



Claude Fabre
Charles Antoine
Nicolas Treps

Introduction à la physique moderne : relativité et physique quantique

◆ Cours
◆ Exercices

Licence
Prépa
CAPES

RESSOURCES



NUMÉRIQUES

DUNOD

Illustration de couverture : © Sakkmasterke - Fotolia.com

Le pictogramme qui figure ci-contre mérite une explication. Son objet est d'alerter le lecteur sur la menace que représente pour l'avenir de l'écrit, particulièrement dans le domaine de l'édition technique et universitaire, le développement massif du photocopillage.

Le Code de la propriété intellectuelle du 1^{er} juillet 1992 interdit en effet expressément la photocopie à usage collectif sans autorisation des ayants droit. Or, cette pratique s'est généralisée dans les établissements

d'enseignement supérieur, provoquant une baisse brutale des achats de livres et de revues, au point que la possibilité même pour

les auteurs de créer des œuvres nouvelles et de les faire éditer correctement est aujourd'hui menacée. Nous rappelons donc que toute reproduction, partielle ou totale, de la présente publication est interdite sans autorisation de l'auteur, de son éditeur ou du

Centre français d'exploitation du droit de copie (CFC, 20, rue des Grands-Augustins, 75006 Paris).



© Dunod, 2015

5 rue Laromiguière, 75005 Paris

www.dunod.com

ISBN 978-2-10-072570-0

Le Code de la propriété intellectuelle n'autorisant, aux termes de l'article L. 122-5, 2° et 3° a), d'une part, que les « copies ou reproductions strictement réservées à l'usage privé du copiste et non destinées à une utilisation collective » et, d'autre part, que les analyses et les courtes citations dans un but d'exemple et d'illustration, « toute représentation ou reproduction intégrale ou partielle faite sans le consentement de l'auteur ou de ses ayants droit ou ayants cause est illicite » (art. L. 122-4).

Cette représentation ou reproduction, par quelque procédé que ce soit, constituerait donc une contrefaçon sanctionnée par les articles L. 335-2 et suivants du Code de la propriété intellectuelle.

PRÉFACE

La physique moderne est née au début du siècle dernier, lorsque les conceptions sur lesquelles était basée la connaissance du monde ont été bouleversées par quelques observations apparemment disparates concernant la lumière. Celle-ci semblait bien comprise et expliquée par les équations établies il y a cent cinquante ans par Maxwell, considérées comme l'aboutissement triomphal de la physique classique. Et pourtant, c'est en se posant des questions profondes sur la lumière, sa vitesse de propagation, la nature de l'onde qui la constitue, et sur certains phénomènes lumineux devenus accessibles à l'expérience – la distribution en fréquence de la lumière des corps chauffés, les propriétés de l'effet photoélectrique, le caractère discret des spectres absorbés ou émis par la matière – que les physiciens ont au début du XX^e siècle complètement révolutionné notre vision du monde. La relativité restreinte, en introduisant l'idée que la vitesse de la lumière ne dépend pas de l'observateur et en admettant la vitesse maximale de propagation de toute information, devait bouleverser notre conception de l'espace et du temps, ainsi que notre vision de l'univers et de la cosmologie lorsqu'elle fut étendue à la description des effets de gravitation dans le cadre de la relativité générale. Par ailleurs, en montrant que la lumière était à la fois une onde et un ensemble de particules (les photons), la théorie quantique nous introduisait dans le monde étrange de la physique de l'infiniment petit, celui où les notions d'onde et de corpuscules sont indissociablement mêlées, nous livrant les clés du monde microscopique. Albert Einstein est au point de départ de ces deux révolutions puisque la relativité restreinte et la physique quantique sont nées de deux articles qu'il a publiés en 1905 – année miraculeuse pour la science – et qu'on lui doit aussi la théorie de la relativité générale établie dix ans plus tard.

Fondamentale pour notre compréhension profonde de la nature, la physique quantique est aussi à la base de toutes les technologies modernes omniprésentes dans notre vie quotidienne. Sans physique quantique, il n'y aurait ni lasers, ni transistors, ni ordinateurs, ni imagerie médicale par résonance magnétique. Dans certains cas, la relativité joue également un rôle essentiel dans notre vie quotidienne. Par exemple, le GPS, système de navigation universel qui nous permet de nous repérer par triangulation n'importe où sur Terre avec une précision de quelques mètres, exploite à la fois la physique quantique nécessaire à la compréhension du fonctionnement des horloges atomiques, et la relativité, restreinte et générale, sans laquelle les mesures du temps nécessaires aux opérations de triangulation seraient entachées de larges erreurs. Les centrales nucléaires, sources d'énergie essentielle de notre monde contemporain, sont un autre exemple de réalisations qui auraient été impossibles sans les connaissances que nous ont apportées les théories de la relativité et de la physique quantique.

Introduction à la physique moderne

Il est paradoxal que ces théories dont l'importance est capitale paraissent ésotériques au public non spécialisé et soient encore considérées comme d'accès difficile dans l'enseignement, au point que leur étude quantitative soit repoussée tardivement dans le cursus des étudiants. Il est vrai que les concepts quantiques et relativistes peuvent être déroutants car ils concernent des phénomènes à des échelles de vitesses ou de distances inhabituelles à la perception de nos sens. Les mathématiques nécessaires à la compréhension de cette physique sont cependant relativement élémentaires, à la portée d'un élève des classes terminales scientifiques.

Les auteurs de ce livre ont trouvé la bonne façon d'expliquer comment ces concepts sont nés, en présentant le caractère logique des déductions ayant amené Einstein et ses collègues à la conclusion que les lois étranges de la relativité et du monde quantique devaient s'imposer. Et ils ont su accompagner les idées physiques du formalisme mathématique simple et rigoureux qui les traduit pour permettre le calcul d'effets physiques fondamentaux. Par-là, ils révèlent aux étudiants et à un public bien plus large la beauté de la théorie, initiant le lecteur à un monde dans lequel les concepts dont il a parfois entendu parler de façon qualitative trouvent leur place naturelle. La présentation en regard, dans un même volume, des deux révolutions qui ont changé notre vision du monde est également une heureuse innovation par rapport à la plupart des ouvrages qui séparent les deux questions. Il est en effet naturel de parler du photon en introduisant des notions de relativité, ne serait-ce que parce que la relation de Planck entre énergie et fréquence du quantum de lumière n'est, comme Louis de Broglie l'a reconnu, que l'une des composantes d'une relation vectorielle relativiste plus générale qui associe également, pour toute particule et pas seulement le photon, l'impulsion à la longueur d'onde de l'onde associée.

En quelque deux cent cinquante pages illustrées de figures très parlantes et émaillées d'encarts précisant des points importants sans couper le fil de la discussion générale, les auteurs ont réalisé la gageure de couvrir un champ immense, en équilibrant une approche théorique simple, la description d'expériences souvent très récentes, et celle d'appareils ou de dispositifs de notre vie quotidienne. Ainsi, le lecteur comprend le lien profond entre recherche fondamentale et innovation et la nécessité de la première comme condition essentielle du développement de la seconde. En ces temps où cette complémentarité n'est pas toujours bien comprise, ce livre joue ainsi un rôle d'éducation salutaire. J'ai également trouvé très positive la démarche consistant à décrire de vraies expériences et non pas des expériences de pensée pour analyser les concepts de base, et l'utilisation résolument moderne du langage de la théorie de l'information pour parler des concepts essentiels de la relativité et de la physique quantique. Tout à fait remarquable, enfin, est le point de vue pris par les auteurs de décrire une science en évolution, ouverte, qui pose encore de nombreuses questions non résolues dans les domaines de l'infiniment petit et de

l'infiniment grand. Il est bon que tout étudiant s'intéressant à ces questions puisse se les poser et qu'il ait entendu parler d'ondes de gravitation, de matière et d'énergie noire. Au moins sur ces plans, il en sait presque autant que les auteurs du livre et les maîtres qui lui enseignent la science connue... et cela peut l'inspirer en lui montrant qu'il reste encore en physique des terres vierges à explorer.

Ce livre sera extrêmement utile aux étudiants comme première approche à la physique moderne, aux enseignants qui y puiseront beaucoup d'idées pédagogiques pour leurs cours, ainsi qu'aux lecteurs curieux qui y trouveront une source d'information très riche pour compléter leur culture scientifique. Claude Fabre, Charles Antoine et Nicolas Treps y ont mis toute leur expérience d'enseignants de la physique et de chercheurs passionnés par la découverte encore si riche du monde des atomes et des photons. Spécialistes de la physique quantique, ils sont particulièrement bien placés pour nous parler des révolutions que l'étude de la lumière a déclenchées il y a un peu plus d'un siècle...

Serge Haroche
Collège de France,
Prix Nobel de Physique 2012

TABLE DES MATIÈRES

Préface	V
Avant-propos	XIII
Introduction	XV

PARTIE 1

RELATIVITÉ

Espace, temps et mouvement en physique	3
Chapitre 1. Mécanique classique et changement de référentiel galiléen	7
1.1 Mécanique newtonienne et principe de relativité galiléenne	8
1.2 Changement de référentiel galiléen	10
Chapitre 2. Lumière classique et changement de référentiel galiléen	13
2.1 Description ondulatoire de la lumière	14
2.2 Lumière et changement de référentiel	16
2.3 Mesures de la vitesse de la lumière	22
Chapitre 3. Principe de relativité restreinte Transformations de Lorentz	31
3.1 Principe de relativité restreinte	31
3.2 Transformations de Lorentz	34
Chapitre 4. Effets relativistes sur le temps et l'espace	39
4.1 Intervalle temporel : relativité du passé, du futur et de la simultanéité	39
4.2 Intervalle d'espace-temps	41
4.3 Longueur propre et contraction des longueurs	43
4.4 Durée propre et dilatation des durées	46
4.5 Les « jumeaux de Langevin »	48
4.6 Des expériences de pensée aux expériences réelles	49
Chapitre 5. Effets relativistes sur les vitesses	55
5.1 Composition des vitesses en relativité	55
5.2 Transformation de Lorentz pour une onde	58
Chapitre 6. Deux exemples de phénomènes relativistes	63
6.1 Le GPS : un laboratoire relativiste	63
6.2 Rayonnement synchrotron	65

Introduction à la physique moderne

Chapitre 7. Dynamique relativiste	69
7.1 Quelques rappels de dynamique classique	69
7.2 Quantité de mouvement relativiste	74
7.3 Énergie relativiste	78
7.4 Collisions relativistes	87
Chapitre 8. Réactions nucléaires et notions de radioactivité	95
8.1 Structure des noyaux atomiques, énergie de liaison	95
8.2 De multiples réactions nucléaires	99
8.3 Radioactivité et dosimétrie	105
Chapitre 9. Les quatre interactions fondamentales, la relativité générale	111
9.1 Interaction gravitationnelle et notion de relativité générale	111
9.2 Interactions électromagnétique, forte et faible	124
9.3 Conclusion : théories d'unification	130

PARTIE 2

MÉCANIQUE QUANTIQUE

Le monde quantique	135
Chapitre 10. Le photon : une introduction à la physique quantique	141
10.1 Le photon : un fait expérimental	142
10.2 Propriétés du photon	145
10.3 Temps et lieu d'arrivée du photon	149
10.4 Interférences lumineuses et photons	153
10.5 Bilan	156
Chapitre 11. Polarisation de la lumière : aspects classiques et quantiques	159
11.1 Description classique de la polarisation	159
11.2 Description quantique	164
Chapitre 12. Systèmes à plusieurs états : notion d'états quantiques	171
12.1 États quantiques	171
12.2 Manipulation et mesure d'états quantiques	174
12.3 Application à la cryptographie	177
12.4 Évolution des états quantiques	183
Chapitre 13. Ondes de matière : une introduction à la fonction d'onde	191
13.1 Onde de matière	191
13.2 Quantification de l'énergie pour une particule confinée	202
13.3 Inégalités de Heisenberg	210

Chapitre 14. Évolution temporelle de la fonction d'onde : équation de Schrödinger d'une particule matérielle	215
14.1 Propriétés de la fonction d'onde	215
14.2 Commentaires	216
14.3 Équation de Schrödinger pour la fonction d'onde	217
14.4 Exemple 1 : la particule libre	219
14.5 Exemple 2 : le puits carré infini	220
14.6 Généralisation	223
14.7 Exemple 3 : marche de potentiel, onde de matière évanescence	224
14.8 Mesure de la vitesse de la particule, notion d'observable	227
Chapitre 15. Les règles générales de la mécanique quantique	235
15.1 À la base de tout : le vecteur d'état	235
15.2 Évolution temporelle	238
15.3 La mesure, lien entre le système physique et l'observateur	239
15.4 Principe d'exclusion de Pauli	246
Chapitre 16. Notions de mécanique quantique « et » relativiste	249
16.1 Équation de Klein-Gordon	250
16.2 Équation de Dirac, théorie quantique des champs	250
16.3 Incompatibilité de la mécanique quantique et de la relativité générale	251
Exercices et problèmes	255
Bibliographie	283
Index	285

AVANT-PROPOS

Cet ouvrage est issu d'un cours enseigné en deuxième année de licence à l'Université Pierre-et-Marie Curie Paris-Sorbonne Universités, ayant pour but d'initier les étudiants aux deux grands « piliers » de la physique actuelle que sont la relativité et la physique quantique. Il présente, au niveau le plus élémentaire possible, les concepts de base de ces deux théories et les illustre par de nombreux résultats expérimentaux récents. Il cherche avant tout à introduire les idées et outils essentiels nécessaires à la compréhension de ces domaines de la physique, en laissant volontairement de côté certains détails techniques ou théoriques qui font l'objet d'enseignements ultérieurs. Il vise également à permettre aux lecteurs d'accéder à des domaines de la physique qui sont à l'origine de recherches extrêmement actives dans les laboratoires de physique du monde entier.

Il s'adresse donc à toutes les personnes désireuses de s'initier à la « physique moderne », et de comprendre les bases et les enjeux de ses développements actuels : en premier lieu les étudiants de licence de physique, mais aussi ceux de licence de mathématique et de chimie, les étudiants des classes préparatoires et d'écoles d'ingénieurs et ceux préparant les concours de recrutement de professeurs de physique et chimie. Il sera certainement aussi très utile à tous les enseignants de physique désireux de renforcer leurs connaissances dans des domaines qui sont maintenant abordés à un niveau élémentaire dans les nouveaux programmes des lycées.

Les différents chapitres de ce manuel comportent des encarts qui ne sont pas indispensables à la compréhension des sujets abordés lors d'une première lecture. Ils permettent de les mettre en perspective ou bien détaillent certains calculs. Enfin, nous proposons à la fin de l'ouvrage des exercices et problèmes qui permettront au lecteur d'appliquer les concepts introduits et de les illustrer par des exemples. Les corrigés détaillés sont téléchargeables sur la page dédiée à l'ouvrage sur dunod.com.

Remerciements

Les trois auteurs tiennent à remercier tout particulièrement leurs collègues - maîtres de conférences, moniteurs et membres administratifs et techniques - qui ont participé au bon déroulement et à la mise au point de cet enseignement : J. Arlandis, M. Bertin, J. Beugnon, D. Brouri, T. Blanchard, C. Claveau, L. Delbes, J.-P. Ferreira, P. Fleury, R. Geiger, J. Joseph, A. Penin, M.-B. Povie, C. Sajus, T. Rybarczyk... et évidemment les étudiants dont les nombreuses questions ont contribué à faire évoluer cet enseignement, en particulier N. Proust et F. Bonnet pour leur relecture attentive du manuscrit. Nos remerciements vont aussi à nos maîtres qui nous ont enseigné et fait aimer cette physique que nous enseignons à notre tour : J.-L. Basdevant, C. Cohen-Tannoudji, A. Messiah, P. Tourrenc, J. Dalibard, B. Linet, L. Nottale, D. Hirondel... et tout particulièrement S. Haroche qui a bien voulu écrire une préface pour cet ouvrage.

INTRODUCTION

La physique s'est élaborée au cours des siècles à partir d'observations de phénomènes naturels, puis d'expérimentations, c'est-à-dire de mesures quantitatives réalisées dans des conditions les plus contrôlées possibles. Celles-ci ont permis de dégager des grandeurs physiques caractéristiques, puis des corrélations quantitatives entre ces grandeurs, qui ont souvent pu souvent être écrites sous forme de lois mathématiques. Ces lois permettent à leur tour de prédire de nouveaux phénomènes et conduisent à de nouvelles expériences. Elles permettent en outre de concevoir et de mettre au point des applications technologiques.

D'abord relatives à des domaines restreints de phénomènes, ces lois ont été au fil des siècles englobées dans des ensembles cohérents de plus en plus vastes, dans lesquels elles apparaissent comme les conséquences d'un nombre de plus en plus réduit de principes de base. La physique est ainsi le résultat d'un processus essentiellement cumulatif : ses développements s'élaborent sur « les épaules des géants », que sont les physiciens des siècles passés, selon la magnifique expression de Newton.

Il arrive qu'une expérimentation conduise à un résultat en contradiction avec les prédictions faites à partir de ces lois. Ce grain de sable dans l'ensemble des explications physiques admises à un moment donné est l'indice qu'il faut modifier d'une manière ou d'une autre les principes de base, voire en changer totalement. Ce sont alors les prémices de ce qu'on appelle une « révolution scientifique ». Les physiciens doivent alors abandonner les certitudes anciennes et imaginer de nouvelles explications du monde. C'est un exercice intellectuel extraordinairement difficile dont seuls sont capables les plus grands esprits.

Les nouveaux principes mis progressivement en place, au prix d'intuitions géniales et de tâtonnements multiples, doivent alors non seulement rendre compte des nouveaux phénomènes observés, mais aussi englober comme cas particulier les phénomènes expliqués par l'ancienne théorie. Ils permettent aussi de prédire des phénomènes nouveaux, et d'élaborer de nouvelles technologies.

Le début du XX^e siècle a été le théâtre pratiquement simultanément de deux de ces révolutions. On donne le nom de « théorie de la relativité » et de « mécanique quantique » aux deux nouveaux cadres conceptuels mis en place à l'issue de ces révolutions. Ces deux théories ont connu un immense succès, car elles ont permis de rendre compte avec une extrême précision de la quasi-totalité des phénomènes physiques connus. Au moment où nous écrivons ces lignes, certains phénomènes restent inexpliqués et font l'objet de recherches intenses, comme la supraconductivité à haute température ou la vitesse de rotation des galaxies. L'explication de ces phénomènes est cherchée

Introduction à la physique moderne

en général dans le cadre de l'une ou l'autre de ces deux théories. Certains cherchent de nouveaux cadres conceptuels. Mais, à notre connaissance, aucune observation expérimentale indiscutablement admise par la communauté des physiciens n'est actuellement en contradiction avec les prédictions de ces théories. Cette affirmation ne sera peut-être plus vraie demain, ou dans un siècle...

Partie 1

Relativité

ESPACE, TEMPS ET MOUVEMENT EN PHYSIQUE

Les lois régissant le mouvement des corps ont été pendant longtemps basées sur des « évidences » : le temps était « absolu », ou universel : il s'écoulait partout de la même manière. Les distances entre différents points d'un solide avaient aussi un caractère absolu, au point de pouvoir servir d'étalon de définition du mètre. Une force appliquée pendant un temps suffisamment long était susceptible, pensait-on, d'accélérer sans limite un corps matériel. La masse d'un corps était aussi considérée comme une caractéristique absolue.

Jusqu'à la fin du XIX^e siècle, il paraissait impensable qu'il en fût autrement. Nous savons maintenant que ces « évidences » étaient des erreurs, et que le temps, l'espace et le mouvement sont en fait des concepts plus subtils : le « tic-tac » d'une horloge n'est pas le même si on le mesure sur une horloge au repos ou en mouvement par rapport à l'observateur. Il en est de même pour le « mètre étalon », dont la longueur n'a pas la même valeur selon qu'il se déplace ou non par rapport à l'appareil de mesure. Nous savons aussi qu'il existe une vitesse limite, celle de la lumière, pour le mouvement de n'importe quel objet matériel. Enfin, la masse d'un corps peut varier et se transformer en énergie utilisable.

Cette mise à mal de nos certitudes est venue paradoxalement d'interrogations sur une autre partie de la physique, l'électromagnétisme, qui inclut l'étude des phénomènes lumineux : l'article d'Einstein considéré comme fondateur de la théorie de la relativité a pour titre « Sur l'électrodynamique des corps en mouvement ». Il propose comme point de départ une nouvelle « évidence », contraire à l'intuition mais indispensable pour la cohérence d'ensemble des lois expliquant les phénomènes physiques : l'existence d'une quantité absolue d'un autre type, la vitesse de propagation de la lumière dans le vide, qui ne dépend pas de l'état de mouvement de l'objet qui lui donne naissance. L'unité de longueur est maintenant liée à ce nouveau point de vue : elle est en effet définie comme la distance parcourue par la lumière pendant un intervalle de temps bien défini.

Au début du XX^e siècle, ce renversement des évidences est apparu plus conceptuel que pratique, car il impliquait des comportements vraiment nouveaux uniquement pour les objets physiques se déplaçant à une très grande vitesse par rapport à nos

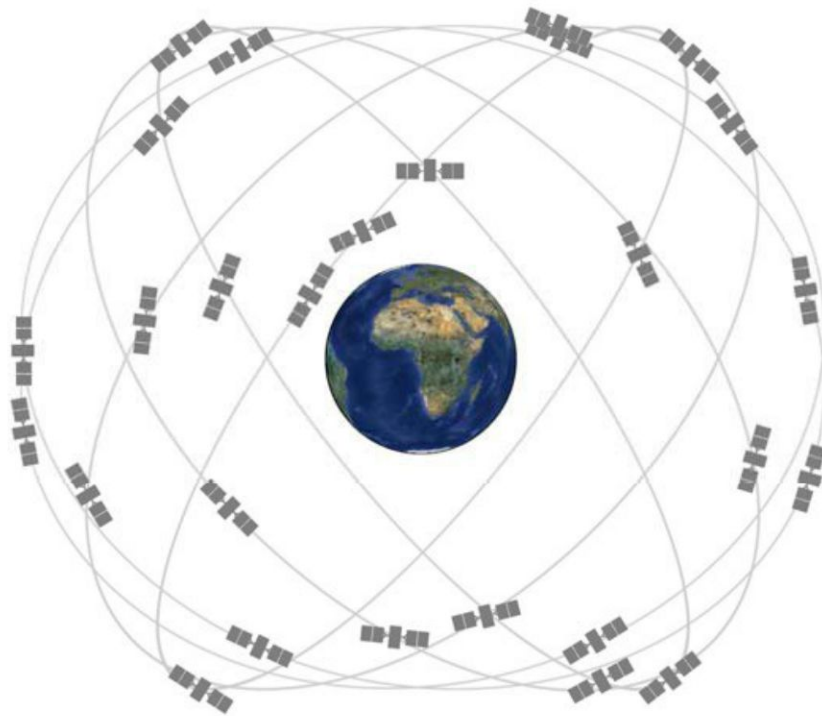


Figure I.1 – Constellation des satellites GPS.

(source libre : <http://www.gps.gov/multimedia/images/>)

appareils de mesure, de l'ordre de $100\,000\text{ km.s}^{-1}$, irréalisable expérimentalement. Ce n'est plus le cas maintenant, au début du XXI^e siècle, et ceci pour deux raisons :

- les progrès des technologies permettent maintenant de réaliser des mesures de distance et de temps d'une extrême précision : les petits écarts qui existent aux faibles vitesses entre les prédictions « classiques » et les prédictions de la théorie de la relativité deviennent accessibles à la mesure. L'exemple le plus frappant de cette situation est le GPS (figure I.1) qui ne permet de mesurer les coordonnées précises d'un point que parce qu'il utilise pour son fonctionnement les équations de la relativité. L'utilisation des « évidences » classiques conduirait à une erreur de positionnement qui dépasserait rapidement plusieurs kilomètres !
- pour étudier la structure intime de la matière, les physiciens ont mis au point des machines, comme les synchrotrons et les anneaux de stockage, qui sont susceptibles de communiquer une vitesse proche de celle de la lumière à des électrons, des protons, des ions, de les faire interagir à très grande énergie et de caractériser avec grande précision les phénomènes qui se produisent (figures I.2 et I.3 associées au *Large Hadron Collider* (LHC), le plus grand accélérateur de particules au monde). La relativité est évidemment le cadre obligé pour la description de ces phénomènes. La première partie de cet ouvrage, consacrée à l'exposé élémentaire des bases de la relativité, sera ainsi l'occasion d'une incursion dans la physique nucléaire, la radioactivité et la physique des particules élémentaires.



Figure I.2 – Vue aérienne du *Large Hadron Collider* situé à la frontière franco-suisse : avec schéma des différents accélérateurs et détecteurs de particules enterrés. Le cercle indique l'emplacement où circulent les particules, qui est enterré à une profondeur de 100 m.

(source : <http://sciencesprings.wordpress.com/2014/07/10/from-bbc-lhc-scientists-to-search-for-fifth-force-of-nature/>)

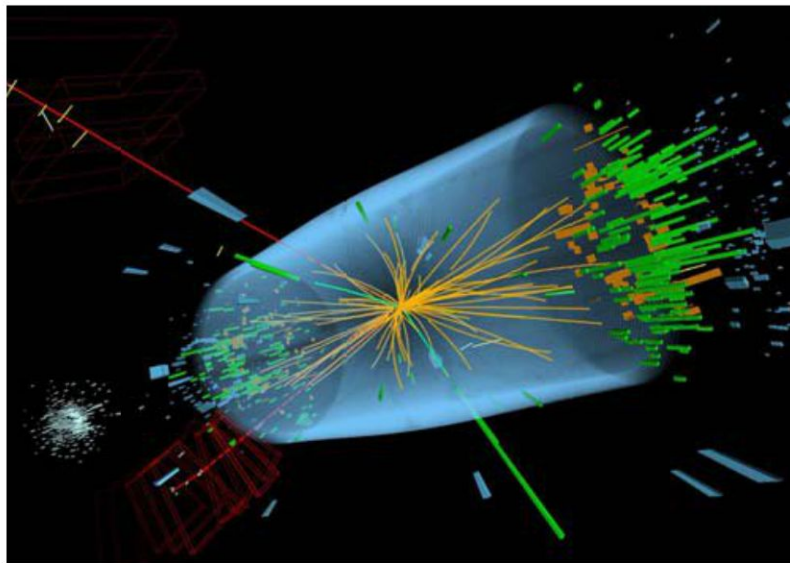


Figure I.3 – Événement typique, obtenu au LHC, sur lequel on a pu mettre en évidence la présence éphémère du mystérieux « boson de Higgs ».

(source : CERN : <https://cdsweb.cern.ch/record/1459462>)

MÉCANIQUE CLASSIQUE ET CHANGEMENT DE RÉFÉRENTIEL GALILÉEN

1

Le but de ce chapitre est de rappeler certains points importants de la mécanique classique concernant les changements de référentiels. Cette problématique est en effet au cœur de la relativité, et l'aspect classique des notions de temps et d'espace est une bonne introduction pour les chapitres suivants.



1.1 La longue élaboration des concepts physiques relatifs au temps, à l'espace et au mouvement

Dès l'Antiquité, les hommes se sont posés des questions sur le temps, l'espace et le mouvement à partir de leur expérience quotidienne. Ils constatent qu'il y a des mouvements extrêmement réguliers dans le ciel, et des mouvements complexes sur la Terre, avec l'omniprésence de la gravité et des frottements. Par exemple, dans le monde méditerranéen, Aristote, au IV^e siècle avant J.-C., jette les bases d'une mécanique qualitative des mouvements terrestres dans un milieu plein et résistant. Il postule l'existence d'un haut et d'un bas absolu, et le fait que tout mouvement nécessite un « moteur » : le mouvement s'arrête lorsque son moteur n'agit plus. Il considère le temps comme donné *a priori*.

Pour les mouvements célestes, Ptolémée, au II^e siècle avant J.-C., postule que la Terre est immobile, et que les corps célestes tournent autour d'elle. Pour rendre compte des observations sur le mouvement complexe des planètes, il met au point le système des épicycles, qui repose sur la composition d'un très grand nombre de mouvements circulaires emboîtés.

Il faut attendre la Renaissance pour que soit contestée cette vision antique du monde. Copernic montre en 1543 que la description du mouvement des planètes est extrêmement simplifiée si on met le Soleil, et non la Terre, au centre de l'Univers.

Cette hypothèse hardie se heurte non seulement au dogme chrétien mais aussi à une objection majeure pour les esprits de l'époque : dans le système de Copernic, la Terre n'est plus immobile mais se déplace à grande vitesse. Ce mouvement, nécessairement « violent » pour la mécanique aristotélicienne, devrait donner naissance à des phénomènes mesurables : par exemple une pierre lancée verticalement devrait partir à grande vitesse dans la direction opposée au mouvement de la Terre.

On voit que les problèmes de choix d'un référentiel permettant d'expliquer le plus simplement possible un phénomène, et de composition des vitesses lorsqu'on change de référentiel, sont depuis toujours au cœur des interrogations sur l'espace, le temps et le mouvement.

Il revient à Galilée d'avoir fait tomber l'objection sur la non-observabilité du mouvement de la Terre dans son ouvrage *Dialogue sur les deux plus grands systèmes du monde*, publié en 1632, dont nous extrayons la citation suivante, extrêmement éclairante : « Enfermez-vous avec un ami dans la cabine principale à l'intérieur d'un grand bateau et prenez avec vous des mouches, des papillons, et d'autres petits animaux volants. Prenez une grande cuve d'eau avec un poisson dedans, suspendez une bouteille qui se vide goutte à goutte dans un grand récipient en dessous d'elle. Avec le bateau à l'arrêt, observez soigneusement comment les petits animaux volent à des vitesses égales vers tous les côtés de la cabine. Le poisson nage indifféremment dans toutes les directions, les gouttes tombent dans le récipient en dessous, et si vous lancez quelque chose à votre ami, vous n'avez pas besoin de le lancer plus fort dans une direction que dans une autre, les distances étant égales, et si vous sautez à pieds joints, vous franchissez des distances égales dans toutes les directions. Lorsque vous aurez observé toutes ces choses soigneusement (bien qu'il n'y ait aucun doute que lorsque le bateau est à l'arrêt, les choses doivent se passer ainsi), faites avancer le bateau à l'allure qui vous plaira, pour autant que la vitesse soit uniforme [c'est-à-dire constante] et ne fluctue pas de part et d'autre. Vous ne verrez pas le moindre changement dans aucun des effets mentionnés et même aucun d'eux ne vous permettra de dire si le bateau est en mouvement ou à l'arrêt... »

Le mouvement, dans la mesure où il est uniforme, ne se distingue donc pas fondamentalement de l'état de repos et ne donne pas lieu à des phénomènes physiques observables. C'est un phénomène qui n'est pas violent, mais qui se perpétue de lui-même, sans besoin de « moteur ».

Le mouvement de la Terre étant donc inobservable, seule subsiste alors l'opposition théologique à l'hypothèse héliocentrique. Celle-ci l'a temporairement emporté, puisque Galilée a été contraint de renoncer en 1633 à son explication. Fort heureusement, les conceptions de Galilée et de Copernic se sont tout de même imposées naturellement comme seule conception rationnelle du monde dans les décennies suivantes, et ont définitivement triomphé avec la publication par Newton, en 1684, de lois mathématiques précises qui s'inscrivent dans le prolongement des conceptions de Galilée et qui permettent de déterminer quantitativement et avec une extrême précision le mouvement des corps terrestres aussi bien que célestes.

1.1 MÉCANIQUE NEWTONNIENNE ET PRINCIPE DE RELATIVITÉ GALILÉENNE

Nous n'allons pas ici détailler les principes de la mécanique classique dont les bases ont été jetées par Newton, mais en rappeler certains aspects qui seront importants dans la suite. Nous nous concentrerons en particulier sur la notion de référentiel, qui

1.1. Mécanique newtonienne et principe de relativité galiléenne

est au cœur des concepts sur l'espace et le temps (encart 1.1). Plus précisément, on appelle référentiel \mathcal{R} la donnée d'un repère orthonormé $Oxyz$ de l'espace physique et d'une horloge H permettant de mesurer l'écoulement du temps dans ce repère.

1.1.1 Référentiels galiléens

Les réflexions de Galilée, exprimées de manière extrêmement imagée dans l'encart 1.1, montrent que les lois de la dynamique sont les mêmes dans différents référentiels (le bateau à l'arrêt ou en mouvement), pourvu que le mouvement relatif entre ces référentiels soit uniforme. Lorsque ce n'est pas le cas, il apparaît des forces nouvelles qui modifient le mouvement des corps. Ainsi, on s'aperçoit que le bateau est en mouvement si son mouvement n'est plus uniforme, par exemple lorsqu'il y a une vague importante qui « agite » le bateau.

On est donc conduit à postuler l'existence d'un ensemble de référentiels, appelés **référentiels galiléens**, pour lesquels est valable le **principe de relativité galiléenne** suivant :

Il existe un ensemble de référentiels, appelés référentiels galiléens, dans lesquels les lois de la dynamique ont la même forme.
Ces référentiels sont en mouvement uniforme les uns par rapport aux autres.

Cela signifie plus précisément que les lois de la dynamique s'écrivent dans deux référentiels galiléens \mathcal{R} et \mathcal{R}' différents avec les mêmes équations, pourvu qu'on écrive ces équations en exprimant les positions des points matériels à l'aide de leurs coordonnées dans le référentiel concerné et en utilisant la même variable temps.

1.1.2 Principe fondamental de la dynamique

Nous l'écrivons ici pour un point matériel M de masse m :

$$m \frac{d\mathbf{v}_M(t)}{dt} = \mathbf{F} \quad (1.1)$$

où \mathbf{F} est la somme vectorielle des forces s'appliquant sur ce point.

Dans cette équation $\mathbf{v}_M(t)$ est la vitesse du point M à l'instant t dans le référentiel galiléen \mathcal{R} dans lequel on se place. Considérons alors un autre référentiel galiléen \mathcal{R}' , animé d'une vitesse constante \mathbf{V} par rapport à \mathcal{R} . Soient O et O' les origines des deux repères associés. On a alors, d'après la relation de Chasles :

$$\overrightarrow{O'M} = \overrightarrow{OM} - \overrightarrow{OO'} \quad (1.2)$$

et donc en dérivant par rapport au temps :

$$\mathbf{v}'_M(t) = \mathbf{v}_M(t) - \mathbf{V} \quad (1.3)$$

La vitesse $\mathbf{v}'_M(t)$ du point dans le nouveau référentiel est la vitesse dans l'ancien référentiel à laquelle on retranche vectoriellement la vitesse du référentiel. On en déduit en dérivant une deuxième fois :

$$m \frac{d\mathbf{v}'_M(t)}{dt} = \mathbf{F} \quad (1.4)$$

Le principe fondamental de la dynamique a donc la même forme dans deux référentiels différents, si l'on postule que l'expression des forces est la même dans les deux référentiels. Les mouvements déduits de cette équation seront donc les mêmes dans des repères galiléens différents. Leur observation ne permettra pas de distinguer un référentiel galiléen d'un autre. En revanche, si on se place dans un référentiel autre qu'un référentiel galiléen, les lois de la dynamique changent : il faut alors ajouter au second membre de la relation précédente d'autres forces, appelées forces d'inertie, comme la force centrifuge ou la force de Coriolis dans le cas d'un mouvement circulaire. Comment savoir si le référentiel dans lequel on se place est galiléen ? Précisons par l'absence d'effets liés à la présence de forces d'inertie. La question se tranche donc au niveau expérimental, et de manière approchée. Par exemple la Terre n'est pas un référentiel galiléen : cela a été démontré expérimentalement par Foucault au XIX^e siècle avec son fameux *pendule de Foucault*. On observe en effet que le plan dans lequel oscille tout pendule pesant tourne lentement autour de la verticale, alors qu'aucun couple vertical n'agit sur le pendule. Cette rotation est la manifestation de la force d'inertie appelée force de Coriolis due au mouvement de rotation de la Terre sur elle-même. En revanche les référentiels liés rigidement à l'ensemble des « étoiles fixes » utilisés en astronomie sont de très bonnes approximations de référentiel galiléen, car on n'y a décelé jusqu'à présent aucune force d'inertie.

