots, \mathbf{r}_N, t) = \alpha_{12}^2 \psi(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \dots, \mathbf{r}_N, t).$$

On a alors $\alpha_{12}^2 = 1$ et donc $\alpha_{12} = \pm 1$. Il en ressort que, nécessairement,

$$\forall \mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \dots, \mathbf{r}_N \quad \psi(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \dots, \mathbf{r}_N, t) = \psi(\mathbf{r}_2, \mathbf{r}_1, \mathbf{r}_3, \dots, \mathbf{r}_N, t),$$

ou

$$\forall \mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \dots, \mathbf{r}_N \quad \psi(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \dots, \mathbf{r}_N, t) = -\psi(\mathbf{r}_2, \mathbf{r}_1, \mathbf{r}_3, \dots, \mathbf{r}_N, t).$$

Nous aboutissons au résultat suivant :

Proposition 13.1. *Pour satisfaire au principe d'indiscernabilité, la fonction d'onde décrivant l'état d'un système de plusieurs particules identiques doit être soit symétrique, soit antisymétrique pour l'échange 2 à 2 des particules.*

13.2 Principe de Pauli : spin, fermions et bosons

Le principe de Pauli enseigne que le caractère symétrique ou antisymétrique de la fonction d'onde, ou plus généralement, lorsque nous réintroduirons la variable de spin, du ket qui décrit un système de plusieurs particules identiques ne doit rien au hasard, mais dépend directement, et uniquement, de la *nature des particules* en présence. Bien que ses fondements soient complexes et sortent largement du cadre de ce cours, son énoncé est très simple :

Définition 13.1. *Les particules du monde physique se séparent en deux familles :*

- *les particules de nombre de spin entier, nommées 'bosons'. Leurs fonctions d'onde sont symétriques pour l'échange 2 à 2 des particules. Le noyau de Deutérium, par exemple, est un boson ;*
- *les particules de nombre de spin demi-entier, nommées 'fermions'. Leurs fonctions d'onde sont antisymétriques pour l'échange 2 à 2 des particules. Les protons, les neutrons, et les électrons sont des fermions.*

C'est tout ? Oui, c'est tout. Voyons sans attendre les considérables conséquences physiques de ce principe sur un système à deux particules.

13.3 Système de deux particules sans spin

Afin de se fixer les idées sur le principe de Pauli, nous étudions en détail un système quantique dans lequel deux particules sont en présence. Pour simplifier l'exposé, nous supposons que les particules n'ont pas de spin, et qu'elles se décrivent entièrement dans un formalisme de fonctions d'onde. Cette petite entorse à la réalité physique est néanmoins très instructive.

13.3.1 Deux particules discernables sans spin dans le fond d'un puits quantique

Rappel : système à une seule particule

Nous considérons un puits quantique unidimensionnel de largeur a . Les états d'énergie pour une particule unique de masse m sont donnés par les fonctions d'onde stationnaires $\phi_1(x)$, $\phi_2(x)$, ... et les énergies associées E_1 , E_2 , ... Ces fonctions sont des états propres du Hamiltonien à une particule

$$\hat{H} = -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2}, \quad 0 \leq x \leq a.$$

Supposons que seuls les deux niveaux de plus basse énergie sont accessibles, et notons 1 et 2 ces niveaux (voir les résultats plus précis du § 10.1.1). Le puits quantique devient ce que l'on appelle un système à deux niveaux, dont le diagramme d'énergie peut se représenter selon le schéma de la figure 13.1. Supposant aussi que la particule n'a pas de spin, le ket $|\psi\rangle$ qui la décrit évolue dans un espace de dimension 2, et sa forme générale est

$$|\psi\rangle = \alpha \phi_1(x) + \beta \phi_2(x), \quad |\alpha|^2 + |\beta|^2 = 1.$$

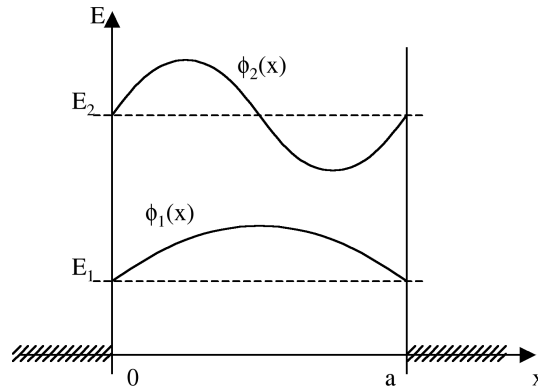


Fig. 13.1. Les deux premiers niveaux d'énergie du puits quantique unidimensionnel.

Paire de particules discernables dans un système à deux niveaux

Plaçons à présent deux particules dans le fond du puits quantique, toujours en ne gardant que les niveaux 1 et 2 accessibles. L'espace dans lequel évolue le ket de ce nouveau système est un espace de fonctions d'onde à deux variables $\psi(x_1, x_2)$, où $(x_1, x_2) \in [0, a]^2$ décrit les positions des particules 1 et 2 respectivement. On rappelle que $\psi(x_1, x_2)$ s'annule sur les bords du domaine $[0, a]$ (voir § 9.1.2).

Si les deux particules sont sans interactions, le Hamiltonien du système s'écrit, en vertu du principe de correspondance, comme la simple somme de deux Hamiltoniens à une particule :

$$\hat{H} = -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x_1^2} - \frac{\hbar^2}{2m} \frac{\partial^2}{\partial x_2^2}, \quad (x_1, x_2) \in [0, a]^2.$$

Supposons pour l'instant que les deux particules sont discernables, de sorte à pouvoir momentanément oublier les contraintes sur la symétrie ou l'antisymétrie de la fonction d'onde. On peut d'emblée, en s'appuyant sur les états propres à une particule $\phi_1(x)$ et $\phi_2(x)$, construire pour le nouvel Hamiltonien des *états propres à deux particules* :

$$\begin{aligned}
|11\rangle : \quad \phi_{11}(x_1, x_2) &= \phi_1(x_1) \phi_1(x_2), & \text{d'énergie } 2E_1, \\
|12\rangle : \quad \phi_{12}(x_1, x_2) &= \phi_1(x_1) \phi_2(x_2), & \text{d'énergie } E_1 + E_2, \\
|21\rangle : \quad \phi_{21}(x_1, x_2) &= \phi_2(x_1) \phi_1(x_2), & \text{d'énergie } E_1 + E_2, \\
|22\rangle : \quad \phi_{22}(x_1, x_2) &= \phi_2(x_1) \phi_2(x_2), & \text{d'énergie } 2E_2.
\end{aligned}$$

L'hypothèse que seuls les deux premiers niveaux du puits sont accessibles à la particule individuelle se traduit ici par le fait que ces quatre états forment une base pour les états accessibles à la paire. On vérifie facilement qu'ils constituent bien une famille orthonormée au sens du produit scalaire $L^2([0, a])$:

$$\langle \phi(x_1, x_2) | \psi(x_1, x_2) \rangle = \int_{[0, a]^2} \overline{\phi(x_1, x_2)} \psi(x_1, x_2) dx_1 dx_2.$$

Ainsi, tout état $|\psi\rangle$ du système à deux particules se décompose selon :

$$|\psi\rangle = \alpha |11\rangle + \beta |12\rangle + \gamma |21\rangle + \delta |22\rangle, \quad |\alpha|^2 + |\beta|^2 + |\gamma|^2 + |\delta|^2 = 1.$$

La fonction d'onde ψ à deux particules associée à $|\psi\rangle$ est

$$\psi(x_1, x_2) = \alpha \phi_{11}(x_1, x_2) + \beta \phi_{12}(x_1, x_2) + \gamma \phi_{21}(x_1, x_2) + \delta \phi_{22}(x_1, x_2).$$

Diagrammes énergétiques

La lecture des formules mathématiques décrivant chaque état de base n'est pas toujours chose agréable. Il est devenu d'usage, en physique quantique, de représenter le 'contenu physique' de chacune de ces fonctions de base sous la forme de *diagrammes énergétiques* (voir les figures 13.2 et 13.3). Ils font apparaître les 2 niveaux accessibles à la particule unique, et font figurer chaque particule sur le niveau qu'elle occupe dans la fonction d'onde. Cette représentation ne vaut cependant que pour les états de base, et ne se généralise pas à une superposition des états $|11\rangle$, $|12\rangle$, $|21\rangle$ et $|22\rangle$. Si l'on cherche ici à se donner une règle heuristique pour dessiner sans calcul préalable les diagrammes des états de base à deux particules, on peut dire qu'il suffit de ranger de toutes les manières possibles les particules 1 et 2 sur les deux niveaux disponibles.

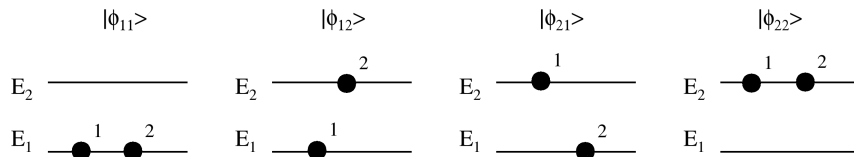


Fig. 13.2. Diagrammes énergétiques représentatifs des 4 états de base du système de deux particules discernables, dans le cas où seuls 2 niveaux sont accessibles à la particule isolée.

13.3.2 Paire de bosons dans le fond d'un puits quantique

Que se passe-t-il quand le principe de Pauli impose que la fonction d'onde soit symétrique ou anti-symétrique pour l'échange des deux particules ? Examinons le cas de deux bosons. La fonction d'onde doit être symétrique :

$$\forall x_1, x_2, \quad \psi(x_1, x_2) = \psi(x_2, x_1).$$

Cherchons, dans l'espace de Hilbert de dimension 4 défini par la base $|11\rangle$, $|12\rangle$, $|21\rangle$ et $|22\rangle$, la forme générale des états vérifiant cette condition. Dans le formalisme des fonctions d'onde, il faut que :

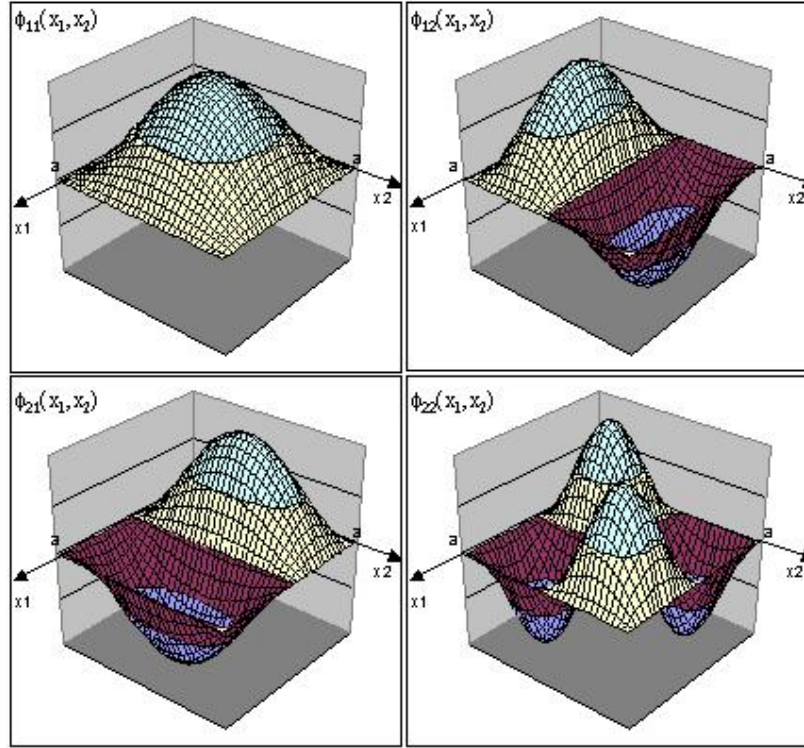


Fig. 13.3. Fonctions d'onde à deux particules représentatives des 4 états de base du système de deux particules discernables.

$$\begin{aligned} \forall x_1, x_2, \quad \psi(x_1, x_2) &= \alpha\phi_{11}(x_1, x_2) + \beta\phi_{12}(x_1, x_2) + \gamma\phi_{21}(x_1, x_2) + \delta\phi_{22}(x_1, x_2) \\ &= \alpha\phi_{11}(x_2, x_1) + \beta\phi_{12}(x_2, x_1) + \gamma\phi_{21}(x_2, x_1) + \delta\phi_{22}(x_2, x_1). \end{aligned}$$

Or, on vérifie facilement que

$$\begin{aligned} \phi_{11}(x_1, x_2) &= \phi_{11}(x_2, x_1), & \phi_{21}(x_1, x_2) &= \phi_{12}(x_2, x_1), \\ \phi_{12}(x_1, x_2) &= \phi_{21}(x_2, x_1), & \phi_{22}(x_1, x_2) &= \phi_{22}(x_2, x_1). \end{aligned}$$

Une condition nécessaire et suffisante pour que la condition de symétrie soit satisfaite est donc $\beta = \gamma$. La fonction d'onde à deux particules s'écrit dans ce cas

$$\psi(x_1, x_2) = \alpha\phi_{11}(x_1, x_2) + \beta[\phi_{12}(x_1, x_2) + \phi_{21}(x_1, x_2)] + \delta\phi_{22}(x_1, x_2).$$

Proposition 13.2. Les fonctions d'onde décrivant une paire de deux 'bosons sans spin' dans le puits quantique unidimensionnel à deux niveaux sont contenues dans un sous espace de dimension 3 de l'espace des fonctions d'onde. Une base orthonormée¹ de ce sous espace des fonctions d'onde symétriques est donnée par

$$|11\rangle, \quad |12^+\rangle = \frac{1}{\sqrt{2}}(|12\rangle + |21\rangle), \quad |22\rangle.$$

La forme générale des états décrivant la paire de bosons est

$$|\psi\rangle = \alpha|11\rangle + \beta|12^+\rangle + \gamma|22\rangle, \quad |\alpha|^2 + |\beta|^2 + |\gamma|^2 = 1.$$

Une représentation graphique de ces états est présentée en figure 13.4, les fonctions d'ondes étant tracées en figure 13.5. Comme les particules sont indiscernables, il n'est plus légitime de les

1. Propriété dont la vérification est laissée au lecteur.

numéroter 1 et 2. On observe que les 3 diagrammes énergétiques pourraient ici être obtenus de manière heuristique en rangeant de toutes les manières possibles deux particules non numérotées et interchangeables sur les deux niveaux disponibles. Une différence essentielle avec le cas discernable vient de ce qu'il n'y a ici plus qu'un seul état dans lequel les particules sont sur des niveaux différents.

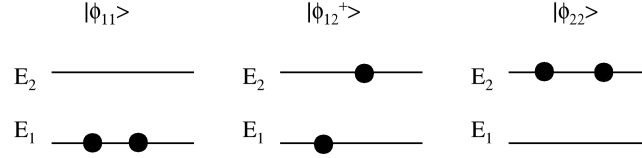


Fig. 13.4. Diagrammes énergétiques représentatifs des 3 états de base du système de deux bosons, dans le cas où seuls 2 niveaux sont accessibles à la particule isolée.

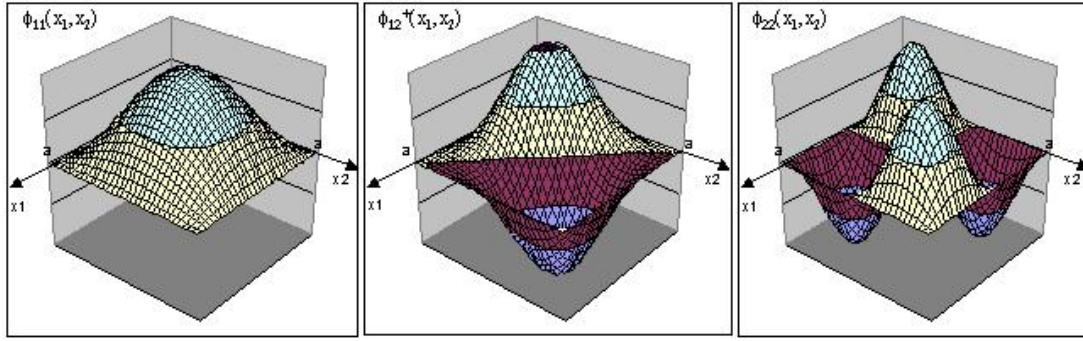


Fig. 13.5. Fonctions d'onde à deux particules représentatives des 3 états de base du système de deux bosons.

Les deux états $|12\rangle$ et $|21\rangle$ se retrouvent en quelque sorte forcés, du fait de l'indiscernabilité, de ne représenter plus qu'un seul état pour le système. Cela se traduit par la perte d'une dimension d'espace, et par la construction de l'état symétrisé $|12^+\rangle$. La relation entre $|12^+\rangle$ et son diagramme est cependant subtile. Dans cet état, il n'est en effet pas légitime de dire que nous avons une particule dans l'état 1 et une particule dans l'état 2. Car si l'on regarde de près la forme de la fonction d'onde associée :

$$\phi_{12}^+(x_1, x_2) = \frac{1}{\sqrt{2}} (\phi_1(x_1)\phi_2(x_2) + \phi_2(x_1)\phi_1(x_2))$$

on constate que chaque particule est dans une situation de superposition quantique à parts égales entre les états 1 et 2. Le diagramme d'énergie ne représente donc pas deux particules sur des niveaux séparés, mais plutôt, se contente d'indiquer quels sont les niveaux d'énergie simultanément accessibles aux deux particules.

Remarquons enfin le comportement grégaire de nos 'bosons sans spin' : pour les fonctions d'onde de la figure 13.5, la diagonale $x_1 = x_2$ est toujours un lieu de valeurs extrémales. La physique statistique enseigne que les bosons sont des particules qui aiment bien prendre des états identiques. Ce phénomène se prépare en fait déjà au niveau quantique.

13.3.3 Paire de fermions dans le fond d'un puits quantique

Dans le cas des fermions, la fonction d'onde $\psi(x_1, x_2)$ est antisymétrique. Cherchons à nouveau, dans l'espace de Hilbert de dimension 4 défini par la base $|11\rangle$, $|12\rangle$, $|21\rangle$ et $|22\rangle$, la forme

générale des états vérifiant cette condition. Dans le formalisme des fonctions d'onde, il faut que

$$\begin{aligned} \forall x_1, x_2 \quad \psi(x_1, x_2) &= \alpha\phi_{11}(x_1, x_2) + \beta\phi_{12}(x_1, x_2) + \gamma\phi_{21}(x_1, x_2) + \delta\phi_{22}(x_1, x_2) \\ &= -\alpha\phi_{11}(x_2, x_1) - \beta\phi_{12}(x_2, x_1) - \gamma\phi_{21}(x_2, x_1) - \delta\phi_{22}(x_2, x_1), \end{aligned}$$

soit

$$\alpha = \delta = 0, \quad \beta = -\gamma.$$

La fonction d'onde à deux fermions s'écrit alors :

$$\psi(x_1, x_2) = \beta [\phi_{12}(x_1, x_2) - \phi_{21}(x_1, x_2)].$$

Proposition 13.3. *Les fonctions d'onde décrivant une paire de deux 'fermions sans spin' dans le puits quantique unidimensionnel à deux niveaux sont contenues dans un sous espace de dimension 1 de l'espace des fonctions d'onde. Un vecteur de base de ce sous espace des fonctions d'onde antisymétriques est donné par :*

$$|12^-\rangle = \frac{1}{\sqrt{2}} (|12\rangle - |21\rangle).$$

La forme générale des états décrivant la paire de fermions est

$$|\psi\rangle = \beta |12^-\rangle, \quad |\beta| = 1.$$

La forme de la fonction d'onde associée à $|12^-\rangle$ est

$$\phi_{12}^- = \frac{1}{\sqrt{2}} (\phi_1(x_1)\phi_2(x_2) - \phi_2(x_1)\phi_1(x_2)).$$

Sa représentation graphique est donnée en figure 13.6. Remarquons que le diagramme énergétique associé est l'unique diagramme que l'on peut construire lorsque que l'on cherche à ranger sur les deux niveaux deux particules non numérotées et interchangeables en s'interdisant d'en mettre deux sur le même niveau. Il s'agit de la même représentation graphique que pour l'état symétrique $|12^+\rangle$. Ce diagramme doit encore se comprendre comme le fait que les niveaux 1 et 2 sont simultanément accessibles à chaque particule. En général, le risque de confusion avec l'état $|12^+\rangle$ n'existera pas, car dans une situation physique donnée, on sait si l'on travaille avec des fermions ou des bosons.

Remarquons par ailleurs que les 'fermions sans spin', non contents de s'éviter sur le diagramme d'énergie, s'évitent également dans l'espace, avec une fonction d'onde qui s'annule le long de la droite $x_1 = x_2$.

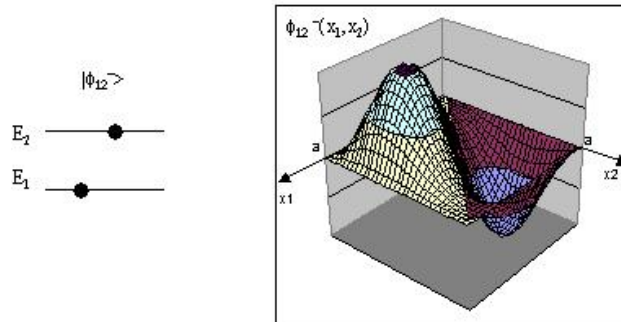


Fig. 13.6. A gauche : diagramme énergétique représentatif de l'unique état du système de deux fermions, dans le cas où seuls 2 niveaux sont accessibles à la particule isolée. A droite : la fonction d'onde à deux particules associée.