Précisons qu'on peut cependant utiliser le référentiel terrestre comme un référentiel galiléen approché pour des mouvements se produisant sur des temps courts devant 24 heures, pour lesquels la force de Coriolis a un effet négligeable.

1.2 CHANGEMENT DE RÉFÉRENTIEL GALILÉEN

1.2.1 Notion d'événement

Un concept fructueux introduit par Einstein est celui d'« événement » : il s'agit d'un phénomène physique bien localisé à la fois dans l'espace et dans le temps. On peut donner comme exemple le choc de deux particules, l'explosion d'une bombe ou le bref flash lumineux émis par une source lumineuse de petite taille. On lui associe

dans un référentiel donné quatre nombres, l'instant t de son apparition et les trois coordonnées d'espace (x, y, z) de sa position M.

On voit sur les exemples donnés qu'un événement E est quelque chose d'intrinsèque, c'est-à-dire qu'on peut le définir par des effets mesurables indépendants du choix de référentiel. Par contre, les quatre coordonnées (t, x, y, z) qui le caractérisent sont définies dans un référentiel donné. Elles ne vont pas être nécessairement les mêmes dans différents référentiels. Il est donc important de connaître les relations qui lient ces quantités dans deux référentiels \mathcal{R} et \mathcal{R}' différents.

1.2.2 Transformations de Galilée

Les transformations de Galilée sont les relations qui lient les coordonnées spatio-temporelles d'un événement donné dans deux référentiels galiléens différents $\mathcal{R} = (Oxyz, H)$ et $\mathcal{R}' = (O'x'y'z', H')$, dans le cadre de la physique classique. Comme nous le verrons plus loin, elles ne sont qu'approximées et valables pour des vitesses faibles par rapport à celle de la lumière. Les transformations valables pour n'importe quelle vitesse relative entre les deux référentiels sont les transformations de Lorentz que nous énoncerons dans le chapitre 3 (Équations 3.1).

Appelons \mathbf{V} la vitesse du référentiel \mathcal{R}' par rapport à \mathcal{R} et supposons qu'à l'instant $t = 0$ les origines O et O' des deux référentiels coïncident. La relation liant les coordonnées spatiales de M dans les deux référentiels est alors, d'après (1.3) :

$$\overrightarrow{O'M} = \overrightarrow{OM} - \mathbf{V}t \quad (1.5)$$

Qu'en est-il du temps dans les deux référentiels ? Celui-ci est mesuré par des phénomènes physiques réguliers dont beaucoup sont de caractère mécanique : mouvement de la Terre, balancier de l'horloge (pendule simple), mouvement harmonique d'une masse tenue par un ressort. Le principe fondamental de la dynamique implique que ces mouvements sont identiques dans tous les référentiels galiléens lorsque les forces qui leur donnent naissance sont les mêmes. Il est donc légitime de postuler que *le temps s'écoule de la même manière dans tous les référentiels galiléens*. De plus, si on choisit de plus la même unité de temps et la même origine des temps dans les deux référentiels, on peut postuler l'identité des coordonnées de temps ($t' = t$) d'un événement donné dans tous les référentiels galiléens. Ici encore, cette identité des temps n'est valable que pour de faibles vitesses relatives entre les deux référentiels. Comme nous le verrons plus loin, elle cesse d'être exacte lorsque V s'approche de la vitesse de la lumière.

Pour simplifier les expressions futures, il est d'usage de se placer dans la configuration géométrique suivante (figure 1.1) : les axes Ox et $O'x'$ des deux référentiels coïncident, et les deux autres restent à tout instant parallèles. Les relations reliant les

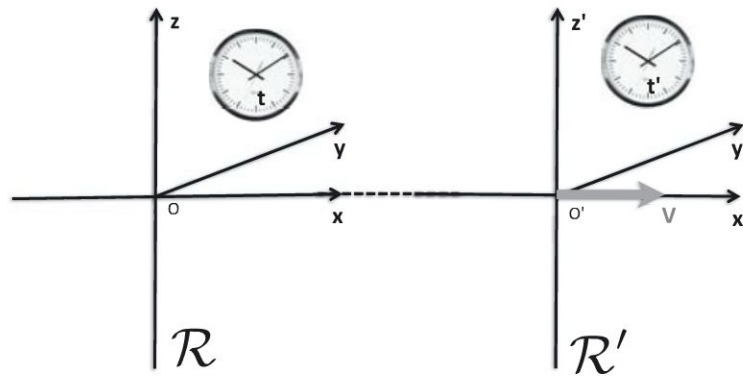


Figure 1.1 – Référentiels galiléens en translation relative à vitesse uniforme V le long de leurs axes Ox .

quatre coordonnées d'un même événement E dans les référentiels \mathcal{R} et \mathcal{R}' s'écrivent alors simplement :

$$\begin{aligned} t' &= t \\ x' &= x - Vt \\ y' &= y \\ z' &= z \end{aligned} \quad (1.6)$$

Dans la suite, nous appellerons « transformations de Galilée » l'ensemble des quatre relations non-relativistes (1.6).

On peut considérer ces transformations sous deux angles, et nous ne nous priverons pas de le faire dans la suite : on peut considérer le référentiel \mathcal{R} comme étant celui de l'observateur, et \mathcal{R}' sera alors celui d'objets physiques en mouvement, ou bien on peut considérer \mathcal{R} comme le référentiel du laboratoire, où se déroulent les phénomènes, et \mathcal{R}' sera alors celui d'un observateur en mouvement.

1.2.3 Loi classique de composition des vitesses

Si le point M se déplace au cours du temps, il a une vitesse \mathbf{v}_M dans le référentiel \mathcal{R} et une vitesse \mathbf{v}'_M dans le référentiel \mathcal{R}' . En dérivant la relation (1.5) par rapport au temps (qui est le même dans les deux référentiels), on obtient la *loi classique de composition des vitesses* :

$$\mathbf{v}'_M = \mathbf{v}_M - \mathbf{V} \quad (1.7)$$

que l'on peut aussi écrire :

$$\mathbf{v}_M = \mathbf{v}'_M + \mathbf{V} \quad (1.8)$$

Cette relation correspond à l'expérience quotidienne. Quand on marche dans un train, la vitesse par rapport au sol est la somme de la vitesse dans le train et de celle du train par rapport au sol : les vitesses du référentiel et du corps mobile s'additionnent. Nous allons voir que cette expression simple et intuitive n'est en réalité vraie que pour les vitesses petites devant celles de la lumière.

LUMIÈRE CLASSIQUE ET CHANGEMENT DE RÉFÉRENTIEL GALILÉEN

2

La lumière joue un rôle essentiel en relativité. Il convient donc d'abord d'en rappeler les propriétés classiques essentielles. Les rappels de ce chapitre serviront aussi d'introduction à l'étude de ses propriétés quantiques, que nous aborderons dans le chapitre 10.



2.1 La longue élaboration des concepts physiques relatifs à la lumière

La vision constitue le principal moyen de contact entre l'homme et le milieu qui l'entoure, mais les phénomènes lumineux courants, comme la formation des images dans la vision ou l'arc-en-ciel, sont en fait extrêmement complexes. Ainsi, il a fallu beaucoup de temps pour dégager des idées simples sur la lumière et des concepts qui soient susceptibles de modélisation. La notion de rayon lumineux s'est dégagée dès l'Antiquité dans les mondes grec (Euclide) et chinois (Mozi), ainsi que les lois de la réflexion de ces rayons sur les miroirs. La compréhension de la vision a posé beaucoup de problèmes, et a fait l'objet de nombreuses supputations. L'opinion dominante, sous l'influence de Pythagore et Euclide, a longtemps été que le rayon visuel se propageait de l'œil vers les objets regardés (qui subsiste encore dans les expressions comme « jeter un coup d'œil » ou « porter son regard sur »). Il a fallu attendre Al Hazen et l'an mille pour que soit admise l'idée que les rayons sont émis des objets lumineux vers l'œil et forment une image à l'intérieur de celui-ci. Pour Al Hazen, la lumière était comme un « corps au bout d'une flèche ».

Les propriétés grossissantes des lentilles ont été découvertes au Moyen-Age, ce qui a permis à des artisans ingénieux d'élaborer empiriquement les premiers instruments d'optique : lunettes de vue, puis au XVI^e siècle lunettes d'approche, ou télescopes. Ces dernières ont été améliorées et utilisées par Galilée et enfin modélisées par Kepler. Au XVII^e siècle, avec la découverte par Snell, puis Descartes, des lois exactes de la réfraction, se met en place ce qu'on appelle maintenant l'optique géométrique, qui permet de comprendre des phénomènes comme l'arc-en-ciel et d'élaborer des instruments d'optique toujours plus perfectionnés. D'autres phénomènes lumineux sont découverts : les interférences, la diffraction et la double réfraction.

Si les phénomènes lumineux sont de mieux en mieux étudiés et caractérisés, la question de la nature de la lumière reste ouverte et fait l'objet de nombreuses spéculations. Huygens, puis au XVIII^e siècle Euler, la tiennent pour l'ébranlement

d'un milieu continu qui se propage de proche en proche. Newton pose la question suivante : « Les rayons lumineux, ne sont-ce pas de fort petits corpuscules élançés ou poussés hors des corps lumineux ? » Le prestige scientifique de Newton fait que cette conception de particules de lumière se propageant dans le vide, avancée comme une simple hypothèse par Newton, prévaut pendant une longue période. Au début du XIX^e siècle Young explique les interférences en s'appuyant sur les hypothèses suivantes : « I) Un éther luminifère, ténu et extrêmement élastique, infiltre l'Univers. II) Des ondulations sont excitées dans cet éther chaque fois qu'un corps devient lumineux. » Quelques années plus tard Fresnel montre que l'hypothèse que la lumière monochromatique est une onde oscillant à une fréquence bien déterminée permet d'expliquer quantitativement aussi bien l'optique géométrique que les phénomènes d'interférences et de diffraction. La conception corpusculaire de la lumière est alors abandonnée.

Les études menées à la même époque par de nombreux physiciens concernant les phénomènes électriques et magnétiques permettent de dégager progressivement une notion extrêmement importante et révolutionnaire : celle de champ, objet physique se manifestant en tout point de l'espace même vide. Au prix d'un extraordinaire effort conceptuel, Maxwell réussit à écrire les équations régissant l'évolution des champs électriques et magnétiques qui rendent compte de l'ensemble des phénomènes connus. Il découvre que ses équations ont des solutions qui ont toutes les propriétés des hypothétiques ondes lumineuses de Young et Fresnel. Selon ses propres termes : « Nous pouvons difficilement éviter la conclusion que la lumière est constituée des ondulations transverses du même milieu qui est la cause des phénomènes électriques et magnétiques ». La conception qui prévaut ainsi à la fin du XIX^e siècle est que la lumière est une onde des champs électriques et magnétiques régie par les équations de Maxwell.

2.1 DESCRIPTION ONDULATOIRE DE LA LUMIÈRE

2.1.1 Caractéristiques de l'onde lumineuse

Considérons un rayon de lumière monochromatique se propageant dans le vide. Dans l'approche ondulatoire, il est décrit en tout point M d'un référentiel \mathcal{R} par une quantité oscillante, que nous appellerons A_r , et qui s'écrit donc en tout point \mathbf{r} et à tout instant t :

$$A_r(\mathbf{r}, t) = A_0 \cos(\omega t - \mathbf{k} \cdot \mathbf{r}) \quad (2.1)$$

où \mathbf{k} est le vecteur d'onde et ω la pulsation de l'onde considérée, qui sont liés à la longueur d'onde λ , à la période T et à la fréquence ν de l'onde par :

$$\lambda = \frac{2\pi}{k} \quad T = \frac{2\pi}{\omega} \quad \nu = \frac{\omega}{2\pi} \quad (2.2)$$

où $k=|\mathbf{k}|$. Pulsation et vecteur d'onde sont liés par la relation, dite de dispersion :

$$\omega = ck \quad (2.3)$$

2.1. Description ondulatoire de la lumière

où c est une constante, homogène à une vitesse, qui vaut précisément :

$$c = 299\,792\,458 \text{ m.s}^{-1} \simeq 3 \times 10^8 \text{ m.s}^{-1}$$

Pour simplifier les calculs trigonométriques, nous utiliserons une description de l'onde par un nombre complexe, dont la partie réelle est l'onde lumineuse :

$$A_r(\mathbf{r}, t) = \text{Re}(A(\mathbf{r}, t)) \quad ; \quad A(\mathbf{r}, t) = A_0 \exp(-i(\omega t - \mathbf{k} \cdot \mathbf{r})) \quad (2.4)$$

Lorsque le temps s'écoule, l'onde se propage dans la direction du vecteur \mathbf{k} . La vitesse de propagation de la phase de cette onde (par exemple la vitesse de propagation de ses maxima) est donnée par :

$$v_\phi = \frac{\omega}{k} = c \quad (2.5)$$

Pour une onde lumineuse monochromatique jaune, par exemple, λ est proche de $0,6 \mu\text{m} = 6 \times 10^{-7} \text{ m}$, T de $2 \text{ fs} = 2 \times 10^{-15} \text{ s}$ et ν de $500 \text{ THz} = 5 \times 10^{14} \text{ Hz}$. Il s'agit donc d'une onde se propageant à une vitesse extrêmement grande, largement supérieure aux vitesses habituelles des objets naturels terrestres, et d'une oscillation extrêmement rapide.

Les effets observables liés à cette onde sont proportionnels à son *intensité* I , égale au carré du module de l'amplitude complexe de l'onde :

$$I(\mathbf{r}, t) = |A(\mathbf{r}, t)|^2 \quad (2.6)$$

En particulier, l'énergie transportée par unité de temps et de surface par la lumière est proportionnelle à cette quantité.

Quelle est la signification physique de la quantité $A(\mathbf{r}, t)$ figurant dans (2.1) et (2.6) ? Dans l'approche de Young et Fresnel, c'est la valeur instantanée de l'oscillation de l'éther luminifère, sans qu'il soit besoin de préciser plus la nature de cette oscillation. Nous verrons dans la partie quantique de cet ouvrage que Schrödinger a introduit de manière analogue une « fonction d'onde » $\psi(\mathbf{r}, t)$ pour décrire une particule quantique sans bien savoir exactement ce que représentait cette fonction. Cela n'a empêché ni Fresnel ni Schrödinger de manipuler ces ondes hypothétiques et d'en déduire des prédictions conformes à l'expérience. Maxwell a montré qu'en fait $A(\mathbf{r}, t)$ était de nature électromagnétique. Plus précisément, l'oscillation lumineuse peut être assimilée au champ électrique de l'onde électromagnétique.

2.1.2 Équation d'onde, principe de superposition

Il est facile de voir que toute onde $A(x, y, z, t)$ de la forme (2.1), quelles que soient son amplitude, sa pulsation et sa direction de propagation obéit à l'équation aux dérivées partielles suivante :

$$\frac{\partial^2 A}{\partial x^2} + \frac{\partial^2 A}{\partial y^2} + \frac{\partial^2 A}{\partial z^2} - \frac{1}{c^2} \frac{\partial^2 A}{\partial t^2} = 0 \quad (2.7)$$

Il s'agit d'une équation *linéaire en A*. Il en résulte une propriété extrêmement importante des ondes, appelée *principe de superposition*, et que nous retrouverons pour la fonction d'onde quantique :

Toute combinaison linéaire de deux ondes est encore une onde possible pour le système, c'est-à-dire une solution de l'équation d'onde.

Considérons alors la somme de deux ondes de même amplitude se propageant dans la même direction Oz, mais présentant une différence de phase ϕ . Elle a pour expression :

$$\begin{aligned} A(\mathbf{r}, t) &= A_0 \exp(-i(\omega t - kz)) + A_0 \exp(-i(\omega t - kz + \phi)) \\ &= 2A_0 \exp\left(-i\left(\omega t - kz + \frac{\phi}{2}\right)\right) \cos \frac{\phi}{2} \end{aligned} \quad (2.8)$$

et son intensité vaut :

$$I(\mathbf{r}) = |A(\mathbf{r}, t)|^2 = 4A_0^2 \cos^2 \frac{\phi}{2} \quad (2.9)$$

L'intensité de l'onde somme est donc maximale si ϕ est nul ou multiple de 2π , c'est-à-dire lorsque les deux ondes sont en phase, et nulle si $\phi = \pi$ à 2π près, c'est-à-dire lorsque les deux ondes sont en opposition de phase. Dans cette dernière configuration les deux ondes sont strictement opposées en chaque point et à tout instant, on est amené à conclure que « *de la lumière ajoutée à de la lumière donne de l'obscurité* », ce qui est conforme aux observations expérimentales dans les phénomènes d'interférence, mais n'est évidemment pas concevable dans le cadre d'une théorie classique corpusculaire de la lumière. Lorsqu'on fait varier la valeur du déphasage ϕ , on observe une variation sinusoïdale de l'intensité mesurée, qu'on appelle franges d'interférence.

2.2 LUMIÈRE ET CHANGEMENT DE RÉFÉRENTIEL

2.2.1 Changement de référentiel pour une onde plane : effet Doppler non relativiste

Comme dans le chapitre précédent sur la dynamique, nous voulons savoir comment un phénomène lumineux donné apparaît quand on l'observe dans des référentiels différents. Du point de vue ondulatoire classique, on considère l'onde lumineuse comme la vibration d'un milieu physique, l'« *éther luminifère* ». La valeur locale et instantanée de l'onde lumineuse est alors définie dans le référentiel privilégié \mathcal{R}_0 dans lequel ce milieu est au repos. On peut supposer que ce référentiel est galiléen.

Considérons alors la même onde dans un référentiel galiléen \mathcal{R}' en mouvement uniforme de vitesse V dans la direction Ox par rapport au référentiel \mathcal{R}_0 de l'éther

2.2. Lumière et changement de référentiel

luminifère. Elle va y être caractérisée par une pulsation ω' et un vecteur d'onde \mathbf{k}' *a priori* différents. Comme il s'agit de la même oscillation, on a :

$$A_0 \exp(-i(\omega t - \mathbf{k} \cdot \mathbf{r})) = A'_0 \exp((-i\omega' t' - \mathbf{k}' \cdot \mathbf{r}')) \quad (2.10)$$

Appelons k'_x, k'_y, k'_z les coordonnées du vecteur \mathbf{k}' dans le référentiel \mathcal{R}' et k_x, k_y, k_z les coordonnées de \mathbf{k} dans le référentiel \mathcal{R}_0 et utilisons la transformation de Galilée (1.6). On a alors :

$$A_0 \exp(-i(\omega t - k_x x - k_y y - k_z z)) = A'_0 \exp(-i(\omega' t - k'_x (x - Vt) - k'_y y - k'_z z)) \quad (2.11)$$

Cette relation devant être vérifiée pour tout point et tout temps, on en déduit les transformations de Galilée pour les caractéristiques d'une onde plane :

$$A'_0 = A_0 \quad \omega' = \omega - k_x V \quad k'_x = k_x \quad k'_y = k_y \quad k'_z = k_z \quad (2.12)$$

Si on introduit le vecteur vitesse \mathbf{V} , aligné selon Ox , du référentiel \mathcal{R}' par rapport à \mathcal{R}_0 , on a les relations vectorielles suivantes :

$$\omega' = \omega - \mathbf{k} \cdot \mathbf{V} \quad ; \quad \mathbf{k}' = \mathbf{k} \quad (2.13)$$

Il y a donc changement de pulsation, ou de fréquence, pour l'onde vue dans les deux référentiels. C'est le phénomène d'*effet Doppler* non relativiste, bien connu pour les ondes sonores : le son perçu par un observateur mobile n'a pas la même hauteur quand celui-ci se rapproche de l'émetteur sonore ou s'en éloigne.

Il n'y a en revanche pas de changement de vecteur d'onde : la longueur d'onde et les plans d'onde, c'est-à-dire les plans de phase constante, restent les mêmes dans les deux référentiels. Mentionnons que les deux propriétés précédentes impliquent que la vitesse de phase de l'onde, donnée en valeur absolue par ω/k n'est pas la même dans les deux référentiels. Nous verrons au chapitre suivant que cette conséquence des transformations de Galilée n'est pas conforme à l'expérience.

2.2.2 Changement de référentiel pour un faisceau lumineux limité : aberration classique

Le paragraphe précédent a traité le cas de l'onde plane, qui est un cas simple, mais éloigné de la réalité, car il entraîne que l'onde a la même intensité en tout point de l'espace, et donc que son énergie totale, intégrée sur tout l'espace, est infinie. Plus proche de la réalité est une onde dont l'intensité est limitée dans le temps et dans l'espace. Considérons le « paquet d'ondes » suivant :

$$A(x, y, z, t) = A_0 \exp(-i(\omega t - kz)) f_0(x, y, t - z/c) \quad (2.14)$$

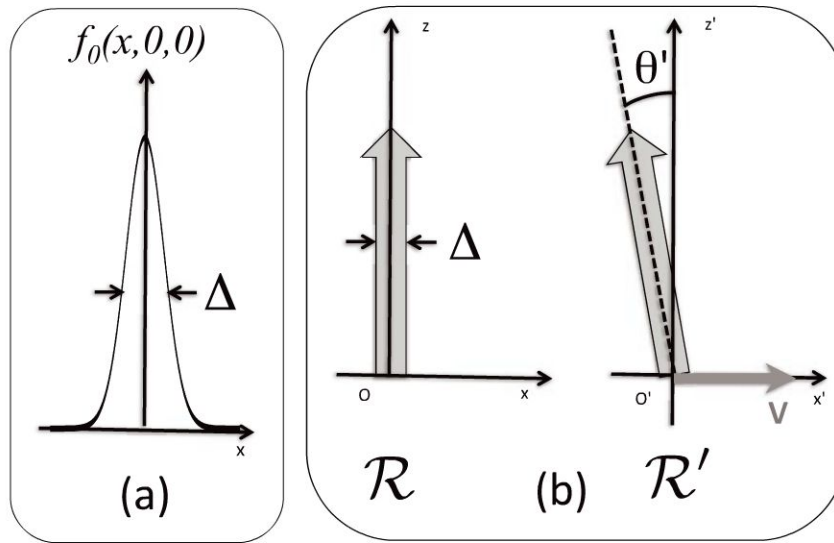


Figure 2.1 – (a) Forme spatiale du « paquet d’ondes lumineuses » ; (b) Propagation de ce paquet d’ondes localisé le long de l’axe Ox vue dans les référentiels \mathcal{R} et \mathcal{R}' . La largeur de la flèche, égale à Δ , donne l’extension latérale du faisceau lumineux.

où f_0 est une fonction en forme de « bosse » à la fois spatiale et temporelle, prenant sa valeur maximale pour la valeur 0 de toutes ses variables, de largeur spatiale Δ et de largeur temporelle Δt (figure 2.1). Il s’agit d’un paquet d’onde localisé spatialement le long de l’axe Oz ayant la forme d’une brève impulsion temporelle se propageant le long de cet axe à la vitesse c . On montre que c’est aussi une solution de l’équation d’onde (2.7), pourvu que son extension spatiale en x et y soit suffisamment grande par rapport à la longueur d’onde pour qu’on puisse négliger les phénomènes de diffraction. L’extension temporelle peut être en revanche quelconque.

Vu dans le référentiel \mathcal{R}' , en translation uniforme selon Ox par rapport à \mathcal{R} , le même paquet d’onde a pour expression, en utilisant (1.6) :

$$A(x', y', z', t') = A_0 \exp(-i(\omega t' - kz')) f_0(x' + Vt', y', t' - z'/c) \quad (2.15)$$

Il s’agit d’un paquet d’ondes centré maintenant en $x' = -Vt', y' = 0, t' = z'/c$. En éliminant le temps entre ces relations, on voit que le maximum d’intensité du paquet se déplace sur la droite d’équation $x' = -(V/c)z', y' = 0$, et non plus $x = 0, y = 0$. Il en résulte que *la direction de propagation du faisceau lumineux change quand l’observateur se déplace par rapport à l’onde*. Il s’agit du phénomène d’*aberration classique*. Plus précisément, l’angle d’aberration par rapport à la direction de propagation initiale Oz est donné par $\tan \theta' = -(V/c)$. Précisons que le terme oscillant est toujours de la forme $\cos(\omega t' - kz')$, c’est-à-dire que les surfaces de phase constante sont des plans perpendiculaires à $O'z'$, donc non orthogonaux à la direction de propagation de l’impulsion.

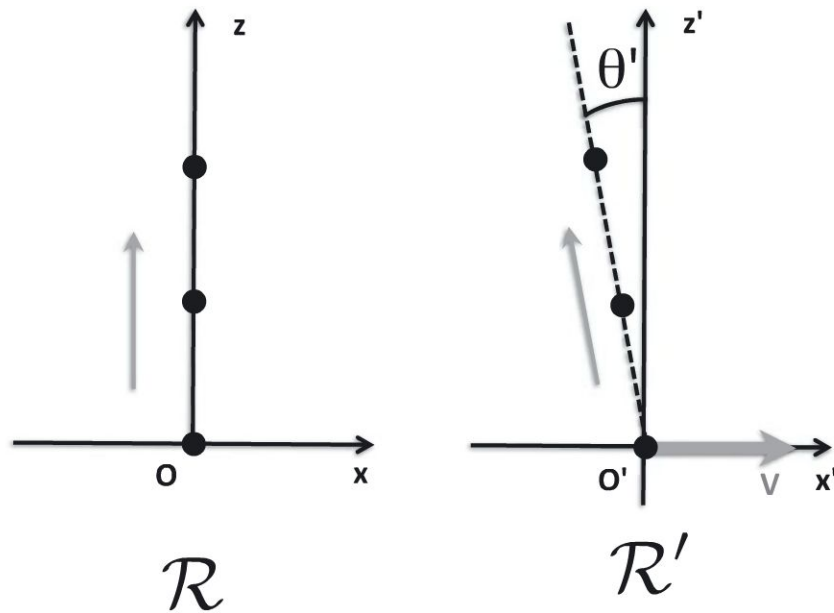


Figure 2.2 – Propagation d'une particule se déplaçant le long de Ox vue dans les référentiels \mathcal{R} et \mathcal{R}' .

Ce phénomène a été découvert en 1725 par l'astronome Bradley, qui avait constaté que les étoiles qu'il observait dans son télescope n'avaient pas une position fixe sur le long terme, mais décrivaient en un an un petit cercle sur la voûte céleste. Ce « mouvement apparent » est une manifestation de l'aberration. Il est dû au fait que la Terre est en mouvement autour du Soleil, et donc que l'angle d'aberration de la lumière venant des étoiles change au cours de l'année.

2.2.3 Changement de référentiel pour une particule classique : loi classique de composition des vitesses

Plaçons-nous un moment dans l'hypothèse de Newton pour la lumière, formée de corpuscules se propageant avec la vitesse vectorielle \mathbf{c} dans le référentiel \mathcal{R} . Considérons le même cas que dans le paragraphe précédent d'un faisceau se propageant le long de l'axe Oz (figure 2.2). La position \mathbf{r} d'un de ces corpuscules dans le référentiel \mathcal{R} est alors donnée par $(x = 0; y = 0, z = ct)$, et dans le référentiel \mathcal{R}' par $(x' = -Vt'; y' = 0; z' = ct')$.

La vitesse du corpuscule dans ce référentiel a donc pour coordonnées : $(c'_x = V; c'_y = 0; c'_z = c)$. On peut finalement écrire pour les vitesses vectorielles \mathbf{c} et \mathbf{c}' des corpuscules de lumière dans les deux référentiels :

$$\mathbf{c}' = \mathbf{c} - \mathbf{V} \quad (2.16)$$

C'est la loi vectorielle classique de composition des vitesses, qui conduit, pour trouver la vitesse dans le nouveau référentiel, à retrancher la vitesse du référentiel à celle

du mobile dans le référentiel initial. Remarquons que la direction de propagation de la lumière dans le nouveau référentiel est la même que dans l'approche ondulatoire du paragraphe précédent : on retrouve le phénomène d'aberration, mais expliqué différemment. Si on remplace les corpuscules de lumière par des gouttes de pluie, il n'y a rien d'étonnant à ce que la direction apparente de la pluie ne soit pas la même si on est immobile ou si on est en mouvement par rapport à celle-ci.

Comme dans le cas de l'approche ondulatoire, la question de l'existence d'un référentiel privilégié se pose : les corpuscules de lumière vont-ils à la vitesse c uniquement dans un référentiel galiléen unique \mathcal{R}_0 , comme c'est le cas d'une particule se déplaçant dans un milieu visqueux, ou vont-ils à la vitesse c dans le référentiel lié la source lumineuse ? Dans cette deuxième éventualité, la vitesse de la lumière émise par un corps en mouvement devrait être la somme de la vitesse de la lumière et de la vitesse du mobile, comme pour une balle lancée par une personne en mouvement. Dans le cas d'un système d'étoiles doubles en rotation l'une autour de l'autre, la vitesse de la lumière émise par les deux étoiles ayant des vitesses différentes ne devrait pas être la même et la lumière issue de l'étoile se rapprochant de nous pourrait alors arriver plus vite que celle issue de l'étoile s'éloignant de nous, ce qui n'est pas conforme aux observations astronomiques, qui semblent donc impliquer l'existence de ce référentiel privilégié \mathcal{R}_0 .

2.2.4 Changement de référentiel pour des champs électrique et magnétique

Nous savons depuis Maxwell que la lumière est une onde électromagnétique, et il est donc important de savoir comment les champs \mathbf{E} et \mathbf{B} se comportent dans un changement de référentiel. L'invariance galiléenne requiert que les effets observables de ces champs, c'est-à-dire les forces électromagnétiques, soient les mêmes dans tout référentiel galiléen. La force en question pour une charge ponctuelle q de vitesse \mathbf{v} étant la force de Lorentz $q(\mathbf{E} + \mathbf{v} \times \mathbf{B})$, on a alors :

$$q(\mathbf{E}' + \mathbf{v}' \times \mathbf{B}') = q(\mathbf{E} + \mathbf{v} \times \mathbf{B}) = q(\mathbf{E} + (\mathbf{v}' + \mathbf{V}) \times \mathbf{B}) \quad (2.17)$$

en utilisant la loi (2.16) $\mathbf{v}' = \mathbf{v} - \mathbf{V}$ de composition des vitesses pour un corpuscule. On en déduit la transformation de Galilée des champs :

$$\mathbf{B}' = \mathbf{B} \quad ; \quad \mathbf{E}' = \mathbf{E} + \mathbf{V} \times \mathbf{B} \quad (2.18)$$

Y a-t-il invariance galiléenne des équations de Maxwell, en tenant compte de la transformation (2.18) des champs ? L'encart 2.2 montre sur un exemple particulier que ce n'est pas le cas. Les phénomènes électromagnétiques peuvent donc être différents dans deux repères galiléens différents. On est par conséquent amené à conclure,

si l'on tient à conserver l'expression galiléenne (1.6) de la formule de changement de référentiel, qu'il existe un référentiel privilégié \mathcal{R}_0 dans lequel sont valables les équations de Maxwell.



2.2 Un dispositif électromagnétique où la force dépend du référentiel

Considérons le dispositif de la figure 2.3 : il se compose d'une charge ponctuelle Q de charge électrique Q à une distance d d'un fil F porteur de charges statiques uniformément réparties. On appelle α la charge par unité de longueur portée par ce fil.

- a) Calculons d'abord la force F exercée sur Q dans le référentiel \mathcal{R} , supposé galiléen, où charges et fils sont au repos. Il s'agit d'un problème d'électrostatique classique. La force est radiale et son module a pour expression :

$$F = \frac{\alpha Q}{2\pi\epsilon_0 d} \quad (2.19)$$

- b) Considérons maintenant le même dispositif dans le référentiel galiléen \mathcal{R}' animé d'une vitesse constante V parallèle au fil. Les charges liées au fil F sont maintenant en mouvement, ce qui donne naissance à un courant $I = -\alpha V$, donc à un champ magnétique tangentiel dont le module vaut, au niveau de Q :

$$B = \frac{\mu_0 I}{2\pi d} \quad (2.20)$$

La charge Q est animée dans le nouveau référentiel d'une vitesse $-V$ et elle est donc sensible au champ magnétique. La force de Lorentz totale qui s'exerce sur Q est la somme d'une force électrostatique dont l'expression est la même dans les deux référentiels, et d'une force magnétique $Q\mathbf{V} \times \mathbf{B}$ elle aussi radiale. Son module vaut donc :

$$F' = \frac{\alpha Q}{2\pi\epsilon_0 d} - QVB = F(1 - \epsilon_0\mu_0 V^2) \quad (2.21)$$

La force n'a pas la même valeur dans les deux référentiels, donc les mouvements provoqués par cette force ne seront pas les mêmes : les lois des phénomènes électriques et magnétiques ne sont pas les mêmes dans deux référentiels galiléens, ce qui est contraire au principe de relativité galiléenne.

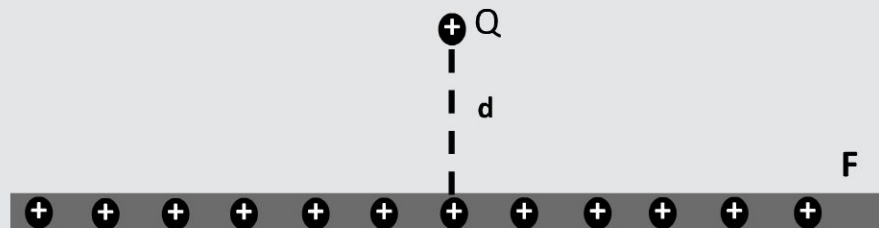


Figure 2.3 – Charge ponctuelle Q placée à une distance d d'un fil uniformément chargé.

La théorie de Maxwell montre que la vitesse de propagation c des ondes électromagnétiques (et donc de la lumière) est donnée par $c = (\epsilon_0 \mu_0)^{-1/2}$. La variation relative de la force que nous venons de trouver est donc en $(1 - V^2/c^2)$, quantité très proche de 1 vu la valeur extrêmement élevée de c , mais non nulle et donc en principe mesurable.

2.3 MESURES DE LA VITESSE DE LA LUMIÈRE

Il semble donc qu'il existe une différence majeure entre les phénomènes dynamiques, qui obéissent au principe de relativité galiléenne, et les phénomènes lumineux et plus généralement électromagnétiques, qui semblent impliquer l'existence d'un référentiel absolu \mathcal{R}_0 qui est en quelque sorte le siège des oscillations des champs électriques et magnétiques. Les ondes lumineuses se comportent apparemment comme les ondes sonores : ces dernières consistant en des vibrations d'un milieu matériel, gazeux ou solide, le référentiel \mathcal{R}_0 dans lequel ce milieu est au repos joue un rôle privilégié. Si on se place dans un référentiel galiléen \mathcal{R} autre que \mathcal{R}_0 , la vitesse de l'onde mesurée dans ce référentiel ne sera plus la même que dans \mathcal{R}_0 , et surtout *elle ne sera pas isotrope* et dépendra des orientations relatives du vecteur d'onde \mathbf{k} de l'onde et de la vitesse \mathbf{V} du référentiel \mathcal{R} par rapport à \mathcal{R}_0 .

Comme toujours en physique, c'est à l'expérience de trancher : la mesure de la vitesse de propagation c de la lumière, et surtout l'évaluation de son éventuelle anisotropie, sont donc essentielles dans ce contexte.

2.3.1 Les premières mesures de la vitesse de la lumière

Mesurer la vitesse de la lumière n'est pas simple en raison de l'extrême rapidité de cette propagation. La première tentative connue en Occident a été réalisée par Galilée vers 1600, qui disposa deux opérateurs munis de lanternes sur deux collines éloignées. Le premier dévoile sa lanterne à un instant t_1 , le second dévoile la sienne dès qu'il commence à voir la lumière de la première lanterne. Le premier note alors le temps t_2 auquel il commence à voir la lumière de la deuxième lanterne. Ce type d'expérience réalisée avec des sources sonores permet de mesurer aisément la vitesse de propagation du son dans l'air. En ce qui concerne la lumière, l'absence de retard $t_2 - t_1$ mesurable constaté par Galilée lui permettait de conclure que la lumière avait une propagation « extrêmement rapide ».

C'est à l'astronome O. Römer que revient le mérite d'avoir fait, en 1676, la première évaluation quantitative de cette vitesse. Les astronomes de l'Observatoire de Paris avaient observé sur de très longues périodes la position du satellite de Jupiter appelé Io, et en particulier, grâce aux horloges perfectionnées dont ils disposaient,

2.3. Mesures de la vitesse de la lumière

noté avec précision les instants t_n d'occultation de Io par Jupiter, séparés d'environ 42 heures. Il s'avéra que ces instants n'étaient pas strictement périodiques comme le prédit la mécanique classique : on observait plus précisément des instants t_n donnés par :

$$t_n = nT + \delta_n \quad (2.22)$$

avec $T = 42,34$ heures et δ_n oscillant autour de zéro avec une amplitude de ± 8 minutes sur une année terrestre. Römer attribua cette variation à la vitesse finie de propagation de la lumière depuis Jupiter jusqu'à l'observateur terrestre (figure 2.4). Appelons R_J le rayon de l'orbite de Jupiter et R_T le rayon de l'orbite terrestre. Au moment de l'année où la Terre est au plus près de Jupiter (point T_1), l'événement (au sens du paragraphe 1.2.1) « $n^{\text{ième}}$ occultation de Io par Jupiter » se produit à la distance $R_J - R_T$ de la terre. Six mois plus tard (point T_2), l'événement « $n'^{\text{ième}}$ occultation de Io par Jupiter » se produit à la distance $R_J + R_T$ (Jupiter ne bouge pratiquement pas sur une année terrestre, sa période de révolution étant de 11,9 ans). Si T est la période de révolution de Io autour de Jupiter, on a donc pour les instants d'occultation mesurés par un observateur terrestre :

$$t_n = nT + \frac{R_J - R_T}{c} \quad ; \quad t'_n = n'T + \frac{R_J + R_T}{c} \quad (2.23)$$

où $n' - n$ est de l'ordre de 100. L'écart maximal par rapport à l'exacte périodicité, c'est-à-dire les 16 minutes mesurées par les astronomes, est donc de $2R_T/c$, temps de propagation de la lumière sur le diamètre terrestre. Cette observation permet donc de déterminer la vitesse c de la lumière si l'on connaît le rayon de l'orbite terrestre. Römer en déduisit pour la vitesse de la lumière : $c \approx 215\,000 \text{ km.s}^{-1}$. En 1694, Halley, disposant d'une mesure plus précise de $R_T \approx 1,5 \times 10^8 \text{ m}$, trouva $c \approx 295\,000 \text{ km.s}^{-1}$, à quelques % de la valeur exacte.

L'amélioration des techniques expérimentales au XIX^e siècle permit aux physiciens A. Fizeau et L. Foucault vers 1850 d'effectuer des mesures de cette vitesse sur des distances de propagation beaucoup plus courtes (de l'ordre de quelques kilomètres), de confirmer et d'affiner cette valeur avec une précision de l'ordre du %.

2.3.2 Quelques ordres de grandeur

La lumière se propage donc à une vitesse proche de $300\,000 \text{ km.s}^{-1}$ qui est largement supérieure à celle de tous les mobiles terrestres : un sprinter a une vitesse de $0,01 \text{ km.s}^{-1}$, un TGV $0,1 \text{ km.s}^{-1}$, un avion de chasse 1 km.s^{-1} , un satellite artificiel 10 km.s^{-1} environ. Les objets astronomiques se déplacent aussi très lentement par rapport à la lumière : 30 km.s^{-1} pour la Terre autour du soleil, 250 km.s^{-1} pour le Soleil autour du centre galactique, soit un rapport v/c n'excédant pas 10^{-3} . Seuls des objets microscopiques comme les particules élémentaires et les ions atteignent

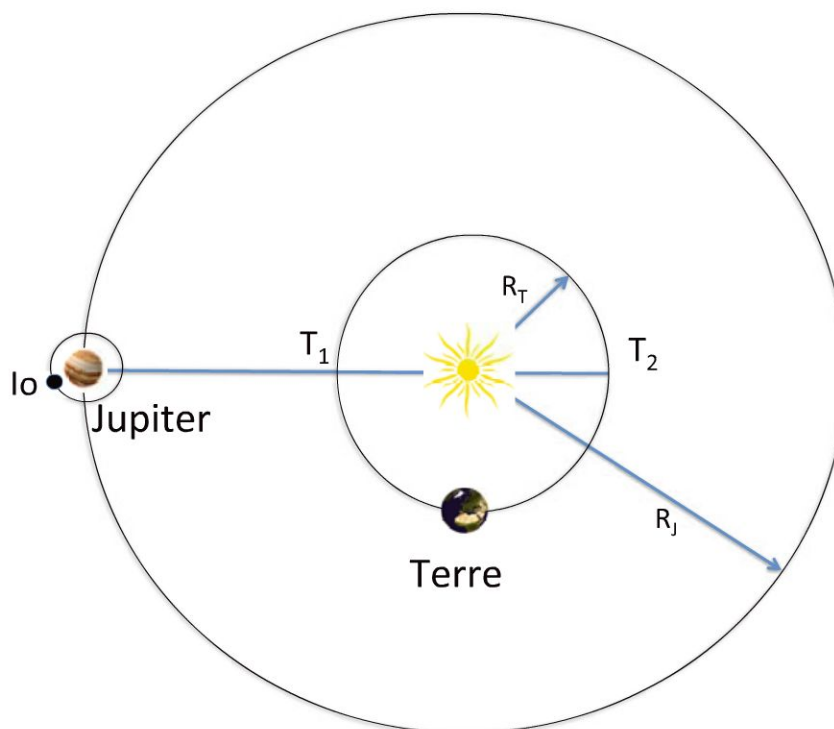


Figure 2.4 – Schéma du système solaire montrant le satellite Io tournant autour de Jupiter et les orbites de Jupiter et de la Terre autour du soleil.

des vitesses proches de la lumière dans les accélérateurs de particules et anneaux de stockage, ou des objets astronomiques extrêmement éloignés comme les quasars.

On peut donc légitimement penser que la vitesse de la Terre par rapport à l'hypothétique référentiel siège des oscillations lumineuses \mathcal{R}_0 est au maximum de l'ordre de un millième de la vitesse de la lumière. Si on veut mettre en évidence le mouvement de la Terre par rapport à ce référentiel, il faudra donc effectuer des mesures de l'anisotropie de c d'une très grande précision. C'est A. Michelson qui le premier, en 1881, eut l'idée d'utiliser les techniques interférométriques de l'optique, qui sont extrêmement sensibles, pour effectuer cette mesure d'anisotropie.

2.3.3 Interféromètre de Michelson

Description. L'interféromètre imaginé par Michelson est schématisé sur la figure 2.5. Son élément central est constitué par une *lame semi-réfléchissante* (L), qui a la propriété de réfléchir et transmettre simultanément la lumière avec des coefficients de réflexion et de transmission égaux. Concrètement, il s'agit d'une lame de verre dont une face a subi un traitement anti-réfléchissant et l'autre face a subi un traitement approprié pour que les puissances des faisceaux réfléchis et transmis soient égales chacune à la moitié de la puissance incidente.

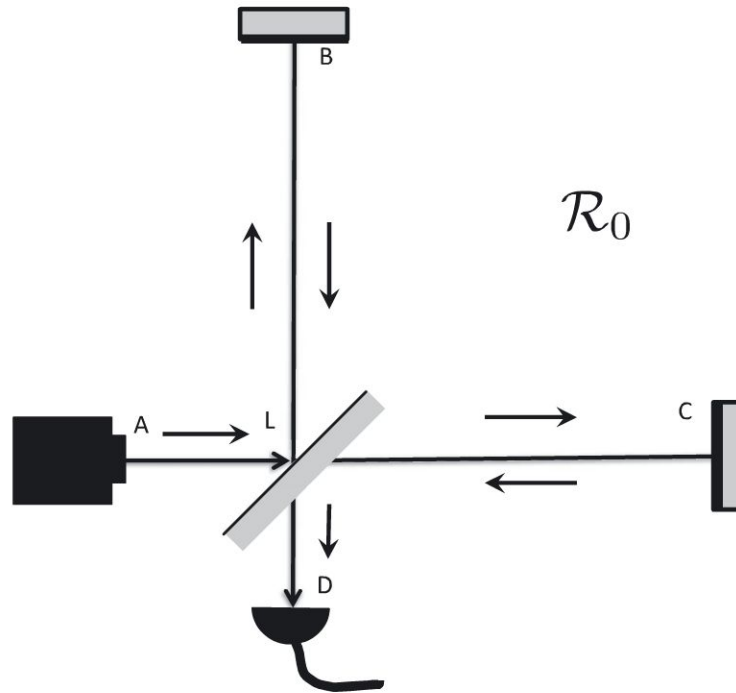


Figure 2.5 – Interféromètre de Michelson immobile dans \mathcal{R}_0 : A est la source lumineuse, (L) une lame semi-réfléchissante, B et C deux miroirs plans et D un détecteur de lumière.

Le dispositif est éclairé selon la direction (AL) par un faisceau de lumière monochromatique de longueur d'onde λ et de période $T = \lambda/c$. Celui-ci est divisé par la lame (L) en deux faisceaux (LB) et (LC) qui se propagent selon des directions orthogonales. Deux miroirs placés en B et C perpendiculairement aux directions de propagation réfléchissent sur eux-mêmes ces faisceaux, qui viennent se recombinaer au même point L de la lame semi-réfléchissante. Un photodétecteur placé sur le faisceau de sortie (LD) en mesure l'intensité. Ce faisceau est la superposition des faisceaux ayant effectué les trajets (ALBL) et (ALCL). Nous avons vu que deux ondes monochromatiques se propageant dans la même direction, d'amplitudes complexes respectives $A_0 \exp(-i(\omega t - kz))$ et $A_0 \exp(-i(\omega t - kz + \phi))$, pouvaient interférer constructivement ou destructivement selon la valeur de la différence ϕ entre leurs phases d'oscillation (section 2.1.2), et qu'on avait :

$$I(\mathbf{r}) = 2A_0^2 \cos^2 \frac{\phi}{2} \quad (2.24)$$

Ainsi, si les longueurs parcourues par les deux ondes (ou les temps de parcours) sont identiques, les ondes sont en phase ($\phi = 0$) et l'intensité mesurée est maximale. En revanche, si les longueurs des deux trajets diffèrent d'une demi-longueur d'onde (ou les temps de parcours d'une demi-période), les deux ondes sont en opposition de phase ($\phi = \pi$). L'interférence est totalement destructive et l'intensité mesurée est nulle. Si on se rappelle que la période optique est de l'ordre de quelques femtosecondes

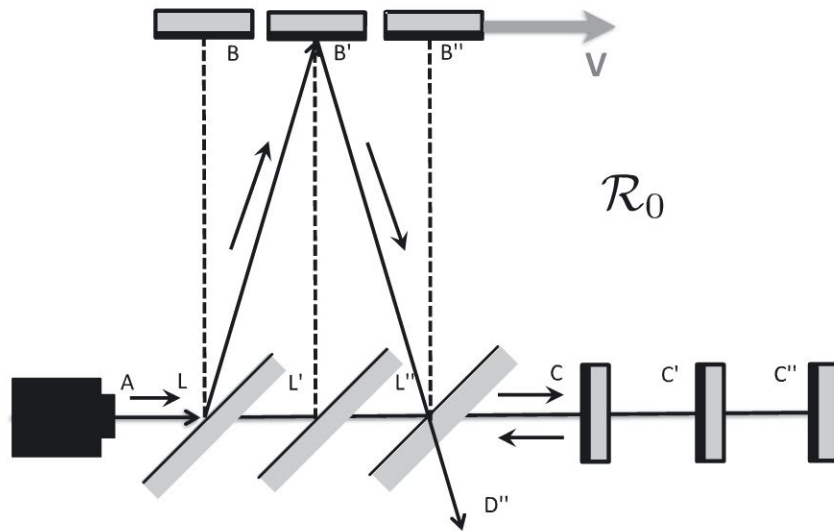


Figure 2.6 – Interféromètre de Michelson en mouvement dans \mathcal{R}_0 : les lettres avec « prime » indiquent la position de la lame et des miroirs au moment où la lumière arrive sur le miroir supérieur ; les lettres avec « seconde » indiquent la position de la lame et des miroirs au moment où la lumière arrive sur la lame semi-réfléchissante.

(10^{-15} s) on voit que le dispositif permet d'apprécier des différences de temps de propagation extraordinairement ténues entre les deux trajets effectués par la lumière.

Cas d'un interféromètre en mouvement Considérons un interféromètre dans lequel les distances LB et LC sont exactement égales à ℓ . Si le référentiel dans lequel l'interféromètre est immobile est le référentiel \mathcal{R}_0 où les équations de Maxwell sont valables, la vitesse de propagation de la lumière y est isotrope, et les temps de parcours des deux faisceaux sont identiques : il y a interférence constructive et l'intensité mesurée en D est maximale.

Considérons maintenant le cas où l'interféromètre est animé d'une vitesse V le long de (AC) par rapport à \mathcal{R}_0 (figure 2.6). Pendant un temps t , tous les points de l'interféromètre se déplacent donc de Vt parallèlement à (AC) . Calculons dans ce référentiel les temps de parcours t_B et t_C des deux faisceaux entre leurs deux passages sur la lame semi-réfléchissante L :

- Pour le trajet (LBL), les faisceaux perpendiculaires au mouvement vont subir le phénomène d'aberration (section 2.2.2) et se propager obliquement dans \mathcal{R}_0 de façon à atteindre toujours le même point B du miroir. Soit t_1 le temps de propagation de L à B. Le point B se déplace de Vt_1 pendant la propagation et on a donc :

$$t_1 = \frac{1}{c} \sqrt{\ell^2 + V^2 t_1^2} \quad (2.25)$$

D'où $t_1 = \ell / \sqrt{c^2 - V^2}$ et $t_B = 2t_1 = 2\ell / \sqrt{c^2 - V^2}$.

- Pendant le temps t_2 du trajet (LC), le miroir C se déplace de Vt_2 , et on a donc $t_2 = (\ell + Vt_2)/c$. De même, pendant le temps t_3 du trajet (CL), la lame L se déplace

2.3. Mesures de la vitesse de la lumière

de Vt_3 , et on a donc $t_3 = (\ell - Vt_3)/c$. Le temps de parcours total vaut alors :

$$t_C = t_2 + t_3 = \ell \left(\frac{1}{c - V} + \frac{1}{c + V} \right) = 2\ell \frac{c}{c^2 - V^2} \quad (2.26)$$

Finalement, la différence des temps de propagation vaut, à l'ordre le plus bas en V/c :

$$t_C - t_B \simeq \frac{\ell V^2}{c^3} \quad (2.27)$$

et le déphasage entre les deux ondes vaut :

$$\phi = 2\pi \frac{t_C - t_B}{T} \simeq 2\pi \frac{\ell V^2}{\lambda c^2} \quad (2.28)$$

Pour un interféromètre de dimension $\ell = 1$ m ce déphasage vaut π pour une lumière jaune et $V/c = 10^{-3}$. Il y a donc annulation totale de l'intensité mesurée si l'interféromètre se déplace à la vitesse de 300 km.s^{-1} par rapport au référentiel \mathcal{R}_0 . En fait, on est capable d'apprécier des variations d'intensité bien plus petites que le passage à l'extinction totale, et donc des différences de phase bien inférieures à π . Le dispositif imaginé par Michelson est donc remarquablement sensible au « vent d'éther », comme on disait au XIX^e siècle, c'est-à-dire au mouvement de l'interféromètre par rapport au référentiel \mathcal{R}_0 de l'éther luminifère.

2.3.4 Résultats expérimentaux

Si l'on fait tourner l'interféromètre dans son plan, on échange tous les 90° les rôles des bras (LBL) et (LCL), et la différence des temps de propagation, donc la phase entre les deux ondes qui interfèrent, doit changer au cours de la rotation. Une première expérience, réalisée en 1881, ne permit pas de voir de variation du signal interférentiel en fonction de cette rotation. Une deuxième expérience, réalisée en 1887 en collaboration avec Morley et utilisant un interféromètre plus grand ($\ell = 10$ m), donna aussi, malgré sa grande sensibilité et le soin extrême apporté à l'expérience, un résultat négatif, ce qui excluait des vitesses V supérieures à quelques km.s^{-1} .

Depuis cette expérience fondatrice, de nombreux physiciens se sont attachés à utiliser les progrès technologiques pour améliorer la sensibilité à la vitesse V de leurs dispositifs expérimentaux. L'utilisation de lasers susceptibles de produire de la lumière extrêmement monochromatique et très stable en intensité a permis, à partir des années 1960, d'augmenter par un facteur important la sensibilité des expériences à une éventuelle anisotropie de la vitesse de la lumière. Toutes, jusqu'à aujourd'hui, se sont soldées par un résultat négatif : aux erreurs (très faibles) d'expérience près, la vitesse de la lumière ne dépend pas du référentiel galiléen dans lequel on la mesure.

L'encart 2.3 décrit brièvement deux de ces expériences.



2.3 Mesures récentes de l'anisotropie de la vitesse de la lumière

1. Décrivons d'abord l'expérience réalisée en 1979 par A. Brillet et J. Hall (*Physical Review Letters* vol 42, p. 549), une des premières à utiliser un laser pour ce type de mesure.

Un laser est constitué par une « cavité Fabry-Perot », c'est-à-dire par deux miroirs distants de L et se faisant exactement face, à l'intérieur de laquelle on a disposé un dispositif amplificateur de lumière. Comme pour le son émis par un tuyau d'orgue ou une corde de guitare, l'onde lumineuse émise par le laser est une onde résonnante pour la cavité Fabry-Perot : elle se renforce à chaque aller-retour, parce qu'elle se retrouve en phase avec elle-même lorsqu'elle a effectué un aller-retour entre les deux miroirs. Cela implique que le temps de son aller-retour dans la cavité est un multiple de la période $T = 1/\nu$ d'oscillation de l'onde :

$$2L/c = pT = p/\nu \quad (2.29)$$

où p est un entier et L la distance entre les miroirs. La fréquence de la lumière émise par un laser dépend donc de la vitesse de la lumière dans le référentiel où le laser est au repos.

L'expérience de Brillet et Hall a consisté à comparer les fréquences émises par deux lasers de même longueur, l'un fixe et l'autre pouvant tourner dans un plan horizontal (figure 2.7). L'éventuelle variation de la vitesse de la lumière en fonction de la direction, Δc , c'est-à-dire son anisotropie, induit une différence entre les fréquences des deux lasers, que l'on peut détecter avec une sensibilité extrême en mesurant le battement entre les deux ondes. Des mesures répétées de la fréquence de ce battement aboutissent à une détermination de la variation de la vitesse de la lumière telle que : $\Delta c/c = (1,5 \pm 2,5) \times 10^{-15}$ donc compatible avec zéro avec une précision relative de 10^{-15} , ce qui constituait un test de l'isotropie de c 4 000 fois plus précis que les mesures précédentes

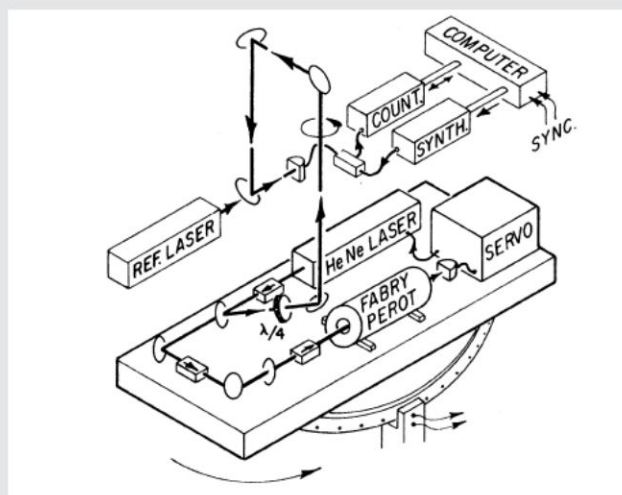


Figure 2.7 - Schéma du dispositif expérimental de comparaison des fréquences d'un laser fixe et d'un laser tournant.

2.3. Mesures de la vitesse de la lumière

2. La mesure a été améliorée d'un facteur 100 en 2010 par C. Eisele, A. Nevsky et S. Schiller (*Physical Review Letters* vol.103, p. 090401), qui ont utilisé le dispositif de la figure 2.8 : il s'agit de quatre miroirs de très grande réflectivité se faisant face deux à deux, constituant deux cavités Fabry-Perot orthogonales. Il existe des techniques dites d'asservissement, qui permettent de forcer un laser à osciller à une fréquence égale à la fréquence de résonance d'une cavité Fabry-Perot, donnée par l'équation (2.29) où L est maintenant la longueur de la cavité Fabry-Perot de référence (sans milieu amplificateur à l'intérieur) et non pas celle de la cavité Fabry-Perot du laser.

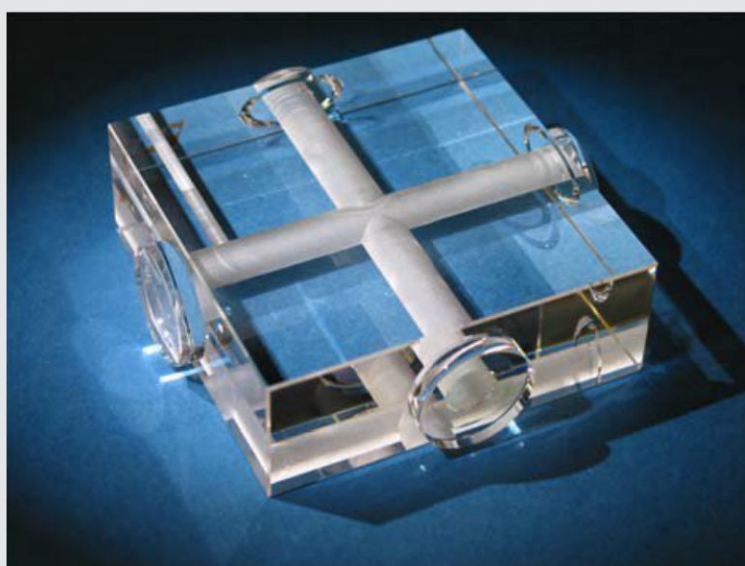


Figure 2.8- Double cavité en verre à faible coefficient de dilatation dont on compare les fréquences de résonance.

Dans l'expérience, deux lasers extrêmement monochromatiques sont asservis chacun sur une des deux cavités Fabry-Perot. La mesure de la différence des fréquences entre les deux lasers permet de remonter à la différence des vitesses de propagation dans deux directions orthogonales. La variation de cette différence lorsqu'on fait tourner le dispositif est une mesure de l'anisotropie de c . La très grande rigidité du dispositif et les progrès dans les techniques optiques ont permis de mesurer une variation relative de vitesse de la lumière entre les deux cavités inférieure à 10^{-17} , correspondant à une vitesse V de déplacement nulle avec une précision de 10 micromètres par heure !

PRINCIPE DE RELATIVITÉ RESTREINTE TRANSFORMATIONS DE LORENTZ

3

Les problèmes soulevés par l'existence hypothétique d'un référentiel absolu \mathcal{R}_0 pour l'électromagnétisme et, en particulier, l'expérience de Michelson, mirent en évidence une inconsistance majeure dans les conceptions théoriques de la fin du XIX^e siècle, plus précisément entre les équations de l'électromagnétisme (les équations de Maxwell) et les équations classiques de changement de référentiel (transformations de Galilée, loi de composition des vitesses) qui pourtant fonctionnent si bien lorsqu'on les applique à la mécanique. Une « révolution scientifique », au sens donné à ce terme dans l'introduction, était donc nécessaire pour sortir de l'impasse. De nombreux physiciens se sont attelés à cette tâche autour des années 1900. Il revient à Albert Einstein d'avoir eu le premier, en 1905, l'audace, l'intuition physique et la profondeur d'analyse nécessaires pour mener jusqu'au bout cette révolution, baptisée « théorie de la relativité restreinte ». C'est ce qu'il fit dans un article publié dans la revue « *Zeitschrift für Physik* »¹ et intitulé « Sur l'électrodynamique des corps en mouvement » : pour résoudre l'inconsistance, il proposa de conserver les lois de l'électrodynamique et de modifier les transformations de Galilée, c'est-à-dire nos conceptions de l'espace et du temps.

Dans le paragraphe suivant, nous énoncerons les principes de base qu'il a proposés, et en déroulerons ensuite les conséquences.

3.1 PRINCIPE DE RELATIVITÉ RESTREINTE

Comme toujours en physique, les principes de base ne peuvent pas se démontrer. Ils sont posés *a priori* et leur validité est fondée sur la conformité de leurs conséquences avec les observations expérimentales. Remarquons aussi que les nouveaux principes doivent être compatibles avec la physique classique, c'est-à-dire redonner les lois de la physique classique et le principe de relativité galiléenne pour des vitesses de déplacement faibles par rapport à la vitesse de la lumière.

1. On trouvera sa traduction en français dans l'ouvrage : Albert Einstein, *œuvres choisies*, t 2, p. 31-58, Editions du Seuil.

3.1.1 Énoncé

- Les lois de la physique ont la même forme dans deux référentiels galiléens \mathcal{R} et \mathcal{R}' différents.
- La lumière se propage dans tout référentiel galiléen \mathcal{R} avec une vitesse c bien déterminée indépendante de l'état de mouvement du corps émetteur.

La première assertion est l'extension à toutes les lois de la physique du principe de relativité galiléenne figurant dans le paragraphe 1.1.1 qui concernait alors uniquement celles de la dynamique. Elle s'applique maintenant aussi aux équations de Maxwell de l'électrodynamique. Elle implique qu'aucune expérience de physique interne à un référentiel galiléen ne permettra de déceler le mouvement de ce référentiel par rapport à un autre. C'est le cas en particulier de l'expérience de Michelson.

Pourquoi ce principe de relativité est-il qualifié de « restreint » ? C'est qu'il ne concerne que les référentiels galiléens et donc que les mouvements uniformes. Einstein a étendu ensuite ce principe au cas des mouvements accélérés : c'est l'objet de ce qu'on appelle la théorie de la relativité générale, que nous évoquerons au chapitre 9. Pour simplifier les expressions, nous qualifierons souvent de « principe de relativité » le principe de relativité restreinte.

3.1.2 Conséquence sur la définition du temps

Le principe de relativité restreinte a une forme extrêmement simple qui permet d'éliminer par une sorte de pirouette le problème posé par le résultat négatif de l'expérience de Michelson. Mais il a de profondes implications sur notre conception du temps, que l'on peut immédiatement voir sur l'exemple suivant.

Considérons l'horloge (H) schématisée sur la figure 3.1. Il s'agit d'une « cavité optique Fabry-Pérot », comme nous en avons déjà rencontré dans l'encart 2.3, constituée de deux miroirs M_1 et M_2 se faisant face et séparés par une distance ℓ . Le miroir M_1 est parfaitement réfléchissant et le miroir M_2 a un coefficient de transmission faible mais non nul. Dans cette cavité un dispositif (analogue à celui existant dans un laser) permet d'entretenir une brève impulsion lumineuse qui fait l'aller-retour entre les deux miroirs à intensité constante. Une faible partie de cette impulsion sort de la cavité à travers le miroir M_2 . On obtient ainsi à la sortie du dispositif laser une succession périodique de brèves impulsions lumineuses séparées de $t = 2\ell/c$ qui constituent le « tic-tac » de l'horloge. Si la distance ℓ est maintenue parfaitement constante, on a ainsi une horloge d'une extrême précision. Ce type d'horloge est actuellement utilisé dans de nombreuses expériences de physique et de métrologie. Il permet de définir le temps du référentiel \mathcal{R} dans lequel la cavité est au repos.

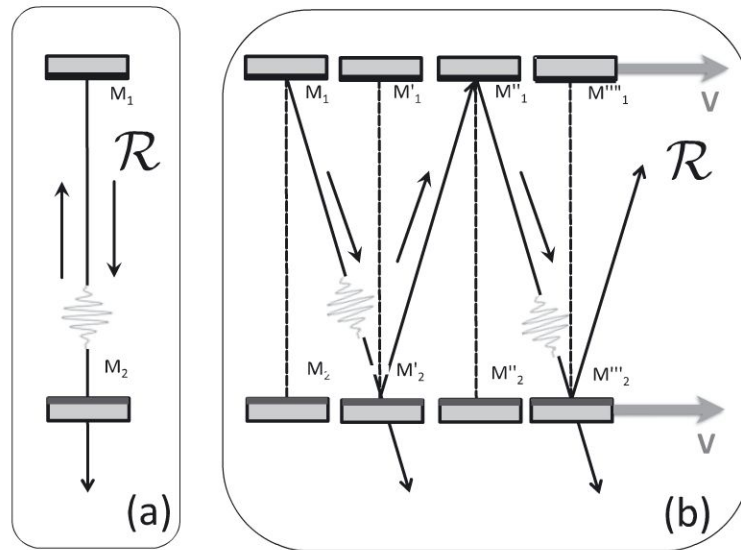


Figure 3.1 – Horloge constituée de deux miroirs M_1 et M_2 se faisant face et émettant ses « tic-tacs » sous la forme de brèves impulsions lumineuses régulièrement espacées. a) L'horloge est au repos par rapport à l'observateur. b) L'horloge est en mouvement par rapport à l'observateur.

Que se passe-t-il si l'horloge (H) est animée d'un mouvement uniforme de vitesse V perpendiculaire à son axe par rapport au référentiel galiléen \mathcal{R}' de l'observateur ? Dans ce référentiel les miroirs se déplacent, et la lumière fait des allers-retours obliques à cause du phénomène d'aberration (figure 3.1b). La distance parcourue par la lumière entre deux réflexions sur les miroirs M_1 et M_2 vaut maintenant une valeur ℓ' qui est donc plus grande que ℓ . Mais le principe de relativité restreinte impose que la vitesse de la lumière est la même dans les référentiels \mathcal{R} et \mathcal{R}' , tous deux galiléens. Les impulsions qui sortent du laser sont donc séparées dans le référentiel \mathcal{R}' d'un temps $t' = 2\ell'/c > t$. La même horloge ne bat pas au même rythme, et donc n'indique pas la même heure, selon qu'on observe les impulsions qu'elle émet dans \mathcal{R} ou dans \mathcal{R}' . Comme $t' > t$, il y a ralentissement de l'horloge lorsqu'elle est en mouvement, et donc « dilatation du temps » défini par une horloge se déplaçant par rapport au temps défini par une horloge immobile.

Il faut donc admettre que *le temps n'est pas absolu mais dépend du référentiel dans lequel on se place*. Il en résulte qu'on doit modifier l'expression 1.6 de la transformation des coordonnées, c'est-à-dire les transformations de Galilée, puisqu'elles comportent la relation $t' = t$.

On pourrait objecter que le raisonnement précédent ne s'applique qu'à l'horloge laser H_1 , et pas à une horloge H_2 fonctionnant différemment et qui donnerait un temps indépendant du référentiel. Mais alors, en comparant dans un référentiel donné les temps donnés par H_1 et H_2 , donc par une mesure interne, on pourrait en déduire la vitesse relative des deux référentiels, ce qui est contraire au principe de relativité.

3.1.3 Conséquence sur la vitesse des objets matériels

Une deuxième conséquence importante du principe de relativité est que la vitesse d'un objet matériel ne peut dépasser celle de la lumière.

Supposons, en effet, qu'un observateur situé en O' puisse se déplacer à une vitesse supérieure à c par rapport au référentiel \mathcal{R} . Alors une impulsion lumineuse émise à l'instant $t = 0$ d'un point A situé sur la partie négative de l'axe Ox ne pourra jamais l'atteindre, et celui-ci n'observera jamais l'arrivée de l'impulsion. Plaçons-nous maintenant dans le référentiel \mathcal{R}' où l'observateur est au repos : l'impulsion lumineuse s'y propage le long de l'axe $O'x$ aussi à la vitesse c et finira par atteindre l'observateur immobile, qui pourra donc par une mesure « interne » déterminer s'il est dans le référentiel \mathcal{R} ou \mathcal{R}' ce qui est contradictoire avec le principe que nous venons de poser.

Le principe de relativité impose donc une contrainte physique à tous les mouvements possibles. Il interdit par exemple l'existence d'un mouvement uniformément accéléré, qui conduit inévitablement à des vitesses tendant vers l'infini. Nous verrons dans le chapitre 7 comment les lois de la dynamique relativiste assurent effectivement cette limitation sur les vitesses.