On peut enfin vérifier sur l'exemple du puits infini unidimensionnel que l'orthogonalité de $|12^-\rangle$ avec chacun des trois états symétriques $|11\rangle$, $|12^+\rangle$ et $|22\rangle$ implique que

Propriété 13.1 (Orthogonalité des états symétriques et antisymétriques). *Les sous espaces des fonctions d'onde symétriques et antisymétriques, qui décrivent respectivement les états de bosons et de fermions, sont orthogonaux.*

Preuve. On omet encore les variables de spin. Soit $\psi = \psi(\mathbf{r}_1, \dots, \mathbf{r}_N)$ antisymétrique et $\phi = \phi(\mathbf{r}_1, \dots, \mathbf{r}_N)$ symétrique. Alors,

$$\begin{aligned} \int_{\mathbb{R}^{3N}} \psi(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N) \phi(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N) d\mathbf{r} &= \int_{\mathbb{R}^{3N}} \psi(\mathbf{r}_2, \mathbf{r}_1, \dots, \mathbf{r}_N) \phi(\mathbf{r}_2, \mathbf{r}_1, \dots, \mathbf{r}_N) d\mathbf{r} \\ &= - \int_{\mathbb{R}^{3N}} \psi(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N) \phi(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N) d\mathbf{r} \end{aligned}$$

par échange des particules 1 et 2. On en déduit immédiatement

$$\int_{\mathbb{R}^{3N}} \psi(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N) \phi(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N) d\mathbf{r} = 0,$$

ce qui donne le résultat escompté. $\square$

13.3.4 Principe de Pauli et équation de Schrödinger

A l'instar de la propriété de normalisation pour les fonctions d'onde, il est important de se demander si l'équation de Schrödinger est compatible avec la conservation au cours du temps du caractère symétrique ou antisymétrique d'une fonction d'onde à plusieurs particules. Dans le cas présent, nous avons pu construire pour les sous espaces symétriques et antisymétriques des bases Hilbertiennes dont les vecteurs sont aussi des états propres du Hamiltonien. Ces vecteurs de base ont la propriété d'être stables dans le temps. En vertu du principe de superposition, on sait alors que les sous espaces symétrique et antisymétrique sont également stables.

Proposition 13.4 (Propriété générale). *Le principe de Pauli et l'équation de Schrödinger sont compatibles :*

- une fonction d'onde initialement symétrique reste symétrique au cours du temps ;
- une fonction d'onde initialement antisymétrique reste antisymétrique au cours du temps.

La démonstration de cette propriété dans le cas général est la suivante. Traitons par exemple le cas des fermions. Soit $\mathcal{S}$ une permutation de $\mathbf{r} = (\mathbf{r}_1, \dots, \mathbf{r}_N) : \mathcal{S}\mathbf{r} = (\mathbf{r}_{s(1)}, \dots, \mathbf{r}_{s(N)})$; alors, pour $\psi^0 \in L^2(\mathbb{R}^{3N})$, $\psi^0(\mathcal{S}\mathbf{r}) = (-1)^{\epsilon(\mathcal{S})} \psi^0(\mathbf{r})$ où $\epsilon(\mathcal{S})$ est la signature de la permutation $\mathcal{S}$. On note $\psi(t)$ la solution de l'équation de Schrödinger $i\hbar\partial_t\psi = -\frac{\hbar^2}{2m}\Delta\psi$ de condition initiale $\psi(0) = \psi^0$, $\tilde{\psi}(t)$ la solution de l'équation de Schrödinger $i\hbar\partial_t\tilde{\psi} = -\frac{\hbar^2}{2m}\Delta\tilde{\psi}$ de condition initiale $\tilde{\psi}(0, \mathbf{r}) = \psi^0(\mathcal{S}\mathbf{r})$, et $\psi_{\mathcal{S}}(t, \mathbf{r}) = \psi(t, \mathcal{S}(\mathbf{r}))$. On vérifie que $\tilde{\psi}(t)$ et $\psi_{\mathcal{S}}(t)$ satisfont l'équation de Schrödinger avec la même condition initiale, et sont donc égales. Ceci montre bien que le caractère antisymétrique initial est préservé au cours du temps.

Cette propriété est valable quelque soit le système physique considéré. La théorie quantique reste donc cohérente. Ouf!²

13.4 Paire d'électrons dans un système à deux états

Nous examinons une situation plus physique où la variable de spin est explicitement prise en compte. On considère ici des électrons, qui sont des fermions, et on les dispose à nouveau dans un système à deux niveaux.

2. Une manière plus abstraite de voir la chose est de remarquer que l'opérateur d'échange de deux particules est un opérateur auto-adjoint, et donc une observable. On montre ensuite facilement que cette observable commute avec le Hamiltonien.

13.4.1 Etats quantiques 'à spin' de l'électron unique

Nous sommes face à une situation nouvelle, dans laquelle il faut rendre compte, au niveau de la description quantique, de l'état de spin de l'électron. Le formalisme quantique propose simplement d'associer à la fonction d'onde $|\psi\rangle$ qui décrit la dynamique de translation de l'électron dans l'espace, un état de spin $|\sigma\rangle$. Cette association est faite par le biais du produit tensoriel $|\psi\rangle \otimes |\sigma\rangle$.

Les choses peuvent se compliquer ensuite lorsque l'on considère les superpositions d'états. Nous ne chercherons cependant pas ici à entrer plus avant dans les détails. Nous admettons que pour un électron introduit seul dans notre puits quantique à 2 niveaux, une base orthonormale d'états quantiques à une particule est alors obtenue en associant les états d'énergie $|\phi_1\rangle$ et $|\phi_2\rangle$ de la particule sans spin avec les états de base $|\uparrow\rangle$ et $|\downarrow\rangle$ de l'espace du spin $\frac{1}{2}$. On obtient la base orthonormée d'un espace de Hilbert de dimension 4 :

$$|\phi_1\rangle \otimes |\uparrow\rangle, \quad |\phi_1\rangle \otimes |\downarrow\rangle, \quad |\phi_2\rangle \otimes |\uparrow\rangle, \quad |\phi_2\rangle \otimes |\downarrow\rangle.$$

Par rapport au cas de la particule sans spin, on peut dire, d'une certaine façon, que la présence du spin *dédoublé* chaque état. A ces états peuvent être associés les diagrammes d'énergie de la figure 13.7.

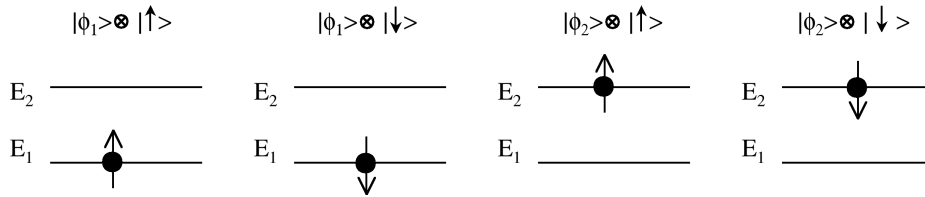


Fig. 13.7. Diagrammes énergétiques des 4 états de base pour une particule de spin $\frac{1}{2}$ placée sur un système à deux niveaux.

13.4.2 Etats quantiques accessibles à la paire d'électrons

Les électrons étant des fermions, le ket qui les représente doit être antisymétrique. La mécanique quantique construit une base de ces kets en associant :

- d'une part une base de fonctions d'onde (à deux particules cette fois) symétriques avec une base d'états de spins (à 2 spins) antisymétriques ;
- d'autre part une base de fonctions d'onde antisymétriques avec une base d'états de spins symétriques.

Nous connaissons déjà les bases symétrique et antisymétrique pour la fonction d'onde. Qu'en est-il pour l'espace de spin ?

Etats triplets et singulet d'un système à deux spins

Pour décrire conjointement un système formé de deux spins $\frac{1}{2}$, une base naturelle est celle qui combine les états $|\uparrow\rangle$ et $|\downarrow\rangle$ de chaque particule. On obtient une base orthonormée d'un espace de Hilbert de dimension 4 :

$$\begin{aligned} |\uparrow\uparrow\rangle &: \text{deux spins dans l'état } |\uparrow\rangle, \\ |\downarrow\downarrow\rangle &: \text{deux spins dans l'état } |\downarrow\rangle, \\ |\uparrow\downarrow\rangle &: \text{premier spin dans l'état } |\uparrow\rangle, \text{ deuxième spin dans l'état } |\downarrow\rangle, \\ |\downarrow\uparrow\rangle &: \text{premier spin dans l'état } |\downarrow\rangle, \text{ deuxième spin dans l'état } |\uparrow\rangle. \end{aligned}$$

De même que pour les fonctions d'onde à 2 particules, les deux derniers états peuvent être combinés pour former un état symétrique, et un état antisymétrique de spins :

$$|\uparrow\downarrow^+\rangle = \frac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle) \quad |\uparrow\downarrow^-\rangle = \frac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle).$$

On construit ainsi une base orthonormée de l'espace à deux spins, formée de trois états symétriques :

$$|\uparrow\uparrow\rangle, \quad |\downarrow\downarrow\rangle, \quad |\uparrow\downarrow^+\rangle \quad (\text{états triplets}),$$

et d'un état antisymétrique :

$$|\uparrow\downarrow^-\rangle \quad (\text{état singulet}).$$

Spins parallèles et anti-parallèles

Une particularité des états triplets et singulets est qu'ils peuvent s'interpréter en terme d'orientation relative des deux spins l'un par rapport à l'autre. Les états triplets correspondent à des états de spins *parallèles*, c'est à dire, si on se les figure comme des petits vecteurs, qui pointent dans la même direction. Le parallélisme selon l'axe z se comprend bien pour les états $|\uparrow\uparrow\rangle$ et $|\downarrow\downarrow\rangle$. Pour l'état $|\uparrow\downarrow^+\rangle$, on montre que le parallélisme ne se joue plus selon z , mais dans le plan (x, y) . Sans rentrer dans les détails, c'est le signe $+$ dans l'expression $\frac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle)$ qui permet de créer un état dans lequel les 2 spins pointent encore dans la même direction.

L'état singulet $|\uparrow\downarrow^-\rangle$ correspond quant à lui à un état de spins *antiparallèles*, c'est à dire, qui pointent dans des directions opposées, ou encore, dont la somme est nulle. C'est le signe $-$ dans l'expression $\frac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle)$ qui est à l'origine de cette orientation relative.³

Etats quantiques complets de la paire d'électrons

Procédons à présent au regroupement des fonctions d'onde et des états de spin. Nous trouvons pour décrire les états quantiques de la paire d'électrons sur un système à 2 niveaux une base orthogonale formée de six niveaux d'énergie, classés ci-dessous par énergie décroissante :

- Energie $2E_2$:

$$|22\rangle \otimes |\uparrow\downarrow^-\rangle \quad (\text{spins antiparallèles}),$$

- Energie $E_1 + E_2$:

$$\begin{aligned} &|12^+\rangle \otimes |\uparrow\downarrow^-\rangle && (\text{spins antiparallèles}), \\ &|12^-\rangle \otimes |\uparrow\uparrow\rangle, \quad |12^-\rangle \otimes |\downarrow\downarrow\rangle, \quad |12^-\rangle \otimes |\uparrow\downarrow^+\rangle, && (\text{spins parallèles}), \end{aligned}$$

- Energie $2E_1$:

$$|11\rangle \otimes |\uparrow\downarrow^-\rangle \quad (\text{spins antiparallèles}).$$

Voyons comment il est possible de rendre compte de ces états au moyens de diagrammes énergétiques. Si nous plaçons de toutes les manières possibles des particules non numérotées de spins $|\uparrow\rangle$ ou $|\downarrow\rangle$ sur les deux niveaux du diagramme d'énergie, tout en s'interdisant de mettre deux particules de même spin sur le même niveau, on obtient les 6 diagrammes de la figure 13.8.

Certains de ces diagrammes peuvent être associés de manière assez évidente à des états de la base. La véritable ambiguïté concerne les états du centre de la deuxième ligne : tous deux désignent

3. Une approche plus rigoureuse demande à considérer, dans cet espace de spin de dimensions 4, l'opérateur 'norme au carré de la somme des spins', qui vaut, dans la base $|\uparrow\uparrow\rangle, |\uparrow\downarrow\rangle, |\downarrow\uparrow\rangle, |\downarrow\downarrow\rangle$:

$$(\widehat{\mathbf{S}_1 + \mathbf{S}_2})^2 = \hbar^2 \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix}$$

On s'aperçoit alors que les états triplets et singulet forment une base d'états propres de $(\widehat{\mathbf{S}_1 + \mathbf{S}_2})^2$, avec les valeurs propres $2\hbar^2$ et 0 respectivement. On retrouve, dans la valeur $2\hbar^2$, la norme au carré $I(I+1)\hbar^2$ d'un spin de nombre de spin $I = 1$.

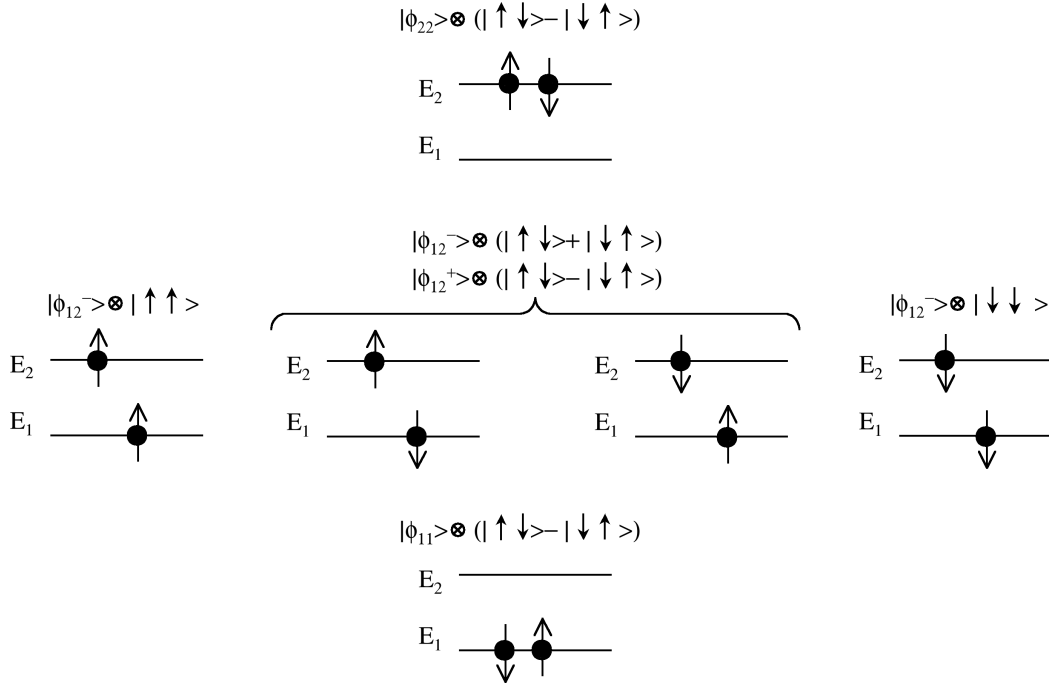


Fig. 13.8. Les 6 diagrammes énergétiques qui peuvent être construits pour deux fermions de spin $\frac{1}{2}$ placés sur un système à deux niveaux. Quatre de ces diagrammes sont directement identifiables aux états de base $|22\rangle \otimes |\uparrow\downarrow\rangle$, $|12^-\rangle \otimes |\uparrow\uparrow\rangle$, $|12^-\rangle \otimes |\downarrow\downarrow\rangle$ et $|11\rangle \otimes |\uparrow\downarrow\rangle$. L'identification avec les états $|12^+\rangle \otimes |\uparrow\downarrow\rangle$ et $|12^-\rangle \otimes |\uparrow\downarrow^+\rangle$ est plus délicate. C'est là une limite de ce système de représentation graphique (la vérité étant dans la forme mathématique des états).

bien une occupation mutuelle des niveaux 1 et 2 par les 2 électrons, mais ne sont pas capables de rendre compte du caractère parallèle ou antiparallèle des spins. Nous touchons là à une limite de ce type de représentation. Bien que ces dessins soient une aide précieuse à la compréhension du contenu physique des états avec lesquels nous travaillons, le détail de l'information se trouve bel et bien dans les formules !

Remarque 13.1. Le lecteur attentif aura remarqué qu'à aucun moment nous n'avons parlé de l'interaction électrostatique répulsive entre les deux électrons. Il est clair que pour la vraisemblance physique de ce modèle, il aurait été indispensable de prendre en compte cet aspect dans le Hamiltonien à deux particules. Notre but n'était cependant pas tant ici d'avoir un modèle vraisemblable, qu'un système simple pour illustrer la mise en pratique concrète du principe de Pauli. Mentionnons toutefois, qu'en dépit de la complexité apportée par la prise en compte des interactions coulombiennes, on constate que pour des objets tels que les atomes et les molécules, tout se passe, en première approximation, comme si il existait bel et bien des niveaux électroniques à une particule, sur lesquels viendraient se ranger un à un les électrons comme nous l'avons fait ici. Cette approximation, très courante, porte le nom d'*approximation orbitale*.

13.5 Règles de remplissage des niveaux électroniques

Il est remarquable, au regard de la lourdeur du formalisme quantique, en particulier lorsqu'on commence à mélanger fonctions d'onde et espaces de spins, que la forme des états de base d'un système à plusieurs électrons puisse être, non seulement décrite, mais également prédite, au moins dans ses grandes lignes, par des diagrammes schématisés. Quelque soit la complexité du système considéré, on constate en fait que les diagrammes prédisant les états électroniques peuvent s'obtenir avec la règle heuristique de construction suivante :

Regle 13.1. *Les électrons se rangent librement dans les niveaux d'énergie, en adoptant chacun un état de spin up ou down, et en s'interdisant uniquement que deux électrons de même spin co-existent sur un même niveau.*

Cette règle permet en particulier de construire dans des systèmes tels que les atomes et les molécules, la forme de l'état électronique à température ambiante. En générale, l'agitation thermique est insuffisante pour faire sortir ces systèmes de leurs états fondamentaux. On obtient ainsi leur structure en 'empilant dans le fond du diagramme d'énergie' les électrons les uns sur les autres, en commençant par le niveau le plus bas et en procédant ensuite par niveaux d'énergie croissante. En dépit de leurs imperfections, et notamment, leur incapacité à expliciter de manière univoque l'orientation relative de deux spins, ces diagrammes sont toujours bien plus lisibles que la forme exacte de la fonction d'onde sous jacente, surtout lorsqu'on passe à des systèmes à 3 particules ou plus. Leur utilisation à fin de raisonnements physiques est très répandue.

13.6 Interaction d'échange et magnétisme

On a vu au § 12.6 que le Hamiltonien associé à la variable de spin était nul en l'absence de champs électromagnétiques. Pourtant, si nous regardons bien le diagramme d'énergie obtenu pour la paire d'électrons, nous nous apercevons que, du fait du principe de Pauli, le niveau d'énergie le plus bas du système est atteint pour une configuration de spins *anti-parallèles*. La configuration parallèle, elle, n'apparaît que pour des niveaux d'énergie supérieurs.

Il n'est donc *pas équivalent énergétiquement* pour les électrons d'avoir leurs spins antiparallèles ou parallèles. Bien qu'il n'existe pas explicitement de Hamiltonien portant directement sur les états de spins, il se crée une sorte d'interaction 'cachée', due au principe de Pauli, qui pénalise ou favorise certaines orientations relatives.

Cette interaction est dite *interaction d'échange*. Elle existe dans les atomes, dans les molécules, mais se fait tout particulièrement sentir dans les corps solides. Dans certains d'entre eux, la différence d'énergie mise en jeu peut être de l'ordre de 0.1 eV. Ceci est considérable⁴, si l'on sait que l'agitation thermique n'est comparativement capable de contrer cette énergie qu'à des températures de l'ordre de 1000 K. L'orientation parallèle est parfois l'orientation privilégiée. A température ambiante, l'interaction d'échange peut alors, si les circonstances sont favorables, être à l'origine d'un alignement massif de spins électroniques. Du fait du moment magnétique porté par ces spins, l'objet tout entier se transforme alors en aimant.

L'interaction d'échange, et donc, le principe de Pauli, sont tout à fait fondamentaux dans ce phénomène. En leur absence, si nous ne devons compter que sur les interactions d'origine magnétique entre spins voisins pour les orienter, il faudrait descendre à des températures proche du zéro absolu pour voir un corps s'auto-aimanter. Le magnétisme spontané des corps est un exemple de phénomène purement quantique qui s'exprime à l'échelle macroscopique.

4. Rappelons qu'une énergie de $E = 0.1$ eV correspond à une température de $T = E/k_B \simeq 1150$ K.

Exercices

Exercice 13.1 (Un modèle d'Etoile à Neutrons).

En fin de vie, lorsqu'elles s'arrêtent de briller, les étoiles dont la masse est supérieure à 1,44 fois la masse du Soleil s'effondrent brutalement sur elles-mêmes. Par effet de rebond, une quantité colossale d'énergie et de matière est alors expulsée dans l'espace : c'est l'explosion de la super-nova. En lieu et place de l'étoile initiale, la matière restante se condense à tel point que de nouvelles réactions entre particules ont lieu, dans lesquelles tous les électrons et tous les protons fusionnent pour former des neutrons. Un nouveau corps prend forme : *l'étoile à neutrons*. Ce corps quasi exclusivement formé de neutrons ne possède plus d'atome. Nous nous proposons ici, à l'aide du principe de Pauli, d'estimer sa densité et sa taille.

Nous assimilons l'étoile à neutrons à une sphère de rayon R , de densité homogène, contenant un nombre N de neutrons à l'état fluide. Cette sphère peut être regardée comme un espace confinant. De fait, l'énergie des neutrons est quantifiée. On admet sans démonstration que, pour une particule sans spin confinée dans un espace de volume V , le nombre de niveaux d'énergie compris entre le niveau fondamental et une énergie donnée E peut être approximé, quand le volume est grand, par :

$$\mathcal{N}(E) = V \frac{4}{3} \pi \frac{(2mE)^{\frac{3}{2}}}{h^3},$$

avec $h = 2\pi\hbar$. On rappelle enfin que les neutrons sont des fermions, de spin $\frac{1}{2}$.