3.2 TRANSFORMATIONS DE LORENTZ

3.2.1 Synchronisation des horloges dans un référentiel galiléen

Nous venons de voir que le temps, qui paraissait comme s'écouler de la même manière dans tous les référentiels, est en fait relatif et non absolu lorsqu'on confronte le fonctionnement réel d'une horloge au principe de relativité. Une horloge permet de définir sans ambiguïté le temps à l'endroit où elle se trouve, par exemple à l'origine O des coordonnées du référentiel \mathcal{R} . Mais qu'en est-il pour le temps mesuré en un autre point M du même référentiel ? L'utilisation d'horloges identiques en ces deux points permet de s'assurer de l'identité de l'unité de temps en O et en M . Il faut alors définir un processus de *synchronisation des horloges*. Il s'agit d'une question pratique d'une extrême importance, qui se pose concrètement par exemple sur le web, où la synchronisation précise des temps à l'échelle terrestre est nécessaire pour la transmission des données.

Il est nécessaire pour cela que les deux horloges échangent un signal, ce qui ne peut se faire à une vitesse supérieure à celle de la lumière, comme nous venons de le voir. Des impulsions de champ électromagnétique sont en général utilisées à cette fin, par exemple dans les horloges radio-pilotées, qui sont connectées par onde radio à une horloge de référence de très grande précision, ou des récepteurs GPS qui sont connectés aux horloges embarquées dans les satellites GPS.

La procédure est la suivante : dans un premier temps l'horloge de référence située en O envoie dans toutes les directions une brève impulsion qui est renvoyée vers O par un miroir situé en M. La mesure du temps t_1 d'aller-retour de l'impulsion permet de déterminer la distance $OM = ct_1$. Dans un deuxième temps l'horloge située en O envoie vers M une impulsion à l'instant $t = 0$ de cette horloge. Cette impulsion est reçue en M où l'opérateur « met à l'heure » son horloge en appelant $t = OM/c = t_1/2$ l'instant où il a reçu l'impulsion. Le fait de tenir ainsi compte du temps de propagation du signal entre les différents points du référentiel permet effectivement de définir un temps commun à tous les points du référentiel.

3.2.2 Transformation des coordonnées spatio-temporelles dans un changement de référentiel galiléen

L'utilisation de règles identiques, d'horloges identiques dans deux référentiels galiléens \mathcal{R} et \mathcal{R}' , et du processus de synchronisation que nous venons de détailler, permet de définir les coordonnées (t, x, y, z) et (t', x', y', z') d'un même événement (section 1.2.1) dans ces deux référentiels. On appelle *transformation de Lorentz* la transformation qui relie (t', x', y', z') à (t, x, y, z) . Nous ne donnerons pas de démonstration de l'expression de cette transformation dans ce cours d'introduction. Elle doit respecter le principe de relativité que nous venons d'énoncer ainsi que des propriétés plus générales, comme l'indépendance par rapport à l'origine des temps et des coordonnées, ainsi que d'isotropie de l'espace.

Le principe de relativité implique que les équations de Maxwell de l'électrodynamique doivent avoir la même forme dans deux repères galiléens. On peut donc aussi trouver les transformations de Lorentz en cherchant les transformations des coordonnées et des champs qui laissent invariantes les équations de Maxwell. C'est de cette manière que Lorentz, avant Einstein, avait donné l'expression des équations 3.1.

Pour simplifier leur écriture, nous nous placerons dans le cas où le référentiel \mathcal{R}' animé d'une vitesse uniforme \mathbf{V} par rapport à \mathcal{R} , est défini comme indiqué dans le paragraphe 1.2.2 : son axe $O'x'$ coïncide avec Ox , direction de la vitesse \mathbf{V} , les deux autres axes restant à tout instant parallèles (figure 1.1). L'origine des temps $t = 0$ est prise à l'instant où les points O et O' coïncident. L'expression de la transformation de Lorentz $\mathcal{L}_x(V)$ est alors dans ce cas :

$$\begin{aligned} t' &= \gamma \left(t - \frac{Vx}{c^2} \right) \\ x' &= \gamma (x - Vt) \\ y' &= y \\ z' &= z \end{aligned} \quad (3.1)$$

avec

$$\gamma = \frac{1}{\sqrt{1 - V^2/c^2}} \quad (3.2)$$

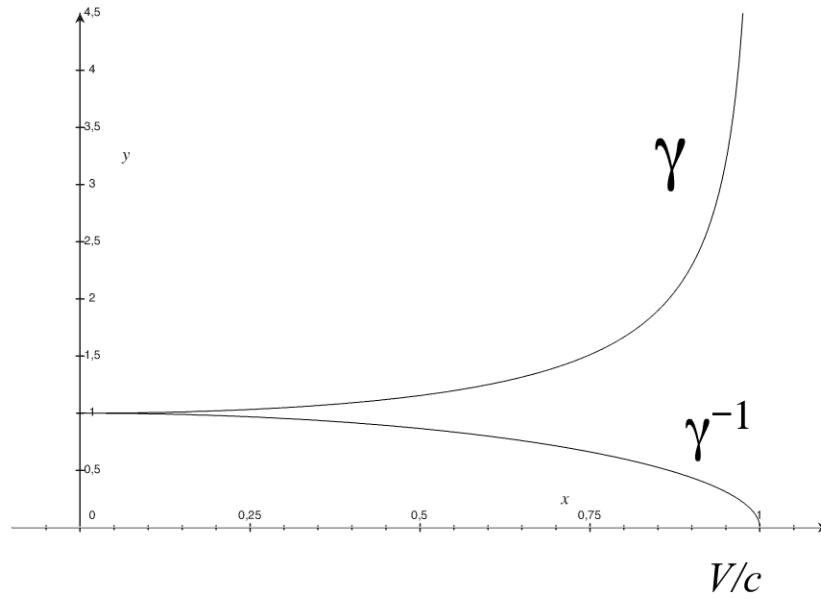


Figure 3.2 – Variation des quantités γ et γ^{-1} en fonction de V/c .

Remarquons que le temps dans le référentiel \mathcal{R}' dépend de V comme on s'y attendait, mais aussi de la position x .

La variation de γ avec V est donnée sur la figure 3.2 : il s'agit d'un paramètre toujours plus grand que 1 qui diverge lorsque V s'approche de c et qui n'est pas défini pour $V > c$, ce qui est conforme aux conclusions du paragraphe 3.2.3. Remarquons que lorsque $V \ll c$, la transformation de Lorentz tend comme il se doit vers la transformation de Galilée introduite au chapitre 1. Plus précisément γ diffère de 1 de moins de 0,1 % tant que $V/c < 0,04 \approx 13\,000 \text{ km.s}^{-1}$. Les écarts à la transformation de Galilée sont donc très faibles sauf lorsqu'on s'approche de très près de la vitesse de la lumière ($\gamma = 2$ pour $V \approx 200\,000 \text{ km.s}^{-1}$). Ils sont néanmoins accessibles aux mesures d'extrême précision, ou accumulées sur des temps très longs. C'est le cas notamment des mesures de localisation de type GPS pour qui les corrections relativistes à la transformation de Galilée jouent un rôle extrêmement important. Nous reviendrons plus en détail sur ce point au chapitre 6.

3.2.3 Propriétés de groupe des transformations de Lorentz

La transformation s'inverse aisément, et on a :

$$\begin{aligned} t &= \gamma \left(t' + \frac{Vx'}{c^2} \right) \\ x &= \gamma (x' + Vt') \\ y &= y' \\ z &= z' \end{aligned} \tag{3.3}$$

On a ainsi $\mathcal{L}_x^{-1}(V) = \mathcal{L}_x(-V)$. On peut, en outre, montrer que la composition de deux transformations $\mathcal{L}_x(V)$ et $\mathcal{L}_x(V')$ est une autre transformation $\mathcal{L}_x(V'')$. L'ensemble des transformations $\{\mathcal{L}_x(V)\}$ forme donc un groupe, le groupe spécial de Lorentz, ayant $\mathcal{L}_x(0)$ pour élément neutre.

Ce groupe est un sous-groupe du groupe de Lorentz total qui comporte toutes les transformations entre coordonnées d'espace-temps de deux référentiels galiléens quelconques. On montre en particulier que le groupe de Lorentz total n'est pas commutatif : $\mathcal{L}_x(V) \circ \mathcal{L}_y(V') \neq \mathcal{L}_y(V') \circ \mathcal{L}_x(V)$.

3.2.4 Représentations dans l'espace-temps

La forme 3.1 des transformations de Lorentz semble compliquée, et on peut en chercher d'autres formulations. La linéarité de ces relations permet une écriture matricielle dans un espace à quatre dimensions, l'*espace-temps* ou *espace de Minkowski*. On définit un *quadrivecteur* position \tilde{E} d'un événement E dans un référentiel \mathcal{R} par ses quatre coordonnées (ct, x, y, z) , toutes homogènes à des longueurs. Au changement de référentiel décrit dans le paragraphe précédent est alors associée une matrice de changement de base :

$$\tilde{E}' = \bar{L}_x(V)\tilde{E} \quad (3.4)$$

$\bar{L}_x(V)$ étant la matrice 4×4 :

$$\bar{L}_x(V) = \begin{bmatrix} \gamma & -\gamma V/c & 0 & 0 \\ -\gamma V/c & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (3.5)$$

On peut aussi simplifier l'écriture des coefficients de la matrice en introduisant la *rapidité* α définie par :

$$th\alpha = V/c \quad (3.6)$$

γ vaut alors $ch\alpha$, d'où l'on déduit :

$$\bar{L}_x(V) = \bar{L}_x\{\alpha\} = \begin{bmatrix} ch\alpha & -sh\alpha & 0 & 0 \\ -sh\alpha & ch\alpha & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (3.7)$$

Il est facile de voir qu'avec ce paramétrage :

$$\bar{L}_x\{\alpha\} \circ \bar{L}_x\{\alpha'\} = \bar{L}_x\{\alpha + \alpha'\} \quad (3.8)$$

3.2.5 Invariants relativistes

Les objets mathématiques qui ont la même forme dans tous les référentiels galiléens sont appelés invariants relativistes. Ils jouent un rôle physique important en raison de leur caractère intrinsèque. Nous en mentionnons ici quelques-uns. Nous en découvrirons d'autres dans la suite de ce cours.

Considérons tout d'abord la quantité :

$$s^2 = c^2 t^2 - x^2 - y^2 - z^2 \quad (3.9)$$

En utilisant (3.7), il est facile de voir que :

$$s'^2 = c^2 t'^2 - x'^2 - y'^2 - z'^2 = (ct \cosh \alpha - x \sinh \alpha)^2 - (x \cosh \alpha - ct \sinh \alpha)^2 - y^2 - z^2 = s^2 \quad (3.10)$$

si l'on se rappelle que $ch^2 \alpha - sh^2 \alpha = 1$.

Cette quantité prend donc la même valeur dans tous les référentiels galiléens. Elle n'est par contre pas toujours positive, et ne peut donc pas être utilisée comme une distance au sens propre. Nous verrons sa signification physique dans le chapitre suivant.

L'équation d'onde (2.7) est invariante par transformation de Lorentz : un calcul analogue à celui que nous venons d'effectuer pour s montre que si (2.7) est vérifiée pour une onde $A(x, y, z, t)$ dans le référentiel \mathcal{R} , $A'(x', y', z', t')$ vérifie de même :

$$\left(\frac{1}{c^2} \frac{\partial^2}{\partial t'^2} - \frac{\partial^2}{\partial x'^2} - \frac{\partial^2}{\partial y'^2} - \frac{\partial^2}{\partial z'^2} \right) A'(x', y', z', t') = 0 \quad (3.11)$$

Une solution de cette équation dans le référentiel \mathcal{R} est l'onde plane $A = A_0 \cos(\omega t - \mathbf{k} \cdot \mathbf{r} + \phi)$, avec $\omega^2/c^2 - \mathbf{k}^2 = 0$. Cette solution a la même forme dans \mathcal{R}' où elle s'écrit $A = A_0 \cos(\omega' t' - \mathbf{k}' \cdot \mathbf{r}' + \phi)$. On a donc :

$$\omega^2/c^2 - \mathbf{k}^2 = \omega'^2/c^2 - \mathbf{k}'^2 \quad (3.12)$$

C'est donc bien un invariant relativiste.

EFFETS RELATIVISTES SUR LE TEMPS ET L'ESPACE

4

Nous allons voir dans ce chapitre les premières conséquences des transformations de Lorentz introduites dans le chapitre précédent. Elles concernent le temps et l'espace, qui sont maintenant fortement liés l'un à l'autre de par la forme même de ces transformations.

Considérons deux événements distincts. Nous les matérialiserons par exemple par de brefs flashes lumineux émis par des lampes de petites dimensions. Nous mettrons le premier, E_0 , à l'origine des coordonnées d'espace-temps dans le référentiel \mathcal{R} . Les relations (3.1) impliquent alors que toutes ses coordonnées d'espace et de temps sont nulles dans n'importe quel référentiel galiléen dont les origines coïncident lorsque $t = t' = 0$. Le second est situé en un temps (t) et en un point M de coordonnées (x, y, z) du même référentiel. Pour simplifier la discussion, restreignons-nous au cas où M est situé sur l'axe Ox ($y = z = 0$). On peut donc représenter l'événement correspondant par un point quelconque du plan (x, ct) de la figure 4.1, l'événement E_0 se situant à l'origine des coordonnées.

4.1 INTERVALLE TEMPOREL : RELATIVITÉ DU PASSÉ, DU FUTUR ET DE LA SIMULTANÉITÉ

Considérons tout d'abord l'intervalle temporel entre l'événement E_0 de coordonnées $(x = 0, t = 0)$ et un événement quelconque de coordonnées (x, t) . Dans le référentiel \mathcal{R} , cet intervalle Δt est égal à $t - 0 = t$. Dans le référentiel \mathcal{R}' , il vaut :

$$\Delta t' = t' - 0 = \gamma(t - Vx/c^2) \quad (4.1)$$

Il dépend donc du référentiel, comme on pouvait s'y attendre après la discussion du chapitre précédent sur la relativité du temps.

Remarquons tout d'abord que *le signe de l'intervalle de temps $\Delta t'$ dépend du référentiel*. Prenons par exemple $t > 0$: dans le référentiel \mathcal{R} l'événement E a lieu *après* l'événement E_0 . Mais si E est situé en un point M tel que $x > c^2 t/V$ alors (4.1) implique que $\Delta t' < 0$: dans le référentiel \mathcal{R}' l'événement E a lieu *avant* l'événement E_0 . On constate donc que, dans le monde relativiste, *les notions de passé et de futur n'ont pas de sens absolu*, puisqu'elles dépendent du référentiel dans lequel on se place.

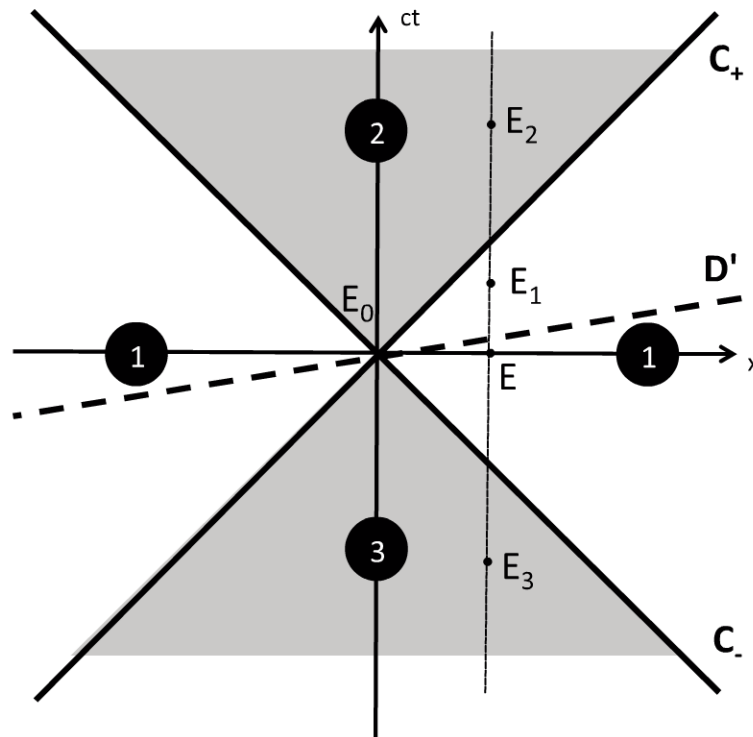


Figure 4.1 – Classification des événements par rapport à l'événement origine E_0 .

Cherchons maintenant l'ensemble des événements qui se produisent au même instant que E_0 . Dans \mathcal{R} , ce sont tous les événements caractérisés par $t = 0$, qui sont donc sur l'axe Ox du schéma de la figure 4.1. Dans \mathcal{R}' , ils sont caractérisés par $t' = 0$, et se trouvent donc sur la droite D' d'équation $ct = x\theta\alpha = xV/c$ de la figure 4.1, car les événements sont repérés sur cette figure par des points dont les coordonnées sont données dans le référentiel \mathcal{R} . On constate que ces ensembles d'événements sont différents dans les deux référentiels : *la notion de simultanéité perd donc, dans le monde relativiste, le caractère absolu qu'elle avait en physique classique*. La relativité de la simultanéité est une des conséquences les plus importantes et les plus troublantes de la transformation de Lorentz.

On peut se faire une idée moins abstraite de cette relativité de la simultanéité en considérant l'expérience de pensée suivante, schématisée sur la figure 4.2 : deux détecteurs sont disposés en deux points A et B de l'axe Ox de coordonnées $-d$ et $+d$ dans le référentiel \mathcal{R} . Une source lumineuse est disposée à l'origine O' des coordonnées du référentiel \mathcal{R}' . Elle se déplace donc à la vitesse V le long de l'axe Ox par rapport aux détecteurs. Elle émet dans les deux directions de l'axe Ox une impulsion lumineuse à l'instant $t = t' = 0$ où les deux référentiels coïncident. Analysons ce qui se passe dans les deux référentiels :

- dans le référentiel \mathcal{R} lié aux détecteurs les deux flashes parcourent des distances d égales et arrivent donc en même temps aux points A et B où ils sont détectés ;

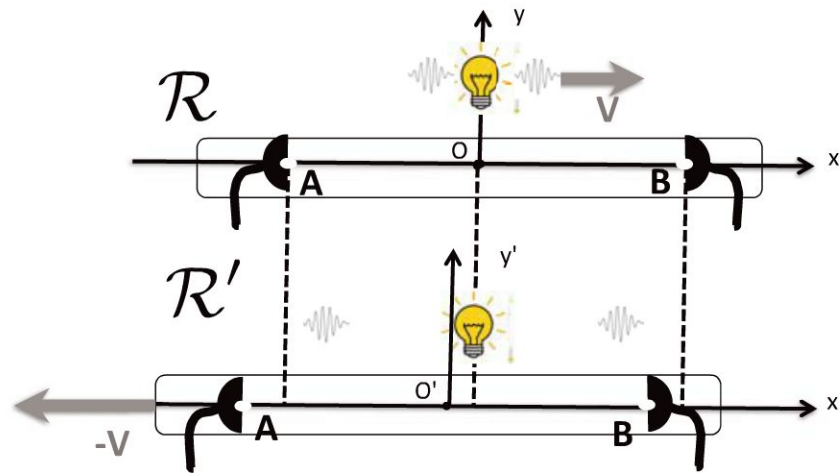


Figure 4.2 - Illustration de la relativité de la simultanéité.

- dans le référentiel \mathcal{R}' lié à la source où la vitesse de la lumière est aussi c , les deux flashes ne seront pas observés en même temps, car pendant le temps de propagation de la lumière, le détecteur A s'est éloigné de la source lumineuse, tandis que le détecteur B s'en est rapproché.

À cause de l'égalité de la vitesse de la lumière dans les deux référentiels, les événements « arrivée de la lumière issue de O en A et B » sont donc bien simultanés dans un référentiel et non simultanés dans l'autre.

4.2 INTERVALLE D'ESPACE-TEMPS

4.2.1 Définition

Dans le paragraphe 3.2.5 du chapitre précédent, nous avons vu que, à la différence de l'intervalle temporel, la quantité s^2 définie en (3.9) ne dépend pas du référentiel dans lequel on la calcule. Nous définirons l'*intervalle d'espace-temps* Δs entre deux événements E_1 et E_2 de coordonnées d'espace-temps respectives (t_1, x_1, y_1, z_1) et (t_2, x_2, y_2, z_2) dans un même référentiel \mathcal{R} par la relation :

$$\Delta s^2 = c^2(t_1 - t_2)^2 - (x_1 - x_2)^2 - (y_1 - y_2)^2 - (z_1 - z_2)^2 \quad (4.2)$$

En ramenant, par changement de l'origine des coordonnées et des temps, un des événements en $(0,0,0,0)$, on peut utiliser la démonstration du paragraphe 3.2.5 pour en déduire que l'*intervalle d'espace-temps* Δs ne dépend pas du référentiel utilisé. Comme toujours en physique, la mise en évidence de quantités intrinsèques, c'est-à-dire invariantes sous certaines transformations, joue un rôle important dans la compréhension des phénomènes en jeu. Nous allons voir dans le paragraphe suivant que c'est bien le cas pour l'intervalle d'espace-temps.

4.2.2 Classification des intervalles d'espace-temps

Considérons à nouveau l'événement E_0 , situé à l'origine des temps et des coordonnées, et un événement E se produisant en un point de l'axe Ox , de coordonnées $(t, x, y = z = 0)$, représentés sur la figure 4.1. Le signe du carré de l'intervalle d'espace-temps $\Delta s^2 = c^2 t^2 - x^2$ entre ces événements permet de définir trois situations physiques bien différentes, correspondant à trois régions différentes du plan (x, ct) de la figure :

- 1) $\Delta s^2 = 0$: *intervalle de genre lumière*
On a alors $x = \pm ct$ (droites C_+ et C_- sur la figure 4.1). Si $t > 0$, le flash lumineux émis en O à $t = 0$ atteint M au moment exact où celui-ci émet son propre flash. Si $t < 0$, c'est le flash émis en M qui atteint O au moment où il émet son flash.
- 2) $\Delta s^2 < 0$: *intervalle de genre espace*
On a alors $|x| > c|t|$ (région 1, et événement E_1 de la figure 3.2) : le flash lumineux émis en O (qui se propage selon la droite C_+) atteint M *après* que celui-ci a émis son propre flash (ou l'inverse). Ce flash arrive « trop tard » : il n'y a pas de « contact », c'est-à-dire de lien physique possible entre les événements E_0 et E_1 , puisque toute transmission ne peut se faire qu'avec une vitesse inférieure ou égale à celle de la lumière.
- 3) $\Delta s^2 > 0$: *intervalle de genre temps*
On a alors $c|t| > |x|$: le flash lumineux émis en O atteint M *avant* que celui-ci n'émette son propre flash (ou l'inverse). Le déclencheur du flash en M a eu un contact avec le déclencheur du flash en O : il y a un lien physique possible, ou transmission d'information, entre les deux événements.

Dans cette dernière situation, on peut faire une distinction supplémentaire entre la région 2 où $\Delta s^2 > 0$ et $t > 0$, et la région 3 où $\Delta s^2 > 0$ et $t < 0$. En effet, dans ces régions le signe de l'instant où se produit l'événement E ne dépend pas du référentiel. Considérons par exemple l'événement E_2 de la région 2. Alors dans un autre référentiel $t' = \gamma(t - vx/c^2)$ sera aussi positif pour les événements E_2 de $x < 0$. Pour ceux de $x > 0$, on a dans la région 2 $x < ct$, et donc $t' > \gamma(t - vt/c) = \gamma t(1 - v/c)$. L'instant t' des événements de la région 2 est donc positif pour n'importe quelle transformation de Lorentz. De même les événements E_3 de la région 3, un raisonnement analogue montre qu'ils se produisent à un instant négatif dans tout référentiel galiléen.

Les régions 2 et 3 définissent donc des propriétés de l'événement E intrinsèques, indépendantes du référentiel : la région 2 est le *futur* de l'événement origine E_0 , lieu de tous les événements pouvant avoir un *lien de causalité avec l'événement origine*, la région 3 le *passé* de l'événement E_0 , où c'est l'inverse. Cette distinction entre le passé et le futur n'est pas possible dans la région 1, où l'antériorité dépend du référentiel, comme nous l'avons vu dans le paragraphe précédent, et dans laquelle il

ne peut y avoir aucun lien de causalité, ni dans un sens ni dans l'autre. Cette région 1 est souvent appelée l'*ailleurs* de E_0 .

Les événements se produisant dans le plan xOy et pas seulement sur l'axe Ox sont décrits dans un espace à trois dimensions (x, y, ct) . Le lieu des événements E qui ont un intervalle d'espace-temps nul avec l'événement origine est un cône de révolution autour de l'axe des temps, qualifié de *cône de lumière*, et les régions 1 (genre espace), 2 (genre temps, futur) et 3 (genre temps, passé) y sont toujours bien définies par les signes de Δs^2 et de t . On ne peut donner de représentation graphique de ces régions pour des événements quelconques de l'espace-temps à quatre dimensions, mais il est d'usage de conserver le nom de cône de lumière à la surface d'équation $\Delta s^2 = 0$ dans cet espace.

4.3 LONGUEUR PROPRE ET CONTRACTION DES LONGUEURS

Considérons un objet matériel rigide aux extrémités duquel sont gravés deux points, et que nous appellerons « règle ». ℓ_0 est la distance entre ces deux points mesurée dans le référentiel dans lequel cet objet est au repos. Cette quantité est appelée *longueur propre* de la règle.

Quelle est la longueur de cette règle quand elle est mise en mouvement, ou, ce qui revient au même, quand l'observateur se déplace par rapport à la règle ? Pour répondre à cette question, il faut d'abord définir précisément le protocole de mesure. On en envisagera ici successivement deux, schématisés sur la figure 4.3 :

- (1) Dans le premier protocole, la règle, au repos dans \mathcal{R} , est disposée entre les points O et A de l'axe Ox de ce référentiel. L'observateur se trouve au point O' du référentiel \mathcal{R}' : il note les instants t'_O et t'_A de son référentiel où sa position coïncide avec les extrémités O et A de la règle. Il en déduit la longueur ℓ' de la règle dans \mathcal{R}' par :

$$\ell' = V(t'_A - t'_O) \quad (4.3)$$

Dans la configuration choisie, on a évidemment $t'_O = 0$. Quant aux abscisses x et x' de l'extrémité de la règle dans les deux référentiels, elles sont liées par la transformation de Lorentz :

$$x = \gamma(x' + Vt') \quad (4.4)$$

avec $x = \ell_0$. Le moment t'_A où l'extrémité de la règle passe en O' est celui où x' vaut zéro soit $t'_A = \ell_0/(\gamma V)$, et la longueur de la règle vue de \mathcal{R}' , donnée par l'expression (4.3), vaut finalement :

$$\ell' = \frac{\ell_0}{\gamma} \quad (4.5)$$

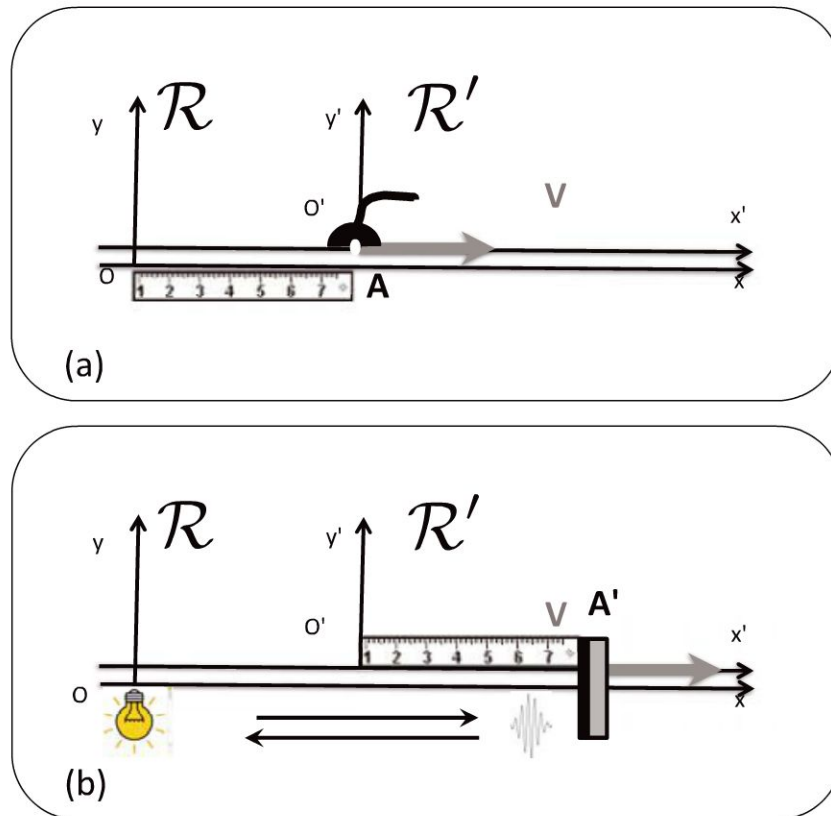


Figure 4.3 – Différentes manières de mesurer la longueur d'un objet en mouvement par rapport à l'observateur : a) le temps de passage de l'extrémité A de la règle est mesuré par une horloge située en O' ; (b) le temps d'arrivée d'une impulsion émise par O et réfléchi par A est mesuré par une horloge située en O et corrigé du temps de propagation.

γ étant supérieur à 1, ℓ' est inférieur à ℓ_0 : c'est le phénomène de *contraction des longueurs*. L'effet ne peut dépendre que de la vitesse relative entre les deux référentiels et la relation (4.5) s'applique donc aussi au cas de la longueur d'un objet mobile mesurée depuis un référentiel fixe.

- (2) Le deuxième protocole est utilisé très fréquemment dans les mesures de distances et de longueurs d'objets lointains : il s'agit de la méthode du radar, qui consiste à envoyer une impulsion de champ électromagnétique (micro-onde ou optique) sur un objet et à mesurer l'instant t où la lumière réfléchi par l'objet revient sur l'émetteur, c'est-à-dire l'instant de l'« écho ». Si l'objet est fixe, la distance de l'objet à l'émetteur vaut alors $ct/2$. Cette technique permet de mesurer la longueur d'une règle située le long de la ligne de visée grâce à la mesure de la différence entre les temps d'arrivée t_C et t_D des échos renvoyés par les deux extrémités de la règle. Si la règle est au repos par rapport à l'opérateur radar on a alors $\ell = c(t_D - t_C)/2$.

4.3. Longueur propre et contraction des longueurs

Qu'en est-il pour une règle en mouvement ? Il faut ici prendre garde au fait que comme les deux extrémités de la règle ne sont pas atteintes par l'impulsion lumineuse aux mêmes instants, la règle s'est déplacée entre ces deux instants : on mesure donc la somme de la longueur de la règle et de son déplacement. Il faut alors retrancher ce déplacement pour obtenir la longueur de la règle. On voit donc que la longueur d'un objet n'a de sens que si on mesure la position de ses extrémités *au même instant dans un référentiel donné*. La simultanéité n'étant pas une notion absolue, comme nous l'avons vu, il en est de même de la longueur des objets.

Précisons le dispositif : la règle, au repos dans \mathcal{R}' , est disposée entre les points O' et A' de l'axe $O'x'$ de ce référentiel. L'observateur se trouve au point O du référentiel \mathcal{R} . Il émet à $t = 0$ une impulsion lumineuse le long de l'axe Ox , dont la position dans \mathcal{R} à l'instant t est donnée par $x_i = ct$. Quant à l'extrémité de la règle où l'on a disposé un miroir pour réfléchir l'impulsion lumineuse, sa position x_A à l'instant t dans \mathcal{R} est donnée par la transformation de Lorentz $x'_A = \ell_0 = \gamma(x_A - Vt)$. L'instant t_A de la rencontre dans \mathcal{R} entre l'impulsion et l'extrémité de la règle correspond à $x_A = x_i$ soit :

$$t_A = \frac{\ell_0}{\gamma(c - V)} \quad (4.6)$$

t_A est la moitié de l'intervalle de temps mesuré par l'observateur entre l'émission et la réception de l'impulsion. La distance parcourue par l'impulsion ct_A est la somme de la longueur de la règle et du déplacement de son extrémité dans le référentiel \mathcal{R} , $ct_A = \ell + Vt_A$, d'où l'on déduit :

$$\ell = \frac{\ell_0}{\gamma} \quad (4.7)$$

On retrouve ainsi, par un autre type de mesure de longueur, le même phénomène de contraction des longueurs.

Les protocoles que nous venons de traiter concernaient une règle disposée parallèlement au mouvement relatif des deux référentiels et ne peuvent s'appliquer pour une règle perpendiculaire au mouvement, disposée le long de la droite Oy par exemple, et il faut donc en envisager d'autres pour mesurer la longueur de la règle dans ce cas. Quel que soit le protocole utilisé, son analyse s'appuiera sur la transformation de Lorentz des coordonnées y , qui sont inchangées dans cette transformation : il n'y aura donc pas d'effet dans cette configuration. *Il n'y a de contraction des longueurs que le long du mouvement.*

Enfin, remarquons qu'il ne faut pas confondre longueur de la règle mesurée dans un référentiel en mouvement par rapport à celle-ci, et vision de la règle par un observateur en mouvement. Dans le premier cas les distances entre points sont évaluées à un instant identique pour tous les points, alors que dans le deuxième cas, la différence des temps de propagation de la lumière depuis les différents points jusqu'à

l'observateur joue un rôle important. Il y a alors déformation apparente de l'objet en mouvement, une sorte d'effet relativiste de perspective.

4.4 DURÉE PROPRE ET DILATATION DES DURÉES

Considérons maintenant l'effet d'un changement de référentiel sur les durées. Pour être plus précis, prenons un objet ponctuel (H) (une horloge) qui émet deux brèves impulsions lumineuses, deux « tic-tacs » de l'horloge. Soient t_1 et t_2 les instants d'émission de ces impulsions. On appelle *durée propre ou temps propre* la différence $\tau_0 = t_2 - t_1$ entre ces deux instants mesurés dans le référentiel dans lequel cet objet est au repos.

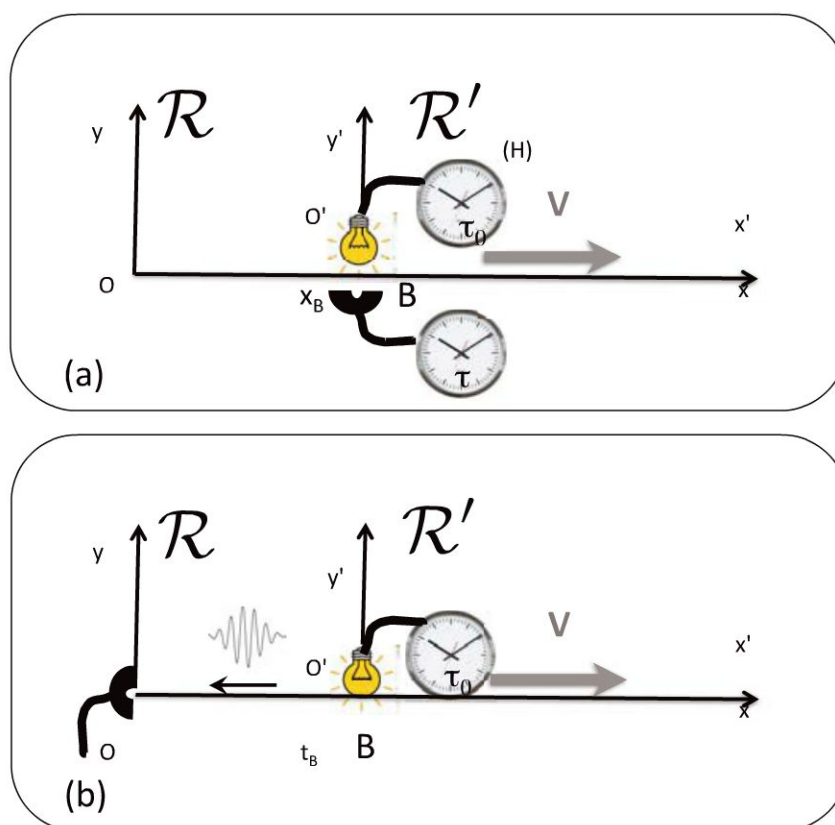


Figure 4.4 – Différentes manières de mesurer la durée entre deux « tops » émis par l'horloge (H) : a) Un observateur en O' observe le passage de la règle; b) Une impulsion lumineuse envoyée de O est réfléchiée par le miroir M situé à l'extrémité de la règle.

On veut maintenant mesurer la durée entre ces deux impulsions lorsque l'objet H a un mouvement de vitesse V par rapport au référentiel \mathcal{R} de l'observateur. On peut ici aussi envisager plusieurs protocoles de mesure (voir Figure 4.4).

- (1) Plaçons nous dans le cas où (H) se trouve à l'origine O' du référentiel \mathcal{R}' . On a disposé à intervalles réguliers dans le référentiel \mathcal{R} et le long de l'axe Ox un

4.4. Durée propre et dilatation des durées

ensemble d'observateurs munis d'horloges. La première impulsion, émise à $t' = 0$, est observée immédiatement en O à l'instant $t = 0$ du référentiel \mathcal{R} . La deuxième impulsion est observée en un point B d'abscisse x_B et à un instant donné t_B du référentiel \mathcal{R} . Utilisons la transformation de Lorentz inverse liant les coordonnées de cet événement dans les référentiels \mathcal{R} et \mathcal{R}' :

$$t_B = \gamma(t'_B + Vx'_B/c^2) \quad (4.8)$$

On a $x'_B = 0$ dans le référentiel \mathcal{R}' . D'autre part t'_B est la durée propre τ_0 et t_B est la durée τ dans le référentiel de l'observateur. On a donc finalement :

$$\tau = \gamma\tau_0 \quad (4.9)$$

- (2) On peut aussi comme dans la section précédente envisager un protocole de mesure de durée utilisant des impulsions lumineuses émises par l'horloge (H) et reçues par l'observateur. (H) est positionnée par exemple à l'origine O' du référentiel \mathcal{R}' , et l'observateur à l'origine O du référentiel \mathcal{R} . L'horloge émet vers l'observateur une impulsion lumineuse le long de $O'x'$ dans la direction des x' négatifs à $t'_1 = 0$ et une autre à $t'_2 = \tau_0$. Cette impulsion se propage à la vitesse c , et sa position x' à tout instant t' ultérieur dans \mathcal{R}' est donnée par $x' = -c(t' - \tau_0)$. L'observateur au point O du référentiel \mathcal{R} reçoit la première impulsion à $t_1 = 0$. Quant à la deuxième, l'équation de son mouvement dans \mathcal{R} est obtenue en utilisant la transformation de Lorentz :

$$\gamma(x - Vt) = -c(\gamma(t - Vx/c^2) - \tau_0) \quad (4.10)$$

Elle atteint l'observateur lorsque $x = 0$, donc à l'instant t_i donné par $t_i = \tau_0/(\gamma(1 - V/c))$. Il faut maintenant tenir compte du fait que la source s'est éloignée entre les deux impulsions d'une longueur $V\tau$ dans le référentiel \mathcal{R} , où τ est le temps qui sépare les deux impulsions dans \mathcal{R} . L'impulsion lumineuse a donc passé dans \mathcal{R} un temps $V\tau/c$ à parcourir ce trajet supplémentaire. Il faut donc le retrancher à t_i pour obtenir l'intervalle de temps τ entre impulsions dans \mathcal{R} corrigé des effets de propagation. On obtient :

$$\tau = t_i - V\tau/c \quad (4.11)$$

soit :

$$\tau = \frac{t_i}{1 + V/c} = \frac{\tau_0}{\gamma(1 - V^2/c^2)} = \gamma\tau_0 \quad (4.12)$$

On retrouve ainsi, par un autre type de mesure de temps basé sur l'enregistrement d'impulsions de champ électromagnétique, le même phénomène de dilatation des temps.

Comme dans le paragraphe précédent, l'effet ne dépend que de la vitesse relative entre les deux référentiels et la relation (4.9) s'applique donc aussi au cas de la durée entre deux « tic-tac » d'une horloge fixe mesurée depuis un référentiel en mouvement par rapport à celle-ci. γ étant supérieur à 1, τ est supérieur à τ_0 : c'est le phénomène de *dilatation des durées*. Il en résulte qu'une horloge bat moins vite lorsqu'un observateur en mouvement en mesure les « tic-tac », comme nous l'avons vu qualitativement dans le chapitre précédent en considérant l'horloge à impulsions lumineuses. Ce phénomène est aussi appelé « ralentissement des horloges ».

4.5 LES « JUMEAUX DE LANGEVIN »

La contraction des longueurs et la dilatation des durées sont des phénomènes « apparents », comme peut l'être un effet de perspective : ils concernent la description d'un même phénomène physique du point de vue de différents observateurs, pas le phénomène lui-même. Mais les transformations de Lorentz ont aussi des conséquences mesurables sur les phénomènes physiques, comme l'a souligné Langevin en considérant l'expérience de pensée dite des « jumeaux de Langevin ».

Considérons deux horloges identiques H_1 et H_2 positionnées initialement à l'origine O d'un référentiel galiléen. H_1 reste immobile en ce point, tandis que H_2 est mise en mouvement, se déplace ensuite à vitesse constante V jusqu'en un point M , rebrousse chemin lorsque M est atteint, effectue le déplacement inverse de M à O à la même vitesse, et est enfin immobilisée au point O où se trouve l'horloge H_1 . Il est alors possible de comparer les temps indiqués par les deux horloges. Langevin avait formulé le problème de manière plus imagée, remplaçant les horloges par deux frères jumeaux dont l'un effectuait un voyage spatial en fusée, le point M par une étoile, et comparé les âges des jumeaux plutôt que les temps indiqués par les horloges.

Faisons l'hypothèse que le temps $2d/V$ pris pour le déplacement OMO est beaucoup plus long que la somme des temps pris par les phases d'accélération, d'inversion de la vitesse et de décélération, et que ces phases de mouvement non uniforme ont une influence négligeable sur le fonctionnement de l'horloge H_2 . Cette hypothèse est corroborée par le traitement exact du problème dans le cadre de la relativité générale, qui est bien au-delà du niveau de cours d'introduction. Nous pouvons alors oublier ces phases et ne prendre en compte que les deux trajets à vitesse uniforme. Au moment de la rencontre finale, l'horloge immobile H_1 indique donc le temps $t_1 = 2d/V$. Qu'en est-il pour l'horloge H_2 ? Dans son référentiel, la longueur OM est affectée par la contraction des longueurs d'un facteur $1/\gamma$. La distance totale parcourue n'est donc que de $2d/\gamma$, et le temps passé dans le référentiel de H_2 pour effectuer le trajet OMO sera $t_2 = 2d/(V\gamma)$: ce sera le temps que l'horloge H_2 indiquera lors de son rendez-vous avec l'horloge H_1 , plus petit que t_1 , la différence entre les temps des

4.6. Des expériences de pensée aux expériences réelles

deux horloges étant de l'ordre de Vd/c^2 . Pour les jumeaux, celui qui a voyagé a vu son temps biologique ralenti : il sera plus jeune que le jumeau sédentaire. Par exemple, au bout du voyage aller-retour de 40 ans effectué à $V = 0,2c$ jusqu'à Proxima du centaure, distant de quatre années-lumière, la différence des âges entre les jumeaux sera de 0,8 années.

Remarquons que ce résultat résulte aussi bien du phénomène de dilatation des temps que de contraction des longueurs : dans le référentiel \mathcal{R} où la longueur du trajet est d , le temps pris par le trajet OM est dilaté par le facteur γ par rapport au temps propre de l'horloge H_2 , et on a donc $\gamma t_2 = 2d/V$.

En outre, l'effet n'est pas réciproque : on ne peut pas inverser les rôles de H_1 et H_2 puisqu'une horloge ne peut pas à la fois avancer et retarder par rapport à l'autre. Les lois de la physique ne sont donc pas les mêmes dans les référentiels liés à chacune des horloges, ce qui semble voler le principe de relativité. En fait, il n'en est rien car si le référentiel lié à H_1 est galiléen, il ne peut en être de même du référentiel lié à H_2 . En effet, l'horloge H_2 connaît des phases d'accélération au départ, de changement de direction de la vitesse ensuite, et de décélération à l'arrivée : elle ne se trouve donc pas à tout instant dans un référentiel galiléen. Le principe de relativité ne s'applique donc pas dans cette configuration.

4.6 DES EXPÉRIENCES DE PENSÉE AUX EXPÉRIENCES RÉELLES : DURÉE DE VIE DES PARTICULES ET HORLOGES EN MOUVEMENT

Les effets relativistes sur le temps et l'espace ont pour l'instant été illustrés par des expériences de pensée avec des objets macroscopiques pour lesquels $V/c < 10^{-3}$ typiquement, c'est-à-dire pour lesquels γ ne diffère de 1 que de quelques 10^{-6} . Pour accéder à des effets mesurables, et donc tester effectivement la validité et l'utilité de la théorie de la relativité restreinte, il faut soit faire des mesures sur des objets microscopiques approchant la vitesse de la lumière, soit faire des mesures d'une extrême précision sur des objets macroscopiques de grande stabilité. Nous allons voir dans ce paragraphe un exemple récent pour chacune de ces situations.

4.6.1 Durée de vie du muon

Le muon est une particule élémentaire qui est une sorte d'« électron lourd » : il a la même charge, mais une masse 207 fois plus grande, que l'électron. Le muon est instable et se désintègre au bout d'un certain temps t pour donner naissance à un électron, facilement détectable, et à deux neutrinos. Comme nous le verrons plus en

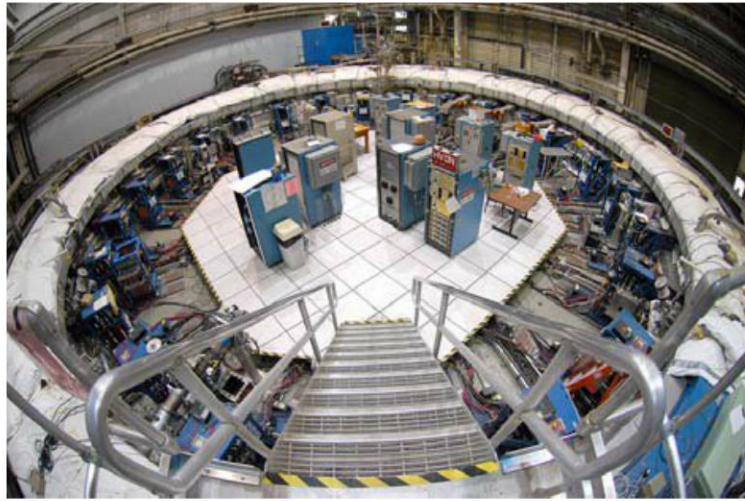


Figure 4.5 – Anneau de stockage de muons pour l'expérience muon g-2 aux États-Unis.
(source : Brookhaven National Laboratory, <http://muon-g-2.fnal.gov>)

détail dans le chapitre 12, ce temps n'est pas fixe mais varie aléatoirement, ce qui est une des caractéristiques quantiques les plus importantes (et les plus troublantes) des phénomènes microscopiques. Néanmoins, lorsqu'on part d'un très grand nombre initial N_0 de muons, le nombre moyen $N(t)$ de muons subsistant au bout d'un temps t obéit à une loi bien définie. Il est donné par :

$$N(t) = N_0 e^{-t/\tau} \quad (4.13)$$

où τ est le temps moyen de désintégration, qui vaut $\tau_0 = 2,198 \pm 0,02 \mu\text{s}$ pour des muons au repos.

Les physiciens ont construit de grands dispositifs appelés « anneaux de stockage » qui permettent de maintenir pendant très longtemps des particules élémentaires chargées sur une trajectoire circulaire avec une vitesse proche de celle de la lumière. Nous décrirons ici, pour être plus précis, une expérience réalisée en 1977 par une collaboration de chercheurs européens (J. Bailey *et al*, *Nature*, vol. 267, p. 301) sur l'anneau de stockage de muons du CERN, analogue à celui de la figure 4.5). L'anneau a un diamètre de 14 m. Des muons dont la vitesse $v = 0,9994c$ correspond à $\gamma = 29,327 \pm 0,004$ sont injectés dans celui-ci à l'instant $t = 0$. 20 compteurs d'électrons disposés autour de l'anneau permettent de mesurer avec précision le nombre de désintégrations $N(t)$ en fonction du temps écoulé après l'injection. Ces données comparées à la loi (4.13) permettent de déterminer expérimentalement le temps moyen de désintégration, qui est trouvé égal à $\tau = 64,37 \pm 0,03 \mu\text{s}$. Le temps de désintégration constitue ainsi une horloge interne pour le muon, dont la caractéristique temporelle est précisément définie dès lors qu'on moyenne les mesures sur un grand nombre de muons. τ_0 est alors le temps propre de cette horloge. On constate que cette horloge a ralenti effectivement par un facteur $\tau/\tau_0 = 29,28 \simeq \gamma$ qui est très important à cause

4.6. Des expériences de pensée aux expériences réelles

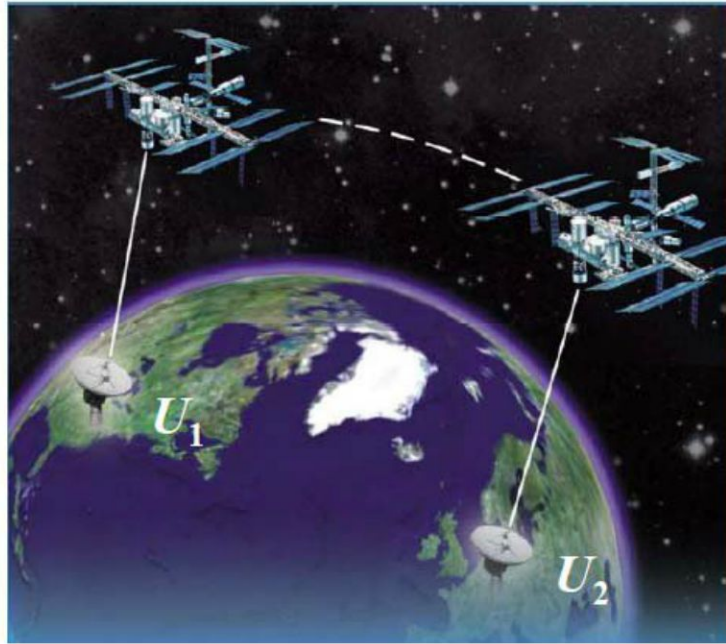


Figure 4.6 - L'horloge atomique ACES-PHARAO, embarquée sur la station spatiale internationale, permettra de synchroniser très précisément des horloges situées sur des continents différents.

(source : CNES : http://smc.cnes.fr/PHARAO/Fr/GP_platform_aces.htm)

de la grande vitesse des muons mesurés. De plus la prévision théorique donnée par la relativité est vérifiée par l'expérience avec une précision relative de l'ordre de 0,2%, bien que le mouvement effectif des muons dans l'anneau soit circulaire donc non uniforme.

4.6.2 Horloges en mouvement

Les horloges les plus précises actuellement sont les « horloges atomiques » : elles consistent en une source de rayonnement électromagnétique monochromatique dont la fréquence ν est asservie sur une transition entre deux niveaux d'énergie E_1 et E_2 d'un atome selon la relation de Bohr (chapitre 13) :

$$h\nu = E_2 - E_1 \quad (4.14)$$

où h est la constante de Planck. L'unité de temps est définie par rapport à une transition entre niveaux de l'atome de Césium : par définition de la seconde, ou du Hertz son inverse, la fréquence ν_0 de la transition entre les deux niveaux de plus basse énergie de l'atome de Césium vaut $\nu_0 = 9\,192\,631\,770$ Hz. Comme tous les atomes de Césium de l'Univers sont strictement identiques (c'est une autre des propriétés quantiques troublantes du monde microscopique) deux horloges à Césium différentes

donneront le même temps, aux erreurs expérimentales près. Des générations de physiciens se sont attachées à réduire au maximum ces erreurs, en utilisant des sources de rayonnement (masers ou lasers) de très grande stabilité et des atomes les plus immobiles possibles, c'est-à-dire les plus froids possibles. La différence entre les temps mesurés par deux horloges atomiques identiques pendant une seconde est en 2011 inférieure à 10^{-17} s dans le meilleur des cas. Sur les 14 milliards d'années, qui sont l'âge de l'Univers, de telles horloges auraient donc un désaccord inférieur à 5 secondes ! On conçoit donc que de tels instruments soient capables d'observer les très faibles effets relativistes qui existent sur des objets qui se déplacent à de faibles vitesses.

Nous allons maintenant décrire une des expériences qui illustrent le mieux le « paradoxe » des jumeaux de Langevin. Elle a été réalisée en 1985 à bord de la navette spatiale Challenger : elle a consisté à comparer les temps donnés par des horloges à Césium de grande stabilité, les unes embarquées à bord de la navette, les autres restées à terre. Lors du retour à terre de la navette, au bout d'une semaine restée sur orbite, on constata un décalage des horloges embarquées de $180 \mu\text{s}$ avec le temps donné par les horloges restées au sol, décalage très largement supérieur à la dérive de ces horloges les unes par rapport aux autres sur le même intervalle de temps. L'effet observé est plus précisément de $-295,02 \pm 0,29$ ps par seconde de vol. On voit donc que les jumeaux de Langevin ne sont pas un simple amusement de théoricien : l'effet peut effectivement être observé lorsqu'on dispose d'instruments suffisamment précis.

La comparaison de cette observation expérimentale avec la théorie n'est pas simple car ni la Terre ni la navette ne constituent des référentiels galiléens. Cependant, la vitesse orbitale de la navette (7700 m.s^{-1}) est bien supérieure à celle de la surface de la Terre (environ $4,6 \text{ m.s}^{-1}$), et on peut considérer en première approximation que les horloges au sol appartiennent à un référentiel galiléen, alors que celles qui sont dans la navette appartiennent à un référentiel non galiléen en mouvement de rotation uniforme par rapport au précédent. La vitesse orbitale de la navette permet de calculer un effet relativiste de dilatation des temps, $\gamma^{-1} - 1 \simeq v^2/2c^2$, valant $3,3 \cdot 10^{-10}$, soit 330 ps par seconde. La différence avec la valeur mesurée a deux origines. Le premier effet est lié au fait que le référentiel de la navette est en rotation : il en résulte un décalage « gyroscopique » de fréquence de l'horloge. Le deuxième effet provient de la relativité générale et nous l'aborderons dans le chapitre 9. Il est dû au fait que la fréquence d'une horloge dépend du champ de gravitation dans lequel elle se trouve : une horloge au sol et une horloge à l'altitude de l'orbite de la navette (330 km), même immobiles, donnent des temps légèrement différents. La détermination précise de l'ensemble de ces effets permet de prédire un décalage temporel en très bon accord avec la valeur mesurée, ce qui constitue une confirmation supplémentaire du phénomène relativiste de dilatation des durées. Par extrapolation, on en déduit que Sergey Krikyalov, qui est à ce jour le cosmonaute resté le plus longtemps en orbite (803 jours cumulés sur

4.6. Des expériences de pensée aux expériences réelles

plusieurs vols) est maintenant revenu sur Terre plus jeune de 20 ms que ses collègues restés à Terre ! Mentionnons finalement le projet spatial ACES-PHARAO, qui a pour but d'installer dans la station spatiale internationale une horloge atomique dans le but de synchroniser avec une très grande précision les horloges de toute la Terre et de mesurer les différents effets de modification des durées mesurées dans des référentiels différents (Figure 4.6).

EFFETS RELATIVISTES SUR LES VITESSES

5

Nous nous intéressons dans ce chapitre aux effets relativistes sur les vitesses, c'est-à-dire à la cinématique relativiste. Ces effets concernent aussi bien les objets matériels macroscopiques ou microscopiques, que la lumière envisagée sous sa double forme ondulatoire ou corpusculaire.

On considèrera donc un objet matériel ponctuel mobile situé en un point M dépendant du temps. Dans le référentiel galiléen \mathcal{R} il a pour coordonnées spatiales $(x(t), y(t), z(t))$. Dans un autre référentiel galiléen \mathcal{R}' ses coordonnées sont $(x'(t'), y'(t'), z'(t'))$. On appellera \mathbf{v} et \mathbf{v}' son vecteur vitesse dans l'un ou l'autre référentiel, qu'il faut se garder de confondre avec la vitesse \mathbf{V} de déplacement de \mathcal{R}' par rapport à \mathcal{R} .

5.1 COMPOSITION DES VITESSES EN RELATIVITÉ

5.1.1 Loi de composition des vitesses

Plaçons-nous dans le cas des transformations de Lorentz spéciales où la vitesse \mathbf{V} est parallèle à l'axe Ox des référentiels \mathcal{R} et \mathcal{R}' (figure 1.1). La composante v'_x selon $O'x'$ de la vitesse du point M dans le référentiel \mathcal{R}' est donnée par :

$$v'_x = \frac{dx'}{dt'} = \frac{dx'}{dt} \frac{dt}{dt'} = \frac{dx'/dt}{dt'/dt} \quad (5.1)$$

La transformation de Lorentz (3.1) permet de calculer ces deux dernières dérivées :

$$\frac{dx'}{dt} = \gamma(v_x - V) \quad \frac{dt'}{dt} = \gamma \left(1 - \frac{Vv_x}{c^2} \right) \quad (5.2)$$

et on a finalement :

$$v'_x = \frac{v_x - V}{1 - Vv_x/c^2} \quad (5.3)$$

On a de même :

$$v'_y = \frac{v_y}{\gamma(1 - Vv_x/c^2)} \quad ; \quad v'_z = \frac{v_z}{\gamma(1 - Vv_x/c^2)} \quad (5.4)$$

Lorsque les vitesses \mathbf{v} et \mathbf{V} sont toutes deux petites devant c , cette relation se réduit comme il se doit à la loi de composition des vitesses en cinématique classique $v'_x = v_x - V$; $v'_y = v_y$; $v'_z = v_z$. Notons aussi qu'à la différence de la loi classique, la loi

relativiste aboutit à une modification des composantes v'_y et v'_z de la vitesse, à cause du facteur dt/dt' , c'est-à-dire de l'effet de dilatation du temps.

La transformation inverse s'écrit :

$$v_x = \frac{v'_x + V}{1 + Vv'_x/c^2} \quad ; \quad v_y = \frac{v'_y}{\gamma(1 + Vv'_x/c^2)} \quad ; \quad v_z = \frac{v'_z}{\gamma(1 + Vv'_x/c^2)} \quad (5.5)$$

5.1.2 Loi de composition de vitesses parallèles

Considérons tout d'abord un mobile M dont la vitesse v par rapport à \mathcal{R} est parallèle à Ox . Les relations (5.3) et (5.4) impliquent que sa vitesse v' par rapport à \mathcal{R}' est aussi parallèle à Ox avec les relations :

$$v' = \frac{v - V}{1 - Vv/c^2} \quad ; \quad v = \frac{v' + V}{1 + Vv'/c^2} \quad (5.6)$$

Si la vitesse v' de M par rapport au référentiel \mathcal{R}' vaut $c/2$, et que ce référentiel est animé d'un mouvement de vitesse $V = c/2$ par rapport à \mathcal{R} , la vitesse v de M par rapport à \mathcal{R} vaut $0,8c$ au lieu de la valeur c que donnerait la cinématique classique. La loi relativiste de composition des vitesses n'est donc pas une simple loi d'addition des vitesses.

Utilisons ici le paramétrage par la rapidité introduit au paragraphe 3.2.4 (relation 3.7). Appelons respectivement α , α' et α_R les rapidités associées aux vitesses v , v' et V . On a alors :

$$th\alpha = \frac{th\alpha' + th\alpha_R}{1 + th\alpha' th\alpha_R} = th(\alpha' + \alpha_R) \quad (5.7)$$

d'après une relation connue de trigonométrie hyperbolique. On a alors :

$$\alpha = \alpha' + \alpha_R \quad (5.8)$$

Les rapidités, et non les vitesses, s'additionnent donc simplement lorsque les mouvements sont parallèles. Cette relation permet aussi de voir que lorsque $V \simeq c$ ou $v' \simeq c$, alors $\alpha' \rightarrow \infty$ ou $\alpha_R \rightarrow \infty$, ce qui entraîne $\alpha \rightarrow \infty$ donc $v \simeq c$: on ne peut pas rajouter une vitesse, aussi grande soit-elle, à la vitesse de la lumière, qui est bien une limite supérieure à toute vitesse possible.

5.1.3 Émission de photons par une source en mouvement

Considérons ici la lumière comme formée de photons, envisagés ici comme des corpuscules classiques se déplaçant à la vitesse c , et plaçons-nous dans le cas où ces

5.1. Composition des vitesses en relativité

photons sont émis par le point O' origine du référentiel \mathcal{R}' dans le plan $x'Oy'$ dans une direction faisant un angle θ' avec la direction du mouvement Ox' . On a donc pour la vitesse \mathbf{v}' du photon dans \mathcal{R}' :

$$v'_x = c \cos \theta' \quad ; \quad v'_y = c \sin \theta' \quad (5.9)$$

Les composantes de la vitesse \mathbf{v} du photon dans le référentiel \mathcal{R} sont alors :

$$v_x = \frac{V + c \cos \theta'}{1 + V \cos \theta'/c} \quad ; \quad v_y = \frac{c \sin \theta'}{\gamma(1 + V \cos \theta'/c)} \quad ; \quad v_z = 0 \quad (5.10)$$

Un calcul simple montre que $v^2 = v_x^2 + v_y^2 + v_z^2 = c^2$: le photon se propage donc bien aussi à la vitesse c dans le référentiel \mathcal{R} , ce qui est conforme au principe de relativité sur l'invariance de la vitesse de la lumière, et reste dans le plan xOy . Sa trajectoire dans \mathcal{R} fait un angle θ avec Ox tel que :

$$\tan \theta = \frac{\sin \theta'}{\gamma(\cos \theta' + V/c)} \quad (5.11)$$

On constate qu'en général $\theta \neq \theta'$, sauf pour $\theta = 0$ ou π . Il en résulte que, à l'exception des photons émis strictement vers l'avant ou vers l'arrière, la lumière émise par un point en mouvement change de direction de propagation lorsqu'on l'observe dans un référentiel au repos (ou de manière équivalente lorsqu'un observateur en mouvement observe une source lumineuse immobile) : c'est la version relativiste exacte du phénomène d'aberration que nous avons rencontré dans sa version classique approchée à la section 2.2.2 et que nous retrouverons quand nous décrirons la lumière comme une onde (section 5.2.2).

La figure 5.1 donne la variation de θ en fonction de θ' pour différentes valeurs du rapport V/c . On constate que si la vitesse V de la source se rapproche de celle de la lumière, l'angle θ est proche de zéro pour toutes les valeurs de θ' , à l'exception de celles qui sont très proches de π . Par exemple, si le photon se propage selon Oy , perpendiculairement au déplacement de la source, alors $\tan \theta = 1/\gamma$: sa direction se rapproche de celle de la direction du mouvement de la source. Si la source se déplace à une vitesse $V \simeq c$ et émet des photons de manière uniforme dans toutes les directions du référentiel \mathcal{R}' , alors l'immense majorité de ces photons se propagera dans des directions très proches de l'axe Ox lorsqu'on les observera dans le référentiel \mathcal{R} (figure 5.2). En particulier, la moitié des photons, émis dans \mathcal{R}' avec des angles $0 \leq \theta' \leq \pi/2$, seront observés dans le référentiel \mathcal{R} à l'intérieur d'un cône très étroit de demi-angle au sommet $1/\gamma$ et dans la direction du mouvement : il y aura *concentration de la lumière autour de la direction de propagation de la source*.

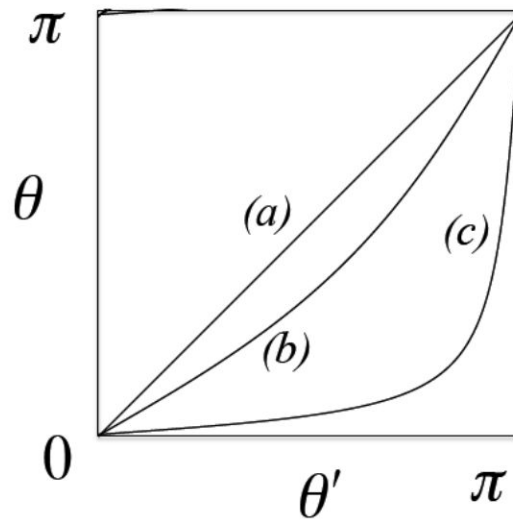


Figure 5.1 – Variation de l'angle de propagation du photon dans \mathcal{R}' en fonction de l'angle de propagation du photon dans \mathcal{R} pour trois valeurs de V/c : 0 (a), 0,5 (b), 0,99 (c).

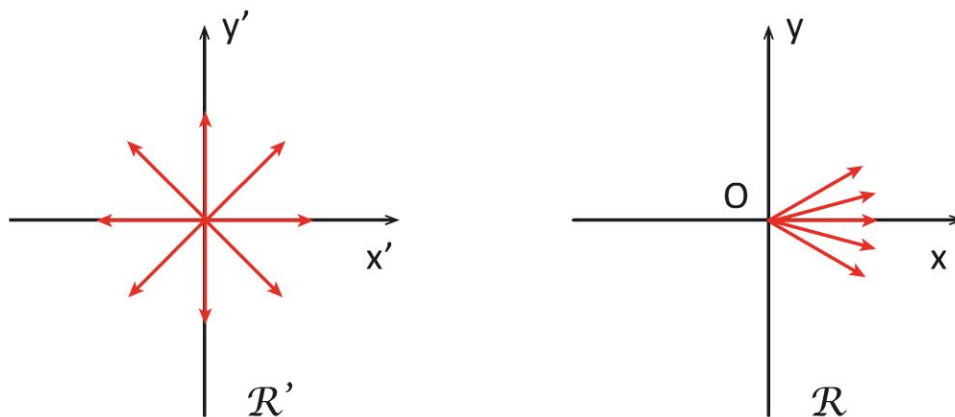


Figure 5.2 – La lumière émise dans \mathcal{R}' , référentiel se déplaçant à la vitesse V par rapport à \mathcal{R} , l'est dans toutes les directions, alors que dans \mathcal{R} elle est émise à l'intérieur d'un cône extrêmement étroit.

5.2 TRANSFORMATION DE LORENTZ POUR UNE ONDE

5.2.1 Invariance de la phase d'une onde

Envisageons maintenant la lumière comme un phénomène ondulatoire : comme nous l'avons vu dans le paragraphe 2.1, une onde lumineuse plane et monochromatique est alors décrite dans un référentiel \mathcal{R} par la quantité complexe oscillante :

$$A(\mathbf{r}, t) = A_0 e^{-i(\omega t - \mathbf{k} \cdot \mathbf{r})} = A_0 e^{-i\Phi(t, \mathbf{r})} \quad (5.12)$$

5.2. Transformation de Lorentz pour une onde

La phase $\Phi(t, \mathbf{r})$ de l'onde peut être considérée comme caractérisant un événement se produisant en un point donné à un instant donné (le maximum de l'onde, son annulation, le nombre d'oscillations...). Elle doit prendre donc la même valeur dans tous les référentiels galiléens : c'est un invariant relativiste, et on doit donc avoir pour tout (t, \mathbf{r})

$$\omega t - \mathbf{k} \cdot \mathbf{r} = \omega' t' - \mathbf{k}' \cdot \mathbf{r}' \quad (5.13)$$

où ω et \mathbf{k} sont la pulsation et le vecteur d'onde de la même onde dans le référentiel \mathcal{R} . En utilisant l'expression (3.1) de la transformation de Lorentz, et en exploitant le fait que la relation (5.13) doit être vérifiée en tout point (x, y, z) et à chaque instant t , on en déduit :

$$\begin{aligned} \omega &= \gamma (\omega' + V k'_x) \\ k_x &= \gamma (k'_x + \omega' V / c^2) \\ k_y &= k'_y \\ k_z &= k'_z \end{aligned} \quad (5.14)$$

ou la transformation inverse :

$$\begin{aligned} \omega' &= \gamma (\omega - V k_x) \\ k'_x &= \gamma (k_x - \omega V / c^2) \\ k'_y &= k_y \\ k'_z &= k_z \end{aligned} \quad (5.15)$$

Considérons alors le *quadrivecteur* $\tilde{\Phi}$ caractérisant la phase de l'onde dans le référentiel \mathcal{R} par ses quatre coordonnées $(\omega/c, k_x, k_y, k_z)$. On a alors :

$$\tilde{\Phi}' = \bar{L}_x(V) \tilde{\Phi} \quad (5.16)$$

où $\bar{L}_x(V)$ est la matrice 4×4 figurant dans l'équation (3.5) : on voit que *pulsation et vecteur d'onde se transforment de la même manière que temps et position*. Ils appartiennent à une même catégorie de quadrivecteurs caractérisés par des propriétés mathématiques identiques vis-à-vis du groupe de transformations de Lorentz. Cette analogie formelle entre propriétés de la lumière envisagée comme particule ponctuelle et comme onde est extrêmement profonde : elle se prolonge en mécanique quantique dans la fameuse *dualité onde-corpuscule* que nous exposerons aux chapitres 10 et 13. Cette analogie a d'ailleurs servi de guide lors de l'élaboration de la théorie quantique.