- (1) Calculer, dans l'état fondamental de l'étoile, l'énergie $E_{\max}$ maximale accessible aux neutrons en fonction du rayon R .
- (2) Montrer que l'énergie cinétique totale des neutrons s'exprime selon :

$$E_{\text{tot}} = \int_0^{E_{\max}} 2E \frac{d\mathcal{N}}{dE} dE.$$

Calculer E_{tot} en fonction de R et d'autres paramètres du problème.

- (3) On peut montrer que l'énergie potentielle de gravitation d'une sphère homogène de masse volumique ρ et de rayon R est

$$E_{\text{grav}} = -\frac{8\pi^2}{15} \mathcal{G} \rho^2 R^5,$$

avec $\mathcal{G}$ la constante de gravitation universelle. En déduire l'énergie totale de l'étoile à neutrons en fonction de R . On notera m la masse d'un neutron.

- (4) Pour quel rayon cette énergie est-elle minimale ?
- (5) Si la masse de l'étoile à neutron est de 2 masses solaires ($M_{\text{soleil}} \approx 2 \times 10^{30}$ kg), quel est son rayon ? Quelle est sa densité ?

Exercice 13.2 (Ferro et Paramagnétisme : le modèle de Curie).

Certains corps solides peuvent, selon les circonstances, développer des propriétés magnétiques à l'échelle macroscopique. C'est notamment le cas des aimants, qui présentent spontanément une aimantation et créent de forts champs magnétiques dans leur voisinage : ces corps sont dits *ferromagnétiques*. C'est également le cas des solides dits *paramagnétiques*, dans lesquels l'aimantation n'est pas spontanée, mais peut être induite par l'application d'un champ magnétique extérieur. Ce phénomène peut servir à amplifier des champs magnétiques : c'est, par exemple, la raison d'être des blocs de fer doux que l'on place à l'intérieur des bobines des transformateurs électriques.

La physique statistique montre que l'apparition de ces aimantations ne peut pas être expliquée par la mécanique classique, et constitue un phénomène purement quantique. Un des principaux mécanismes pouvant être mis en jeu est l'orientation massive de dipôles magnétiques microscopiques contenus dans le solide. Ces dipôles sont induits conjointement par le moment orbital et le spin d'électrons piégés dans les couches profondes de certains atomes (chrome, cobalt, ...).

Nous considérons ici un solide contenant en son sein un nombre N de tels dipôles, que nous supposons tous identiques. On admet que les moments cinétiques $\mathbf{J}_i$ de ces dipôles ($1 \leq i \leq N$) se comportent comme des spins, tous de même nombre quantique I , et on pose ici pour simplifier $I = \frac{1}{2}$.

Dipôles indépendants

- (1) Les dipôles, supposés sans interaction les uns avec les autres, sont placés dans un champ magnétique $\mathbf{B}$ orienté selon l'axe z . Ecrire, dans la base $(|\uparrow\rangle, |\downarrow\rangle)$, la matrice du Hamiltonien $\hat{H}_i$ d'interaction du dipôle i avec $\mathbf{B}$. Quels sont les états d'énergie? (*On introduira le rapport gyromagnétique γ*)
- (2) En vertu des principes de physique statistique, ces états propres sont occupés au prorata du facteur de Boltzman $e^{-E/k_B T}$, avec k_B la constante de Boltzman, T la température, et E l'énergie de l'état considéré. Calculer, pour l'assemblée de dipôles, la valeur moyenne $\langle m_z(T, B) \rangle$ de leur composante selon l'axe z .
- (3) Tracer l'allure de $\langle m_z(T, B) \rangle$ en fonction de B pour différentes valeurs de la température T .
- (4) On définit une *susceptibilité* magnétique χ selon

$$\chi = \left. \frac{\partial \langle m_z \rangle}{\partial B} \right|_{B=0}.$$

Calculer χ . Comment χ dépend-il de la température? La dépendance en $\frac{1}{T}$ obtenue est dite *loi de Curie*.

Dipôles en interaction d'échange

En réalité, les dipôles interagissent entre eux via l'interaction d'échange. Nous admettons que le Hamiltonien total du système devient :

$$\hat{H} \simeq \sum_{i=1}^N \hat{H}_i - \sum_{i=1}^N \frac{\sigma}{\gamma \hbar^2} \hat{J}_{z,i} \langle m_z \rangle,$$

avec σ une constante énergétique dite *énergie d'échange*.

- (5) Montrer que cela revient à placer des dipôles indépendants dans un champ magnétique effectif B_{eff} que l'on déterminera en fonction de B et $\langle m_z \rangle$.
- (6) En déduire une équation permettant de calculer $\langle m_z \rangle$ en fonction de B et de la température T . Proposer une méthode graphique de résolution.
- (7) Montrer de manière graphique que :
 - selon le signe de σ , et selon la température T , il peut apparaître une aimantation dans le solide même quand $B = 0$;
 - il existe une température T_c (dite *température de Curie*) au delà de laquelle l'aimantation spontanée devient impossible. Calculer T_c en fonction des données du problème. *Application numérique* : $\sigma = 0.3 \text{ eV}$;
 - pour $T < T_c$, la solution $\langle m_z \rangle = 0$ est instable.

La liaison chimique : des molécules diatomiques aux corps solides

Le phénomène de quantification, dont un des effets est de fixer une valeur d'énergie minimale pour les systèmes physiques usuels (l'énergie fondamentale), est à l'origine d'interactions particulières entre les atomes. En se rapprochant les uns des autres, les noyaux atomiques parviennent en effet parfois à abaisser en commun l'énergie de leurs nuages électroniques : c'est le phénomène de *liaison*, dite *chimique* ou *ionique* selon les circonstances.

Ce chapitre constitue une brève introduction aux domaines de la *chimie quantique* et de la *physique des solides*. Nous traitons dans un premier temps le mécanisme de formation de la liaison chimique. Nous expliquons en particulier pourquoi elle ne peut se former qu'entre certains atomes "compatibles". Ensuite, nous montrons que le processus de modification des états électroniques concerne également les corps solides. Des structures particulières d'états électroniques peuvent alors être observées, qui expliquent notamment les propriétés de *conduction électrique* ou la *transparence optique* de certains objets.

14.1 La molécule diatomique : liaison chimique et liaison ionique	225
14.1.1 Description quantique du système	225
14.1.2 L'approximation des Combinaisons Linéaires d'Orbitales Atomiques (CLOA)	226
14.1.3 Cas de la combinaison de deux états : traitement CLOA	227
14.1.4 Orbitales liante et antiliante	228
14.1.5 Hybridation de deux niveaux atomiques proches en énergie : la liaison chimique covalente	230
14.1.6 Hybridation de niveaux atomiques peu compatibles : liaison ionique . .	231
14.1.7 Compatibilité des orbitales atomiques	233
14.1.8 Exemples	234
14.2 Structure électronique des corps solides : théorie des bandes	235
14.2.1 Approximation CLOA	235
14.2.2 Le cas particulier des solides cristallins	237
14.2.3 propriétés électroniques des solides	241
Exercices	245

14.1 La molécule diatomique : liaison chimique et liaison ionique

14.1.1 Description quantique du système

Nous considérons deux noyaux d'atomes A et B disposés dans l'espace l'un à côté de l'autre. Nous nommons l'ensemble *molécule* AB, sans aucun *a priori* sur les éventuelles conséquences physiques de ce rapprochement. Nous notons n_A et n_B les nombres de charge de ces atomes, et $\mathbf{r}_A$

et $\mathbf{r}_B$ leurs positions. Vis à vis d'un électron qui viendrait à passer près de la molécule, le potentiel d'interaction Coulombienne résultant est (voir la figure 14.1) :

$$V(\mathbf{r}) = V_A(\mathbf{r}) + V_B(\mathbf{r}) = -\frac{n_A e^2}{4\pi\epsilon_0} \frac{1}{|\mathbf{r} - \mathbf{r}_A|} - \frac{n_B e^2}{4\pi\epsilon_0} \frac{1}{|\mathbf{r} - \mathbf{r}_B|}.$$

Nous nous intéressons aux états d'énergie accessibles à un électron qui se trouverait placé dans ce potentiel, et tout particulièrement les états liés (d'énergie négative), car ce sont eux qui déterminent la structure électronique de la molécule. Comme nous cherchons les états d'un électron seul, on omet dans la suite la variable de spin, et on ne considère que la fonction d'onde spatiale. Le Hamiltonien s'écrit alors

$$\hat{H} = -\frac{\hbar^2}{2m} \Delta + V_A(\mathbf{r}) + V_B(\mathbf{r}) = \hat{h} + \hat{V}_A + \hat{V}_B.$$

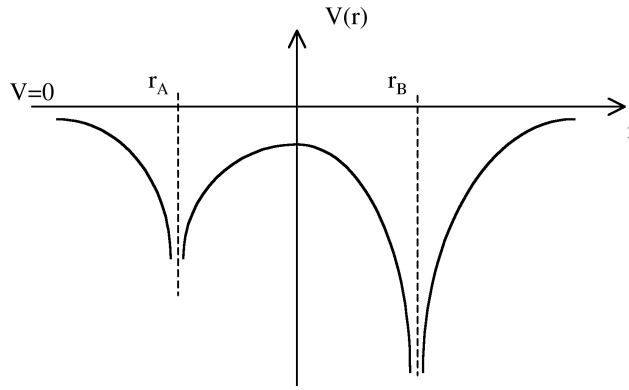


Fig. 14.1. Le double trou de potentiel créé par la juxtaposition des noyaux A et B . Sur ce schéma, $n_B > n_A$. Le puits de l'atome B s'en trouve plus large et plus profond.

14.1.2 L'approximation des Combinaisons Linéaires d'Orbitales Atomiques (CLOA)

Pour étudier les états liés de la molécule AB , il est commode de se servir de l'information que l'on connaît pour les atomes isolés A et B . On va ainsi décrire la partie spatiale de la fonction d'onde associée à la molécule AB dans la base formée par les états propres (liés) $\{|\phi_{A,n}\rangle, |\phi_{B,n}\rangle\}_{n \in \mathbb{N}}$ des atomes A et B . On emploie le terme d'*orbitales* pour désigner les états liés (atomiques ou moléculaires). On peut donc dire que l'on cherche ici à obtenir des orbitales moléculaires comme des combinaisons linéaires des orbitales atomiques (voir la figure 14.2).

L'*approximation des combinaisons linéaires d'orbitales atomiques* (CLOA) consiste à résoudre le problème sur une base *tronquée* $\{|\phi_{A,n}\rangle, |\phi_{B,m}\rangle\}_{(n,m) \in \Sigma}$, $\Sigma \subset \mathbb{N}^2$ étant de cardinal fini. Un critère permettant de choisir les vecteurs introduits dans cette base consiste, en se fixant *a priori* une gamme d'énergie dans laquelle on décide de rechercher des orbitales moléculaires, à ne retenir que les orbitales atomiques dont l'énergie tombe dans cette même gamme. Nous pourrions justifier cette manière de pratiquer *a posteriori*.

Mathématiquement, une telle approche se place dans le cadre plus général des méthodes de *synthèse modale* : pour trouver une base de Galerkin pour résoudre un problème donné, on divise ce problème en sous-problèmes plus facilement solubles, et on se sert des solutions de ces sous-problèmes comme base de Galerkin sur laquelle on va résoudre le problème initial. La base ainsi construite est parfois appelée "base réduite", dans la mesure où c'est une base de fonctions souvent de très petite taille par rapport à une discrétisation plus conventionnelle, mais toutefois très bien adaptée à la résolution du problème. Pour de plus amples développements théoriques sur la synthèse modale, on renvoie à [6, Chapitre 7]. Remarquons toutefois que la même la base non tronquée $\{|\phi_{A,n}\rangle, |\phi_{B,n}\rangle\}_{n \in \mathbb{N}}$ n'est pas une base complète de $L^2(\mathbb{R}^3)$.

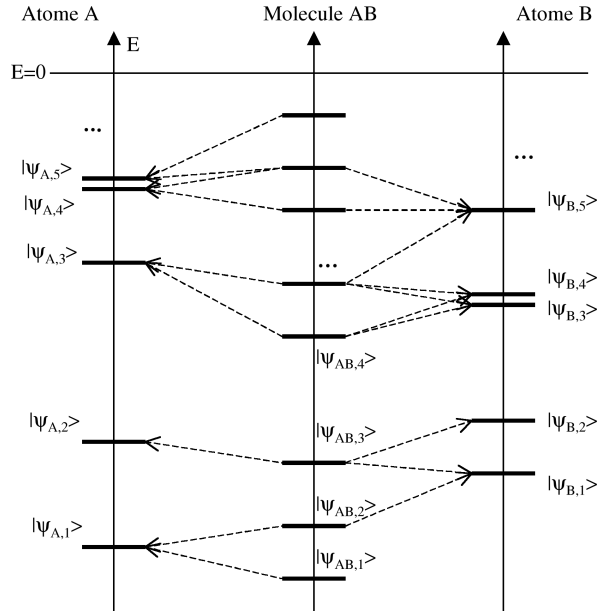


Fig. 14.2. L'approximation CLOA : les diagrammes énergétiques de l'atome A isolé, de la molécule AB , et de l'atome B isolé sont représentés de manière juxtaposée. On constate en pratique que chaque orbitale $|\phi_{AB}\rangle$ de la molécule s'approxime comme une combinaison linéaire de niveaux atomiques énergétiquement proches, désignés par les flèches en pointillés. Par exemple, l'état noté $|\phi_{AB,3}\rangle$ est quasiment exclusivement une combinaison linéaire de $|\phi_{A,2}\rangle$, $|\phi_{B,1}\rangle$ et $|\phi_{B,2}\rangle$. Il ne s'agit ici que d'un dessin de principe. Des schémas plus précis sont donnés dans la suite du cours.

14.1.3 Cas de la combinaison de deux états : traitement CLOA

Nous considérons le cas particulier où les atomes A et B possèdent chacun un état (notés respectivement $|\phi_A\rangle$ et $|\phi_B\rangle$), isolés dans leurs diagrammes énergétiques respectifs, et avec, en revanche, des énergies associées E_A et E_B proches l'une de l'autre. L'approximation CLOA assure que la base de fonctions $\{|\phi_A\rangle, |\phi_B\rangle\}$ est suffisante pour rechercher de manière semi-quantitative toutes les orbitales moléculaires dont l'énergie est proche de E_A et E_B . On cherche donc les orbitales de la forme :

$$|\phi\rangle = c_A|\phi_A\rangle + c_B|\phi_B\rangle, \quad c_A, c_B \in \mathbb{C}. \quad (14.1)$$

Avant de poursuivre, précisons quelques propriétés de $|\phi_A\rangle$ et $|\phi_B\rangle$. En tant qu'états d'énergie des atomes indépendants, ils vérifient :

$$\begin{cases} (\hat{h} + \hat{V}_A)|\phi_A\rangle = E_A|\phi_A\rangle, \\ (\hat{h} + \hat{V}_B)|\phi_B\rangle = E_B|\phi_B\rangle. \end{cases}$$

Nous supposons par ailleurs que les fonctions d'onde qui représentent $|\phi_A\rangle$ et $|\phi_B\rangle$ sont à valeurs réelles (ce qui est toujours possible pour les états propres du Hamiltonien atomique).

Energies des Orbitales moléculaires

On cherche $|\phi\rangle$ état propre de $\hat{H}$ pour une énergie E à déterminer :

$$\hat{H}|\phi\rangle = E|\phi\rangle.$$

Cette écriture ne peut être que formelle, car, la CLOA étant une méthode approchée, cette identité ne peut être vérifiée exactement. Il faut donc entendre cette relation comme une égalité *approximative* : on demande que l'égalité devienne vraie en *formulation faible* sur l'espace engendré par $|\phi_A\rangle$ et $|\phi_B\rangle$. Ceci signifie que

$$\begin{cases} \langle \phi_A | \hat{H} | \phi \rangle = \langle \phi_A | E | \phi \rangle, \\ \langle \phi_B | \hat{H} | \phi \rangle = \langle \phi_B | E | \phi \rangle. \end{cases}$$

Utilisant la décomposition (14.1), on arrive à

$$\begin{cases} c_A \langle \phi_A | \hat{H} | \phi_A \rangle + c_B \langle \phi_A | \hat{H} | \phi_B \rangle = c_A \langle \phi_A | E | \phi_A \rangle + c_B \langle \phi_A | E | \phi_B \rangle, \\ c_A \langle \phi_B | \hat{H} | \phi_A \rangle + c_B \langle \phi_B | \hat{H} | \phi_B \rangle = c_A \langle \phi_B | E | \phi_A \rangle + c_B \langle \phi_B | E | \phi_B \rangle. \end{cases}$$

Cette écriture peut se simplifier :

- (1) étant donné que les fonctions d'onde associées à $|\phi_A\rangle$ et $|\phi_B\rangle$ sont à valeur réelles, et étant donné que le Hamiltonien $\hat{H}$ est un opérateur différentiel à coefficients réels, on a $\langle \phi_A | \hat{H} | \phi_B \rangle = \langle \phi_B | \hat{H} | \phi_A \rangle \in \mathbb{R}$. Quitte à changer le signe de $|\phi_A\rangle$ par exemple, on peut supposer que

$$\beta = \langle \phi_A | \hat{H} | \phi_B \rangle = \langle \phi_B | \hat{H} | \phi_A \rangle \leq 0.$$

La constante $\beta \in \mathbb{R}$ est homogène à une énergie ;

- (2) on fait l'approximation que les propriétés des orbitales atomiques des atomes isolés sont à peu près préservées lorsque les atomes forment une molécule. Ceci est vrai au premier ordre, et on pourrait prendre en compte des termes de corrections (ce qui n'est pas très important pour une première compréhension du phénomène de liaison). On suppose donc

$$\langle \phi_A | \hat{H} | \phi_A \rangle \approx E_A, \quad \langle \phi_B | \hat{H} | \phi_B \rangle \approx E_B,$$

et

$$\langle \phi_A | \phi_B \rangle = \langle \phi_B | \phi_A \rangle \approx 0.$$

Avec ces approximations, on obtient

$$\begin{cases} E_A c_A + \beta c_B = E c_A, \\ \beta c_A + E_B c_B = E c_B. \end{cases}$$

Le recherche des états d'énergie pour l'électron situé au voisinage des noyaux atomiques A et B se résume ainsi à un problème de recherche de vecteurs propres (c_A, c_B) et de valeurs propres E dans un espace de dimension 2, pour la matrice symétrique à coefficient réels :

$$\begin{pmatrix} E_A & \beta \\ \beta & E_B \end{pmatrix}.$$

Tant que $\beta \neq 0$ ou $E_A \neq E_B$, il existe pour cette matrice deux sous espaces propres de dimension 1, orthogonaux entre eux, et associés à deux valeurs propres E_1, E_2 :

$$E_1 = \frac{E_A + E_B}{2} - \frac{\sqrt{(E_A - E_B)^2 + 4\beta^2}}{2}, \quad E_2 = \frac{E_A + E_B}{2} + \frac{\sqrt{(E_A - E_B)^2 + 4\beta^2}}{2}. \quad (14.2)$$

Le calcul des vecteurs propres associés ne pose ensuite plus de difficulté de principe. Nous interrompons toutefois les calculs, et nous discutons la signification physique de ce résultat.

14.1.4 Orbitales liante et antiliante

Dans la molécule AB, les niveaux d'énergie atomiques $|\phi_A\rangle$ et $|\phi_B\rangle$ perdent leur qualité d'états propre du Hamiltonien. Ils se combinent pour créer deux nouveaux états d'énergie, que nous notons $|\phi_1\rangle$ et $|\phi_2\rangle$ associés aux nouvelles énergies E_1 et E_2 . Dans ces états, du fait du mélange de $|\phi_A\rangle$

et $|\phi_B\rangle$, l'électron se retrouve à *cheval*, ou plutôt *réparti* sur les deux atomes de la molécule. Ce mélange des états $|\phi_A\rangle$ et $|\phi_B\rangle$ est dit phénomène d'*hybridation*.

Lorsque les états $|\phi_A\rangle$ et/ou $|\phi_B\rangle$ sont initialement porteurs d'électrons, une grande partie de la physique du rapprochement des atomes A et B se joue dans la manière dont E_1 et E_2 se comparent aux énergies initiales E_A et E_B . Dans ce qui suit, nous supposons, quitte à échanger les atomes A et B, que $E_A < E_B$.

Propriété 14.1. Les énergies E_1 , E_2 , E_A et E_B sont rangées dans l'ordre suivant :

$$E_1 \leq E_A \leq E_B \leq E_2.$$

Propriété 14.2. Les énergies E_1, E_2 sont réparties de manière symétrique autour des énergies E_A, E_B , au sens où

$$\frac{E_A + E_B}{2} = \frac{E_1 + E_2}{2}.$$

La première propriété provient simplement de la minoration $\sqrt{(E_A - E_B)^2 + 4\beta^2} \geq E_A - E_B \geq 0$. On a ainsi $E_2 \geq E_B$ et $E_1 \leq E_A$. La seconde propriété résulte directement de la définition (14.2) des énergies E_1, E_2 .

On représente les orbitales atomiques $|\phi_A\rangle$ et $|\phi_B\rangle$ et les orbitales moléculaires $|\phi_1\rangle$ et $|\phi_2\rangle$ dans le diagramme énergétique de la figure 14.3. La nouvelle orbitale $|\phi_1\rangle$ a une énergie plus basse que les niveaux atomiques. Elle offre une situation énergétiquement plus favorable pour un électron. Elle est dite *orbitale liante*. La nouvelle orbitale $|\phi_2\rangle$ a une énergie plus haute que les niveaux atomiques. Elle offre une situation énergétiquement moins favorable pour un électron. Elle est dite *orbitale anti-liante*.