5.2.2 Transformation du vecteur d'onde : aberration

Revenons maintenant au problème du changement de référentiel pour la lumière déjà abordé au paragraphe 5.1.3 à l'aide d'une approche corpusculaire. Nous décrivons maintenant la lumière sous la forme d'une onde plane. Considérons le cas où, dans le référentiel \mathcal{R} , elle a pour pulsation ω et se propage dans la direction Oy . Dans le référentiel \mathcal{R}' , elle a une pulsation ω' et un vecteur d'onde de module ω'/c . Elle se propage dans le plan $x'O'y'$ dans une direction faisant un angle θ' par rapport à l'axe Ox' . Les coordonnées du vecteur d'onde dans ce référentiel sont donc :

$$k'_x = (\omega'/c) \cos \theta' \quad ; \quad k'_y = (\omega'/c) \sin \theta' \quad (5.17)$$

Les composantes du vecteur d'onde \mathbf{k} du photon dans le référentiel \mathcal{R} sont alors :

$$k_x = \gamma(\omega'/c)(\cos \theta' + V/c) \quad ; \quad k_y = k'_y = (\omega'/c) \sin \theta' \quad (5.18)$$

Dans le référentiel \mathcal{R} la même onde se propage donc dans le plan xOy dans une direction faisant un angle θ par rapport à l'axe Ox tel que :

$$\tan \theta = \frac{\sin \theta'}{\gamma(\cos \theta' + V/c)} \quad (5.19)$$

L'onde se propage selon une direction différente dans les deux référentiels, c'est le phénomène d'aberration que nous avons déjà vu pour une onde classique. Remarquons qu'on trouve pour l'angle une expression identique à (5.11) démontrée pour une particule de vitesse c . Une fois de plus, approche ondulatoire et approche corpusculaire de la lumière aboutissent à des conclusions physiques identiques.

5.2.3 Transformation de la pulsation : effet Doppler relativiste

L'expression de la transformation de la pulsation peut aussi s'écrire de manière vectorielle :

$$\omega = \gamma(\omega' + \mathbf{k}' \cdot \mathbf{V}) = \gamma\omega' \left(1 + \frac{\mathbf{e}_{\mathbf{k}'} \cdot \mathbf{V}}{c} \right) \quad (5.20)$$

où $\mathbf{e}_{\mathbf{k}'}$ est le vecteur unitaire dans la direction du vecteur d'onde \mathbf{k}' . C'est l'expression relativiste exacte de l'effet Doppler, effet qui a déjà été abordé dans le chapitre 2 et dont l'expression approchée classique est donnée par l'équation (2.13).

- Si l'onde se propage dans \mathcal{R}' parallèlement à la vitesse \mathbf{V} mais de O' vers O , $\mathbf{e}_{\mathbf{k}'} \cdot \mathbf{V} = -V$. On a alors pour les longueurs d'onde dans les deux référentiels :

$$\lambda = \lambda' \sqrt{\frac{1 + V/c}{1 - V/c}} \quad (5.21)$$

5.2. Transformation de Lorentz pour une onde

Cette dernière relation indique que $\lambda > \lambda'$ pour V positif. Il y a décalage de la longueur d'onde vers les grandes valeurs, c'est-à-dire vers le rouge dans le cas d'un rayonnement visible, lorsque la source lumineuse s'éloigne de l'observateur : c'est le fameux *red shift* qui joue un rôle très important en astrophysique (encart 5.1). Si l'observateur se rapproche de la source, il y a décalage vers le bleu (*blue shift*). Par exemple, un observateur à bord d'une voiture extrêmement rapide peut voir un feu de circulation comme vert alors qu'il est en fait rouge dans son propre référentiel. Il suffit pour cela qu'il roule à une vitesse $v = c/4$!

- Si l'onde se propage dans \mathcal{R}' perpendiculairement à la vitesse \mathbf{V} , alors $\omega' = \gamma\omega \neq \omega$. Il y a dans cette configuration non seulement aberration, mais aussi décalage Doppler de la pulsation, alors que ce n'est pas le cas pour l'effet Doppler classique. Cet effet Doppler transverse est uniquement dû à l'effet de dilatation du temps entre les deux référentiels.



ENCART 5.1 L'effet Doppler dans les observations astronomiques

Les astronomes disposent sur leurs télescopes ou radiotélescopes de spectromètres extrêmement précis qui leur permettent de mesurer la longueur d'onde du rayonnement qu'ils observent. Grâce à la formule (5.20) de l'effet Doppler ils peuvent remonter à la vitesse des objets célestes, ce qui constitue un outil extrêmement puissant de connaissance de l'Univers, comme nous allons le voir sur les deux exemples suivants.

1. Les radioastronomes Penzias et Wilson ont constaté en 1964 que la Terre baignait dans un rayonnement micro-onde venant de toutes les directions, appelé « fond diffus cosmologique ». Celui-ci est la trace actuelle du Big Bang qui s'est produit il y a 13,7 milliards d'années et a donné naissance à un rayonnement isotrope en expansion continue. Des mesures extrêmement précises du spectre en fréquence de ce rayonnement effectuées par des satellites (COBE, PLANCK, ...) ont montré que le spectre observé n'est pas le même lorsqu'on le mesure dans deux directions opposées de la voûte céleste : en particulier la longueur d'onde correspondant au maximum d'émission diffère légèrement dans deux directions opposées. Il y a donc anisotropie du fond cosmologique, qui est due à la variation du décalage Doppler en fonction des directions relatives de l'onde et de la vitesse de l'observateur. La mesure du décalage Doppler dans la direction de l'Univers où cet effet est maximal donne la direction et l'amplitude de la vitesse de l'observateur par rapport au référentiel où le Big Bang s'est produit. On a ainsi trouvé une vitesse de l'ensemble du système solaire de 627 km.s^{-1} dans la direction de l'amas de l'Hydre. Ce décalage Doppler ne rend pas compte de la totalité des observations : l'écart observé par rapport au décalage Doppler, extrêmement faible, est attribué à la structure de l'Univers au moment de la génération de ce fond diffus cosmologique : on peut donc ainsi « voir » l'Univers tel qu'il était peu de temps après le Big Bang.

2. Comme nous l'avons déjà souligné, les atomes d'un élément donné sont strictement identiques dans tout l'Univers, et ils émettent ou absorbent le rayonnement à des longueurs d'onde λ_0 bien déterminées, liées à la position de leurs niveaux d'énergie. Le spectre d'un élément en constitue en quelque sorte la signature. Or les astronomes observent dans leurs télescopes des spectres globalement identiques à ceux des éléments existant sur terre, mais à des longueurs d'onde observées souvent décalées par rapport aux valeurs mesurées sur terre. Ce décalage est évidemment dû à l'effet Doppler et la mesure de la quantité $Z = (\lambda - \lambda_0)/\lambda_0$ permet de remonter à la composante de la vitesse relative entre la Terre et l'objet observé sur la ligne de visée.

La mesure des décalages Doppler Z de raies spectrales émises par les étoiles de notre Galaxie permet d'en déterminer la distribution des vitesses. On constate par exemple que le Soleil a une vitesse de 370 km.s^{-1} par rapport au trou noir situé au centre de notre galaxie. On peut aussi observer des objets étranges extrêmement lointains, aux confins de l'Univers, les « quasars » et les « sursauts gamma », qui présentent de très forts décalages Doppler vers le rouge. Des valeurs de Z supérieures à 6 ont été mesurées sur certains de ces objets, correspondant à des vitesses d'éloignement de l'ordre de $0,95c$, pour lesquelles la formule relativiste de l'effet Doppler, et non pas son approximation classique, est indispensable. On voit ainsi que l'expansion de l'Univers aboutit à des vitesses relatives entre ses points les plus éloignés qui sont proches de celle de la lumière.

DEUX EXEMPLES DE PHÉNOMÈNES RELATIVISTES

6

Nous détaillons dans ce chapitre deux exemples de dispositifs dans lesquels la relativité ne constitue pas une faible correction mais joue un rôle essentiel. Le premier exemple est un objet qui fait maintenant partie de notre environnement quotidien, puisqu'on le trouve, entre autres, dans beaucoup d'automobiles et de téléphones portables : il s'agit du système GPS de positionnement. Le second n'est pas aussi répandu en raison de son coût et de sa taille, mais on en trouve maintenant de nombreux exemplaires à travers le monde : il s'agit des sources de rayonnement synchrotron très utilisées par les physiciens, les chimistes, les biologistes, les archéologues... pour caractériser très précisément la structure des molécules contenues dans leurs échantillons.

6.1 LE GPS : UN LABORATOIRE RELATIVISTE

La constance de la vitesse de la lumière a une conséquence pratique importante : lier de manière fondamentale mesures de distances et mesures de temps. En effet, dans tous les référentiels galiléens il suffit de multiplier le temps de propagation de la lumière entre deux points pour obtenir la valeur de leur distance. Ainsi repère-t-on la position des étoiles en années-lumière. Rappelons aussi que l'étalon de longueur, le mètre, n'est plus une unité de base du système SI, car la définition du mètre est dérivée de l'étalon de temps, la seconde, définie par l'horloge à Césium : le mètre est la longueur parcourue par la lumière en $(1/299\,799\,458)^{\text{e}}$ de seconde. Pour positionner un point dans l'espace, il suffit donc de mesurer les temps de propagation de la lumière entre ce point et des points de positions connues. C'est le principe du GPS, que nous allons maintenant détailler.

6.1.1 Description du système de positionnement

Le GPS (*Global Positioning System*) est un système de positionnement développé par les USA qui est maintenant accessible au grand public. La Russie a un système analogue, appelé GLONASS, et l'Europe est en train de déployer un système de nouvelle génération, appelé GALILEO.

La « constellation GPS » est constituée d'une trentaine de satellites situés à 20 000 km d'altitude sur des orbites circulaires judicieusement disposées et parcourues en 12 h (Figure 1 de l'introduction à la relativité dans cet ouvrage). Chacun contient une horloge atomique à Rubidium synchronisée avec celles des autres satellites. Ainsi, le temps et la position de ces satellites sont connus avec une extrême précision. Chaque satellite émet vers la Terre un signal micro-onde (aux alentours de 1,5 GHz) formé d'une succession de brèves impulsions, accompagnées de signaux codés permettant d'identifier le numéro j du satellite et l'instant d'émission t_j de l'impulsion. Le récepteur, situé au point de coordonnées (x, y, z) à déterminer, reçoit les signaux émis par les différents satellites qu'il a dans son horizon et enregistre le temps d'arrivée t_{jR} de l'impulsion issue du satellite j par rapport à sa propre horloge. Il a en mémoire des données très précises concernant les satellites (et constamment remises à jour) qui lui permettent de connaître très précisément la trajectoire $x_j(t), y_j(t), z_j(t)$ du satellite numéro j . Malheureusement, l'horloge incluse dans le récepteur est de précision bien moindre que celle des satellites, et en particulier elle n'est pas synchronisée : le récepteur ne connaît donc le temps d'arrivée qu'à une constante additive inconnue près : $t_{jR} = T_0 + \delta t_{jR}$, où δt_{jR} est mesuré et T_0 est inconnu.

L'impulsion électromagnétique s'est propagée à la vitesse de la lumière entre émetteur et récepteur et on a donc :

$$(x - x_j(t_j))^2 + (y - y_j(t_j))^2 + (z - z_j(t_j))^2 = c^2(T_0 + \delta t_{jR} - t_j)^2 \quad (6.1)$$

Cette équation comporte quatre inconnues : x, y, z, T_0 . Pour en déduire sa position en longitude, latitude et altitude, le récepteur a donc besoin des signaux provenant de quatre satellites différents, qui fourniront quatre équations de propagation dont la résolution donnera les valeurs des quatre quantités inconnues. Dans de bonnes conditions de réception et pour des temps de mesure suffisamment longs, l'incertitude sur la position trouvée est de l'ordre d'une dizaine de mètres, ce qui nécessite une précision absolue sur la mesure des temps de propagation de l'ordre de la dizaine de nanosecondes.

6.1.2 Effets relativistes

Les satellites se déplacent à une vitesse de $3\,800 \text{ m.s}^{-1}$, soit $\gamma - 1 \simeq 0,8 \times 10^{-10}$. Il semblerait donc que les effets relativistes soient négligeables pour des objets aussi lents. Il n'en est rien en raison de la très grande précision nécessaire pour obtenir un positionnement précis. En effet, le décalage relativiste entre le temps au sol t et le temps t' du satellite vaut $t' - t = (\gamma - 1)t$. Il est de 10 ns au bout de deux minutes, 300 ns au bout d'une heure, $7 \mu\text{s}$ au bout d'un jour, correspondant à des erreurs de positionnement respectives de 3 m, 100 m et 2 km. La prise en compte de

l'effet relativiste de dilatation du temps est donc absolument indispensable dans la détermination de la position.

Il existe aussi un deuxième effet relativiste qui intervient dans le positionnement, mais cette fois-ci de relativité générale. Nous l'avons déjà rencontré dans le paragraphe 4.6.2 sur les horloges jumelles : l'horloge au sol et l'horloge embarquée sont situées à des altitudes h très différentes par rapport au centre de la Terre, et voient donc des valeurs très différentes du potentiel $\Phi = GM/h$ du champ de gravitation (où M est la masse de la Terre et G la constante de gravitation). Or, comme nous le verrons au chapitre 9, le temps dépend aussi du potentiel de gravitation dans lequel se trouve immergée l'horloge, selon la formule $(t - t')/t = (\Phi - \Phi')/c^2 \simeq g(h - h')/c^2 \simeq 3 \times 10^{-9}$ où g est l'accélération de la pesanteur à la surface de la Terre. On voit que l'effet de relativité générale est même légèrement plus grand que celui de relativité restreinte (et en fait dans un sens opposé). Au bout d'une journée il aboutit à un décalage des horloges de $-45 \mu\text{s}$. L'effet total à prendre en compte est de l'ordre de $-38 \mu\text{s}$ par jour.

Si en revanche on connaît aussi bien la position du récepteur que des satellites, la mesure des temps de propagation des signaux venant de satellites se trouvant dans des directions différentes permet de remonter à la vitesse de la lumière dans ces directions et fournit un test « gratuit » (c'est-à-dire ne nécessitant pas de développer de montage expérimental spécifique) de l'isotropie de la vitesse de la lumière, avec une précision relative de 10^{-9} (P. Wolf, G. Petit, *Physical Review A* t. 56, p. 4405 1997).

6.2 RAYONNEMENT SYNCHROTRON

6.2.1 Principe de fonctionnement

On montre en électromagnétisme que toute charge en mouvement accéléré émet un champ électromagnétique. Par exemple, une charge dans un fil mise en mouvement sinusoïdal à la pulsation ω par un courant alternatif rayonne dans toutes les directions une onde électromagnétique à la même pulsation : c'est le principe de fonctionnement d'une antenne. De manière analogue, les électrons dont la trajectoire est courbée par un champ magnétique subissent une accélération centripète et émettent donc un champ : c'est ce qu'on appelle le *rayonnement synchrotron*. La figure 6.2 donne une vue schématique du synchrotron SOLEIL mis en service en 2007 dans la région parisienne. Des paquets d'électrons, produits et accélérés dans la partie intérieure à l'anneau circulaire, sont ensuite injectés dans l'anneau, qui a 350 mètres de diamètre, et où ils circulent à une vitesse V constante et proche de c caractérisée par un facteur γ égal à 5 400. On insère sur le trajet des électrons des *onduleurs* (figure 6.2), constitués par une succession d'aimants permanents alternativement nord et sud produisant un champ magnétique perpendiculaire à la trajectoire et oscillant



Figure 6.1 – Schéma général du synchrotron SOLEIL.

(source libre : http://fr.wikipedia.org/wiki/Synchrotron_soleil)

spatialement avec une période spatiale p de 5 cm. Les électrons sont donc mis en mouvement oscillant par ce champ magnétique et donnent naissance à un rayonnement synchrotron qui est ensuite utilisé dans les différentes expériences installées sur les différentes « lignes de lumière » que l'on voit sur les bras tangents à l'anneau.

6.2.2 Détermination des caractéristiques du rayonnement émis

Appelons \mathcal{R} le référentiel du laboratoire et \mathcal{R}' le référentiel animé d'une vitesse V dans la direction des onduleurs. Dans ce dernier référentiel, les électrons sont animés d'un mouvement oscillant autour d'une position moyenne fixe. Ils rayonnent un champ électromagnétique dans toutes les directions de \mathcal{R}' à la fréquence ν' de leur mouvement oscillant. Dans \mathcal{R}' les électrons voient passer à la vitesse $-V$ des aimants qui sont distants de $p/\gamma \approx 10 \mu\text{m}$. Il faut, en effet, tenir compte de l'effet de contraction des longueurs dans le changement de référentiel, qui est ici très important vu la grande vitesse des électrons. La période de leur mouvement est donc $p/\gamma V$, et la fréquence $\nu' = \gamma V/p$.

Le rayonnement synchrotron est évidemment utilisé dans le référentiel \mathcal{R} du laboratoire, et il faut donc effectuer maintenant la transformation de Lorentz de \mathcal{R}' à \mathcal{R}

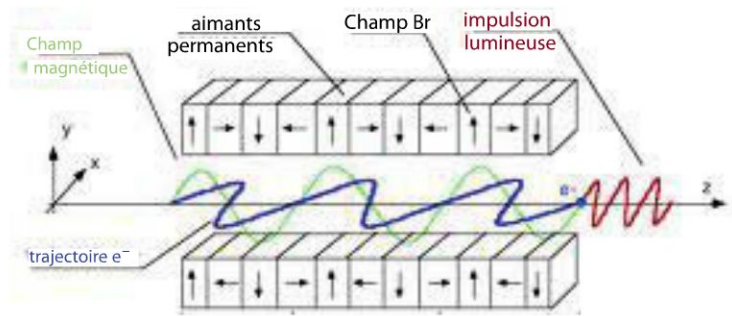


Figure 6.2 – Schéma d'un onduleur.

sur l'onde émise pour en connaître les caractéristiques. On a alors un double effet relativiste :

- l'effet d'aberration qui modifie la distribution spatiale du rayonnement : comme nous l'avons vu dans le paragraphe 5.2, l'essentiel du rayonnement est émis dans le référentiel \mathcal{R} dans un cône étroit autour de la vitesse moyenne des électrons, c'est-à-dire de l'axe de l'onduleur, avec un demi-angle au sommet de l'ordre de $\gamma^{-1} \simeq 2.10^{-4} \text{rd}$. Il s'agit donc d'un faisceau de rayonnement extrêmement directif ;
- l'effet Doppler qui change la fréquence de l'onde. Dans la direction de la vitesse du référentiel qui est celle du maximum d'émission, la formule (5.20) donne pour la fréquence ν de l'onde dans \mathcal{R} :

$$\nu = \gamma \nu' \left(1 + \frac{V}{c}\right) \simeq 2\gamma^2 \frac{c}{p} \quad (6.2)$$

Pour SOLEIL, cette fréquence vaut $3,5.10^{17} \text{ Hz}$, ce qui correspond à une longueur d'onde λ de $0,8 \text{ nm}$: il s'agit donc d'un rayonnement X.

Les synchrotrons tels que SOLEIL, qui sont maintenant extrêmement nombreux dans le monde, produisent donc un rayonnement extrêmement directif et monochromatique dans le domaine des rayons X, dont on peut changer la longueur d'onde en changeant l'énergie, donc la vitesse des électrons. λ étant de l'ordre de grandeur des dimensions des atomes, des molécules ou des mailles cristallines, on conçoit que ce rayonnement constitue une extraordinaire sonde de ces objets microscopiques, sans rivale en termes de brillance et de monochromaticité. C'est la raison pour laquelle il est un instrument de prédilection pour les chercheurs de nombreuses disciplines : physique des atomes et des cristaux, mais aussi chimie, biologie, pharmacie et médecine, qui utilisent donc de manière routinière les effets relativistes que nous venons de détailler.

DYNAMIQUE RELATIVISTE

7

Nous avons vu précédemment comment la description du mouvement (c'est-à-dire la *cinématique*) devait être modifiée pour satisfaire au principe de relativité. Dans ce chapitre, nous étudions comment doivent être reformulées les lois de la *dynamique*, c'est-à-dire, essentiellement, les liens entre forces et mouvement.

En fait, si les lois de l'électromagnétisme et, par suite, l'expression de la force de Lorentz, sont en accord avec les transformations de Lorentz et le principe de relativité, nous allons voir qu'il n'en est pas de même pour les lois de la mécanique newtonienne, et qu'il nous faut revoir les définitions de l'énergie et de la quantité de mouvement si l'on veut conserver les principes usuels de la dynamique, à savoir le principe fondamental de la dynamique (PFD) et le théorème de l'énergie cinétique.

La fin de ce chapitre, consacrée aux collisions relativistes, nous permettra d'illustrer ces définitions par quelques exemples génériques.

7.1 QUELQUES RAPPELS DE DYNAMIQUE CLASSIQUE

Ces rappels concernent la dynamique *classique* des objets matériels et des champs électromagnétiques, et les formules considérées ne sont donc, *a priori*, pas relativistes. Nous mettons néanmoins l'accent sur le fait que les lois de Maxwell respectent le principe de relativité de Lorentz.

Nous insistons également sur les lois de conservation (masse, quantité de mouvement, énergie) sur lesquelles nous nous appuierons, dans les paragraphes suivants de ce chapitre, pour introduire les notions de quantité de mouvement et d'énergie relativistes.

7.1.1 Dynamique newtonienne

Comme on l'a vu au chapitre 1, la quantité de mouvement \mathbf{p} d'une particule ponctuelle de masse m et de vitesse \mathbf{v} est définie, en mécanique classique, par la relation suivante :

$$\mathbf{p} = m\mathbf{v} \quad (7.1)$$

La façon dont cette particule évolue dans l'espace-temps galiléen est alors régie par le principe fondamental de la dynamique (PFD) :

$$\frac{d\mathbf{p}}{dt} = m\mathbf{a} = \mathbf{F} \quad (7.2)$$

où $\mathbf{a} = d\mathbf{v}/dt$ est l'accélération de la particule et où \mathbf{F} est égale à la somme des forces \mathbf{F}_i qui s'appliquent sur elle :

$$\mathbf{F} = \sum_i \mathbf{F}_i \quad (7.3)$$

L'énergie cinétique (classique) de cette particule est définie par la relation :

$$E_c = \frac{1}{2}m\mathbf{v}^2 = \frac{\mathbf{p}^2}{2m} \quad (7.4)$$

Elle est liée au travail des forces \mathbf{F}_i par le théorème de l'énergie cinétique :

$$W_C(\mathbf{F}) = E_c(B) - E_c(A) \quad (7.5)$$

où $W_C(\mathbf{F})$ est le travail de la force \mathbf{F} « le long du chemin C » :

$$W_C(\mathbf{F}) = \int_C \mathbf{F} \cdot d\mathbf{l} = \sum_i \int_C \mathbf{F}_i \cdot d\mathbf{l} \quad (7.6)$$

et où $E_c(B)$ et $E_c(A)$ représentent les valeurs que prend l'énergie cinétique de la particule aux deux extrémités A et B du chemin.

Nous verrons aux paragraphes 7.2 et 7.3 comment les expressions *relativistes* de la quantité de mouvement et de l'énergie cinétique diffèrent de celles rappelées ici.

7.1.2 Constantes du mouvement classiques

En mécanique classique, certaines quantités ne varient pas dans le temps et sont donc appelées des *constantes du mouvement*. Par exemple, la masse d'un corps est une constante du mouvement :

$$m = \text{constante} \quad (7.7)$$

De même, pour un ensemble de corps pouvant réagir (chimiquement par exemple) et ainsi changer de *structure*, la masse totale de ces corps est une constante du mouvement (loi de Lavoisier) :

$$M = \sum_i m_i = \text{constante} \quad (7.8)$$

7.1. Quelques rappels de dynamique classique

Si l'ensemble de particules, de masse m_i et de vitesse \mathbf{v}_i , est isolé, la quantité de mouvement totale est également conservée :

$$\mathbf{p} = \sum_i m_i \mathbf{v}_i = \text{constante} \quad (7.9)$$

Enfin, si les particules de cet ensemble isolé sont soumises à une force \mathbf{F} dérivant d'un potentiel extérieur U (c'est-à-dire telle que $\mathbf{F} = -\mathbf{grad}(U)$, vecteur dont les trois composantes sont $-\partial U/\partial x$, $-\partial U/\partial y$ et $-\partial U/\partial z$), l'énergie totale, c'est-à-dire la somme de l'énergie cinétique totale E_c et de l'énergie potentielle U , est également une constante du mouvement :

$$E_{\text{totale}} = E_c + U = \text{constante} \quad (7.10)$$

Chacune de ces constantes du mouvement peut en fait être reliée à une propriété particulière des lois physiques qu'on appelle « symétrie » (encart 7.1). Ces symétries (parfois subtiles à appréhender), et les constantes du mouvement associées, sont généralement activement recherchées lors de la résolution d'un problème car elles permettent souvent de prédire la forme mathématique des solutions et de connaître l'état final d'un système sans s'intéresser au détail de son évolution intermédiaire.

Précisons que les constantes du mouvement ne sont pas à confondre avec les *invariants relativistes* introduits au paragraphe 3.3.5.



ENCART 7.1 Lien entre symétries et lois de conservation

Il existe un lien étroit entre les lois de conservations (les constantes du mouvement) et l'existence de « symétrie » des lois physiques associées, comme par exemple, la symétrie par translation dans le temps ou dans l'espace. Bien entendu, ces symétries particulières ne doivent pas être confondues avec les symétries usuelles des objets ou des formes géométriques (comme les multiples symétries du cercle par exemple).

Le théorème de Noether précise alors que :

« À chaque symétrie des lois de la physique correspond une loi de conservation (une constante du mouvement), et inversement ».

Ainsi, à la *symétrie par translation dans l'espace* (c'est-à-dire à l'indépendance des lois de la physique vis-à-vis de l'origine spatiale des coordonnées), correspond la loi de conservation de la quantité de mouvement.

On peut le montrer en considérant un système de particules dont l'énergie potentielle totale U est indépendante de la position globale du système (c'est-à-dire de son centre de gravité). Dans ce cas, les dérivées spatiales de U sont nulles et aucune force extérieure ne s'applique sur ce système. Le Principe Fondamental

de la Dynamique nous indique alors que la dérivée temporelle de la quantité de mouvement totale est nulle :

$$\mathbf{F} = -\nabla U = \mathbf{0} = \frac{d\mathbf{p}}{dt} \quad (7.11)$$

ce qui signifie que les trois composantes de \mathbf{p} sont des constantes du mouvement.

Une autre façon de comprendre cette relation est de considérer un système isolé de deux particules, repérées dans un référentiel galiléen par les vecteurs position \mathbf{r}_1 et \mathbf{r}_2 , dont le potentiel d'interaction ne dépend que de la distance entre ces deux particules $U(\mathbf{r}_1 - \mathbf{r}_2)$. Dans ce cas, on a :

$$\frac{d\mathbf{p}_1}{dt} = -\nabla_{\mathbf{r}_1} U(\mathbf{r}_1 - \mathbf{r}_2) = +\nabla_{\mathbf{r}_2} U(\mathbf{r}_1 - \mathbf{r}_2) = -\frac{d\mathbf{p}_2}{dt} \quad (7.12)$$

et la dérivée temporelle de la quantité de mouvement totale $\mathbf{p}_1 + \mathbf{p}_2$ est nulle.

De la même manière, les lois de la physique sont indépendantes du choix de l'origine du temps, ainsi que de l'orientation dans l'espace des axes de coordonnées. Il y a *symétrie par translation dans le temps et rotation dans l'espace*. On peut montrer qu'à ces deux symétries sont associées deux lois de conservation : celle de l'énergie et celle du moment cinétique total.

La plupart des lois de la physique sont également invariantes par inversion du signe des charges (conjugaison de charge « C »), par remplacement de t en $-t$ (inversion du temps « T »), par symétrie des coordonnées par rapport à une réflexion dans un miroir (parité « P »). Ce sont des symétries plus subtiles, et qui peuvent être violées dans certaines interactions comme l'interaction faible (figure 9.10). On montre en revanche que le produit des trois transformations PCT est conservé par toutes les interactions (pour plus d'informations, consulter les sites en.wikipedia.org/wiki/Noether%27s_theorems ou www-cosmosaf.iap.fr/Noether_et_le_Lagrangien.htm)

7.1.3 Et les lois de l'électromagnétisme ?

L'électromagnétisme forme l'autre grand pilier de la physique classique. Or, nous avons vu aux chapitres 1 et 2 que, si les équations de la mécanique classique étaient bien invariantes par changement de référentiel galiléen, il n'en était plus de même pour les équations de Maxwell et la force de Lorentz.

En revanche, et cela a joué un rôle fondateur dans la construction de la théorie de la relativité restreinte, les équations de l'électromagnétisme sont invariantes par transformation de Lorentz, à condition de transformer non seulement les coordonnées (ct, x, y, z) mais également les composantes (E_x, E_y, E_z) et (B_x, B_y, B_z) des champs

7.1. Quelques rappels de dynamique classique

électromagnétiques \mathbf{E} et \mathbf{B} (encart 7.2) par les expressions :

$$\begin{cases} E'_x = E_x \\ E'_y = \gamma(E_y - \beta c B_z) \\ E'_z = \gamma(E_z + \beta c B_y) \end{cases} \quad \text{et} \quad \begin{cases} B'_x = B_x \\ B'_y = \gamma(B_y + \beta E_z/c) \\ B'_z = \gamma(B_z - \beta E_y/c) \end{cases} \quad (7.13)$$

où $\beta = V/c$.

On remarque alors que les composantes de \mathbf{E} et \mathbf{B} ne se transforment pas exactement comme (ct, x, y, z) : ce ne sont pas les éléments d'un quadri-vecteur mais ceux d'un *tenseur*, le tenseur champ électromagnétique.

Précisons enfin que l'on peut associer à \mathbf{E} et \mathbf{B} une densité d'énergie \mathcal{E} (qui, lorsqu'elle est absorbée par un corps, chauffe celui-ci) et une densité de quantité de mouvement \mathcal{P} ($\|\mathcal{P}\| = \mathcal{E}/c$), qui se traduit par ce que l'on appelle une *pression de radiation* sur les objets éclairés (qui oriente la queue des comètes à l'opposé de la direction du Soleil et modifie la vitesse des atomes irradiés par des faisceaux laser par exemple).



7.2 Invariance relativiste des potentiels électromagnétiques

On définit en électromagnétisme les potentiels scalaire et vecteur V et \mathbf{A} , à partir desquelles on peut exprimer les champs électromagnétiques \mathbf{E} et \mathbf{B} :

$$\mathbf{B} = \text{rot}(\mathbf{A}) \quad \text{et} \quad \mathbf{E} = -\nabla V - \partial \mathbf{A} / \partial t \quad (7.14)$$

Si l'on impose à ces potentiels de satisfaire en plus à la relation $\text{div}(\mathbf{A}) + \frac{1}{c^2} \partial V / \partial t = 0$ (condition de jauge de Lorentz) on montre que les équations vérifiées par ces potentiels sont :

$$\left(\frac{\partial^2}{c^2 \partial t^2} - \frac{\partial^2}{\partial x^2} - \frac{\partial^2}{\partial y^2} - \frac{\partial^2}{\partial z^2} \right) V = \rho / \epsilon_0 \quad \left(\frac{\partial^2}{c^2 \partial t^2} - \frac{\partial^2}{\partial x^2} - \frac{\partial^2}{\partial y^2} - \frac{\partial^2}{\partial z^2} \right) \mathbf{A} = \mu_0 \mathbf{J} \quad (7.15)$$

où ρ et \mathbf{J} sont respectivement la charge volumique et le vecteur courant volumique.

On montre que ces équations sont invariantes par changement de référentiel Galiléen pourvu que l'on impose que les quantités $(\rho c, \mathbf{J})$ et $(V/c, \mathbf{A})$ se transforment de la même façon que les coordonnées spatio-temporelles (ct, x, y, z) , c'est-à-dire par les mêmes transformations de Lorentz, par exemple :

$$\begin{cases} \frac{V'}{c} = \gamma \left(\frac{V}{c} - \beta A_x \right) \\ A'_x = \gamma \left(A_x - \beta \frac{V}{c} \right) \\ A'_y = A_y \\ A'_z = A_z \end{cases} \quad (7.16)$$

Utilisant les relations (7.14) pour revenir aux champs \mathbf{E} et \mathbf{B} , on obtient finalement les transformations (7.13).

7.2 QUANTITÉ DE MOUVEMENT RELATIVISTE

7.2.1 Définition de la quantité de mouvement relativiste

Si la quantité de mouvement totale d'un système physique est bien conservée en mécanique classique (section 7.1.2), l'encart 7.3 ci-dessous montre, sur un exemple, qu'elle ne l'est plus si l'on utilise la loi de composition *relativiste* des vitesses établie au chapitre 5. Un constat s'impose alors : afin de préserver le principe de conservation de la quantité de mouvement totale, il faut modifier la définition de celle-ci. Comment ? En tenant compte, comme indiqué dans l'encart, de l'effet de *saturation des vitesses* à travers le facteur γ .

Ainsi, par définition, la quantité de mouvement d'un objet ponctuel de masse m se déplaçant à la vitesse \mathbf{v} par rapport à un référentiel galiléen est égale à :

$$\mathbf{p} = \gamma_v m \mathbf{v} \quad \text{avec} \quad \gamma_v = \frac{1}{\sqrt{1 - \mathbf{v}^2/c^2}} \quad (7.17)$$

Attention à ne pas confondre γ_v , où \mathbf{v} est la vitesse de l'objet ponctuel dans le référentiel \mathcal{R} , et le facteur γ qui intervient dans la transformation de Lorentz, où V est la vitesse relative d'un référentiel par rapport à l'autre.

Aux faibles vitesses, le facteur γ_v peut se développer en puissances de $(\mathbf{v}/c)^2$: $\gamma_v = 1 + \frac{1}{2}\mathbf{v}^2/c^2 + \dots$, et l'on retrouve bien l'expression classique de \mathbf{p} :

$$\mathbf{p} \simeq m \mathbf{v} \quad \text{si} \quad \|\mathbf{v}\| \ll c \quad (7.18)$$

à des corrections relativistes près.

Même si l'exemple détaillé en encart ne joue pas le rôle d'une démonstration, on peut tout de même souligner que cette définition est satisfaisante du point de vue du principe de relativité. En effet, la quantité de mouvement classique $p_{\text{classique}} = m dx/dt$ s'exprime en fonction de la distance dx , mesurée dans \mathcal{R} , pendant la durée dt , mesurée également dans \mathcal{R} , alors que la quantité de mouvement relativiste $p_{\text{relativiste}} = \gamma m dx/dt = m dx/d\tau$ fait intervenir la durée propre $d\tau$ du mouvement dans le référentiel lié à l'objet en mouvement, $d\tau$ étant la seule durée qui ait un sens, en relativité restreinte, pour définir la vitesse d'un objet le long d'une trajectoire.



7.3 Mise en défaut de la quantité de mouvement classique

Il est particulièrement instructif de constater que la loi de conservation de la quantité de mouvement totale (par exemple, avant et après un choc élastique entre deux particules) conduit à une incompatibilité entre la loi de composition relativiste des vitesses (obtenue au chapitre précédent) et l'expression classique usuelle de la quantité de mouvement $\mathbf{p} = m \mathbf{v}$.

7.2. Quantité de mouvement relativiste

En effet, considérons deux particules de même masse m , lancées suivant la même direction (Oz), en sens opposés, par deux observateurs A et B. A est immobile dans le référentiel \mathcal{R} , et B, immobile dans le référentiel \mathcal{R}' , se déplace parallèlement à (Ox) à la vitesse v . A envoie sa particule, à la vitesse u_0 par rapport à lui-même, suivant l'axe (Oz). B envoie sa particule, également à la vitesse u_0 par rapport à lui-même, suivant l'axe (Oz'). Il y a alors choc élastique de telle façon que les deux particules rebrousse chemin dans leur référentiel propre (figure 7.1).

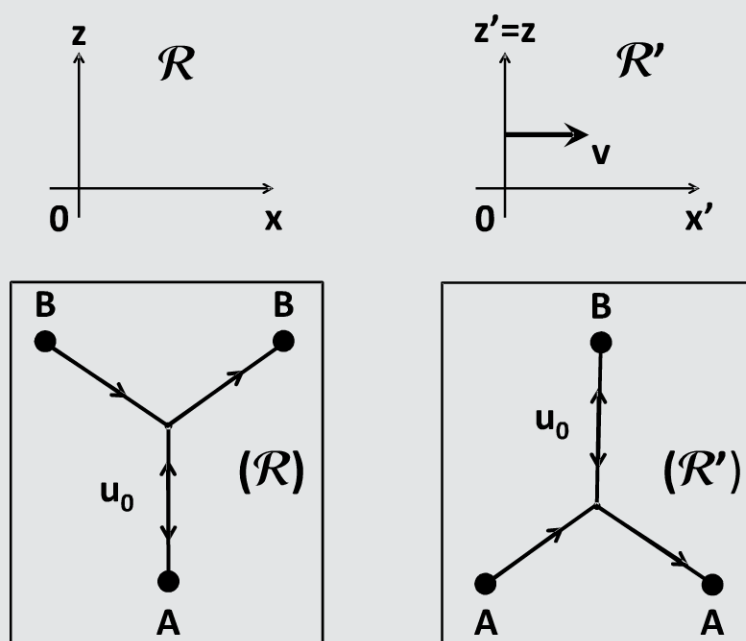


Figure 7.1 - Choc élastique « symétrique » de deux particules.

Intéressons nous à la quantité de mouvement totale du système, avant et après le choc.

Si l'on considère l'expression classique $\mathbf{p} = \sum_i m_i \mathbf{v}_i$, on a alors :

$$(\mathbf{p}_A + \mathbf{p}_B)_{\text{avant le choc}} = m(\mathbf{v}_A + \mathbf{v}_B)_{\text{avant le choc}} \quad (7.19)$$

Or $\mathbf{v}_A = u_0 \mathbf{e}_z$ dans \mathcal{R} et $\mathbf{v}'_B = -u_0 \mathbf{e}_{z'}$ dans \mathcal{R}' . La loi de composition relativiste des vitesses (dont on ne remet pas en cause la validité, établie au chapitre 5) nous donne alors l'expression de la composante suivant l'axe (Oz) de \mathbf{v}'_B dans \mathcal{R} :

$$v_{Bz} = \frac{v'_{Bz}}{\gamma_v (1 + v'_{Bx} v/c^2)} = -\frac{u_0}{\gamma_v} \quad (7.20)$$

puisque $v'_{Bx} = 0$ et $v'_{Bz} = -u_0$.

Ainsi :

$$(\mathbf{p}_{Az} + \mathbf{p}_{Bz})_{\text{avant le choc}} = mu_0 \left(1 - \frac{1}{\gamma_v}\right) \quad (7.21)$$

Par le même raisonnement, on a après le choc (par symétrie $u_0 \longleftrightarrow -u_0$) :

$$(\mathbf{p}_{Az} + \mathbf{p}_{Bz})_{\text{après le choc}} = mu_0 \left(\frac{1}{\gamma_v} - 1 \right) \quad (7.22)$$

d'où la violation de la conservation de la quantité de mouvement classique totale :

$$(\mathbf{p}_A + \mathbf{p}_B)_{\text{avant le choc}} \neq (\mathbf{p}_A + \mathbf{p}_B)_{\text{après le choc}} \quad (7.23)$$

contrairement au cas *purement* non-relativiste où la loi de composition *classique* des vitesses ($v_{Bz} = -u_0$) redonne bien la loi de conservation $(\mathbf{p}_A + \mathbf{p}_B)_{\text{avant le choc}} = \mathbf{0} = (\mathbf{p}_A + \mathbf{p}_B)_{\text{après le choc}}$.

Considérons maintenant l'expression suivante pour la quantité de mouvement totale d'un système de particules :

$$\mathbf{p} = \sum_i \gamma_i m_i \mathbf{v}_i \quad (7.24)$$

où les facteurs $\gamma_i = 1/\sqrt{1 - v_i^2/c^2}$ rendent compte de l'effet relativiste de saturation des vitesses pour chacune des particules.

Alors :

$$(\mathbf{p}_{Az})_{\text{avant le choc}} = \frac{mv_{Az}}{\sqrt{1 - v_A^2/c^2}} = \frac{mu_0}{\sqrt{1 - u_0^2/c^2}} \quad (7.25)$$

et :

$$(\mathbf{p}_{Bz})_{\text{avant le choc}} = \frac{mv_{Bz}}{\sqrt{1 - v_B^2/c^2}} = -\frac{mu_0}{\gamma_v \sqrt{1 - v_B^2/c^2}} \quad (7.26)$$

avec :

$$v_B^2 = (v_{Bz})^2 + (v_{Bx})^2 = \frac{u_0^2}{\gamma_v^2} + v^2 \quad (7.27)$$

d'où :

$$1 - v_B^2/c^2 = (1 - v^2/c^2)(1 - u_0^2/c^2) = \frac{(1 - u_0^2/c^2)}{\gamma_v^2} \quad (7.28)$$

et par conséquent :

$$(\mathbf{p}_{Az} + \mathbf{p}_{Bz})_{\text{avant le choc}} = 0 \quad (7.29)$$

De même, par symétrie $u_0 \longleftrightarrow -u_0$, on a $(\mathbf{p}_{Az} + \mathbf{p}_{Bz})_{\text{après le choc}} = 0$, et l'on retrouve bien ainsi la loi de conservation de la quantité de mouvement totale :

$$(\mathbf{p}_A + \mathbf{p}_B)_{\text{avant le choc}} = (\mathbf{p}_A + \mathbf{p}_B)_{\text{après le choc}} \quad (7.30)$$

C'est cette définition de \mathbf{p} que l'on considèrera par la suite.

7.2.2 Principe fondamental de la dynamique relativiste

Par analogie avec le lien entre quantités de mouvement classique $\mathbf{p}_{classique} = m\mathbf{v}$ et relativiste $\mathbf{p}_{relativiste} = \gamma m\mathbf{v}$, on peut réécrire le PFD sous une forme relativiste en faisant intervenir le temps (propre) mesuré dans le référentiel lié à l'objet en mouvement : $\frac{d\mathbf{p}}{d\tau} = \sum_i \mathbf{f}_i$, où les forces \mathbf{f}_i sont les *forces relativistes* associées aux forces classiques \mathbf{F}_i .

L'expression de ces forces \mathbf{f}_i peut s'obtenir simplement en considérant un référentiel galiléen particulier : le référentiel propre \mathcal{R}' de l'objet à un instant donné. La vitesse de l'objet étant nulle dans ce référentiel \mathcal{R}' , il n'y a aucun effet relativiste (au moins pendant un bref intervalle de temps) et le PFD s'exprime de façon classique sous la forme $\frac{d\mathbf{p}'}{dt'} = \sum_i \mathbf{F}'_i$, où les forces \mathbf{F}'_i sont les forces classiques usuelles. En utilisant le lien entre la coordonnée temporelle t' (qui est aussi égale au temps propre τ ici) et la coordonnée t , on peut montrer¹ que les forces relativistes \mathbf{f}_i s'expriment simplement en fonction des forces classiques \mathbf{F}_i sous la forme $\mathbf{f}_i = \gamma \mathbf{F}_i$. Comme $\frac{d\mathbf{p}}{d\tau} = \gamma \frac{d\mathbf{p}}{dt}$, le PFD garde alors une forme particulièrement simple en relativité restreinte :

$$\frac{d\mathbf{p}}{dt} = \sum_i \mathbf{F}_i \quad (7.31)$$

pour un objet ponctuel de masse m soumis, dans \mathcal{R} , à des forces \mathbf{F}_i , ces forces conservant leur expression classique habituelle.

On peut aussi faire apparaître explicitement le vecteur accélération usuel $\mathbf{a} = \frac{d\mathbf{v}}{dt}$ en dérivant $\mathbf{p} = \gamma m\mathbf{v}$ par rapport au temps t :

$$\frac{d\mathbf{p}}{dt} = \gamma m \mathbf{a} + m\mathbf{v} \frac{d\gamma}{dt} = \sum_i \mathbf{F}_i \quad (7.32)$$

qui ne redonne l'expression (presque classique) $\gamma m \mathbf{a} = \sum_i \mathbf{F}_i$ que pour des mouvements dont la norme de la vitesse $v = \|\mathbf{v}\|$ reste constante.

Un exemple important est le mouvement d'une charge q évoluant dans un champ électromagnétique (\mathbf{E}, \mathbf{B}) . Son mouvement obéit alors à l'équation suivante :

$$\frac{d\gamma m\mathbf{v}}{dt} = \gamma m \frac{d\mathbf{v}}{dt} + m\mathbf{v} \frac{dv}{dt} \frac{v}{c^2 (1 - v^2/c^2)^{3/2}} = q(\mathbf{E} + \mathbf{v} \wedge \mathbf{B}) \quad (7.33)$$

qui diffère de l'expression classique usuelle par une correction relativiste d'un facteur γ et par un terme supplémentaire dépendant de la variation temporelle de la norme de la vitesse $\frac{dv}{dt}$.

1. L'expression relativiste du PFD et des forces agissant sur un objet dans un référentiel quelconque peut s'obtenir beaucoup plus directement en utilisant la notion de quadrivecteur détaillée à la section 7.3.5.

7.3 ÉNERGIE RELATIVISTE

7.3.1 Définition de l'énergie cinétique relativiste

Elle est définie, comme dans le cas classique, par le théorème de l'énergie cinétique (Equation 7.5). On montre dans l'encart 7.4 que cela conduit à prendre pour l'énergie cinétique relativiste l'expression :

$$E_c = mc^2 (\gamma_v - 1) \quad (7.34)$$

Aux faibles vitesses, le facteur $\gamma_v - 1 = \frac{1}{2}v^2/c^2 + \dots$ redonne bien l'expression classique :

$$E_c \simeq \frac{1}{2}mv^2 \quad \text{si } \|\mathbf{v}\| \ll c \quad (7.35)$$

à des corrections relativistes près.

Au contraire, aux grandes vitesses (c'est-à-dire pour v approchant c), la quantité $\frac{1}{2}mv^2$ « sature » à $\frac{1}{2}mc^2$ alors que l'énergie cinétique relativiste continue à augmenter indéfiniment :

$$\lim_{v \rightarrow c} (mc^2 (\gamma_v - 1)) = +\infty \quad (7.36)$$



7.4 Dérivation de l'expression de l'énergie cinétique relativiste

Considérons le mouvement parallèle à l'axe (Ox), dans un référentiel donné, d'un objet de masse m initialement au repos. Le théorème de l'énergie cinétique s'écrit alors :

$$\Delta E_c = \int F dx \quad (7.37)$$

c'est-à-dire :

$$\Delta E_c = E_c = \int \frac{d}{dt} (\gamma_v mv) dx = \int \frac{d}{dt} (\gamma_v mv) v dt \quad (7.38)$$

Exprimons v en fonction de γ_v en inversant la relation

$$\gamma_v = 1/\sqrt{1 - v^2/c^2} \quad : \quad v = \frac{c}{\gamma_v} \sqrt{\gamma_v^2 - 1}. \quad (7.39)$$

Ce qui conduit à :

$$\frac{d}{dt} (\gamma_v mv) = mc \frac{\gamma_v}{\sqrt{\gamma_v^2 - 1}} \frac{d\gamma_v}{dt} \quad (7.40)$$

d'où :

$$W(\mathbf{F}) = \int mc \frac{\gamma_v}{\sqrt{\gamma_v^2 - 1}} \frac{d\gamma_v}{dt} \frac{c}{\gamma_v} \sqrt{\gamma_v^2 - 1} dt = mc^2 \int d\gamma_v = mc^2 (\gamma_v - 1) \quad (7.41)$$

7.3.2 Application : accélérateur électrostatique de particules

Considérons à titre d'exemple d'application d'utilisation de l'énergie cinétique relativiste le problème d'une particule chargée (masse m , charge q), initialement au repos à l'origine du référentiel \mathcal{R} et soumise à un champ électrostatique $\mathbf{E} = E\mathbf{e}_x$ uniforme. Le théorème de l'énergie cinétique relativiste s'écrit :

$$W(qE\mathbf{e}_x) = \int qEdx = \Delta E_c = mc^2(\gamma_v - 1) \quad (7.42)$$

avec $V(x) = Ex$ le potentiel électrostatique au point d'abscisse x . On obtient ainsi : $\gamma_v = 1 + \frac{qEx}{mc^2}$ ou :

$$v = \frac{c}{\gamma_v} \sqrt{\gamma_v^2 - 1} = \frac{c}{1 + \frac{qEx}{mc^2}} \sqrt{\left(1 + \frac{qEx}{mc^2}\right)^2 - 1} \quad (7.43)$$

à comparer à l'expression classique (obtenue selon le même raisonnement) :

$$v_{classique} = \sqrt{\frac{2qEx}{m}}. \quad (7.44)$$

Comme attendu, l'expression relativiste de la vitesse sature à c , la vitesse de la lumière, alors que l'expression classique n'a aucune limite supérieure (figures 7.2 et 7.3). Cet effet de saturation est constamment vérifié dans les accélérateurs de particules.

Ce type d'accélérateur est utilisé dans le domaine de la médecine ou de l'industrie (utilisant les rayons X produits lors du brusque freinage des électrons issus de l'accélérateur pour faire de l'imagerie par exemple).

7.3.3 Énergie interne d'un système composite

Considérons un système composite, c'est-à-dire formé de sous-parties en interaction mutuelle, par exemple un noyau formé de nucléons, un atome ou une molécule formés de noyaux et d'électrons, un gaz ou un solide formés de molécules. L'énergie interne E_{int} d'un système composite est l'énergie totale du système lorsqu'il est globalement au repos et non soumis à des forces extérieures. Elle caractérise son état, c'est-à-dire l'agencement de ses composants et l'intensité des interactions entre ceux-ci.

L'énergie interne joue un rôle important dans le premier principe de la thermodynamique et dans la description des réactions chimiques. Il y a variation d'énergie interne lors d'une transformation du système qui change l'agencement de ses composants (lors de réactions chimiques par exemple), mais aussi lorsque sa température

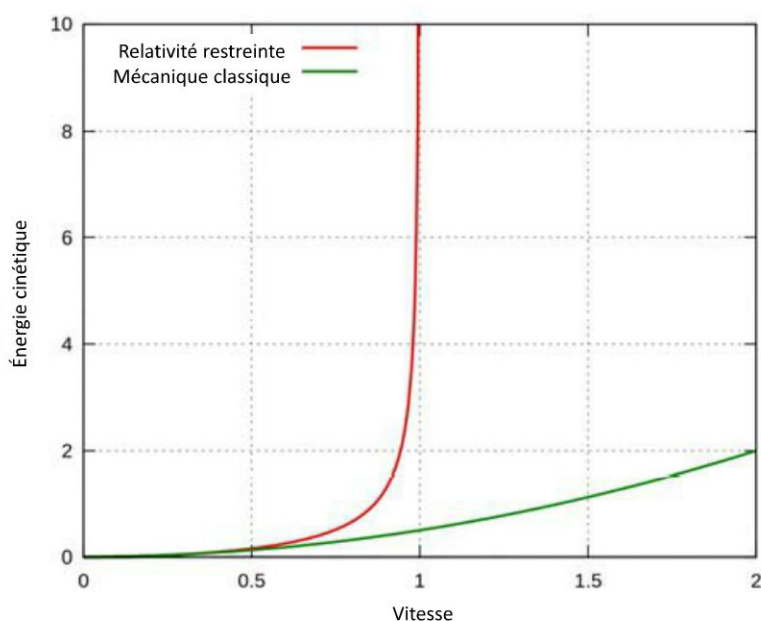


Figure 7.2 - Variation de l'énergie cinétique (relativiste et classique) en fonction de v/c .
(source libre : http://en.wikipedia.org/wiki/Tests_of_relativistic_energy_and_momentum)

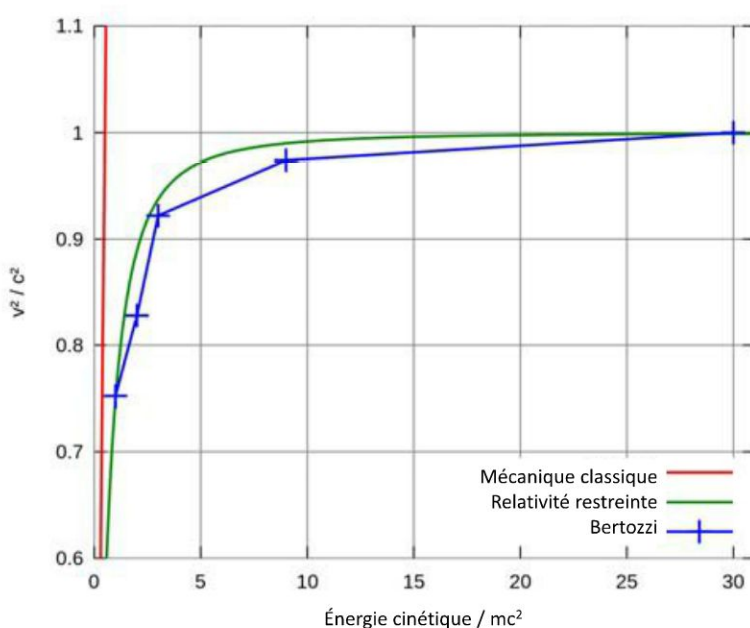


Figure 7.3 - Résultats de l'expérience de Bertozzi en 1964 démontrant l'effet de saturation des vitesses d'électrons accélérés.

(source libre : http://en.wikipedia.org/wiki/Tests_of_relativistic_energy_and_momentum)

varie (il y a variation de l'énergie cinétique liée à l'agitation thermique). Du point de vue de la physique classique, la masse totale du système est conservée dans de telles transformations du système. En chimie, c'est la fameuse loi de Lavoisier.

Prenons l'exemple de la réaction chimique de formation de H_2O à partir de H_2 et O_2 : elle se produit par réorganisation des liaisons entre les atomes d'hydrogène et d'oxygène qui constituent le système composite, et entraîne donc une variation de son énergie interne (mesurée à $1,2 \times 10^8 J$ pour une mole de H_2 et une demie mole de O_2). La conservation de l'énergie totale au cours de la transformation implique que la réaction s'accompagne soit d'un fort dégagement de chaleur (augmentation de l'énergie d'agitation thermique) soit d'une explosion, c'est-à-dire d'une grande énergie cinétique de la vapeur d'eau produite, utilisée par exemple dans la propulsion des fusées. Par contre, la loi de Lavoisier nous indique que la masse totale des composants est conservée dans une telle réaction chimique.

D'un point de vue expérimental, on a par contre observé que lors des réactions nucléaires, qui changent l'agencement des nucléons et donc l'énergie interne du système, il n'y avait pas conservation de la masse totale. Nous y reviendrons plus en détail dans le chapitre 8.

Les physiciens ont été amenés à poser qu'à toute variation de l'énergie interne est associée une variation de la masse totale du système, c'est-à-dire de sa capacité à être mis en mouvement par une force extérieure. Plus précisément, la relation entre variation d'énergie interne ΔE_{int} et variation de masse associée ΔM s'écrit :

$$\Delta E_{int} = \Delta M c^2 \quad (7.45)$$

ce qui aboutit, en prenant un choix commode de l'origine des énergies à la célèbre loi d'Einstein d'équivalence entre masse et énergie d'un corps au repos :

$$E_{int} = M c^2 \quad (7.46)$$

où M est la masse totale du système composite qui intervient par exemple dans l'expression (7.17) de la quantité de mouvement, et caractérise l'inertie du système. Cette expression a fait couler beaucoup d'encre depuis sa première formulation en 1905 par Einstein. Bien qu'il n'y ait encore à ce jour aucune démonstration rigoureuse satisfaisante², cette expression est justifiée par ses conséquences expérimentales et par sa cohérence au sein de l'ensemble de la théorie de la relativité. On trouvera dans l'encart 7.5 la description d'une expérience de pensée qui rend plausible dans un cas particulier simple cette relation.

Il est important de noter que la formule (7.46) est valable pour un système au repos : l'équivalence masse-énergie est le seul effet relativiste qui se manifeste à vitesse nulle. Il n'en est donc que plus étonnant, même si la variation de masse est la plupart du temps extrêmement faible : lors de la réaction chimique de formation de l'eau

2. Lire à ce sujet l'article de E. Hecht « How Einstein confirmed $E_0 = mc^2$ », *American Journal of Physics*, **79** p. 591-600 (2011).

à partir des molécules de dioxygène et de dihydrogène, extrêmement exothermique, que nous avons évoquée plus haut, la variation d'énergie interne est de $8,2 \times 10^{-19}$ J, donc celle de masse est de $0,9 \times 10^{-35}$ kg lorsqu'on fait réagir deux molécules de H_2 avec une molécule de O_2 , soit une variation relative de quelques 3×10^{-10} . Elle a donc échappé aux mesures des chimistes. Elle est par contre parfaitement observable dans les réactions nucléaires : lors de la réaction de fusion d'un noyau de deutérium avec un noyau de tritium qui produit un noyau d'hélium et un neutron libre, la variation d'énergie interne est de $2,8 \times 10^{-12}$ J, celle de masse est donc de 3×10^{-29} kg, soit une variation relative de 4×10^{-3} , donc plus grande par sept ordres de grandeur.

ENCART 7.5 Expérience de pensée montrant qu'une variation d'énergie interne entraîne une variation de la masse du système composite

Considérons un système physique C constitué par un corps composite de masse M . Son état, et donc son énergie interne, est susceptible de changer lorsqu'il absorbe de la lumière. Par exemple, il s'échauffe et sa température change, ou bien il change de niveau d'énergie. Nous allons voir que sa masse doit aussi changer lors d'un tel changement d'énergie interne.

Considérons le dispositif de la figure 7.4, où le corps C se trouve à l'origine des coordonnées et où deux sources lumineuses S_1 et S_2 , situées sur l'axe Oy d'un référentiel \mathcal{R} en deux points symétriques par rapport à O , émettent au même instant deux impulsions lumineuses identiques d'énergie E_{lum} en direction de O . Le corps C absorbe alors simultanément ces deux impulsions et change donc d'énergie interne. Pour les besoins de la démonstration, il faut savoir qu'une impulsion de lumière d'énergie E_{lum} transporte aussi une quantité de mouvement \mathbf{p}_{lum} dans la direction de sa propagation et de module $p_{lum} = E_{lum}/c$.

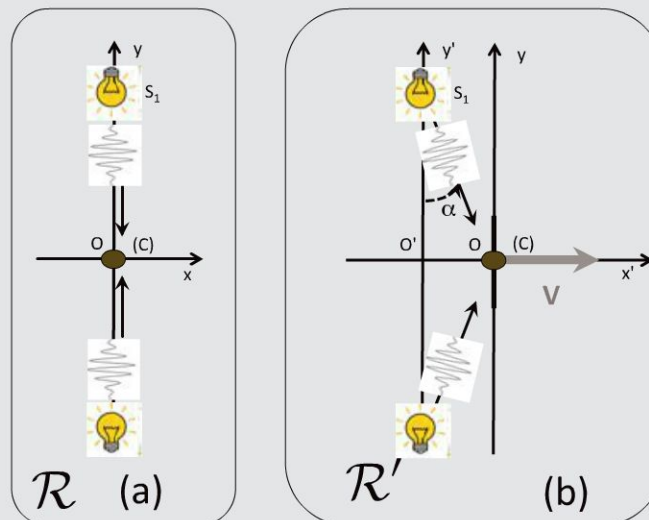


Figure 7.4 – Un corps C de masse M absorbe deux impulsions lumineuses venant de directions opposées (a) dans le référentiel où sources lumineuses et corps sont au repos (b) dans un référentiel où l'observateur a une vitesse $-V$ par rapport au corps.

Plaçons-nous d'abord dans le référentiel \mathcal{R} où le corps et les deux sources lumineuses sont immobiles. Son énergie interne passe de E_{int1} à E_{int2} , et sa quantité de mouvement de $\mathbf{p}_1 = \mathbf{0}$ à \mathbf{p}_2 . Dans ce référentiel, la quantité de mouvement totale \mathbf{p}_{lum} des deux impulsions se propageant en sens inverse est nulle. La conservation de la quantité de mouvement dans cette transformation s'écrit alors :

$$\mathbf{p}_2 = \mathbf{p}_{lum} + \mathbf{p}_1 \quad (7.47)$$

donc $\mathbf{p}_2 = \mathbf{0}$: le corps C n'est pas mis en mouvement lors de l'absorption. La conservation de l'énergie totale, réduite dans \mathcal{R} à la seule énergie interne implique alors :

$$E_{int2} = E_{int1} + 2E_{lum} \quad (7.48)$$

Plaçons-nous maintenant dans un référentiel \mathcal{R}' en mouvement le long de Ox de vitesse $-V$ par rapport à \mathcal{R} . Dans ce référentiel, les sources lumineuses et le corps C se déplacent à la vitesse V , et à cause du phénomène d'aberration que nous avons étudié dans le paragraphe 5.2.2, les deux impulsions se propagent selon une direction qui fait un angle α non nul avec l'axe $O'y'$. Leurs quantités de mouvement ne sont donc plus exactement opposées, et leur somme acquiert une composante non nulle le long de l'axe Ox . On suppose maintenant que le mouvement est extrêmement lent, de sorte qu'on peut utiliser les lois classiques de changement de référentiel et l'expression classique de la quantité de mouvement, qui vaut donc MV pour le corps C dans le référentiel \mathcal{R}' . L'angle α est dans ce cas très petit, et égal à V/c (voir section 5.2.2). La conservation de la composante selon Ox de la quantité de mouvement lors de l'absorption des impulsions s'écrit alors :

$$p_{2x} = MV + 2p_{lum} \sin \alpha = MV + 2E_{lum}/c \times V/c = MV + (E_{int2} - E_{int1})V/c^2 \quad (7.49)$$

La conservation de la quantité de mouvement dans le référentiel \mathcal{R} nous a montré que la vitesse du corps C n'est pas changée lors de l'absorption. C'est donc aussi vrai dans \mathcal{R}' , où elle reste à la valeur V . On est donc tenté d'écrire $p_{2x} = MV$, ce qui est incompatible avec (7.49). Il faut donc se résoudre à envisager que la masse M varie dans la transformation et passe de la valeur M à la valeur $M + \Delta M$. (7.49) implique donc :

$$(M + \Delta M)V = MV + (E_{int2} - E_{int1})V/c^2 \quad (7.50)$$

D'où finalement :

$$\Delta M = (E_{int2} - E_{int1})/c^2 \quad (7.51)$$

7.3.4 Énergie totale relativiste

En relativité restreinte comme en mécanique classique, l'énergie totale est définie comme la somme des différents types d'énergie : interne, cinétique et potentielle :

$$E_{totale} = E_{int} + E_{cin} + E_{pot} \quad (7.52)$$

c'est-à-dire :

$$E_{totale} = Mc^2 + Mc^2(\gamma_v - 1) + E_{pot} = \gamma_v Mc^2 + E_{pot} \quad (7.53)$$

pour un système (composite ou non) de masse M . Dans cette expression, le terme E_{pot} fait, par exemple, référence à l'énergie potentielle de gravitation, ou, de façon plus générale, à l'énergie potentielle associée à toute force extérieure agissant sur le système.

Il est à noter que, parfois, et bien que cela puisse être source de confusion, on introduit la notion de « masse en mouvement », ou « masse relativiste », $M' = \gamma_v M$, l'énergie totale s'écrivant alors $M'c^2 + E_{pot}$.

L'expression de l'énergie totale relativiste est très souvent utilisée sous la forme suivante (faisant le lien entre E_{totale} et \mathbf{p}) :

$$(E_{totale} - E_{pot})^2 - \mathbf{p}^2 c^2 = M^2 c^4 \quad (7.54)$$

puisque $\mathbf{p}^2 c^2 = \gamma_v^2 M^2 \mathbf{v}^2 c^2 = (1 - \gamma_v^2) M^2 c^4$.

Plus simplement, pour un système de masse M non soumis à des forces extérieures, on obtient l'importante relation :

$$\boxed{E_{totale}^2 - \mathbf{p}^2 c^2 = M^2 c^4} \quad (7.55)$$

qui reste valable pour des particules de masse nulle (comme le photon) pour lesquelles $E = pc$.

7.3.5 Invariants relativistes et notion de quadrivecteur

Les quantités M et c ne dépendant pas du référentiel choisi, contrairement aux expressions de E_{totale} et \mathbf{p} , la relation précédente conduit à l'existence d'une quantité invariante par changement de référentiel :

$$E_{totale}^2 - \mathbf{p}^2 c^2 = E_{totale}'^2 - \mathbf{p}'^2 c^2 = M^2 c^4 = \text{invariant} \quad (7.56)$$

De la même façon, on a déjà vu (section 3.3.5) que le carré de l'intervalle d'espace-temps Δs^2 gardait la même valeur quelque soit le référentiel choisi pour exprimer Δx et Δt . On dit que c , M , Δs^2 et $E^2 - \mathbf{p}^2 c^2$ sont des *invariants relativistes*.

Remarque

Dans le cas de la lumière (ou de façon générale, pour toute particule de masse nulle) l'invariant $E^2 - \mathbf{p}^2 c^2$ est nul.

Cette notion d'*invariance de valeur*, quelque soit le référentiel, peut se généraliser à la notion d'*invariance de forme* (qu'on appelle la *covariance*) lorsqu'une quantité peut être définie indépendamment du référentiel. Cela en est ainsi, par exemple, des vecteurs usuels à trois dimensions qui peuvent être définis indépendamment du système de coordonnées choisi pour exprimer leurs trois composantes. Dans l'espace-temps quadri-dimensionnel de la relativité restreinte, on ajoute une composante temporelle aux vecteurs spatiaux usuels pour former des *quadrivecteurs* (paragraphe 3.3.4). Ainsi, le quadrivecteur position se lit : (ct, x, y, z) . Ces quadrivecteurs forment un espace vectoriel, appelé *espace de Minkowski*, mais les règles de calcul (par exemple celle du produit scalaire) y sont différentes de l'espace vectoriel usuel à trois dimensions (encart 7.6).

Pour qu'il y ait bien covariance, on peut montrer alors que les quatre composantes de ces quadrivecteurs doivent obéir aux mêmes transformations de Lorentz que les coordonnées d'espace-temps dans un changement de référentiel. Très souvent, les relations obtenues permettent de retrouver – presque sans calcul – les effets usuels de mécanique relativiste, comme l'effet Doppler ou la loi de composition des vitesses.

Exemple

Quadri-vitesse (définie comme la dérivée de la quadriposition (ct, x, y, z) par rapport au temps propre τ) :

$$(\gamma_v c, \gamma_v v_x, \gamma_v v_y, \gamma_v v_z) \quad (7.57)$$

dont la transformation par changement de référentiel permet de retrouver la loi de composition des vitesses.

En effet, soit un objet ponctuel mobile de vecteur vitesse \mathbf{v} dans le référentiel galiléen \mathcal{R} et de vecteur vitesse \mathbf{v}' dans le référentiel galiléen \mathcal{R}' en déplacement uniforme par rapport à \mathcal{R} à la vitesse $\mathbf{V} = V\mathbf{e}_x$ suivant l'axe (Ox) . Soient aussi γ_v , $\gamma_{v'}$ et γ_V les facteurs de Lorentz associés respectivement aux vitesses \mathbf{v} , \mathbf{v}' et V . On passe alors des coordonnées $(\gamma_v c, \gamma_v v_x, \gamma_v v_y, \gamma_v v_z)$ aux coordonnées $(\gamma_{v'} c, \gamma_{v'} v'_x, \gamma_{v'} v'_y, \gamma_{v'} v'_z)$ par la transformation de Lorentz suivante :

$$\begin{cases} \gamma_{v'} c &= \gamma_v \gamma_V \left(\gamma_v c - \frac{V}{c} \gamma_v v_x \right) \\ \gamma_{v'} v'_x &= \gamma_v \gamma_V \left(\gamma_v v_x - \frac{V}{c} \gamma_v c \right) \\ \gamma_{v'} v'_y &= \gamma_v v_y \\ \gamma_{v'} v'_z &= \gamma_v v_z \end{cases} \quad (7.58)$$

dont la première relation $\gamma_{v'} = \gamma_V \gamma_v (1 - Vv_x/c^2)$, reportée dans les autres équations, redonne bien les formules de composition relativiste des vitesses vues au chapitre 5.

Exemple

Quadri-impulsion (définie comme la quadri-vitesse multipliée par la masse m) :

$$(\gamma_v mc, \gamma_v mv_x, \gamma_v mv_y, \gamma_v mv_z) \quad (7.59)$$

Exemple

Quadrivecteur d'onde :

$$(\omega/c, k_x, k_y, k_z) \quad (7.60)$$

dont la transformation par changement de référentiel permet de retrouver les formules des effets Doppler et d'aberration vues au chapitre 5.



7.6 Espace-temps de Minkowski et tenseur métrique

L'ensemble des quadrivecteurs forme un espace vectoriel (à 4D) appelé espace de Minkowski, où le produit scalaire de l'espace usuel à 3D est remplacé par le *pseudo-produit scalaire* :

$$(\mathbf{a}, \mathbf{b}) = a_0 b_0 - a_1 b_1 - a_2 b_2 - a_3 b_3 \quad (7.61)$$

La norme usuelle d'un vecteur est alors transformée en une *pseudo-norme* telle que :

$$\|\mathbf{a}\|^2 = (\mathbf{a}, \mathbf{a}) = a_0^2 - a_1^2 - a_2^2 - a_3^2 \quad (7.62)$$

Exemple

Δs^2 est la pseudo-norme du quadrivecteur $(c\Delta t, \Delta x, \Delta y, \Delta z)$.

Une pseudo-norme au carré peut être nulle ou négative sans que le quadrivecteur correspondant soit nul (cf, par exemple, la quadri-impulsion d'un photon). L'espace-temps de Minkowski n'est donc pas euclidien, comme l'espace habituel des positions.

On peut réécrire cette pseudo-norme (au carré) sous la forme d'une multiplication matricielle :

$$\Delta s^2 = (c\Delta t, \Delta x, \Delta y, \Delta z) \begin{pmatrix} 1 & & & \\ & -1 & & \\ & & -1 & \\ & & & -1 \end{pmatrix} \begin{pmatrix} c\Delta t \\ \Delta x \\ \Delta y \\ \Delta z \end{pmatrix} \quad (7.63)$$

D'une façon générale, en termes de composantes, le pseudo-produit scalaire des quadrivecteurs \mathbf{a} et \mathbf{b} s'écrit :

$$(\mathbf{a}, \mathbf{b}) = \sum_{\mu, \nu=0,1,2,3} g_{\mu\nu} a^\mu b^\nu \quad (7.64)$$

où $g_{\mu\nu}$ est l'élément (μ, ν) de la matrice diagonale apparaissant dans l'expression du ds^2 ci-dessus.