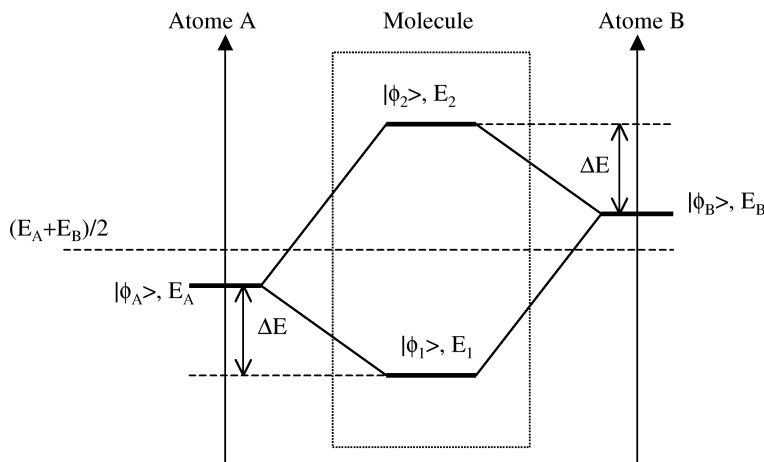


Fig. 14.3. Eclatement des niveaux atomiques des atomes A et B en deux orbitales moléculaires. L'axe d'énergie du diagramme énergétique de la molécule n'est pas représenté.

On peut se demander ce qui détermine la différence d'énergie notée $\Delta E = E_2 - E_1$. Posant ¹

$$x = \frac{E_A - E_B}{\beta} \geq 0,$$

on a

$$\Delta E = -\beta \frac{-x + \sqrt{x^2 + 4}}{2} = |\beta| \frac{2}{x + \sqrt{x^2 + 4}}.$$

1. Pour mémoire, les signes de $|\phi_A\rangle$ et $|\phi_B\rangle$ ont été choisis pour imposer la convention $\beta \leq 0$.

On voit que le caractère stabilisant de l'orbitale liante (et de même, le caractère déstabilisant de l'orbitale antiliante) est proportionnel à l'énergie $|\beta| = \left| \langle \phi_A | \hat{H} | \phi_B \rangle \right|$, avec une constante de proportionnalité d'autant plus grande que les énergies atomiques E_A et E_B sont initialement proches.

14.1.5 Hybridation de deux niveaux atomiques proches en énergie : la liaison chimique covalente

Maintenant que nous connaissons les énergies des orbitales moléculaires, cherchons à déterminer leur forme. Le calcul général n'a pas grand intérêt, et il est plus instructif de se concentrer sur quelques situations extrêmes. Dans ce paragraphe, nous considérons le cas où les niveaux atomiques qui se mélangent ont initialement la même énergie : $E_A = E_B$.

Forme des orbitales

Les énergies des orbitales moléculaires sont ici simplement :

$$E_1 = E_A + \beta, \quad E_2 = E_B - \beta.$$

Utilisant toujours la forme générale des orbitales moléculaires (14.1), on obtient, pour le vecteur propre associée à E_1 , le système d'équation d'inconnues c_A et c_B :

$$\begin{cases} E_A c_A + \beta c_B = E_1 c_A = (E_A + \beta) c_A, \\ \beta c_A + E_A c_B = E_1 c_B = (E_A + \beta) c_B. \end{cases}$$

On voit donc que $c_A = c_B$. On en déduit la forme normalisée d'un représentant de l'orbitale liante (moyennant l'approximation d'une quasi-orthogonalité de $|\phi_A\rangle$ et $|\phi_B\rangle$)

$$|\phi_1\rangle = \frac{1}{\sqrt{2}} (|\phi_A\rangle + |\phi_B\rangle) \quad (\text{orbitale liante, énergie } E_1).$$

On calcule de même l'état propre associé à l'énergie E_2 :

$$|\phi_2\rangle = \frac{1}{\sqrt{2}} (|\phi_A\rangle - |\phi_B\rangle) \quad (\text{orbitale anti-liante, énergie } E_2).$$

Il est remarquable que ces deux orbitales fassent chacune intervenir les états atomiques $|\phi_A\rangle$ et $|\phi_B\rangle$ avec des poids égaux. Un électron qui se situe dans l'état liant $|\phi_1\rangle$ se retrouve ainsi simultanément présent sur les deux atomes. On dit qu'il y a *mise en commun* de cet électron. Le diagramme d'énergie obtenu est présenté en figure 14.4.

Liaison chimique covalente

Considérons le cas où les niveaux $|\phi_A\rangle$ et $|\phi_B\rangle$ sont initialement chacun porteurs d'un électron. Lors du rapprochement, la formation de l'orbitale liante permet à ces deux électrons d'abaisser progressivement leur énergie. La répartition schématique obtenue (à l'aide de la règle de remplissage) est donnée en figure 14.5.

La *paire* d'électrons se retrouve *mise en commun* entre les deux atomes. Il en résulte un abaissement de l'énergie électronique totale d'une quantité $2|\beta|$, qui peut valoir en pratique jusqu'à 1 eV. Comparativement, l'énergie d'agitation thermique $k_B T$ à la température ambiante de $T = 300$ K n'est que de 0.026 eV. Cette situation, extrêmement stable et favorable, pousse les deux atomes à rester en permanence accrochés l'un à l'autre : c'est la *liaison chimique covalente*. Elle est très difficile à casser sans important apport extérieur d'énergie.

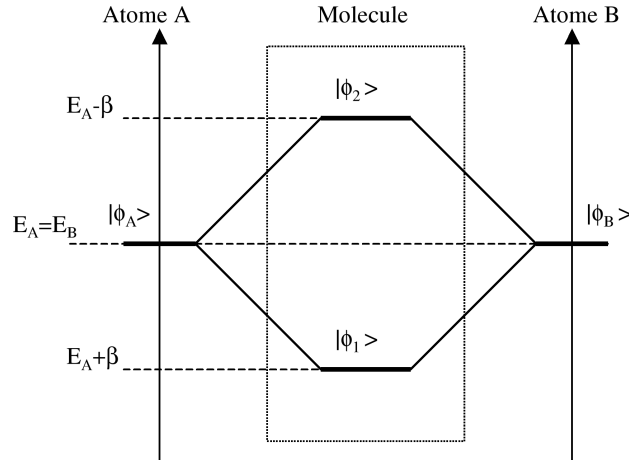


Fig. 14.4. Diagramme énergétique de la molécule issue de l'hybridation de deux orbitales atomiques de même énergie.

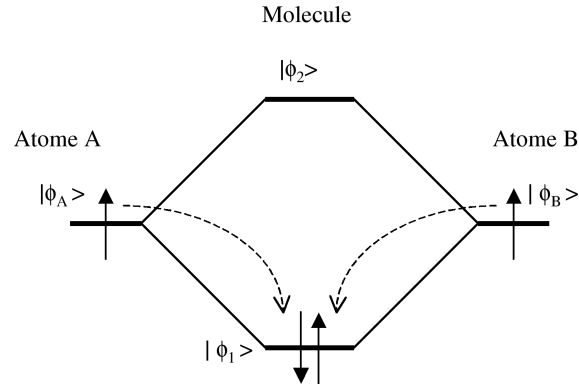


Fig. 14.5. Liaison chimique : mise en commun des électrons de chaque atome sur l'orbitale moléculaire liante.

14.1.6 Hybridation de niveaux atomiques peu compatibles : liaison ionique

Nous considérons ici le cas où la différence $E_B - E_A$ entre les niveaux d'énergie devient grande :

$$x = \frac{E_A - E_B}{\beta} \gg 1.$$

Nous supposons toujours que $E_B > E_A$.

Forme des orbitales

Les énergies des orbitales liante et antiliante s'expriment en fonction de β et x selon

$$E_1 = E_A + \beta \frac{-x + \sqrt{x^2 + 4}}{2}, \quad E_2 = E_B - \beta \frac{-x + \sqrt{x^2 + 4}}{2}.$$

Comme $x \gg 1$, on peut effectuer un développement limité en $\epsilon = \frac{1}{x}$ des énergies E_1 , E_2 :

$$E_1 = E_A + \beta\epsilon + \beta O(\epsilon^3), \quad E_2 = E_B - \beta\epsilon + \beta O(\epsilon^3).$$

Les énergies moléculaires E_1 , E_2 restent ainsi très proches des énergies atomiques. Procédons au calcul de la forme des orbitales $|\phi_1\rangle$ et $|\phi_2\rangle$. Les équations sur c_A et c_B s'écrivent ici, pour l'énergie E_1 :

$$\begin{cases} E_A c_A + \beta c_B = E_1 c_A \simeq (E_A + \beta\epsilon) c_A, \\ \beta c_A + E_B c_B = E_1 c_B \simeq (E_A + \beta\epsilon) c_B. \end{cases}$$

De ces deux équations on tire, à l'ordre 1 en ϵ , que $c_B = \epsilon c_A + O(\epsilon^2)$. On en déduit la forme, normalisée à l'ordre 1 en ϵ , de l'orbital liante :

$$|\phi_1\rangle \simeq |\phi_A\rangle + \epsilon |\phi_B\rangle \quad (\text{orbitale liante, énergie } E_1).$$

On calcule de même pour l'énergie E_2 :

$$|\phi_2\rangle \simeq |\phi_B\rangle - \epsilon |\phi_A\rangle \quad (\text{orbitale anti-liante, énergie } E_2).$$

Non seulement les énergies, mais également les orbitales moléculaires elles-même restent très proches des états atomiques $|\phi_A\rangle$ et $|\phi_B\rangle$. Ces derniers se mélangent en fait à peine. L'état liant $|\phi_1\rangle$ est ainsi essentiellement l'état $|\phi_A\rangle$, perturbé au premier ordre par $|\phi_B\rangle$, et l'état antiliant $|\phi_2\rangle$ est l'état $|\phi_B\rangle$ perturbé au premier ordre par $|\phi_A\rangle$. Un électron se trouvant dans une de ces orbitales moléculaires reste donc essentiellement localisé sur l'un ou l'autre des atomes. Il n'y a presque *pas de mise en commun*. Le diagramme d'énergie correspondant est présenté en figure 14.6.

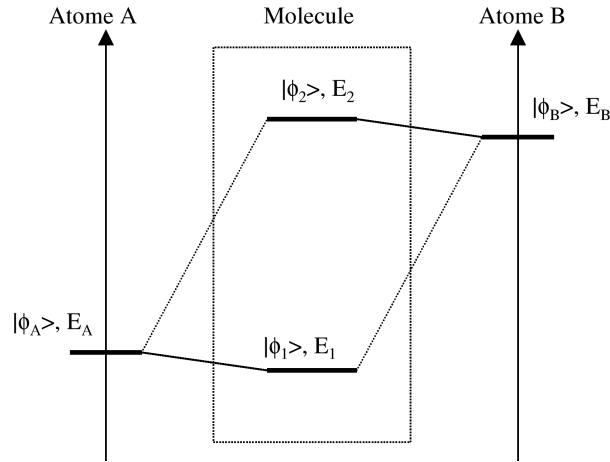


Fig. 14.6. Diagramme énergétique de la molécule issue de l'hybridation de deux orbitales atomiques très décalées en énergie.

Liaison ionique

Considérons à nouveau le cas où, avant d'être rapprochés, les niveaux $|\phi_A\rangle$ et $|\phi_B\rangle$ sont chacun porteurs d'un électron. La paire d'électrons vient se placer dans l'orbitale liante. Comme elle correspond à peu de choses près à l'état $|\phi_A\rangle$, cela revient à dire que les deux électrons se regroupent sur l'atome A : l'atome A gagne un électron, et l'atome B en perd un. La molécule ressemble ainsi à la simple juxtaposition de deux ions A^- et B^+ .

Que se passe-t-il à présent si l'on cherche à casser la molécule en séparant les deux atomes ? Il est peu probable qu'un électron retourne se placer sur l'état $|\phi_B\rangle$. L'énergie à reprendre est souvent de l'ordre de 1 eV, ce qui est extrêmement coûteux. Une solution bien plus rentable pour la paire d'électrons consiste à rester sur l'atome A , et de continuer à ne peupler que l'état $|\phi_A\rangle$.

Le coût énergétique de la séparation de la molécule n'est plus que $2\epsilon|\beta|$, soit en pratique quelque chose de l'ordre de 0.1 eV. Les deux ions A^- et B^+ se séparent, sans se restituer leurs électrons (voir la figure 14.7).

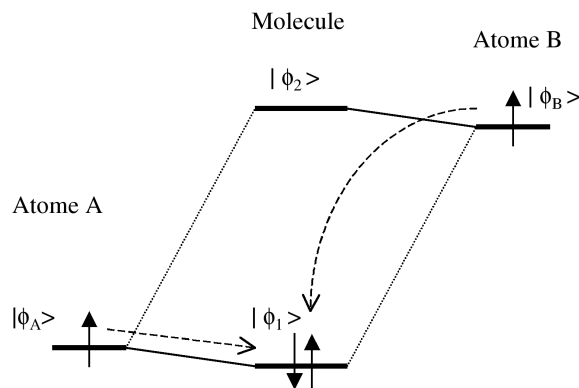


Fig. 14.7. Liaison ionique : l'électron de l'atome B 'tombe' dans l'orbitale liante de la molécule. La différence essentielle avec la liaison chimique vient de ce qu'ici, cette orbitale liante est en première approximation l'état $|\phi_A\rangle$ de l'atome A .

Un nouveau rapprochement des deux ions reformera la molécule, mais en ne mettant plus en jeu que l'énergie $2\epsilon|\beta|$, typiquement 10 fois inférieure à l'énergie de la liaison chimique. Ce type de liaison assez fragile dans laquelle un atome donne un électron à l'autre est nommé *liaison ionique*. Elle peut garder une certaine solidité à température ambiante. Néanmoins, une légère perturbation de l'environnement de la molécule (sa dilution dans l'eau, notamment) peut suffire à la dissocier. D'une certaine manière, la liaison ionique est une liaison chimique ratée...

14.1.7 Compatibilité des orbitales atomiques

Nous voyons que lors de la formation d'une molécule, deux états atomiques ne se mélangent réellement que s'ils sont initialement suffisamment proches en énergie l'un de l'autre. Nous trouvons là une justification *a posteriori* de l'approximation CLOA.

Par ailleurs même dans le cas extrême où $E_A = E_B$, l'hybridation n'est pas systématique. Le fait que le terme $\beta = \langle \phi_A | \hat{H} | \phi_B \rangle$ soit non nul, et surtout, ait une valeur significative, loin d'aller de soit, est en fait un phénomène spécifique d'une certaine *compatibilité* des états $|\phi_A\rangle$ et $|\phi_B\rangle$ à pouvoir se combiner. Cette compatibilité dépend pour beaucoup des propriétés de symétrie des fonctions d'onde qui décrivent $|\phi_A\rangle$ et $|\phi_B\rangle$, et il n'est pas rare qu'entre deux états quelconques, β soit si petit qu'il puisse être considéré comme nul au regard des énergies caractéristiques du problème. Dans le cas particulier des molécules diatomiques, on peut même avoir β exactement égal à zéro!

Malgré la profusion des états électroniques en présence, la plupart des orbitales moléculaires des molécules diatomiques sont ainsi :

- soit des états atomiques qui n'ont pas trouvé à s'hybrider ;
- soit des hybrides mettant en jeu un seul état de l'atome A et un seul état de l'atome B (le seul qui leur soit compatible, tant du point de vue de la condition $E_A \approx E_B$, que du point de vue $\beta \neq 0$).

Dans les molécules à trois atomes et plus, des situations d'hybridation plus complexes peuvent survenir, dans lesquelles trois orbitales atomiques ou plus se combinent avec des poids comparables. Il n'en reste pas moins qu'une compatibilité préalable des atomes est nécessaire pour garantir la stabilité de l'ensemble.

14.1.8 Exemples

Molécules homonucléaires

Ce sont les molécules qui rassemblent deux atomes de même nature (atomes A et B identiques). Les niveaux atomiques mis en présence sont alors les mêmes de part et d'autre de la molécule. En général, ces niveaux s'hybrident par paires, chaque niveau de l'atome A s'alliant à son homologue de l'atome B pour former les nouvelles orbitales moléculaires.

Le schéma de la figure 14.8 vaut pour tous les atomes, et montre de manière schématique comment fusionnent les 5 premiers niveaux atomiques. Les chimistes ont développé pour désigner ces orbitales atomiques et moléculaires toute une nomenclature que nous indiquons ici.

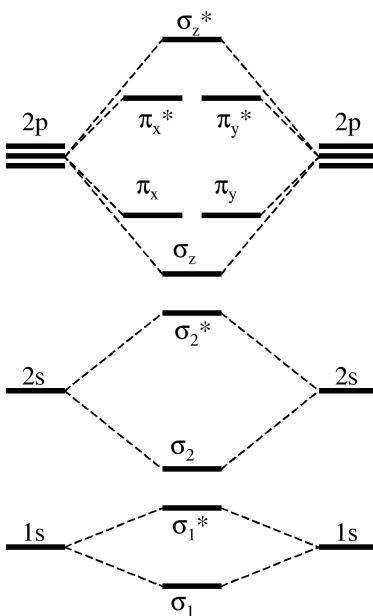


Fig. 14.8. Les premières orbitales moléculaires hybridées de la molécule homo-nucléaire. Les hybridations sont systématiques entre niveaux atomiques homologues de chaque atome. Ce diagramme de principe vaut quelque soit la molécule considérée.

Sur la base de ce schéma, étudions par exemple la formation de la molécule d'hydrogène. On peut dessiner pour cette molécule le diagramme de la figure 14.9. Les deux électrons mis en jeu tombent dans l'orbitale liante σ_1 , et abaissent leurs énergies : il y a donc bien liaison chimique. Nous avons confirmation théorique de l'existence de la molécule H_2 .

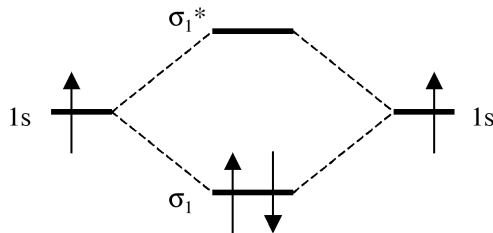


Fig. 14.9. Diagramme énergétique de la liaison chimique dans la molécule H_2 .

Essayons nous au même exercice pour créer une molécule diatomique avec des atomes d'hélium (2 électrons par atome). Du fait des règles de remplissage, les quatre électrons du système sont obligés de peupler non seulement l'orbitale liante, mais également l'orbitale antiliante (voir la figure 14.10). L'énergie électronique totale de la molécule reste ainsi très proche de celle des atomes séparés. Si l'on tient compte par ailleurs de la répulsion électrostatique entre les noyaux atomiques, de l'agitation thermique et de l'intérêt, au plan entropique, de laisser les atomes séparés, il apparaît que cette molécule n'est pas stable. Nous comprenons ainsi qualitativement pourquoi la molécule He_2 n'existe pas.

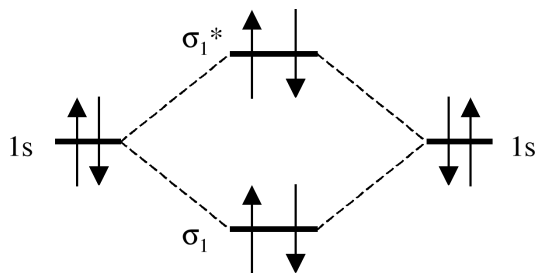
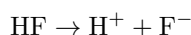


Fig. 14.10. Diagramme énergétique de la molécule He_2 . L'orbitale antiliante, qui reçoit deux électrons, empêche l'abaissement énergétique nécessaire pour la liaison chimique.

Molécules hétéronucléaires

Nous donnons ici l'exemple de la molécule d'acide fluorhydrique HF. Cette molécule, à l'état de gaz dans les conditions normales de température et de pression, se dissout facilement dans l'eau en se dissociant de manière totale selon :



L'atome de fluor porte à l'origine 9 électrons, et celui d'hydrogène un seul. Dans le diagramme de la figure 14.11, aucune des orbitales atomiques du fluor ne se trouve directement en face du niveau 1s de l'hydrogène. Dans le rapprochement, quasiment aucun état atomique n'est ainsi modifié. On note juste une discrète hybridation entre un des niveaux 2p du fluor et le niveau 1s de l'hydrogène. L'orbitale moléculaire liante dans laquelle l'électron de l'hydrogène descend n'est ainsi autre qu'une des orbitales 2p du fluor. Cet électron reste piégé sur le fluor lors de la séparation des deux atomes, pour un coût énergétique modeste. Ceci explique la facile dissociation de la molécule en ions F^- et H^+ .

14.2 Structure électronique des corps solides : théorie des bandes

Lorsqu'un corps solide se forme, ce ne sont en général pas un, mais des milliards d'atomes qui sont mis côte à côte.

14.2.1 Approximation CLOA

Considérons pour commencer le cas simple où les atomes rapprochés sont tous les mêmes. Cette situation correspond par exemple au cas du graphite ou du diamant (corps uniquement formés d'atomes de carbone), au cas des métaux purs (aluminium, cuivre, fer, zinc, plomb, or, platine, . . .), ainsi qu'au cas de certains semiconducteurs utilisés dans l'électronique, comme le silicium ou le germanium. De même que dans la molécule diatomique, les niveaux homologues des divers atomes

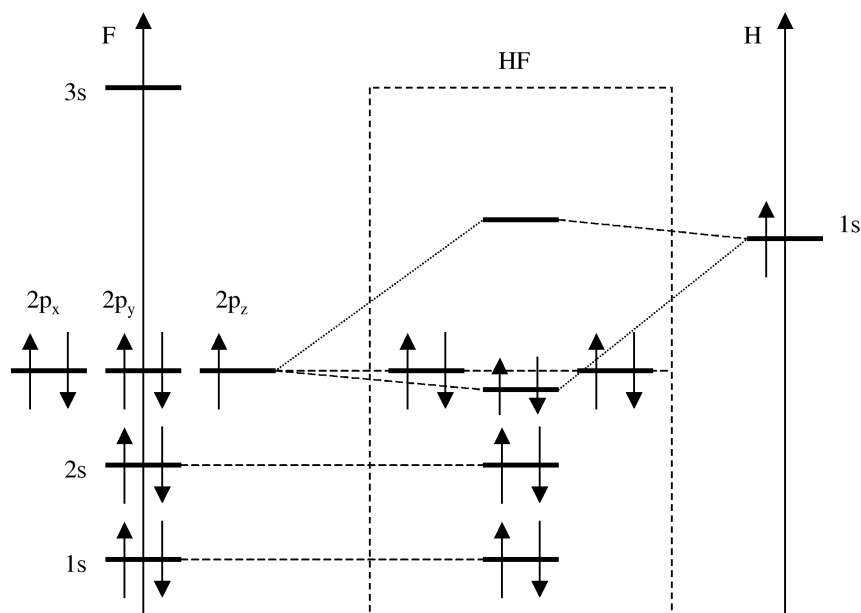


Fig. 14.11. Diagramme énergétique de la molécule d'acide fluorhydrique HF. Elle ne doit son existence qu'à un léger abaissement d'un des niveaux 2p du fluor, par discrète hybridation avec l'orbitale 1s de l'Hydrogène, pourtant déjà éloignée en énergie.

se mélangent entre eux pour former de nouveaux états d'énergie électroniques. La différence vient cependant du fait que ce ne sont pas 2, mais N états qui s'hybrident simultanément, et que ce ne sont pas 2 orbitales, mais N orbitales qui sont formées à chaque fois, avec N souvent de l'ordre de $N_A \sim 10^{23}$! Si l'on considère les énergies de ces orbitales, on constate par ailleurs qu'elles ne sont pas beaucoup plus dispersées que les orbitales liante et antiliante de la molécule diatomique. On obtient ainsi schématiquement le diagramme énergétique de la figure 14.12.

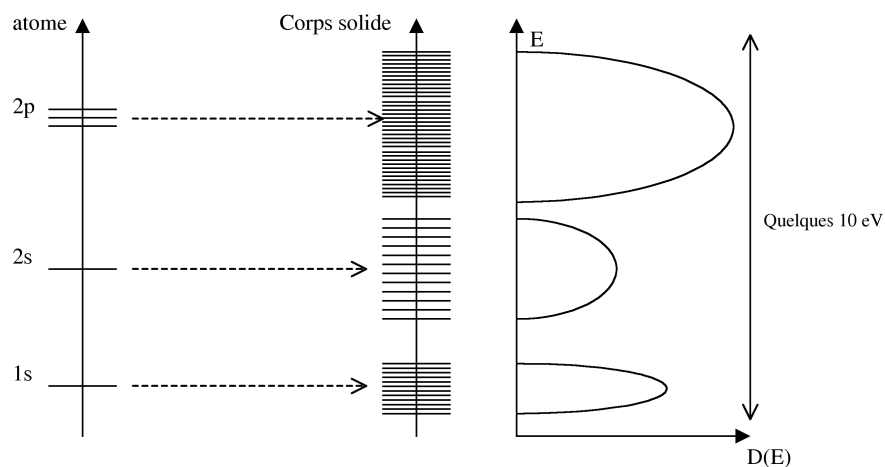


Fig. 14.12. Formation de *bandes d'énergie* par hybridation des orbitales atomiques dans un corps formé d'un seul type d'atome.

Bandes d'énergies

Par hybridation, chaque niveau atomique donne naissance à un petit continuum d'énergie, d'extension limitée : la *bande d'énergie*. Une bande résultant de l'hybridation de 1 état par atome contient exactement N orbitales, une bande résultant de l'hybridation de 2 états par atome contient exactement $2N$ orbitales, etc.

Densité d'état $\mathcal{D}(E)$

Plutôt que de considérer séparément les niveaux d'énergie de toutes les orbitales, on a coutume de définir une fonction $\mathcal{D}(E)$ dite densité d'état, telle que le nombre d'orbitale ayant une énergie comprise entre E_1 et E_2 soit égal à $\int_{E_1}^{E_2} \mathcal{D}(E) dE$. Sur cette densité d'états, chaque bande prend l'allure d'une courbe en cloche.

D'une manière générale, le détail des structures de bande dépend du solide étudié. La forme précise, la largeur, et la hauteur des bandes dans la diagramme d'énergie est en effet très sensible, non seulement à la nature de l'atome dont le solide est constitué, mais également à la manière dont ces atomes sont répartis dans l'espace. C'est ainsi que la graphite et le diamant, pourtant deux solides formés de carbone pur, ont des structures de bandes différentes.

Pour les solides formés d'atomes différents, ces structures de bandes sont également observées. Pour autant, selon la compatibilité chimique des niveaux d'énergie des atomes en présence, les bandes peuvent parfois être plus fines, beaucoup plus séparées les unes des autres, et leurs énergies plus calquées sur celles des niveaux atomiques sous jacents que dans le cas des corps purs.

14.2.2 Le cas particulier des solides cristallins

L'approximation CLOA est la seule méthode permettant de donner une première approximation de la structure de bandes dans un solide *quelconque*. Dans le cas particuliers des solides *cristallins*, il existe en revanche un certain nombre de résultats mathématiques permettant de calculer la forme générale des états d'énergie électronique et la structure de bande associée.

Définition 14.1 (Réseau de Bravais et solide cristallin). *Un solide cristallin est un corps dont les atomes sont disposés selon un motif qui se répète périodiquement à travers l'espace. Plus précisément, pour un cristal donné, il existe trois vecteurs indépendants $\mathbf{a}, \mathbf{b}, \mathbf{c} \in \mathbb{R}^3$ (dits vecteurs de base) tels que quelque soit l'atome que l'on trouve en un point donné $\mathbf{r}_0$ de l'espace, cet atome se répète en tous les autres points $\mathbf{r}_0 + n_1\mathbf{a} + n_2\mathbf{b} + n_3\mathbf{c}$, $(n_1, n_2, n_3) \in \mathbb{Z}^3$. Le réseau périodique défini par les vecteurs $\mathbf{a}, \mathbf{b}, \mathbf{c}$ est appelé réseau de Bravais (voir la figure 14.13), et est noté $\mathcal{R}$:*

$$\mathcal{R} = \{n_1\mathbf{a} + n_2\mathbf{b} + n_3\mathbf{c}, \quad (n_1, n_2, n_3) \in \mathbb{Z}^3\}.$$

Il définit la périodicité de l'arrangement atomique du cristal.

Ceci n'est pas qu'une définition mathématique. Cette propriété s'applique réellement à tout un ensemble de corps, dont nombre sont d'origine naturelle. Ils se forment sur des temps géologiques, au bénéfice de conditions particulières de températures et de pression qui favorisent la mise en ordre périodique des atomes. C'est le cas par exemple des cristaux de quartz, de mica et de feldspath qui constituent le granit, roche formée dans les profondeurs de l'écorce terrestre, ainsi que de beaucoup de pierres précieuses. Il s'agit alors tout à fait de l'image traditionnelle du cristal, qui peut naturellement prendre la forme d'un polyèdre avec des faces planes de diverses orientations. Certains, plus rares, peuvent également être synthétisés : le diamant artificiel, par exemple, utilisé pour la fabrication d'outils de meulage n'est cependant produit qu'à l'état de fine poudre. Le silicium, également, seul élément dont on sait produire des cristaux de plusieurs dizaines de centimètres. La plupart des métaux sont quant à eux constitués, à l'échelle microscopique, d'un micro-assemblage de petits cristaux collés les uns aux autres (les *grains*, collés entre eux par des *joints de grains*, chers à la métallurgie physique).

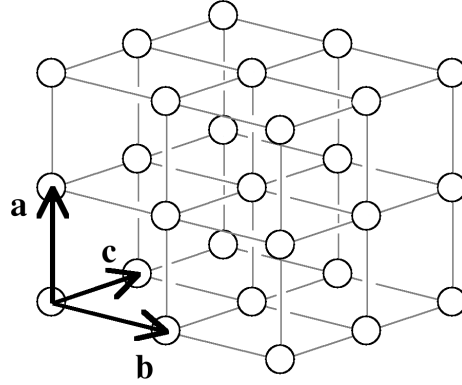


Fig. 14.13. Réseau de Bravais pour un cristal à maille parallélépipédique rectangle (on parle de réseau *rhomboédrique*).

L'étude théorique des cristaux se montre ainsi riche d'enseignements sur de nombreux métaux et alliages. La périodicité de l'arrangement atomique dont il est question ici s'arrête bien évidemment sur les bords du corps considéré. Cela interfère peu avec les propriétés du cristal, du moment que celui-ci est de taille suffisamment grande. Dans ce qui suit, par simplicité, le cristal sera ainsi réellement supposé de taille infinie. Un cristal est alors caractérisé par un motif générateur, la *cellule unité*.

Définition 14.2 (Cellule unité). *Etant donné un réseau de Bravais $\mathcal{R}$, on peut partitionner tout l'espace en considérant un polyèdre semi-ouvert $\mathcal{Q}$, appelé cellule unité, et ses images par le réseau $\mathcal{R}$:*

$$\bigcup_{\mathbf{R} \in \mathcal{R}} (\mathcal{Q} + \mathbf{R}) = \mathbb{R}^3, \quad \forall (\mathbf{R}, \mathbf{S}) \in \mathcal{R}^2, \quad \mathbf{S} \neq \mathbf{R} \Rightarrow (\mathcal{Q} + \mathbf{R}) \cap (\mathcal{Q} + \mathbf{S}) = \emptyset.$$

Il existe une infinité de tels polyèdres².

Nous considérons ici la cellule unité particulière

$$\mathcal{Q} = \left\{ x\mathbf{a} + y\mathbf{b} + z\mathbf{c}, \quad (x, y, z) \in \left[-\frac{1}{2}, \frac{1}{2}\right]^3 \right\}.$$

La connaissance du réseau $\mathcal{R}$ et de la disposition des atomes dans le domaine $\mathcal{Q}$ suffisent à définir complètement la répartition de tous les atomes du cristal (voir la figure 14.14).

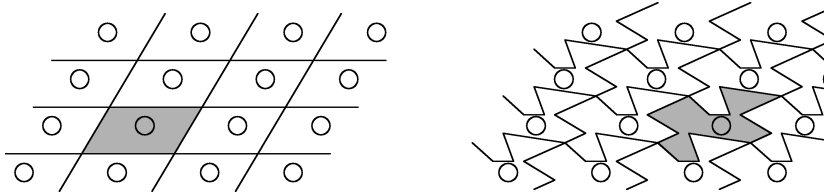


Fig. 14.14. Différents choix possibles de cellule unité pour un même réseau bi-dimensionnel. Le choix d'une cellule de forme simple (à gauche) est une commodité pratique, mais pas une obligation.

2. Il y a une infinité de cellules unités, certes... Mais certaines constructions sont univoques : c'est le cas de la cellule de Wigner-Seitz, qui a de plus le mérite de conserver les symétries de $\mathcal{R}$. Nous ne prétendons surtout pas que le choix que nous ferons par la suite est le meilleur !

Il est utile pour la suite de définir un réseau obtenu à partir du réseau de Bravais : le *réseau réciproque*.

Définition 14.3 (Réseau réciproque). *Le réseau réciproque $\mathcal{R}'$ d'un réseau de Bravais $\mathcal{R}$ est l'ensemble*

$$\mathcal{R}' = \{n_1 \mathbf{a}' + n_2 \mathbf{b}' + n_3 \mathbf{c}', \quad (n_1, n_2, n_3) \in \mathbb{Z}^3\},$$

où les vecteurs $\mathbf{a}'$, $\mathbf{b}'$ et $\mathbf{c}'$, homogènes à des vecteurs d'onde, sont définis par les relations :

$$\begin{cases} \mathbf{a} \cdot \mathbf{a}' = 2\pi, & \mathbf{b} \cdot \mathbf{b}' = 2\pi, & \mathbf{c} \cdot \mathbf{c}' = 2\pi, \\ \mathbf{a} \cdot \mathbf{b}' = \mathbf{a} \cdot \mathbf{c}' = \mathbf{b} \cdot \mathbf{a}' = \mathbf{b} \cdot \mathbf{c}' = \mathbf{c} \cdot \mathbf{a}' = \mathbf{c} \cdot \mathbf{b}' = 0. \end{cases}$$

La propriété essentielle sous jacente à cette définition, est que le réseau réciproque regroupe l'ensemble des vecteurs d'onde $\mathbf{k}$ tels que :

$$\forall \mathbf{R} \in \mathcal{R}, \quad e^{i\mathbf{k} \cdot \mathbf{R}} = 1.$$

Cas du réseau rhomboédrique

Afin de simplifier l'analyse, on se limite dans la suite au cas du réseau rhomboédrique, c'est à dire, au cas où les vecteurs de base du réseau de Bravais définissent un repère orthogonal de l'espace :

$$\mathbf{a} = L_x \mathbf{e}_x, \quad \mathbf{b} = L_y \mathbf{e}_y, \quad \mathbf{c} = L_z \mathbf{e}_z.$$

Dans ce cas, le réseau réciproque est simplement :

$$\mathcal{R}' = \left\{ n_1 \frac{2\pi}{L_x} \mathbf{e}_x + n_2 \frac{2\pi}{L_y} \mathbf{e}_y + n_3 \frac{2\pi}{L_z} \mathbf{e}_z, \quad (n_1, n_2, n_3) \in \mathbb{Z}^3 \right\}.$$

On note ZDB ("zone de Brillouin") la cellule unité³ de $\mathcal{R}'$:

$$\text{ZDB} = \left\{ x\mathbf{a}' + y\mathbf{b}' + z\mathbf{c}', \quad (x, y, z) \in \left[-\frac{1}{2}, \frac{1}{2} \right]^3 \right\}.$$

Etats d'énergie électroniques

Pour un électron introduit dans le potentiel Coulombien périodique du cristal, le Hamiltonien est de la forme

$$\hat{H}_{\text{per}} = -\frac{\hbar^2}{2m} \Delta + \hat{V}_{\text{per}},$$

où V_{per} est un potentiel $\mathcal{R}$ -périodique, *i.e.* tel que pour tout $R \in \mathcal{R}$, $V_{\text{per}}(\mathbf{r} + R) = V_{\text{per}}(\mathbf{r})$ presque partout (en $\mathbf{r}$). D'un point de vue mathématique, on ne précise pas pour l'instant dans quel espace fonctionnel se trouve V_{per} .

Théorème 14.1 (Ondes de Bloch). *Une famille libre d'états propres de l'opérateur Hamiltonien peut être construite au moyens de fonctions d'onde $\phi(\mathbf{r})$ de la forme :*

$$\phi(\mathbf{r}) = \psi(\mathbf{r}) e^{i\mathbf{k} \cdot \mathbf{r}},$$

avec $\mathbf{k} \in \text{ZDB}$, et $\psi(\mathbf{r})$ une fonction $\mathcal{R}$ -périodique de l'espace. Ces états généralisent les ondes de De Broglie dans le vide, et sont dits 'ondes de Bloch'.

3. On appelle cette cellule la (première) zone de Brillouin dans les ouvrages de physique.

Preuve. On présente ici un argument formel (Pour une démonstration complète, on renvoie à [29]). Pour $\psi \in L^2_{\text{loc}}(\mathbb{R}^3)$ par exemple, on introduit la translation $\tau_{\mathbf{R}}$ de vecteur $\mathbf{R}$:

$$\hat{\tau}_{\mathbf{R}}\psi(\mathbf{r}) = \psi(\mathbf{r} - \mathbf{R}).$$

Du fait de la périodicité du potentiel, on peut montrer que $\hat{\tau}_{\mathbf{R}}$ et le Hamiltonien $\hat{H}_{\text{per}}$ commutent, quel que soit le vecteur $\mathbf{R} \in \mathcal{R}$ choisi. Il est par ailleurs facile de voir que les opérateurs $\hat{\tau}_{\mathbf{R}}$ commutent 2 à 2. Il est donc possible de trouver une base commune de vecteurs propres au Hamiltonien $\hat{H}_{\text{per}}$ et à tout l'ensemble des translations $\hat{\tau}_{\mathbf{R}}$ ($\mathbf{R} \in \mathcal{R}$), sous la forme de fonctions ϕ solutions de⁴

$$\begin{cases} \hat{H}_{\text{per}}\phi = E\phi, \\ \hat{\tau}_{\mathbf{R}}\phi = \lambda(\mathbf{R})\phi, \quad \forall \mathbf{R} \in \mathcal{R}. \end{cases} \quad (14.3)$$

Entant donnés deux vecteurs $\mathbf{R}, \mathbf{S} \in \mathcal{R}$, la seconde condition ci-dessus implique $\lambda(\mathbf{R} + \mathbf{S}) = \lambda(\mathbf{R})\lambda(\mathbf{S})$ (car $\hat{\tau}_{\mathbf{R}+\mathbf{S}} = \hat{\tau}_{\mathbf{R}} \circ \hat{\tau}_{\mathbf{S}} = \hat{\tau}_{\mathbf{S}} \circ \hat{\tau}_{\mathbf{R}}$), et $|\lambda(\mathbf{R})| = 1$ (car $\|\hat{\tau}_{\mathbf{R}}\phi\| = \|\phi\|$, en norme $L^2(\mathcal{Q})$ par exemple). Ainsi, il existe $\mathbf{k} \in \mathbb{R}^3$ tel que $\lambda(\mathbf{R}) = e^{-i\mathbf{k} \cdot \mathbf{R}}$. Le système (14.3) devient ainsi

$$\begin{cases} \hat{H}_{\text{per}}\phi = E\phi, \\ \hat{\tau}_{\mathbf{R}}\phi = e^{-i\mathbf{k} \cdot \mathbf{R}}\phi, \quad \forall \mathbf{R} \in \mathcal{R}. \end{cases}$$

Posant alors $\psi(\mathbf{r}) = \phi(\mathbf{r})e^{-i\mathbf{k} \cdot \mathbf{r}}$, on vérifie facilement que pour tout vecteur $\mathbf{R}$ du réseau de Bravais, $\hat{\tau}_{\mathbf{R}}\psi(\mathbf{r}) = \psi(\mathbf{r})$, c'est à dire que ψ est $\mathcal{R}$ -périodique. Les états propres communs au Hamiltonien et à tous les opérateurs de translation sont donc de la forme

$$\phi(\mathbf{r}) = \psi(\mathbf{r})e^{i\mathbf{k} \cdot \mathbf{r}}$$

avec ψ une fonction $\mathcal{R}$ -périodique.

On peut naturellement indiquer le vecteur d'onde par $\mathbf{k} \in \mathbb{R}^3$. Pour éviter les redondances, il faut toutefois se restreindre à $\mathbf{k} \in \text{ZDB}$, quitte à translater $\mathbf{k}$ d'un vecteur $\mathbf{q}$ du réseau réciproque selon

$$\exists \mathbf{q} \in \mathcal{R}', \quad \mathbf{k}' = \mathbf{k} + \mathbf{q} \in \text{ZDB}.$$

En effet,

$$\phi_{\mathbf{k}'}(\mathbf{r}) = \psi(\mathbf{r})e^{i\mathbf{k}' \cdot \mathbf{r}} = \tilde{\psi}(\mathbf{r})e^{i\mathbf{k}' \cdot \mathbf{r}},$$

et $\tilde{\psi}(\mathbf{r}) = \psi(\mathbf{r})e^{-i\mathbf{q} \cdot \mathbf{r}}$ est bien $\mathcal{R}$ -périodique. □

Structure de bandes

Pour un vecteur d'onde $\mathbf{k} \in \text{ZDB}$ donné, la recherche des ondes de Bloch associées conduit pour la partie périodique $\psi(\mathbf{r})$ à l'équation :

$$\hat{G}_{\mathbf{k}}\psi = \left(-\frac{\hbar^2}{2m}\Delta - \frac{i\hbar^2}{m}\mathbf{k} \cdot \nabla + \left(V_{\text{per}}(\mathbf{r}) + \frac{\hbar^2|\mathbf{k}|^2}{2m} \right) \right) \psi = E\psi.$$

Ainsi, les propriétés spectrales de $\hat{H}_{\text{per}}$ sont déterminées par celles de $\hat{G}_{\mathbf{k}}$, pour tous les $\mathbf{k} \in \text{ZDB}$, au sens où

$$\sigma(\hat{H}_{\text{per}}) = \bigcup_{\mathbf{k} \in \text{ZDB}} \sigma(\hat{G}_{\mathbf{k}}).$$

On peut vérifier que $\hat{G}_{\mathbf{k}}$ est auto-adjoint d'inverse compact (en essence, parce que les termes supplémentaires sont des perturbations de $\hat{H}_0 = -\frac{\hbar^2}{2m}\Delta$, qui est d'inverse compact pour des problèmes posés sur des ensembles bornés).

On montre alors que pour $\mathbf{k} \in \text{ZDB}$ fixé, le spectre de $\hat{G}_{\mathbf{k}}$ est une suite croissante de réels $(\epsilon_n(\mathbf{k}))_{n \in \mathbb{N}^*}$, homogènes à des énergies. On obtient le spectre de $\hat{H}_{\text{per}}$ en rassemblant les

4. On se souviendra pour cela de l'interprétation du spectre comme "ensemble des vecteurs propres généralisés".

spectres des $\hat{G}_{\mathbf{k}}$ ($\mathbf{k} \in \text{ZDB}$). Comme on peut montrer par ailleurs que $\mathbf{k} \mapsto \epsilon_n(\mathbf{k})$ est continue pour tout $n \geq 1$, on en déduit que $I_n = \{\epsilon_n(\mathbf{k}), \mathbf{k} \in \text{ZDB}\}$ est un intervalle de $\mathbb{R}$, noté $[\mu_n, \nu_n]$. Finalement,

$$\sigma(\hat{H}_{\text{per}}) = \bigcup_{n \geq 1} [\mu_n, \nu_n].$$

On retrouve ainsi que le spectre de $\hat{H}_{\text{per}}$ est composé de bandes (voir la figure 14.15). De tels diagrammes permettent de discuter du comportement des électrons.

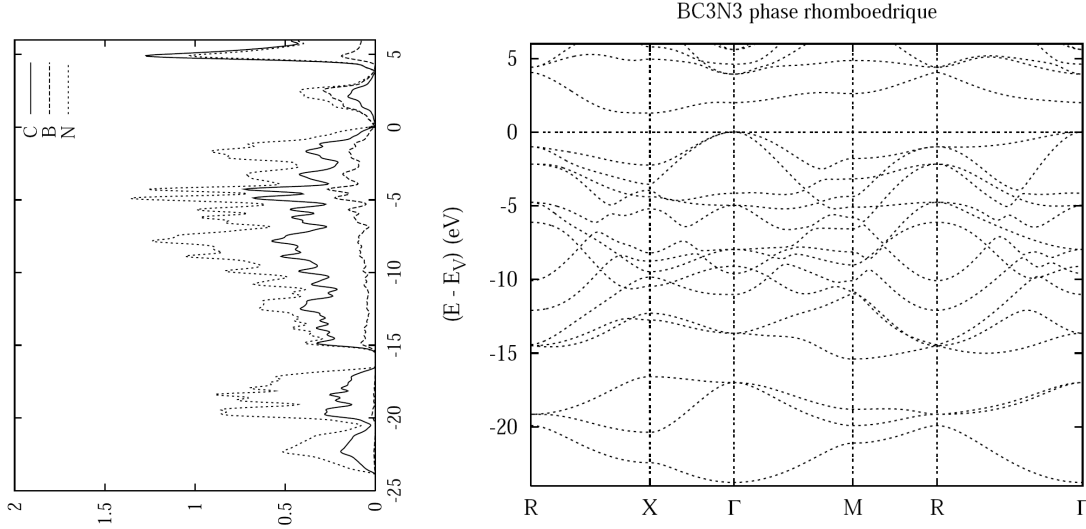


Fig. 14.15. Structure de bandes (droite) et densité d'états associée (gauche) pour un composé BC₃N₃ en cellule rhomboédrique [3]. Sur la structure de bandes sont représentées les courbes $\epsilon_n(\mathbf{k})$ obtenues le long d'une ligne brisée qui explore les divers recoins de la zone de Brillouin. La courbe la plus basse correspond à $n = 0$, celle immédiatement supérieure à $n = 1$, etc. Ces courbes ont tendance à ne plus évoluer de manière strictement monotone au delà de $n = 4$. Ici, la plupart des bandes se recouvrent les unes les autres, et la densité d'états présente l'allure d'une courbe très accidentée, avec de nombreux pics. A noter toutefois, sur la densité d'états, un 'trou' aux alentours de -15 eV. Cette particularité est appelée *gap* d'énergie. Elle est à l'origine des propriétés électroniques et optiques des matériaux.

14.2.3 propriétés électroniques des solides

Remplissage des bandes

La structure de bandes du spectre énergétique des électrons dans un solide est à l'origine de propriétés électroniques et/ou optiques remarquables. Du fait que, dans le solide qui ne contient qu'un seul type d'atomes, chaque bande contient exactement un nombre d'états égal à un multiple du nombre total N d'atomes, et que par ailleurs, le nombre total d'électrons qui viennent remplir ces bandes est également multiple de N , il est possible selon les circonstances que le remplissage s'arrête exactement en limite de bande.

Sur les schémas de la figure 14.16, nous représentons la manière dont les électrons viennent remplir les bandes d'énergie dans le lithium (3 électrons par atome) et dans le graphite (6 électrons par atome).

Dans le cas du lithium, le nombre total d'électrons dans le solide est $3N$ si on a N atomes dans le solide. La bande la plus basse comporte N orbitales, et du fait du spin, peut accueillir $2N$ électrons. La bande immédiatement supérieure accueille les N électrons restant, et se retrouve ainsi à moitié remplie.

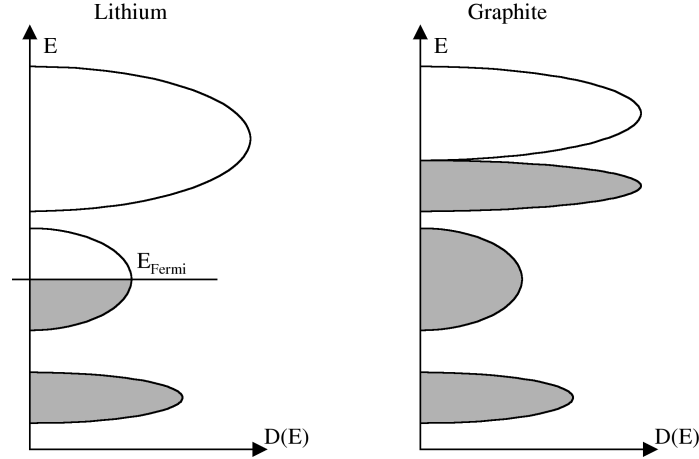


Fig. 14.16. Remplissage des bandes d'énergie par les électrons dans le lithium et le Graphite. Le remplissage commence par les orbitales les plus basses en énergie, à raison de 2 électrons par orbitale, et se poursuit progressivement vers les orbitales immédiatement supérieures. Les bandes, ou portions de bandes ainsi occupées apparaissent en grisé.

Dans le cas du carbone, la structure de bande est légèrement différente. La bande supérieure connaît une anomalie, qui la sépare en deux sous bandes. La sous bande inférieure porte exactement N orbitales. Les $6N$ électrons viennent ainsi remplir en entier les 2 bandes du bas et cette sous bande. Le remplissage s'arrête exactement au ras de la 2^{ème} sous bande.

Matériaux conducteurs, matériaux isolants

Considérons la structure électronique qui a été obtenue dans le lithium. Le remplissage des électrons s'arrête à mi-bande, à un niveau nommé énergie de Fermi. Le niveau de Fermi ϵ_F est défini par la relation

$$\frac{2}{|\text{ZDB}|} \sum_{n=1}^{+\infty} \int_{\text{ZDB}} \mathbf{1}_{\epsilon_F \geq \epsilon_n(\mathbf{k})} d\mathbf{k} = N_{\text{cel}},$$

où N_{cel} est le nombre d'électrons par cellule. Cette énergie se situe en plein continuum, et il existe, juste au dessus d'elle, un foisonnement d'orbitales libres. Il est ainsi possible de communiquer au système d'électrons de très petites quantités d'énergie. Il suffit pour cela de prendre des électrons juste en dessous du niveau de Fermi, et de les faire passer juste au dessus, en les excitant.

Communiquer une très faible quantité d'énergie, c'est précisément ce que l'on fait lorsque l'on cherche à créer un courant électrique, c'est à dire à imprimer aux électrons un mouvement d'ensemble. Dans le cas du lithium, le remplissage des bandes est tel que ces faibles montants d'énergie peuvent être acceptés. Le lithium est ainsi un matériau *conducteur* électrique.

La situation est en revanche très différente lorsque le remplissage des bandes s'arrête sur une bande pleine. Il peut en effet se produire qu'un vide existe entre la dernière bande remplie et la bande suivante (voir la figure 14.17) : on parle de *gap* d'énergie. Il est alors impossible de fournir au système d'électrons une énergie inférieure à la largeur ΔE du gap. Dans certains corps ΔE peut être de l'ordre de quelques eV. C'est peine perdue que de chercher à y créer des courants électriques de manière 'douce', les électrons refusant systématiquement l'apport d'énergie proposé. Le matériau est dit *isolant* électrique⁵.

Il existe des situations intermédiaires, dans laquelle un gap existe, mais de largeur plus modeste, typiquement inférieure ou égale à 1 eV (1,21 eV dans le silicium). Les matériaux sont alors

5. Un courant électrique peut toujours néanmoins être créé dans les isolants en appliquant de très fortes tensions. Il ne s'agit alors plus de conduction électrique proprement dite, mais de *claquage*, qui s'accompagne généralement d'un endommagement du matériau là où le courant est passé.

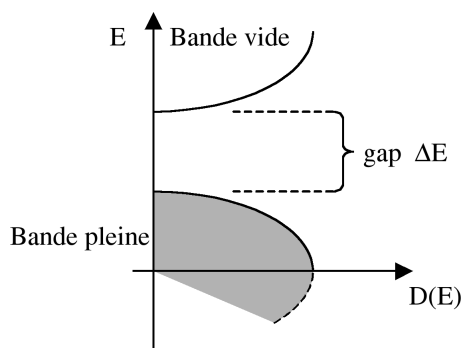


Fig. 14.17. Le gap d'énergie : le remplissage des orbitales électroniques s'arrête exactement sur un bord de bande. L'orbitale d'énergie immédiatement supérieure est distante d'un intervalle ΔE .

dits *semi-conducteurs*. Il est possible, par des sollicitations extérieures modérées, ou par des modifications chimiques mineures lors de leur fabrication, de leur faire prendre sur commande des propriétés de conducteurs électriques ou d'isolants. C'est là ce qui fait leur intérêt dans les machines électroniques.

Transparence optique des isolants à grand gap

De même qu'il est difficile à un isolant d'accepter de l'énergie pour produire un courant électrique, il est aussi difficile à ce même isolant d'accepter d'absorber un photon. Toujours à cause de la présence du gap, seuls les photons d'énergie $\hbar\omega \geq \Delta E$ ont de réelles possibilités de pouvoir être absorbés (voir la figure 14.18).

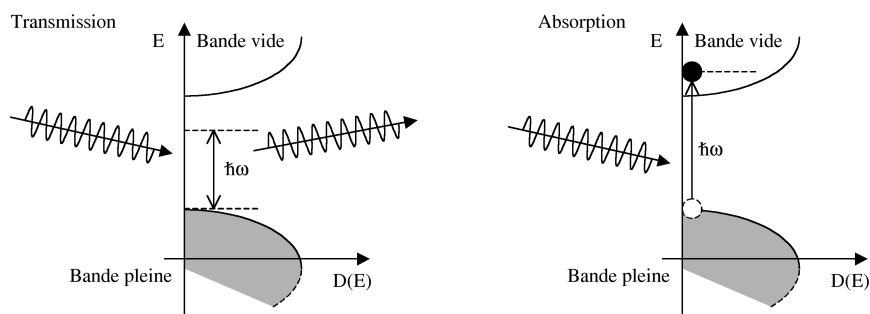


Fig. 14.18. Absorption ou transmission d'un rayonnement dans un isolant en fonction de l'énergie du photon incident. A gauche, l'énergie $\hbar\omega$ est inférieure au gap. Aucune absorption ne peut avoir lieu. A droite, l'énergie est suffisante pour pouvoir porter un électron de la bande pleine dans un des états d'énergie de la bande supérieure. L'absorption a lieu.

Dans le diamant, par exemple, il existe un gap de 7 eV de largeur. Les photons de la lumière visible ont eux des énergies typiquement comprises entre 1.5 eV (lumière rouge) et 3 eV (lumière violette). Ces énergies sont trop faibles pour être absorbées, et nous comprenons ainsi l'exceptionnelle transparence optique du diamant. Les premiers rayonnements absorbables correspondent à des photons de 7 eV, ce qui se situe déjà loin dans l'ultraviolet. Il en va de même pour le quartz et le verre, qui sont respectivement les formes cristallisée et amorphe de la silice (dioxide de silicium SiO_2). Le verre ne laisse passer que la lumière visible, alors que le quartz, dont le gap énergétique est plus grand, se laisse aussi traverser par le rayonnement ultraviolet. Cela explique pourquoi il est impossible de bronzer derrière une vitre, et pourquoi les lampes à bronzer (lampes UV) doivent être fabriquées en quartz.

A l'opposé, les matériaux qui possèdent des gaps plus faibles, et en particulier, tous les matériaux conducteurs, ne peuvent pas être transparents à la lumière.⁶

6. Une nuance peut cependant être apportée : s'il est vrai que l'absence de gap fait que tout rayonnement est susceptible d'être absorbé, le mécanisme d'absorption réclame cependant qu'une certaine épaisseur de matériau soit traversée pour que l'absorption soit totale. Cette épaisseur varie selon les longueurs d'onde. De quelques microns pour la lumière visible, elle passe à plusieurs centimètres pour les rayons X , voire plusieurs mètres pour les rayons γ . C'est ainsi que les minces films d'or dont sont recouvertes les visières de certains avions restent globalement transparents à la lumière visible, en dépit de leur conductivité, et qu'il est illusoire de vouloir se cacher derrière une feuille de papier d'aluminium pour se protéger de rayonnements issus de sources radioactives.

Exercices

Exercice 14.1 (Un modèle simple de liaison chimique).

Préliminaire

On considère, dans l'espace tridimensionnel (x, y, z) , la boîte quantique parallélépipédique $(x, y, z) \in [0, a]^3$. Le potentiel est nul à l'intérieur, et on rappelle que la fonction d'onde d'une particule de masse m confinée dans la boîte doit s'annuler sur les bords de celle-ci.

- (1) Ecrire l'équation de Schrödinger indépendante du temps qui définit la partie spatiale $\phi(x, y, z)$ d'un état stationnaire d'énergie E .
- (2) On admet que ces états d'énergie peuvent être recherchés sous la forme factorisée

$$\phi(x, y, z) = \phi_x(x)\phi_y(y)\phi_z(z).$$

Calculer la forme générale de ces états, et en déduire l'énergie de l'état fondamental.

Liaison entre deux boîtes

- (1) On considère deux boîtes séparées, qui représenteront pour nous deux atomes, chacune de dimensions $a \times a \times a$ avec $a = 1 \text{ \AA}$, et contenant chacune un électron dans son état fondamental. Calculer l'énergie électronique totale des deux boîtes.
- (2) On fait se toucher les deux boîtes, de sorte à ce qu'elles ne forment plus qu'une seule grande boîte de dimensions $a \times a \times 2a$. Quelle est la nouvelle énergie électronique ?
- (3) Quel est le gain énergétique du rapprochement (*énergie de liaison*) ? Faire l'application numérique.

Exercice 14.2. On peut montrer directement qu'il n'y a pas de spectre discret pour le solide périodique. Ceci montre au passage que l'espace de Hilbert usuel n'est pas le bon espace à considérer, car on ne peut pas avoir d'états liés.

- (1) Soit $\lambda \in \sigma_d(\hat{H})$ et $E = \text{Ker}(\hat{H} - \lambda \text{Id})$ (de dimension N). Montrer que si $f \in E$, alors $f_{\mathbf{p}} \in E$ avec $f_{\mathbf{p}}(x) = f(x + p)$ si $\mathbf{p} \in \mathcal{R}$.
- (2) En considérant une base $(u_i)_{i=1, \dots, N}$ de E , en déduire qu'on peut écrire $f_{\mathbf{p}}(x) = \sum_{i=1}^N \alpha_{\mathbf{p}, i} u_i$ avec $\sum_{i=1}^N |\alpha_{\mathbf{p}, i}|^2 = 1$.
- (3) Montrer que pour tout $\epsilon > 0$, il existe $R > 0$ tel que

$$\forall 1 \leq i, j \leq N, \quad \left| \int_{B_R} u_i u_j - \delta_{ij} \right| \leq \epsilon.$$

- (4) En calculant $\int_{B_R} |f_{\mathbf{p}}|^2$ de deux manières différentes, obtenir une contradiction lorsque $|\mathbf{p}| \rightarrow +\infty$.

Exercice 14.3 (Cassage des molécules diatomiques par exposition à la lumière).

Il arrive qu'une molécule éclairée en lumière visible ou ultraviolette voie certaines de ses liaisons chimiques se rompre. En prenant l'exemple de la molécule H_2 , pouvez-vous proposer une explication pour ce mécanisme de dissociation ? Exprimer la fréquence de rayonnement minimale à utiliser en fonction de l'énergie de liaison de la molécule.

Exercice 14.4 (Formation d'une bande d'énergie dans une chaîne atomique).

On considère une chaîne linéaire de N atomes identiques, et on cherche, dans le cadre de l'approximation CLOA, à calculer les orbitales moléculaires résultant de l'hybridation des N exemplaires d'un de leurs niveaux atomiques. On note $|\varphi_i\rangle$ l'orbitale atomique portée par l'atome i ($1 \leq i \leq N$). On suppose que :

$$\langle \varphi_i | \varphi_j \rangle = 0 \quad \text{si } i \neq j, \quad \langle \varphi_i | \hat{H} | \varphi_j \rangle = \begin{cases} E & \text{si } i = j, \\ -\beta & \text{si } i = j \pm 1, \\ 0 & \text{sinon.} \end{cases}$$

On cherche les orbitales moléculaires sous la forme

$$|\psi\rangle = \sum_{i=1}^N c_i |\varphi_i\rangle.$$

- (1) Ecrire dans la base $\{|\varphi_1\rangle, |\varphi_2\rangle, \dots, |\varphi_N\rangle\}$ la matrice $N \times N$ représentant le Hamiltonien $\hat{H}$ du système.
- (2) Vérifier que les N vecteurs propres de cette matrice peuvent être recherchés sous la forme :

$$c_i = \text{Cste} \times \sin(k \times i),$$

pour des nombres k sans dimension à déterminer. Quelles sont les énergies E_k associées ?

- (3) Représenter l'allure de la répartition des énergies E_k sur un diagramme énergétique. Quelle est la largeur de la bande obtenue ?
- (4) Chaque niveau $|\varphi_i\rangle$ est initialement porteur d'un électron. Quel est le gain énergétique de la formation de la chaîne ?

A

Constantes physiques fondamentales

A.1 Constantes physiques les plus utilisées

Le Tableau A.1 rappelle des constantes utiles en physique statistique et quantique.

Tableau A.1. Principales constantes physiques utiles en physique statistique et quantique.

Grandeur physique	Notation usuelle	Valeur
Nombre d'Avogadro	$\mathcal{N}_A$	$6,02 \times 10^{23}$
Constante de Boltzmann	k_B	$1,381 \times 10^{-23} \text{ J/K}$
Constante de Planck	$\hbar$	$1,054 \times 10^{-34} \text{ Js}$
Charge de l'électron	e	$-1,602 \times 10^{-19} \text{ C}$
Masse de l'électron	m_e	$9,11 \times 10^{-31} \text{ kg}$
Masse du proton	m_p	$1,672 \times 10^{-27} \text{ kg}$
Masse du neutron	m_n	$1,674 \times 10^{-27} \text{ kg}$
Permittivité diélectrique du vide	ϵ_0	$8,854 \times 10^{-12} \text{ F/m}$
Electron-Volt	eV	$1,602 \times 10^{-19} \text{ J}$
Constante de gravitation	$\mathcal{G}$	$6,67 \times 10^{-11} \text{ m}^3\text{kg}^{-1}\text{s}^{-2}$
Vitesse de la lumière	c	$2,998 \times 10^8 \text{ m/s}$

A.2 Unités atomiques

On travaille usuellement, en chimie quantique, dans le système d'unités où

$$m_e = 1, \quad e = 1, \quad \hbar = 1, \quad \frac{1}{4\pi\epsilon_0} = 1.$$

Dans ce cas, l'unité de masse est $9,11 \times 10^{-31} \text{ kg}$, l'unité de longueur est le rayon de Bohr $a_0 = 5,29 \times 10^{-11} \text{ m}$, l'unité de temps est $2,42 \times 10^{-17} \text{ s}$, et l'unité d'énergie est le Hartree $\text{Ha} = 4,36 \times 10^{-18} \text{ J} = 27,2 \text{ eV} = 627 \text{ kcal/mol}$. Ce changement d'unité permet de travailler avec des valeurs intuitives des grandeurs physiques : ainsi, les distances moyennes typiques entre un électrons et le noyau auquel il est rattaché sont de l'ordre du rayon de Bohr, et les énergies des états fondamentaux (en valeur absolue) de l'ordre de quelques Ha pour les petits systèmes.

B

Analyse numérique des schémas d'intégration

Nous rappelons ici (brièvement) des notions importantes pour l'analyse numérique des schémas d'intégration des EDO. Le lecteur intéressé par plus de détails pourra par exemple se référer à [9]. On considère ici une méthode numérique à un pas, pour un pas de temps constant Δt et des temps $t_n = n\Delta t$:

$$y^{n+1} = \Phi_{\Delta t}(t_n, y^n), \quad (\text{B.1})$$

dont le flot $\Phi_{\Delta t}$ est une approximation du flot ϕ_t de l'EDO

$$\forall t \in [0, T], \quad \dot{y} = f(t, y(t)), \quad y(0) = y^0 + \eta, \quad (\text{B.2})$$

où η est un petit paramètre précisant l'approximation de la condition initiale. Ainsi, y^n est une approximation au temps t_n de la solution exacte $y(t_n)$ de (B.2).

On suppose par la suite que le champ de force f est globalement Lipschitzien. Dans ces conditions, le problème de Cauchy (B.2) est bien posé sur des temps $t \geq T$, $T < +\infty$ fixé, et on peut étudier théoriquement la convergence des schémas numériques. Une étude dans un cas plus général demande un traitement spécifique au système considéré.

B.1 Consistance et ordre

Définition B.1 (Consistance). Soit $T > 0$ fixé et $N = T/\Delta t$ un entier. On dit que la méthode numérique (B.1) est consistante avec l'équation différentielle (B.2) si

$$\sum_{n=1}^N |y(t_n + \Delta t) - y(t_n) - \Delta t \Phi_{\Delta t}(t_n, y(t_n))| \rightarrow 0$$

lorsque $\Delta t \rightarrow 0$.

La quantité $|y(t_{n+1}) - y(t_n) - \Delta t \Phi_{\Delta t}(t_n, y(t_n))|$ est l'erreur de consistance locale, et représente l'erreur intrinsèque d'approximation du schéma numérique. La définition précédente signifie que l'erreur de consistance est négligable pour Δt suffisamment petit.

En pratique, pour vérifier la consistance, on a souvent recours à la notion d'ordre d'un schéma :

Définition B.2 (Ordre). Soit $\Delta t > 0$. On dit que la méthode numérique (B.1) est d'ordre p si

$$|y(t_n + \Delta t) - y(t_n) - \Delta t \Phi_{\Delta t}(t_n, y(t_n))| \leq K \Delta t^{p+1}$$

pour $\Delta t \leq \Delta t_0$ suffisamment petit.

Ainsi, une méthode d'ordre 1 ou supérieur est consistante (si on peut montrer que la constante K dans la définition B.2 est uniforme).

Exemple B.1 (Schéma d'Euler explicite). Pour le schéma d'Euler explicite et f autonome (ne dépendant pas explicitement du temps), on a, par un développement de Taylor,

$$y(t_n + \Delta t) - y(t_n) - \Delta t \Phi_{\Delta t}(y(t_n)) = \frac{\Delta t^2}{2} y''(t_n) + O(\Delta t^3) = \frac{\Delta t^2}{2} f'(y(t_n)) f''(y(t_n)) + O(\Delta t^3).$$

Ceci montre que le schéma d'Euler explicite est d'ordre 1. Dans le cas où le champ de force f est à dérivées bornées, on obtient également la consistance de la méthode numérique.

B.2 Convergence

Définition B.3 (Stabilité). On dit qu'une méthode numérique est stable s'il existe une constante $M > 0$ telle que pour toutes suite y^n , z^n , ϵ^n ($n \geq 0$) vérifiant

$$\begin{cases} y^{n+1} = y^n + \Delta t \Phi_{\Delta t}(t_n, y^n), \\ z^{n+1} = z^n + \Delta t \Phi_{\Delta t}(t_n, y^n) + \epsilon^n, \end{cases}$$

on ait

$$\max_{0 \leq n \leq N} |z^n - y^n| \leq M \left(|y^0 - z^0| + \sum_{n < N} |\epsilon^n| \right).$$

La stabilité précise la robustesse du schéma face à l'accumulation des erreurs locales (ou d'arrondi). C'est une notion pertinente sur un temps $T > 0$ donné.

Remarque B.1. En particulier, une instabilité éventuelle "en temps long" ($T \rightarrow +\infty$, Δt fixé) ne contredit pas la consistance, qui est une convergence trajectorielle dans le cas limite $\Delta t \rightarrow 0$ pour T fixé.

Définition B.4 (Convergence). On dit que la méthode numérique est convergente si

$$\max_{0 \leq n \leq N} |y(t_n) - y^n| \rightarrow 0$$

lorsque $\eta \rightarrow 0$, $\Delta t \rightarrow 0$.

On montre facilement qu'une méthode stable et consistante est convergente.

B.3 Stabilité linéaire en temps long

En pratique, on ne sait caractériser facilement la stabilité de manière simple que dans le cas de champs de force globalement Lipschitziens. Ceci revient en fait à caractériser la limite de stabilité pour des champs de force linéaires. Dans ce cas,

$$y^{n+1} = A(\Delta t) y^n$$

en supposant toujours le champ de force f autonome.

On voit alors facilement qu'une méthode est stable sur des temps longs si et seulement si les valeurs propres de la matrice $A(\Delta t)$ sont de module inférieur ou égal à 1 ($|\rho(A(\Delta t))| \leq 1$). Considérant une autre solution, partant de z^0 et ne subissant pas de perturbations à chaque pas de temps, et pour une norme euclidienne par exemple, on a en effet

$$|y^{n+1} - z^{n+1}| \leq |\rho(A(\Delta t))| |y^n - z^n|,$$

et donc

$$|y^N - z^N| \leq |\rho(A(\Delta t))|^N |y^0 - z^0|.$$

Pour un potentiel général V régulier, ayant un minimum local en q_0 , on a le développement de Taylor en q_0 :

$$V(x) = V(q_0) + \frac{1}{2}(q - q_0)^T \cdot \nabla^2 V(q_0) \cdot (q - q_0) + O(|x - q_0|^3).$$

La stabilité linéaire s'étudie alors en considérant le potentiel

$$V_{\text{lin}}(x) = V(q_0) + \frac{1}{2}(q - q_0)^T \cdot \nabla^2 V(q_0) \cdot (q - q_0),$$

qui conduit en effet à un champ de forces linéaire

$$-\nabla V_{\text{lin}}(q) = -\nabla^2 V(q_0) \cdot (q - q_0).$$

On peut ensuite se ramener à l'étude de la stabilité linéaire en dimension 1 au travers d'un changement de base diagonalisant la matrice symétrique réelle $\nabla^2 V(q_0)$.

Elements de théorie spectrale

On présente dans ce chapitre les fondements de la théorie spectrale, en insistant tout particulièrement sur les opérateurs auto-adjoints. On prendra garde au fait que nombre de raisonnements ou de propriétés vrais dans le cas d'un espace de dimension finie, ne le sont plus pour des espaces de dimension infinie ! On peut citer par exemple le fait que toutes les normes ne sont pas équivalentes, que les opérateurs linéaires ne sont pas forcément continus, qu'une application peut être injective sans être surjective (et réciproquement) - et donc en particulier, ce n'est pas parce qu'une application a un noyau réduit à $\{0\}$ que l'on peut l'inverser ! Ce chapitre est donc à lire et relire tranquillement... jusqu'à ce que les affirmations ci-dessus vous semblent honteusement contingentes aux espaces de dimension finie !

La théorie spectrale est un outil fondamental pour l'étude des systèmes quantiques. Citons deux applications particulièrement importantes :

- (i) la résolution de l'équation de Schrödinger par décomposition spectrale (on présente un cas particulier en Section 10.4.3) ;
- (ii) l'étude qualitative voire quantitative de l'état fondamental du système (voir Chapitre 14) ;
- (iii) la caractérisation possible des résultats de la mesure, décrite par une observable auto-ajointe (voir Chapitre 11).

C.1 Opérateurs sur des espaces de Banach

On rappelle qu'un espace de Banach est un espace vectoriel normé complet (pour la distance issue de sa norme), et qu'un espace de Hilbert est un espace de Banach muni d'un produit scalaire. Pour deux espaces de Banach E, F , un opérateur $T : E \rightarrow F$ est une application linéaire de E vers F . L'espace vectoriel $\mathcal{L}(E, F)$ des opérateurs linéaires de E vers F est muni de la norme

$$\|T\|_{E \rightarrow F} = \sup_{u \in E \setminus \{0\}} \frac{\|Tu\|_F}{\|u\|_E}. \quad (\text{C.1})$$

Lorsque cette quantité est finie, on a un opérateur borné :

Définition C.1 (Opérateur borné). *Un opérateur T est un opérateur borné si $\|T\|_{E \rightarrow F} < +\infty$.*

Un opérateur borné est un opérateur continu. Bien sûr, tous les opérateurs ne sont pas bornés. Pour de tels opérateurs, on introduit la notion de domaine $D(T)$, qui est le sous-espace vectoriel de E pour lequel T peut être défini et prend ses valeurs dans F . Si $D(T) \neq E$, l'opérateur est dit non-borné. Si $D(T)$ est dense dans E , l'opérateur est à domaine dense.

Remarque C.1. Un opérateur borné sur un sous-espace dense s'étend de manière unique en un opérateur borné sur tout l'espace.

Dans le cas général, la notion d'extension d'un opérateur est définie par rapport au domaine (non-nécessairement dense) de l'opérateur que l'on cherche à étendre.

Définition C.2 (Extension d'un opérateur). *On dit que T_2 est une extension de T_1 si $D(T_1) \subset D(T_2)$ et si les deux opérateurs coïncident sur le domaine de T_1 .*

On définit enfin la notion d'opérateur fermé :

Définition C.3 (Opérateur fermé). *On dit qu'un opérateur T de E dans F de domaine $D(T)$ est fermé si son graphe*

$$\Gamma(T) = \{(u, Tu), \quad u \in D(T)\}$$

est un fermé de $E \times F$. Un opérateur est dit fermable s'il possède une extension fermée.

Ainsi, T est fermé si pour toute suite (u_n) de $D(T)$ telle que $u_n \rightarrow u$ dans E et $Au_n \rightarrow v$ dans F , on a $u \in D(T)$ et $Au = v$.

Exemple C.1. On considère $E = F = L^2(]0, 1[)$ et $Tu = u'$. Alors $D(T) = H^1(]0, 1[)$. Si $u_n \rightarrow u$ dans $L^2(]0, 1[)$ et $u'_n \rightarrow v$ dans $L^2(]0, 1[)$, alors (u_n) est une suite de Cauchy dans $H^1(]0, 1[)$, qui est complet. Ceci montre que $u \in H^1(]0, 1[)$ et $v = u'$ au sens de $L^2(]0, 1[)$.

En revanche, si on prenait comme domaine $D(T) = C^1(]0, 1[)$, alors la suite

$$u_n(x) = \left(\left(x - \frac{1}{2} \right) + \frac{1}{n^2} \right)^{1/2}$$

converge au sens de $H^1(]0, 1[)$ vers

$$u(x) = \left| x - \frac{1}{2} \right|,$$

qui n'est pas dans $D(T)$. Ceci montre que l'opérateur T , avec ce domaine, n'est pas fermé (le domaine de définition est trop petit). En revanche, il est fermable quitte à étendre le domaine de définition à $H^1(]0, 1[)$.

Remarque C.2. Tout opérateur n'est pas fermable. On peut méditer par exemple le cas de $T : L^2(\mathbb{R}_+) \rightarrow \mathbb{R}$ de domaine $D(T) = C_0^\infty(\mathbb{R}_+)$ et donné par $Tu = u(0)$. En effet, la suite $u_n(x) = \exp(-nx)$ converge vers $u = 0$ au sens de $L^2(\mathbb{R}_+)$, et $Tu = 0$; alors que $Tu_n = 1$ pour tout n .

C.2 Adjoint d'un opérateur

Définition C.4 (Adjoint d'un opérateur). *Soit $\mathcal{H}$ un espace de Hilbert, et T un opérateur à domaine dense sur $\mathcal{H}$. Soit $D(T^*)$ l'ensemble défini par*

$$D(T^*) = \{u \in \mathcal{H}, \quad \exists v \in \mathcal{H}, \quad \forall w \in D(T), \quad \langle Tw, u \rangle = \langle w, v \rangle\}.$$

L'opérateur linéaire T^* sur $\mathcal{H}$ de domaine $D(T^*)$ et défini par

$$T^*u = v$$

est appelé l'adjoint de T . Notons que la définition de $T^*u = v$ a un sens par le théorème de Riesz (qui donne l'unicité d'un tel v lorsqu'il existe).

Définition C.5 (Opérateur symétrique). *On dit qu'un opérateur T à domaine dense est symétrique si T^* est une extension de T , autrement dit si*

$$\forall (u, v) \in D(T)^2, \quad \langle Tu, v \rangle = \langle u, Tv \rangle.$$

La grande majorité des opérateurs donnés par la physique sont formellement symétriques (*i.e.* symétriques sur un espace de fonctions suffisamment régulières, tel que $C_0^\infty(\mathbb{R}^3)$). Cette notion n'est pas suffisante dans la majorité des cas, et on voudra souvent avoir affaire à des opérateurs auto-adjoints.

Définition C.6 (Opérateur auto-adjoint). *Un opérateur est auto-adjoint si $T^* = T$. Un opérateur T est essentiellement auto-adjoint si T^* admet une unique extension auto-adjointe.*

Remarque C.3. Un opérateur T est auto-adjoint si et seulement s'il est symétrique et $D(T) = D(T^*)$. En pratique, il est souvent plus aisé de vérifier la symétrie que le caractère auto-adjoint, cette dernière notion demandant en plus de montrer l'égalité des domaines $D(T)$ et $D(T^*)$.

Remarque C.4. On peut montrer (voir [29, Section VIII.2]) qu'un opérateur symétrique est toujours fermable, mais qu'il peut avoir plusieurs extensions auto-adjointes...

C.3 Spectre d'un opérateur

On se place par la suite dans le cas où $E = F = \mathcal{H}$, $\mathcal{H}$ étant un espace de Hilbert. L'image d'un opérateur T est l'ensemble

$$\text{Im}(T) = \{Tu, u \in D(T)\}$$

Définition C.7 (Inverse d'un opérateur). *Un opérateur R est l'inverse d'un opérateur T si $D(R) = \text{Im}(T)$, $D(T) = \text{Im}(R)$ et*

$$TR = \text{Id}_{\text{Im}(T)}, \quad RT = \text{Id}_{\text{Im}(R)}.$$

On déduit de la définition ci-dessus que l'inverse d'un opérateur est unique s'il est défini, et on note un tel opérateur T^{-1} . Trouver l'inverse d'un opérateur revient à résoudre l'équation $Tu = f$ pour tout $f \in \text{Im}(T)$. Un critère d'existence d'inverse de T est alors que $Tu = 0$ implique $u = 0$ (T étant alors bijectif $\mathcal{H} \rightarrow \text{Im}(T)$). L'existence d'un inverse se résume ainsi à l'injectivité de $T : D(T) \rightarrow \mathcal{H}$.

Définition C.8 (Ensemble résolvant et spectre). *Soit T un opérateur linéaire non-borné, fermé, de domaine $D(T)$. L'ensemble résolvant $\rho(T)$ de T est l'ensemble des $\lambda \in \mathbb{C}$ tels que*

- (1) *l'opérateur $T - \lambda \text{Id}$ a un inverse $(T - \lambda \text{Id})^{-1}$, c'est-à-dire que $(T - \lambda \text{Id})u = f$ a une unique solution pour tout $f \in \text{Im}(T - \lambda \text{Id})$ (**injectivité**);*
- (2) *le domaine de $(T - \lambda \text{Id})^{-1}$ (*i.e.* l'image de $D(T)$ par $T - \lambda \text{Id}$) est dense dans $\mathcal{H}$ (**presque surjectivité**);*
- (3) *l'opérateur $(T - \lambda \text{Id})^{-1}$ est borné de $(T - \lambda \text{Id})(D(T))$ (muni de la topologie induite de $\mathcal{H}$) dans $\mathcal{H}$ (**continuité de l'inverse**).*

Le complémentaire de $\rho(T)$ dans $\mathbb{C}$ est le spectre de T , et on le note $\sigma(T)$.

On montre que l'ensemble résolvant est un ouvert de $\mathbb{C}$, et par conséquent, le spectre $\sigma(T)$ est un fermé de $\mathbb{C}$.

Remarque C.5 (Bornes sur le spectre d'un opérateur borné). Pour un opérateur borné, on peut montrer que

$$\sigma(T) \subset \{\lambda \in \mathbb{C}, |\lambda| \leq \|T\|\}.$$

Un élément $\lambda \in \mathbb{C}$ est dans le spectre de T si une des conditions de la Définition C.8 n'est pas remplie. Ceci conduit à la classification suivante des éléments de $\sigma(T)$:

Définition C.9. *Le spectre $\sigma(T)$ est la réunion disjointe de trois sous-ensembles :*

- le spectre ponctuel, noté $\sigma_p(T)$, qui est l'ensemble des $\lambda \in \sigma(T)$ pour lesquels $T - \lambda \text{Id}$ n'est pas injectif. De telles valeurs de λ sont appelées des valeurs propres de T ;
- le spectre résiduel, noté $\sigma_{\text{res}}(T)$, qui est l'ensemble des $\lambda \in \sigma(T)$ pour lesquels $(T - \lambda \text{Id})^{-1}$ est de domaine non dense dans $\mathcal{H}$. On a alors défaut de presque surjectivité : $\text{Im}(T - \lambda \text{Id}) \neq \mathcal{H}$;
- le spectre continu, noté $\sigma_{\text{cont}}(T)$, st l'ensemble des $\lambda \in \sigma(T)$ pour lesquels $(T - \lambda \text{Id})^{-1}$ est un opérateur non borné de domaine dense dans $\mathcal{H}$ (défaut de continuité).

On peut éventuellement considérer une autre partition du spectre en termes de spectre discret $\sigma_d(T)$, qui est l'ensemble des valeurs propres isolées, de multiplicité finie (*i.e.* $\dim \text{Ker}(T - \lambda \text{Id}) < +\infty$), et son complémentaire, appelé spectre essentiel $\sigma_{\text{ess}}(T)$. Le spectre essentiel est donc l'union des spectres résiduels et continus, ainsi que des valeurs propres de multiplicité infinie et des points d'accumulation des valeurs propres de multiplicité finie. On a une caractérisation des éléments du spectre essentiel *via* le

Lemme C.1 (Weyl). *Soit T un opérateur auto-adjoint sur un espace de Hilbert $\mathcal{H}$. Alors $\lambda \in \sigma_{\text{ess}}(T)$ si et seulement s'il existe une suite $(u_n)_{n \geq 0}$ de $\mathcal{H}$ telle que*

$$\|u_n\| = 1, \quad u_n \rightharpoonup 0, \quad \|(T - \lambda)u_n\| \rightarrow 0.$$

La suite (u_n) est appelée suite de Weyl de T pour la valeur λ .

Ce Lemme montre que les valeurs $\lambda \in \sigma_{\text{ess}}(T)$ sont des "quasi-valeurs propres" au sens où il existe une suite (u_n) d'éléments de norme 1 telle que $Au_n - \lambda u_n \rightarrow 0$.

Remarque C.6 (Lien entre spectre ponctuel d'un opérateur et spectre résiduel de son adjoint). Dans le cas où $\lambda \in \sigma(T) \setminus \sigma_p(T)$, on distingue le cas $\text{Im}(\lambda - T)$ dense (spectre continu) ou non (spectre résiduel). Ainsi, par définition, $\lambda \in \sigma_{\text{res}}(T)$ si et seulement s'il existe $v \neq 0$ tel que

$$\forall u \in D(T), \quad \langle (\lambda - T)u, v \rangle = 0.$$

Donc, pour $\lambda \in \sigma(T) \setminus \sigma_p(T)$, $\lambda \in \sigma_{\text{res}}(T)$ si et seulement si $\text{Ker}(\bar{\lambda} - T^*) \neq \{0\}$, ce qui revient à $\bar{\lambda} \in \sigma_p(T^*)$.

Le spectre peut donc être précisé dans le cas d'un opérateur auto-adjoint (voir la Remarque C.6) :

Théorème C.1. *Pour un espace de Hilbert séparable¹ $\mathcal{H}$, le spectre d'un opérateur auto-adjoint est contenu dans la droite réelle, et son spectre résiduel est vide.*

La première assertion est à rapprocher du fait que les matrices symétriques réelles sont diagonalisables et à valeurs propres réelles. Pour un opérateur auto-adjoint, le spectre essentiel comprend donc le spectre continu, les valeurs propres plongées dans le spectre continu, et les valeurs propres isolées de multiplicité infinie.

Exemples

Exemple de spectre résiduel.

On considère l'espace de Hilbert

$$l^2(\mathbb{C}) = \left\{ (u_n)_{n \geq 0} \left| \sum_{n=0}^{+\infty} |u_n|^2 < +\infty \right. \right\},$$

et l'opérateur de *shift* à droite $T : l^2(\mathbb{C}) \rightarrow l^2(\mathbb{C})$ de domaine $D(T) = l^2(\mathbb{C})$ donné par

1. Rappelons qu'un espace de Hilbert $\mathcal{H}$ est séparable s'il possède un sous-ensemble dénombrable dense, *i.e.* s'il existe $\mathcal{E} \subset \mathcal{H}$ dénombrable tel que $\text{Vect}(\mathcal{E}) = \mathcal{H}$.

$$(Tu)_n = \begin{cases} 0 & \text{si } n = 0, \\ u_{n-1} & \text{sinon.} \end{cases}$$

On a bien sûr $\sigma_p(T) = \emptyset$. Par ailleurs, $\text{Im}(\lambda - T) = \{(0, u), u \in l^2(\mathbb{C})\}$ n'est pas dense dans $\mathbb{C}$, on s'attend donc à ce que T possède un spectre résiduel. Nous allons préciser le spectre de T en calculant son adjoint et en utilisant la Remarque C.6. Pour trouver l'adjoint de T , on écrit

$$\langle Tu, v \rangle = \sum_{n \geq 0} u_n v_{n+1} = \langle u, T^* v \rangle.$$

L'opérateur T^* est donc l'opérateur de *shift* à gauche $(T^*u)_n = u_{n+1}$, de domaine $l^2(\mathbb{C})$ également. On vérifie facilement que $\|T\| = \|T^*\| = 1$, d'où

$$\sigma(T), \sigma(T^*) \subset \{|\lambda| \leq 1\}.$$

Par ailleurs, tout $\lambda \in \mathbb{C}$ tel que $|\lambda| < 1$ est valeur propre de T^* associé au vecteur propre $(u_n) = (\lambda^n)$. On montre également que T^* ne peut pas avoir de valeur propre de module supérieur ou égal à 1 (pour des questions de normalisation du vecteur propre). Ainsi,

$$\sigma_p(T^*) = \{|\lambda| < 1\},$$

et donc, utilisant la Remarque C.6,

$$\sigma_{\text{res}}(T) = \{|\lambda| < 1\}.$$

Comme le spectre est fermé, on a également

$$\sigma(T) = \sigma(T^*) = \{|\lambda| \leq 1\}.$$

Il est délicat de caractériser la partie $\{|\lambda| = 1\}$ des deux spectres (on pourra consulter [29, Section VI.3]).

Exemple de spectre continu.

Soit $T : L^2(\mathbb{R}^3) \rightarrow L^2(\mathbb{R}^3)$ défini par

$$Tu(x) = xu(x),$$

de domaine

$$D(T) = \{u \in L^2(\mathbb{R}^3) \mid xu \in L^2(\mathbb{R}^3)\}.$$

On montre tout d'abord que $\sigma(T) \subset \mathbb{R}$. En effet, soit $\lambda \in \mathbb{C} \setminus \mathbb{R}$, alors si $(\lambda - T)u = v$, on a

$$v(x) = (\lambda - T)^{-1}u = \frac{u(x)}{\lambda - x}.$$

Ceci montre que $(\lambda - T)^{-1}$ est continu et de domaine $L^2(\mathbb{R}^3)$. On montre en fait que c'est un opérateur borné, de borne $\|(\lambda - T)^{-1}\| = |\text{Im}(\lambda)|^{-1}$.

Montrons à présent $\sigma(T) = \sigma_c(T) = \mathbb{R}$. Soit $\lambda \in \mathbb{R}$. On remarque en premier lieu que $\lambda \notin \sigma_p(T)$, car $(x - \lambda)u(x) = 0$ implique $u = 0$ au sens de $L^2(\mathbb{R}^3)$. Ensuite, on peut se rendre compte que T est auto-adjoint, donc que son spectre résiduel est vide. On conclut alors en trouvant une suite de Weyl pour λ : par exemple, $u_n = n^{3/2}\tilde{u}(n(x - \lambda))$ avec $\tilde{u} \in C_0^\infty(\mathbb{R}^3)$ à support dans $B(0, 1)$ et de norme $\|\tilde{u}\|_{L^2(\mathbb{R}^3)} = 1$, est telle que $\|u_n\|_{L^2(\mathbb{R}^3)} = \|\tilde{u}\|_{L^2(\mathbb{R}^3)} = 1$, et

$$\|(T - \lambda)u_n\|_{L^2(\mathbb{R}^3)}^2 = \int_{\mathbb{R}^3} |x - \lambda|^2 |\tilde{u}(n(x - \lambda))|^2 n \, dx = \frac{1}{n^2} \int_{\mathbb{R}^3} |y|^2 |\tilde{u}(y)|^2 \, dy \rightarrow 0,$$

et $u_n \rightharpoonup 0$ car, pour $w \in L^2(\mathbb{R}^3)$,

$$\left| \int_{\mathbb{R}^3} \overline{u_n} w \right| = \left| \int_{|x - \lambda| \leq \frac{1}{n}} \overline{u_n} w \right| \leq \|\tilde{u}\|_{L^2(\mathbb{R}^3)} \left(\int_{|x - \lambda| \leq \frac{1}{n}} |w|^2 \right)^{1/2} \rightarrow 0.$$

C.4 Diagonalisation des opérateurs autoadjoints compacts

Définition C.10 (Opérateur compact). On dit qu'un opérateur $T : E \rightarrow F$ est compact si l'image de la boule unité de E est d'adhérence compacte², i.e. de toute suite bornée $(u_n)_{n \geq 0}$ de E , on peut extraire une sous-suite $(Au_{n_k})_{k \geq 0}$ convergeant dans F .

Bien sûr, un opérateur compact est un opérateur borné. On dispose d'un théorème de diagonalisation des opérateurs compacts :

Théorème C.2 (Diagonalisation des opérateurs compacts). Soit $\mathcal{H}$ un espace de Hilbert séparable de dimension infinie et $T : \mathcal{H} \rightarrow \mathcal{H}$ un opérateur borné auto-adjoint compact. Alors il existe une suite $(\mu_n)_{n \geq 0}$ de réels tendant vers 0, et une base hilbertienne $(e_n)_{n \geq 0}$ de $\mathcal{H}$ telle que

$$\sigma(T) = (\mu_n) \cup \{0\}, \quad \text{avec } \forall n \in \mathbb{N}, T e_n = \mu_n e_n.$$

Par ailleurs, pour tout $\lambda \in \sigma(T) \setminus \{0\}$, l'espace propre $E_\lambda = \text{Ker}(\lambda - T)$ est de dimension finie.

Remarque C.7. Le spectre des opérateurs auto-adjoints compacts peut être de cardinal fini (on dit qu'on a un opérateur de rang fini), auquel cas il est de la forme $\sigma(T) = \sigma_p(T) = \{\mu_1, \dots, \mu_n\} \cup \{0\}$. Si T n'est pas de rang fini, il est possible qu'il soit injectif si son spectre ponctuel est de la forme $\{\mu_1, \dots, \mu_n, \dots\}$ avec $\mu_n \rightarrow 0$ lorsque $n \rightarrow +\infty$. Dans ce cas, $\sigma_c(T) = \{0\}$.

En pratique, on utilisera surtout le

Corollaire C.1. Soit $\mathcal{H}$ est un espace de Hilbert séparable de dimension infinie, et T un opérateur auto-adjoint non-borné sur $\mathcal{H}$. S'il existe $\lambda \in \rho(T) \cap \mathbb{R}$ tel que $(\lambda - T)^{-1}$ soit un opérateur compact, alors il existe une suite $(\nu_n)_{n \geq 0}$ de réels non nuls telle que $|\nu_n| \rightarrow +\infty$, et une base hilbertienne $(e_n)_{n \geq 0}$ de $\mathcal{H}$ telle que

$$\sigma(T) = \{\nu_n\} \cup \{0\},$$

avec

$$\forall n \in \mathbb{N}, T e_n = \nu_n e_n.$$

Par ailleurs, pour tout $n \geq 0$, l'espace propre $E_n = \text{Ker}(\nu_n - T)$ est de dimension finie.

C.5 Décomposition des opérateurs auto-adjoints quelconques

On note $\mathcal{P}$ l'ensemble des projecteurs orthogonaux sur $\mathcal{H}$, i.e. $P \in \mathcal{P}$ si P est un opérateur tel que $P = P^* = P^2$.

Définition C.11 (Famille spectrale). Une famille spectrale est une famille $(P_\lambda)_{\lambda \in \mathbb{R}}$ de projecteurs orthogonaux sur $\mathcal{H}$ tels que

- (1) $P_\lambda \cdot P_\mu = P_{\inf(\lambda, \mu)}$;
- (2) $s - \lim_{\mu \rightarrow \lambda, \mu > \lambda} P_\mu = P_\lambda$;
- (3) $s - \lim_{\lambda \rightarrow -\infty} P_\lambda = 0$, $s - \lim_{\lambda \rightarrow +\infty} P_\lambda = \text{Id}$;

La limite forte (*strong limit*) $s - \lim$ est à entendre au sens suivant : $s - \lim_{n \rightarrow +\infty} T_n = T$ si et seulement si pour tout $x \in \mathcal{H}$, $T_n x \rightarrow T x$.

On notera par la suite $\mathcal{S}$ l'ensemble des familles spectrales. Pour une famille spectrale donnée, on peut définir un opérateur T de domaine

$$D = \left\{ x \in \mathcal{H}, \int_{-\infty}^{+\infty} |\lambda|^2 d\|P_\lambda x\|^2 < +\infty \right\}$$

2. On parle d'ensemble relativement compact lorsque l'adhérence de cet ensemble est compact. Cela revient à dire que l'ensemble en question est contenu dans un compact.

par

$$\forall x \in D, \quad Tx = \int_{-\infty}^{+\infty} \lambda dP_\lambda x.$$

La mesure borélienne $\mu_\phi(\Omega) = (\phi, P_\Omega \phi)$ est noté $d(\phi, P_\lambda \phi)$. Les intégrales ci-dessus sont des intégrales au sens de Riemann : pour f continue, et un intervalle $-\infty < a < b < +\infty$,

$$\int_a^b f(\lambda) dP_\lambda x = \lim_{K \rightarrow +\infty} \sum_{k=1}^K f(\lambda_k) (P_{\lambda_{k+1}} - P_{\lambda_k}) x$$

pour une partition de $[a, b]$ de pas tendant vers 0 lorsque $K \rightarrow +\infty$. L'intégrale est bien définie car $\lambda \mapsto P_\lambda x$ est une fonction à variations bornées. On considère ensuite la limite $a \rightarrow -\infty$ et $b \rightarrow +\infty$ de ces intégrales. On vérifie que Tx est bien défini si $x \in D(T)$ et que les opérateurs ainsi construits sont linéaires auto-adjoints non-bornés. On définit ainsi une application $\mathcal{S} \rightarrow \mathcal{A}$, où on a noté $\mathcal{A}$ l'ensemble des opérateurs auto-adjoints non-bornés.

Théorème C.3. *L'application F définit une bijection entre les familles spectrales et les opérateurs auto-adjoints.*

Les opérateurs auto-adjoints quelconques sont donc caractérisés de manière équivalente par leur mesure spectrale, qui est une mesure borélienne à valeur projecteurs. Pour un opérateur T auto-adjoint, on considère

$$P^T : \mathcal{B}(\mathbb{R}) \rightarrow \mathcal{P}$$

telle que $|P^T(B)\psi|^2$ soit la probabilité que T prenne ses valeurs dans B . On a en fait

$$P^T(B) = \int_B dP_\lambda.$$

Notons également que dans ce formalisme

$$\text{Id} = \int_{-\infty}^{+\infty} dP_\lambda, \quad T = \int_{-\infty}^{+\infty} \lambda dP_\lambda.$$

Exemple C.2. Dans le cas d'un opérateur auto-adjoint compact, on a un spectre purement ponctuel, et dans ce cas

$$P_\lambda = \sum_{\lambda_k \leq \lambda} P_{E_k}$$

où P_{E_k} est le projecteur orthogonal sur le sous-espace propre E_k associé à la valeur propre λ_k .

Exemple C.3. On présente ici la décomposition explicite de certains opérateurs quantiques :

- pour l'opérateur $\hat{x}_k$, on a $(P_\lambda^{\hat{x}_k} \cdot u)(x) = \mathbf{1}_{x_k \leq \lambda} u(x)$;
- pour l'opérateur $\hat{p}_k$, on a $P_\lambda^{\hat{p}_k} \cdot u = \mathcal{F}^{-1} \mathbf{1}_{\xi_k \leq \lambda} \mathcal{F} u$;
- pour le Hamiltonien $H_0 = -\frac{1}{2} \Delta$ de la particule libre, on a $H_0 u = (2\pi)^{-3/2} \int_0^{+\infty} \frac{1}{2} |k|^2 \hat{u}(k) e^{ik \cdot x} dk$,
et donc $P_\lambda^{H_0} \cdot u = \mathcal{F}^{-1} \mathbf{1}_{\frac{1}{2} |\xi|^2 \leq \lambda} \mathcal{F} u$.

Dans ce cadre, la nomenclature du spectre devient plus parlante :

Théorème C.4. *Soit $\mathcal{H}$ un espace de Hilbert séparable, T un opérateur auto-adjoint sur $\mathcal{H}$ et $(P_\lambda)_{\lambda \in \mathbb{R}}$ la famille spectrale qui lui est associée. Alors*

- $\lambda \in \sigma_p(T)$ si et seulement si $\lim_{\epsilon \rightarrow 0, \epsilon > 0} P_{\lambda - \epsilon} \neq P_\lambda$;
- $\lambda \in \sigma_c(T)$ si et seulement si $\lim_{\epsilon \rightarrow 0, \epsilon > 0} P_{\lambda - \epsilon} = P_\lambda$ et $P_{\lambda - \epsilon} \neq P_{\lambda + \epsilon}$ pour tout $\epsilon > 0$.

Ainsi, les éléments du spectre ponctuel sont les points où la mesure spectrale “saute”, alors qu'elle varie continûment aux points du spectre continu.

D

Quelques relations de commutation entre observables

Relations de commutations les plus importantes

$$\begin{aligned} [\hat{x}, \hat{p}_x] &= i\hbar, & [\hat{x}, \hat{p}_y] &= 0, & [\hat{x}, \hat{p}_z] &= 0, \\ [\hat{y}, \hat{p}_x] &= 0, & [\hat{y}, \hat{p}_y] &= i\hbar, & [\hat{y}, \hat{p}_z] &= 0, \\ [\hat{z}, \hat{p}_x] &= 0, & [\hat{z}, \hat{p}_y] &= 0, & [\hat{z}, \hat{p}_z] &= i\hbar, \\ [\hat{L}_x, \hat{L}_y] &= i\hbar \hat{L}_z, & [\hat{L}_y, \hat{L}_z] &= i\hbar \hat{L}_x, & [\hat{L}_z, \hat{L}_x] &= i\hbar \hat{L}_y. \end{aligned}$$

Autres relations de commutation

$$\begin{aligned} [\hat{L}_x, \hat{x}] &= 0, & [\hat{L}_x, \hat{y}] &= i\hbar \hat{z}, & [\hat{L}_x, \hat{z}] &= -i\hbar \hat{y}, \\ [\hat{L}_y, \hat{x}] &= -i\hbar \hat{z}, & [\hat{L}_y, \hat{y}] &= 0, & [\hat{L}_y, \hat{z}] &= i\hbar \hat{x}, \\ [\hat{L}_z, \hat{x}] &= i\hbar \hat{y}, & [\hat{L}_z, \hat{y}] &= -i\hbar \hat{x}, & [\hat{L}_z, \hat{z}] &= 0, \\ [\hat{L}_x, \hat{x}^2] &= 0, & [\hat{L}_x, \hat{y}^2] &= 2i\hbar \hat{y} \hat{z}, & [\hat{L}_x, \hat{z}^2] &= -2i\hbar \hat{y} \hat{z}, \\ [\hat{L}_y, \hat{x}^2] &= -2i\hbar \hat{x} \hat{z}, & [\hat{L}_y, \hat{y}^2] &= 0, & [\hat{L}_y, \hat{z}^2] &= 2i\hbar \hat{x} \hat{z}, \\ [\hat{L}_z, \hat{x}^2] &= 2i\hbar \hat{x} \hat{y}, & [\hat{L}_z, \hat{y}^2] &= -2i\hbar \hat{x} \hat{y}, & [\hat{L}_z, \hat{z}^2] &= 0, \\ [\hat{L}_x, \hat{p}_x] &= 0, & [\hat{L}_x, \hat{p}_y] &= i\hbar \hat{p}_z, & [\hat{L}_x, \hat{p}_z] &= -i\hbar \hat{p}_y, \\ [\hat{L}_y, \hat{p}_x] &= -i\hbar \hat{p}_z, & [\hat{L}_y, \hat{p}_y] &= 0, & [\hat{L}_y, \hat{p}_z] &= i\hbar \hat{p}_x, \\ [\hat{L}_z, \hat{p}_x] &= i\hbar \hat{p}_y, & [\hat{L}_z, \hat{p}_y] &= -i\hbar \hat{p}_x, & [\hat{L}_z, \hat{p}_z] &= 0, \\ [\hat{L}_x, \hat{p}_x^2] &= 0, & [\hat{L}_x, \hat{p}_y^2] &= 2i\hbar \hat{p}_y \hat{p}_z, & [\hat{L}_x, \hat{p}_z^2] &= -2i\hbar \hat{p}_y \hat{p}_z, \\ [\hat{L}_y, \hat{p}_x^2] &= -2i\hbar \hat{p}_x \hat{p}_z, & [\hat{L}_y, \hat{p}_y^2] &= 0, & [\hat{L}_y, \hat{p}_z^2] &= 2i\hbar \hat{p}_x \hat{p}_z, \\ [\hat{L}_z, \hat{p}_x^2] &= 2i\hbar \hat{p}_x \hat{p}_y, & [\hat{L}_z, \hat{p}_y^2] &= -2i\hbar \hat{p}_x \hat{p}_y, & [\hat{L}_z, \hat{p}_z^2] &= 0, \\ [\hat{L}_x^2, \hat{x}^2 + \hat{y}^2 + \hat{z}^2] &= 0, & [\hat{L}_y^2, \hat{x}^2 + \hat{y}^2 + \hat{z}^2] &= 0, & [\hat{L}_z^2, \hat{x}^2 + \hat{y}^2 + \hat{z}^2] &= 0, \\ [\hat{L}_x^2, \hat{p}_x^2 + \hat{p}_y^2 + \hat{p}_z^2] &= 0, & [\hat{L}_y^2, \hat{p}_x^2 + \hat{p}_y^2 + \hat{p}_z^2] &= 0, & [\hat{L}_z^2, \hat{p}_x^2 + \hat{p}_y^2 + \hat{p}_z^2] &= 0, \\ [\hat{L}^2, \hat{L}_x] &= 0, & [\hat{L}^2, \hat{L}_y] &= 0, & [\hat{L}^2, \hat{L}_z] &= 0, \\ [\hat{L}_x, \hat{V}] &= i\hbar (\hat{y} \hat{F}_z - \hat{z} \hat{F}_y), & [\hat{L}_y, \hat{V}] &= i\hbar (\hat{z} \hat{F}_x - \hat{x} \hat{F}_z), & [\hat{L}_z, \hat{V}] &= i\hbar (\hat{x} \hat{F}_y - \hat{y} \hat{F}_x). \end{aligned}$$

Bibliographie

- [1] M. P. ALLEN ET D. J. TILDESLEY, *Computer Simulation of Liquids* (Oxford Science Publications, 1987).
- [2] R. BALIAN, *Du microscopique au macroscopique (cours de physique statistique de l'École Polytechnique)* (Ellipses, 1982).
- [3] E. BETRANHANDY, *Proposition et études ab initio des stabilités relatives de nouveaux matériaux par géomimétisme structural : modélisation des propriétés physico-chimiques*, Thèse de doctorat (Université Sciences et Technologies - Bordeaux I, 2005).
- [4] HAÏM BRÉZIS, *Analyse fonctionnelle : théorie et applications* (Dunod, 1999).
- [5] E. CANCÈS, F. LEGOLL, ET G. STOLTZ, Theoretical and numerical comparison of sampling methods for molecular dynamics, *accepted for publication in M2AN* (2007).
- [6] ERIC CANCÈS, CLAUDE LE BRIS, ET YVON MADAY, *Méthodes mathématiques en chimie quantique*, volume 53 of *Mathématiques & Applications* (Springer-Verlag, Berlin, Heidelberg, 2006).
- [7] PHILIPPE CHARTIER ET ERWAN FAOU, Intégration symplectique des systèmes hamiltoniens intégrables : comportement en temps long, *Cours de DEA Niveau II* (2003). (disponible à l'adresse <http://www.irisa.fr/ipso/perso/faou/publis/NotesDEA03.ps>).
- [8] O. COUSSY ET F. CHEVOIR. Physique des états de la Matière. Cours de première année, ENPC.
- [9] M. CROUZEIX ET A.L. MIGNOT, *Analyse numérique des équations différentielles* (Masson, 1984).
- [10] B. DIU, C. GUTHMAN, D. LEDERER, ET B. ROULET, *Physique statistique* (Hermann, 1989).
- [11] A. DOUADY, Déterminisme et indéterminisme dans un modèle mathématique, In *Chaos et déterminisme*, A. Dahan-Dalmenico and P. Arnoux, Eds. (Le Seuil, 1992), pp. 11–18.
- [12] S. DUANE, A. D. KENNEDY, B. J. PENDLETON, ET D. ROWETH, Hybrid Monte-Carlo, *Phys. Lett. B* **195**(2) (1987) 216–222.
- [13] R. DAUTRAY ET J.L. LIONS, *Analyse mathématique et calcul numérique pour les sciences et les techniques. I. Modèles physiques* (Masson, 1987).
- [14] H. FLYVBERG ET H. G. PETERSEN, Error estimates on averages of correlated data, *J. Chem. Phys.* **91** (1989) 461–466.
- [15] D. FRENKEL ET B. SMIT, *Understanding Molecular Simulation* (Academic Press, 2005).
- [16] C. J. GEYER, Practical Markov chain Monte Carlo (with discussion), *Stat. Sci.* **7**(4) (1992) 473–511.
- [17] W. R. GILKS, S. RICHARDSON, ET D. J. SPIEGELHALTER, *Markov Chain Monte Carlo in practice* (Chapman and Hall, 1996).
- [18] E. HAIRER, C. LUBICH, ET G. WANNER, Geometric numerical integration illustrated by the symplectic method, *Acta Numerica* **12** (2003) 399–450.

- [19] ERNST HAIRER, CHRISTIAN LUBICH, ET GERHARD WANNER, *Geometric Numerical Integration : Structure-Preserving Algorithms for Ordinary Differential Equations*, volume 31 of *Springer Series in Computational Mathematics* (Springer-Verlag, Berlin, Heidelberg, 2006).
- [20] J.-P. HANSEN ET I. R. McDONALD, *Theory of Simple Liquids* (Academic Press, 1986).
- [21] W. K. HASTINGS, Monte Carlo sampling methods using Markov chains and their applications, *Biometrika* **57** (1970) 97–109.
- [22] J. ISRAELASHVILI, *Intermolecular and surface forces* (Academic Press, 1992).
- [23] L. LANDAU ET E. LIFCHITZ, *Physique statistique* (Ellipses, 1994).
- [24] B. J. LEIMKUHLER ET S. REICH, *Simulating Hamiltonian dynamics*, volume 14 of *Cambridge monographs on applied and computational mathematics* (Cambridge University Press, 2005).
- [25] S.-K. MA, *Statistical Mechanics* (World Scientific, 1985).
- [26] N. METROPOLIS, A. W. ROSENBLUTH, M. N. ROSENBLUTH, A. H. TELLER, ET E. TELLER, Equations of state calculations by fast computing machines, *J. Chem. Physics* **21**(6) (1953) 1087–1091.
- [27] S. P. MEYN ET R. L. TWEEDIE, *Markov chains and stochastic stability*, Communications and control engineering series (Springer-Verlag, London, New York, 1993).
- [28] J. VON NEUMANN, *Mathematical foundations of quantum mechanics*, Princeton landmarks in mathematics and physics (Princeton university press, 1996).
- [29] M. REED ET B. SIMON, *Methods of modern mathematical physics*, volume I - IV (Academic Press, 1980).
- [30] G. ROBERTS ET R. TWEEDIE, Exponential convergence of langevin diffusions and their discrete approximations, *Bernoulli* **2** (1997) 314–363.