La matrice $g_{\mu\nu}$ porte le nom de *tenseur métrique*, et l'ensemble des quatre signes (+ − − −) de ses termes diagonaux $g_{\mu\mu}$ porte le nom de *signature de la métrique*.

En relativité générale, que l'on abordera au chapitre 9, les facteurs $g_{\mu\nu}$ n'ont plus une forme aussi simple.

7.4 COLLISIONS RELATIVISTES

7.4.1 Méthode générale : lois de conservation

En mécanique non relativiste, les collisions entre particules se divisent en deux catégories : élastiques et inélastiques. Dans les collisions élastiques, l'énergie et la quantité de mouvement totales sont conservées. Dans les collisions inélastiques, seule la quantité de mouvement totale est conservée. L'énergie (non relativiste) n'est pas conservée, une partie de l'énergie cinétique pouvant se transformer en énergie interne (en chaleur par exemple). De même, une partie de l'énergie interne peut se convertir en énergie cinétique (comme cela se passe dans une explosion ou lors de la propulsion d'un véhicule).

En mécanique relativiste, du fait que l'énergie interne est incluse dans la définition de l'énergie totale, celle-ci est conservée dans les deux types de collisions, élastiques et inélastiques :

$$\left(\sum_i \gamma_i m_i c^2 \right)_{\text{avant}} = \left(\sum_j \gamma'_j m'_j c^2 \right)_{\text{après}} \quad (7.65)$$

$$\left(\sum_i E_{\text{totale},i} \right)_{\text{avant}} = \left(\sum_j E_{\text{totale},j} \right)_{\text{après}} \quad (7.66)$$

où la somme s'effectue sur l'ensemble des particules présentes, ainsi que les trois composantes de la quantité de mouvement totale :

$$\left(\sum_i \gamma_i m_i \mathbf{v}_i \right)_{\text{avant}} = \left(\sum_j \gamma'_j m'_j \mathbf{v}'_j \right)_{\text{après}} \quad (7.67)$$

$$\left(\sum_i \mathbf{p}_i \right)_{\text{avant}} = \left(\sum_j \mathbf{p}'_j \right)_{\text{après}} \quad (7.68)$$

Ces quatre relations ainsi que la considération des invariants (7.56) pour chacune des particules impliquées dans la collision permettent alors de déterminer les masses

et vitesses des particules après la collision en fonction des masses et vitesses des particules avant la collision (voir les exemples ci-dessous).

D'autres quantités sont aussi conservées : la charge totale, par exemple, ainsi que d'autres *nombres quantiques* caractérisant les particules (nombre baryonique, isospin, saveur...), mais nous ne les détaillerons pas ici.

Les trois cas de collisions inélastiques relativistes présentées ci-après sont chacun représentatifs d'une classe entière de réactions : une fusion de deux particules en une seule et son opposé, la désintégration d'une particule en deux particules, et enfin un exemple de diffusion de deux particules.

7.4.2 Fusion de l'hydrogène « 2 particules → 1 particule »

Sans entrer dans les détails sur la notion de fusion nucléaire, qui sera vue au chapitre suivant, nous pouvons examiner un exemple typique de collision relativiste intervenant dans un processus bien connu de fusion nucléaire : celle de l'hydrogène qui se produit à l'intérieur des étoiles (tel notre soleil).

En effet, l'équation bilan de la fusion d'un noyau de deutérium et d'un noyau de tritium conduisant à la production d'un noyau d'hélium et d'un neutron est particulièrement simple (quatre corps seulement) et s'écrit :



Pour donner une idée du calcul, simplifions à l'extrême le traitement de cette collision inélastique et considérons la réaction de fusion suivante : deux particules de même masse se percutent et donnent naissance à une troisième particule.

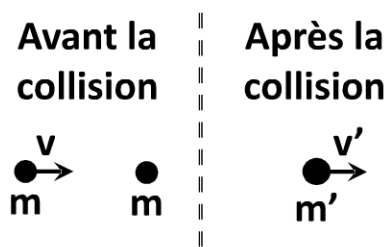


Figure 7.5 – Illustration d'une collision inélastique où deux particules de masse m fusionnent pour ne former qu'un seul corps de masse m' .

Avant la collision, on a donc une masse m au repos et une autre masse m venant la percuter à la vitesse v . Après la collision inélastique, on n'a plus qu'une seule masse m' allant à la vitesse v' (selon l'axe défini par la vitesse v , voir la figure 7.5). Le problème est donc à une seule dimension d'espace, ce qui simplifie beaucoup les calculs.

Commençons par le calcul de cette collision dans le cadre de la mécanique classique. La masse m' est alors simplement égale à la somme des deux masses m , et l'on

obtient :

$$\begin{cases} p_{\text{avant}} = mv \\ E_{\text{avant}} = \frac{1}{2}mv^2 \end{cases} \quad \text{et} \quad \begin{cases} p_{\text{après}} = m'v' = 2mv' \\ E_{\text{après}} = \frac{1}{2}m'v'^2 = mv'^2 \end{cases} \quad \text{puisque} \quad m' = 2m \quad (7.70)$$

La loi de conservation de la quantité de mouvement s'écrit alors $mv = 2mv'$, d'où :

$$v' = v/2 \quad (7.71)$$

ce qui, reporté dans l'expression de l'énergie cinétique, conduit à la relation suivante :

$$E_{\text{cin,avant}} = \frac{1}{2}mv^2 \neq E_{\text{cin,après}} = \frac{1}{4}mv^2 \quad (7.72)$$

traduisant le fait que la collision n'est pas élastique : une partie Q de l'énergie cinétique (énergie externe) s'est transformée en énergie interne (énergie stockée, par exemple pour un corps étendu, sous forme de chaleur, de déformations, de rotation...).

Il y a bien sûr conservation de l'énergie *totale* lors de cette collision :

$$E_{\text{avant}} = \frac{1}{2}mv^2 = E_{\text{après}} = \frac{1}{4}mv^2 + Q \quad (7.73)$$

où Q est l'énergie d'excitation du système de masse $m' = 2m$.

Passons au calcul relativiste. Les différentes formes de l'énergie (cinétique et interne) sont contenues alors dans l'expression de E , et l'on a :

$$\begin{cases} p_{\text{avant}} = \gamma_v mv \\ E_{\text{avant}} = (1 + \gamma_v) mc^2 \end{cases} \quad \text{et} \quad \begin{cases} p_{\text{après}} = \gamma_{v'} m' v' \\ E_{\text{après}} = \gamma_{v'} m' c^2 \end{cases} \quad (7.74)$$

Les lois de conservation de la quantité de mouvement et de l'énergie conduisent alors respectivement aux deux relations suivantes :

$$v' = \frac{\gamma_v mv}{\gamma_{v'} m'} \quad \text{et} \quad (1 + \gamma_v) m = \gamma_{v'} m' \quad (7.75)$$

d'où l'expression de la vitesse v' :

$$v' = \frac{\gamma_v}{1 + \gamma_v} v \quad (7.76)$$

qui est, contrairement au cas classique, supérieure à $v/2$ puisque γ_v est supérieur à 1.

Quant à la masse m' , elle s'obtient grâce à l'invariant relativiste liant E et p :

$$m'^2 c^4 = E^2 - p^2 c^2 = (1 + \gamma_v)^2 m^2 c^4 - \gamma_v^2 m^2 v^2 c^2 \quad (7.77)$$

c'est-à-dire :

$$m' = m \sqrt{1 + 2\gamma_v + \gamma_v^2 - \gamma_v^2 v^2/c^2} = 2m \sqrt{\frac{1 + \gamma_v}{2}} \geq 2m \quad (7.78)$$

qui traduit simplement le fait que l'énergie de masse $m'c^2$ est constituée de la somme des deux énergies de masse mc^2 ainsi que de l'énergie interne d'excitation Q .

Cet exemple simple permet de mettre en évidence la propriété fondamentale de non-conservation de la masse lors de collisions inélastiques.

7.4.3 Désintégration du pion Π^0 « 1 particule \rightarrow 2 particules »

Un pion Π^0 peut se désintégrer en deux photons γ d'énergies E_1 et E_2 comme indiqué sur le schéma de la figure 7.6. Les directions d'émission de ces photons sont notées θ_1 et θ_2 . Le problème est donc à deux dimensions d'espace. On cherche, par exemple, à établir la relation entre la masse m du pion (supposée inconnue) et les quatre grandeurs mesurables E_1 , E_2 , θ_1 et θ_2 .

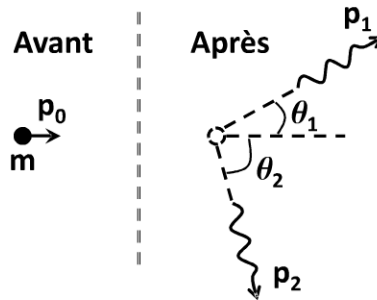


Figure 7.6 - Désintégration d'une particule en deux particules : schéma représentant les trajectoires.

Les relations de conservation de l'énergie et de la quantité de mouvement totales s'écrivent :

$$\mathbf{p}_0 = \mathbf{p}_1 + \mathbf{p}_2 \quad (7.79)$$

$$E_0 = E_1 + E_2 \quad (7.80)$$

où E_0 est l'énergie du pion, et \mathbf{p}_0 , \mathbf{p}_1 , \mathbf{p}_2 les quantités de mouvement des trois particules.

Une façon simple d'éliminer \mathbf{p}_0 et E_0 de ces deux équations est alors d'utiliser la relation fondamentale entre \mathbf{p}_0 et E_0 :

$$\begin{aligned} m^2 c^4 &= E_0^2 - \mathbf{p}_0^2 c^2 \\ &= (E_1 + E_2)^2 - (\mathbf{p}_1 + \mathbf{p}_2)^2 c^2 \end{aligned} \quad (7.81)$$

Comme $E_1 = p_1 c$ et $E_2 = p_2 c$ pour les photons, on obtient :

$$\begin{aligned} m^2 c^4 &= (E_1 + E_2)^2 - [(E_1 \cos \theta_1 + E_2 \cos \theta_2)^2 + (E_1 \sin \theta_1 - E_2 \sin \theta_2)^2] \\ &= 2E_1 E_2 (1 - \cos \theta_1 \cos \theta_2 + \sin \theta_1 \sin \theta_2) \end{aligned} \quad (7.82)$$

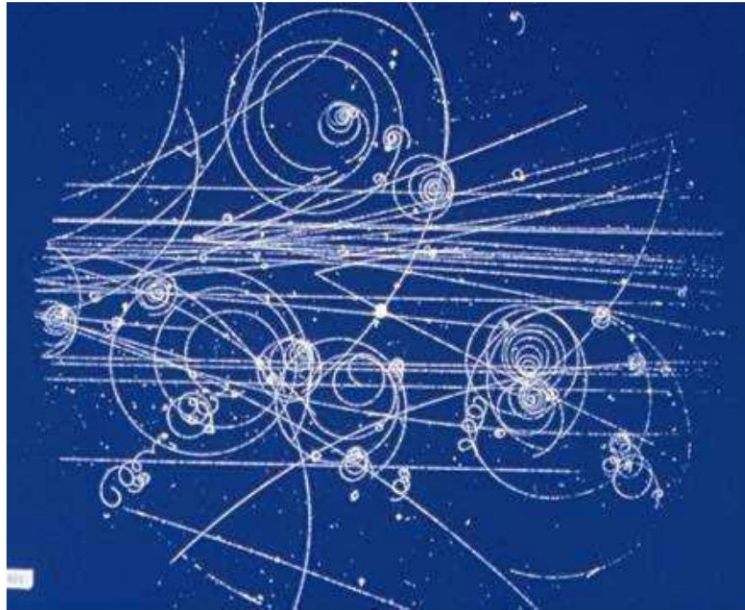


Figure 7.7 – Désintégration d'un pion négatif d'une énergie de 16 milliards d'électronvolts dans un jet de particules. Cet événement a été enregistré en 1970, avec la première chambre à bulles à hydrogène liquide en service au CERN à Genève. Le diamètre de cette chambre est de 30 cm. Sur la photo, les traces rectilignes indiquent le parcours des particules neutres, tandis que les traces courbes marquent la trajectoire de particules chargées en présence d'un champ magnétique.

(source : CERN : <http://cds.cern.ch/record/39474>)

d'où :

$$m = \sqrt{2E_1 E_2 (1 - \cos(\theta_1 + \theta_2))} / c^4 \quad (7.83)$$

Exemple

On mesure $E_1 = E_2 = 67 \text{ MeV}$ dans une expérience où l'on a choisi $\theta_1 = \theta_2 = \pi/2$. La masse du pion s'en déduit :

$$m(\pi^0) = 2E_1/c^2 = 134 \text{ MeV}/c^2 \quad (7.84)$$

Le pion π^0 étant électriquement neutre, il est particulièrement difficile à observer expérimentalement (comme dans une chambre à bulles par exemple, figure 7.7). Ce sont les photons (ainsi que les paires électrons-positrons) produits lors de sa désintégration qui permettent de déduire sa présence.

7.4.4 Désintégration du neutron « 1 particule \rightarrow n particules »

Le neutron est stable quand il entre dans la composition d'un noyau, mais instable quand il est libre. On observe alors qu'il se désintègre au bout d'un temps aléatoire dont la durée moyenne est de 900 s et qu'il donne naissance à deux particules : un

proton et un électron. L'expérience montre qu'on recueille des électrons dont l'énergie E_e peut prendre toutes les valeurs possibles, supérieures à son énergie au repos $m_e c^2$, où m_e est la masse de l'électron.

Les relations de conservation de l'énergie et de la quantité de mouvement s'écrivent dans ce cas, avec des notations évidentes :

$$E_n = E_p + E_e \quad ; \quad \mathbf{p}_n = \mathbf{p}_p + \mathbf{p}_e \quad (7.85)$$

Le neutron étant initialement au repos, les quantités de mouvement \mathbf{p}_p et \mathbf{p}_e du proton et de l'électron sont égales et opposées, et on a en outre $E_n = m_n c^2$. L'énergie E_e de l'électron vaut $E_e = \sqrt{m_e^2 c^4 + p_e^2 c^2}$. On a aussi $E_e = E_n - \sqrt{m_p^2 c^4 + p_p^2 c^2} = m_n c^2 - \sqrt{m_p^2 c^4 + p_e^2 c^2}$. Nous avons donc deux relations reliant les deux inconnues E_e et p_e aux quantités fixes et connues m_n , m_p et m_e . La résolution de ce système de deux équations va donner en particulier une valeur bien définie, que nous ne calculerons pas, à l'énergie E_e , ce qui contredit les résultats expérimentaux que nous avons évoqués au début.

Ce problème a été l'objet de maintes spéculations dans les débuts de la physique nucléaire, certains allant même jusqu'à rejeter les lois de conservation de l'énergie et de la quantité de mouvement pour les réactions nucléaires ! C'est W. Pauli qui a trouvé la solution en 1930, en faisant l'hypothèse que la désintégration du neutron donne naissance non pas à deux particules mais à trois, le proton, l'électron et une hypothétique particule non détectée dans les expériences. On rajoute ainsi de nouvelles relations de conservation qui aboutissent à une énergie variable de l'électron. Cette particule a été appelée ultérieurement « neutrino » (« petit neutron ») par E. Fermi. Elle joue un rôle très important dans de nombreuses réactions nucléaires. On retrouvera en particulier cette particule dans le chapitre 12 (section 12.4.2).

7.4.5 Diffusion Compton « 2 particules → 2 particules »

Bien qu'il ait une masse nulle, un photon transporte une énergie et une quantité de mouvement non nulles, comme le confirme l'effet Compton que nous allons maintenant présenter.

Un photon d'énergie hc/λ est diffusé, selon un angle θ , par un électron supposé au repos. Après la collision, le photon a l'énergie hc/λ' et l'électron a acquis la vitesse \mathbf{v} , dont la direction fait un angle φ par rapport à la direction de propagation du photon incident (figure 7.8). Dans l'expérience initiale de Compton, c'est la longueur d'onde λ' de la lumière diffusée qui est mesurée en fonction de l'angle θ .

La diffusion Compton est un exemple particulièrement instructif de collisions relativistes mélangeant des particules massives et non-massives.

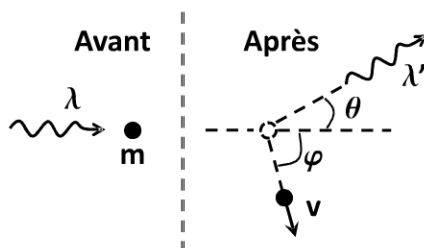


Figure 7.8 – Diffusion Compton : schéma des trajectoires.

Avant la collision, la quadri-impulsion totale s'écrit :

$$\left(\frac{h}{\lambda} + mc, \frac{h}{\lambda}, 0, 0 \right) \quad (7.86)$$

et après la collision elle devient :

$$\left(\frac{h}{\lambda'} + \gamma_v mc, \frac{h}{\lambda'} \cos \theta + \gamma_v mv \cos \varphi, \frac{h}{\lambda'} \sin \theta - \gamma_v mv \sin \varphi, 0 \right) \quad (7.87)$$

avec $\gamma_v = 1/\sqrt{1 - v^2/c^2}$.

La conservation de la quadri-impulsion conduit alors, après quelques lignes de calcul, au résultat suivant :

$$\lambda' - \lambda = \frac{h}{mc} (1 - \cos \theta) \quad (7.88)$$

donnant la variation de la longueur d'onde en fonction de l'angle de diffusion.

Particulièrement importante en astrophysique, la diffusion Compton *inverse* est une diffusion d'électrons par des photons (du milieu interstellaire par exemple). Par rapport au calcul précédent, l'échange des rôles de l'électron et du photon dans la collision n'est en rien problématique puisqu'en se plaçant dans le référentiel propre de l'électron on retrouve le cas de diffusion directe étudié ici.

RÉACTIONS NUCLÉAIRES ET NOTIONS DE RADIOACTIVITÉ

8

Le chapitre précédent a mis en évidence un lien fondamental entre masse et énergie : une variation ΔE de l'énergie interne d'un système se traduit par une variation ΔM de la masse de ce même système :

$$\Delta E = \Delta M \cdot c^2 \quad (8.1)$$

Cette équivalence est l'un des points centraux de la physique des noyaux atomiques, appelée *physique nucléaire*, qui regroupe l'étude de la structure et de la stabilité des noyaux et celle des réactions nucléaires (naturelles ou artificielles), les rayonnements émis (c'est-à-dire la *radioactivité*) et leurs effets sur les milieux biologiques (c'est-à-dire la *dosimétrie*).

Les applications de la physique nucléaire sont extrêmement nombreuses et quotidiennes : énergie et armes nucléaires, techniques d'imagerie et de traitement en médecine nucléaire, datation isotopique en géologie et archéologie...

8.1 STRUCTURE DES NOYAUX ATOMIQUES, ÉNERGIE DE LIAISON

8.1.1 Rappels sur les constituants d'un noyau atomique

La matière est constituée de molécules, c'est-à-dire de groupements d'atomes qui s'associent au moyen de liaisons (covalentes, ioniques, hydrogène...). On connaît, en 2015, 118 types d'atomes différents, allant de l'hydrogène ${}^1_1\text{H}$ (le plus léger) à l'ununoctium ${}^{294}_{118}\text{Uuo}$ (le plus lourd), et plusieurs autres éléments sont en passe d'être *synthétisés* dans des accélérateurs de particules (français, allemands, russes, américains...).

Bien que le mot *atome* (qui dérive du grec ancien *atomos*) signifie *insécable*, les atomes ne peuvent pas être considérés comme des constituants élémentaires indestructibles de la matière. Ils sont formés d'un noyau central très dense, entouré d'électrons. La taille typique du noyau est de quelques femtomètres ($1 \text{ fm} = 10^{-15} \text{ m}$), très inférieure à la distance moyenne électron-noyau qui est de l'ordre de quelques angströms ($1 \text{ Å} = 10^{-10} \text{ m}$). Les noyaux atomiques sont eux-mêmes constitués de

protons et de *neutrons*, qui sont les deux formes possibles d'une autre particule appelée *nucléon*. Les nucléons sont également sous-structurés, sous forme d'assemblages de *quarks* que nous aborderons au prochain chapitre.

Les masses du proton et du neutron sont très proches :

$$\begin{aligned} m_p &\simeq 1,672622 \cdot 10^{-27} \text{ kg} \\ m_n &\simeq 1,674927 \cdot 10^{-27} \text{ kg} \end{aligned} \quad (8.2)$$

contrairement à leurs charges électriques qui sont très différentes : le neutron n'a pas de charge électrique (d'où son nom) alors que le proton a une charge électrique positive, égale à l'inverse de la charge d'un électron.

$$q_p = +e \simeq 1,602177 \cdot 10^{-19} \text{ C} \quad (8.3)$$

$$q_n = 0 \text{ C} \quad (8.4)$$

Les protons ayant la même charge électrique, ils ont tendance à se repousser du fait de l'interaction coulombienne. Cependant, à l'échelle du noyau, celle-ci est bien moins intense que l'*interaction forte* qui attire mutuellement les nucléons et assure la stabilité des noyaux atomiques. L'interaction forte sera brièvement décrite au prochain chapitre.

Les atomes étant électriquement neutres, ils possèdent le même nombre de protons que d'électrons. Ce nombre est noté Z : c'est à la fois le *nombre de charge* du noyau et le *numéro atomique* de l'élément considéré. Le nombre total de nucléons d'un noyau est quant à lui appelé *nombre de masse*, et il est noté A .

On représente alors un atome sous la forme ${}_Z^AX$, où A et Z sont les nombres de masse et de charge du noyau de l'élément dont le symbole chimique est « X ». On emploie aussi le terme *nucléide* pour désigner un élément, ou un noyau, défini par un couple (A, Z) .

Citons comme exemples, l'élément le plus léger, l'hydrogène ${}_1^1H$: 1 nucléon (1 proton, 0 neutron) et 1 électron ; et l'élément le plus lourd, l'ununoctium ${}_{118}^{294}Uuo$: 294 nucléons (118 protons, 176 neutrons) et 118 électrons.

Il est à noter que, lorsque l'on connaît bien la signification des symboles chimiques, on peut se contenter d'écrire AX puisque l'information « Z » est déjà contenue dans le symbole chimique de l'élément, reflétant son nombre d'électrons, et donc de protons.

Deux éléments ayant le même nombre de charge Z mais différant par leur nombre de masse sont appelés des *isotopes*.

Par exemple, parmi les 15 isotopes connus du carbone (allant de 8C à ${}^{22}C$), les deux plus abondants naturellement sont le carbone-12 (à 98,93 %) et le carbone-13 (à 1,07 %), des traces de carbone-14 étant également présentes dans la matière environnante.

8.1. Structure des noyaux atomiques, énergie de liaison

Bien qu'ayant des propriétés chimiques et biologiques semblables, puisqu'ils ont le même nombre d'électrons dans leur nuage électronique, deux isotopes peuvent avoir des propriétés physiques très différentes (en particulier, la façon dont leur noyau se désintègre). Cette différence dans le mode de désintégration est le point de départ de la méthode de datation nucléaire (*datation au carbone-14* par exemple).

Il existe actuellement des milliers d'isotopes des 118 éléments découverts à ce jour.

Les masses atomiques étant de l'ordre de 10^{-27} kg, une unité de masse plus appropriée a été introduite : l'*unité de masse atomique*, définie aujourd'hui comme un douzième de la masse d'un atome de carbone-12 :

$$1u = \frac{1}{12}m({}_{6}^{12}\text{C}) \simeq 1,66054 \cdot 10^{-27} \text{ kg} \quad (8.5)$$

qui correspond à l'énergie de masse suivante (en eV) :

$$1u \cdot c^2 \simeq 931,5 \text{ MeV} \quad (8.6)$$

La définition du nombre d'Avogadro N_A est également basée sur l'élément carbone : c'est le nombre d'atomes contenus dans 12 g de carbone-12, c'est-à-dire :

$$N_A = \frac{12 \cdot 10^{-3}}{1,992647 \cdot 10^{-26}} = 6,02214 \cdot 10^{23} \quad (8.7)$$

Comme énoncé plus haut, la masse d'un atome est très proche de la somme des masses de ses nucléons :

$$m({}_Z^AX) \simeq Zm_p + (A - Z)m_n \simeq A \cdot u \quad (8.8)$$

(et dans le cas du carbone-12, ces deux masses sont rigoureusement égales par définition de u). Cette approximation n'est cependant pas valable lorsque l'on s'intéresse aux énergies de liaison nucléaires.

8.1.2 Défaut de masse, énergie de liaison

Expérimentalement, on constate que la somme des masses des constituants d'un atome est toujours supérieure à la masse de cet atome : il existe donc un *défaut de masse* Δm défini par la relation :

$$\Delta m({}_Z^AX) = [Zm_p + (A - Z)m_n] - m({}_Z^AX) \quad (8.9)$$

Par exemple, pour le carbone-12, on obtient la valeur suivante pour son défaut de masse :

$$\Delta m({}_{12}^{12}\text{C}) = 6m_p + 6m_n - m({}_{12}^{12}\text{C}) \quad (8.10)$$

$$= 1,59 \cdot 10^{-28} \text{ kg} \simeq 8 \cdot 10^{-3} \times m({}_{12}^{12}\text{C}) \quad (8.11)$$

La relation d'Einstein $\Delta E = \Delta M \cdot c^2$, détaillée dans le paragraphe 7.3.3 du chapitre précédent, nous indique alors qu'à cette variation de masse $\Delta m \left({}^A_Z X \right)$ peut être associée une variation d'énergie interne, notée conventionnellement $B(A, Z)$:

$$B(A, Z) = \Delta m \left({}^A_Z X \right) \cdot c^2 \quad (8.12)$$

et appelée *énergie de liaison* du noyau.

L'énergie de liaison $B(A, Z)$ est en fait la différence entre l'énergie de masse des constituants du noyau séparés et éloignés les uns des autres à l'infini et l'énergie de masse de ces mêmes constituants lorsqu'ils sont liés entre eux au sein du noyau. Cette *énergie potentielle* de formation est donc également l'énergie qu'il faut fournir à un noyau pour le dissocier en ses constituants élémentaires, c'est-à-dire en neutrons et protons.

Par exemple, l'énergie de liaison d'un noyau de carbone-12 vaut :

$$E_{\text{liaison}} \left({}^{12}_6\text{C} \right) = B(12, 6) \simeq 1,48 \cdot 10^{-17} \text{ J} \simeq 92,16 \text{ MeV}$$

L'énergie de liaison *nucléaire* $B(A, Z)$ ne doit pas être confondue avec l'énergie de liaison *atomique*, ou *énergie d'ionisation*, E_I , qui correspond à l'énergie à fournir à un atome pour lui arracher un (ou des) électron(s).

Pour rendre compte de la stabilité des noyaux, on utilise en fait plutôt l'*énergie de liaison par nucléon* B/A qui traduit l'énergie moyenne qu'il faut fournir à un noyau pour lui arracher un nucléon. Ainsi, selon cette approche (qualifiée de *modèle de la goutte liquide*), plus B/A est important, plus le noyau correspondant est stable.

Il existe actuellement 80 éléments qui ont au moins un isotope stable, et il y a en tout 254 *nucléides* ${}^A_Z X$ stables connus.

Un noyau est dit *instable* si son assemblage de nucléons ne perdure pas dans le temps (du moins à l'échelle de l'âge de l'Univers). Cette modification, appelée désintégration radioactive, est *aléatoire* mais peut être caractérisée par une *durée de vie* moyenne (section 8.3). Il existe plusieurs milliers de nucléides instables dont les durées de vie s'étalent sur près de 30 ordres de grandeurs, de quelques fractions de seconde à quelques milliards d'années.

La désintégration radioactive d'un noyau peut prendre plusieurs formes : division en plusieurs noyaux plus légers (*fission*, *réaction α*), transformation d'un neutron en proton (*réactions β*)... Pour les noyaux légers comme ceux des isotopes de l'hydrogène, il peut y avoir des réactions de *fusion* conduisant à la création d'un noyau plus lourd avec dégagement d'une grande quantité d'énergie.

8.2 DE MULTIPLES RÉACTIONS NUCLÉAIRES

8.2.1 Lois de conservation, énergie de réaction

Une réaction nucléaire est un processus au cours duquel un ou plusieurs noyaux atomiques subissent une transformation, c'est-à-dire un changement de structure. Il peut s'agir d'une *collision inélastique* entre deux corps (entre deux noyaux, par exemple, ou entre un noyau et une particule : électron, neutron, proton...) ou d'une *désintégration radioactive* (dans ce cas, un seul noyau, appelé *noyau père*, donne naissance à un ou plusieurs *noyaux fils*). Dans tous les cas, il y a transformation de masse en énergie : la relation d'Einstein $\Delta E = \Delta M \cdot c^2$ est à la base de toute réaction nucléaire.

On distingue généralement trois étapes distinctes lors d'une réaction nucléaire : *avant*, *pendant* et *après* la réaction. Avant et après la réaction, les corps concernés (noyaux et particules) sont supposés être assez éloignés les uns des autres pour qu'il n'y ait plus aucune interaction entre eux. L'interaction n'est supposée avoir lieu que pendant un bref moment, que nous ne détaillerons pas.

Nous nous concentrons plutôt sur les divers types de réactions possibles et sur les relations qui existent entre les *réactifs* (les corps avant la réaction) et les *produits* (les corps après la réaction). Ces relations découlent principalement des quatre *lois de conservation* suivantes, les deux premières ayant déjà été étudiées dans le chapitre précédent :

1. conservation de la *quantité de mouvement totale* :

$$\sum \mathbf{p}_{\text{réactifs}} = \sum \mathbf{p}_{\text{produits}} \quad (8.13)$$

2. conservation de l'*énergie totale* :

$$\sum E_{\text{réactifs}} = \sum E_{\text{produits}} \quad (8.14)$$

3. conservation du *nombre de masse total* :

$$\sum A_{\text{réactifs}} = \sum A_{\text{produits}} \quad (8.15)$$

4. conservation du *nombre de charge total* :

$$\sum Z_{\text{réactifs}} = \sum Z_{\text{produits}} \quad (8.16)$$

Il existe également d'autres lois de conservation (qui ne seront pas détaillées dans ce cours), comme celle du moment cinétique par exemple, ou celles d'autres nombres caractéristiques du noyau comme les nombres leptonique, baryonique ou d'isospin.

Les réactions nucléaires peuvent se mettre sous une forme similaire aux réactions chimiques :



où X_1 et X_2 sont les réactifs et X_3 et X_4 (et $X_5, X_6...$ si besoin) les produits de la réaction. La quantité Q est l'énergie (de masse) de la réaction.

Les lois de conservation des nombres de masse et de charge conduisent alors aux relations :

$$A_1 + A_2 = A_3 + A_4 \quad (8.18)$$

$$Z_1 + Z_2 = Z_3 + Z_4 \quad (8.19)$$

ce qui permet d'équilibrer les réactions lorsque certains nombres A_i ou Z_i sont inconnus.

L'énergie (de masse) Q de la réaction est par définition égale à l'énergie de masse libérée par la réaction :

$$Q = \left(\sum m_{\text{réactifs}} - \sum m_{\text{produits}} \right) \cdot c^2 \quad (8.20)$$

où $\left(\sum m_{\text{réactifs}} - \sum m_{\text{produits}} \right)$ est le défaut de masse de la réaction.

Pour la réaction 8.17, Q s'écrit :

$$Q = \left[m \left({}_{Z_1}^{A_1} X_1 \right) + m \left({}_{Z_2}^{A_2} X_2 \right) - m \left({}_{Z_3}^{A_3} X_3 \right) - m \left({}_{Z_4}^{A_4} X_4 \right) \right] \cdot c^2 \quad (8.21)$$

qu'on peut également exprimer en fonction des énergies de liaison respectives :

$$Q = B(A_3, Z_3) + B(A_4, Z_4) - B(A_1, Z_1) - B(A_2, Z_2) \quad (8.22)$$

en prenant garde à l'ordre des termes.

Du fait de la conservation de l'énergie totale (somme de l'énergie de masse mc^2 et de l'énergie cinétique $E_c = (\gamma - 1)mc^2$) :

$$\sum \left(m_{\text{réactifs}} c^2 + E_{c, \text{réactifs}} \right) = \sum \left(m_{\text{produits}} c^2 + E_{c, \text{produits}} \right) \quad (8.23)$$

on montre facilement que Q est aussi égale à la variation d'énergie cinétique de la réaction :

$$Q = E_c(X_3) + E_c(X_4) - E_c(X_1) - E_c(X_2) \quad (8.24)$$

Lorsque Q est positive, il y a donc transformation d'une partie de la masse totale initiale en énergie cinétique des noyaux et particules produits. On dit que la réaction est *exoénergétique*. Elle est alors *spontanée* et il y a un gain d'énergie de liaison (totale) lors de la réaction.

Lorsque Q est négative, la réaction est dite *endoénergétique*. Dans ce cas, une partie de l'énergie cinétique totale initiale se transforme en énergie de masse des produits. Une telle réaction ne peut donc se produire que si les réactifs (X_1 et X_2) ont une énergie cinétique relative suffisamment importante (collision avec des particules rapides par exemple).

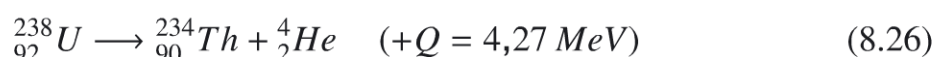
Il est à remarquer que, dans le cas où l'un des noyaux (réactifs ou produits) est dans un *état excité*, avec une énergie interne supplémentaire E^* , il faut prendre en compte cette énergie dans l'expression de Q . Par exemple, un noyau excité ${}_{Z_1}^{A_1}X_1^*$ peut se désexciter sous la forme ${}_{Z_1}^{A_1}X_1^* \longrightarrow {}_{Z_1}^{A_1}X_1 (+Q)$ en émettant un photon d'énergie $Q = E^* = h\nu$. D'une façon générale, l'énergie libérée par une réaction nucléaire se retrouve sous trois formes : énergie cinétique des corps massifs produits, énergie $h\nu$ des photons émis, énergie interne d'excitation.

8.2.2 Réactions α , β et γ

Réaction α . Certains noyaux instables se désintègrent en émettant un noyau d'hélium-4 ${}^4_2\text{He}$, aussi appelé *particule α* :



Par exemple, la désintégration radioactive de l'uranium-238 :



est responsable d'une part importante de la production naturelle d'hélium sur Terre.

Les particules α produites dans de telles réactions ont une énergie cinétique de l'ordre de 5 MeV , c'est-à-dire une vitesse de l'ordre de 5% de c , ce qui implique d'utiliser la relativité restreinte pour rendre compte de leur mouvement.

Du fait de leur masse ($m({}^4_2\text{He}) \simeq 4,0026 u$) et de leur charge ($+2 e$) relativement importantes, les particules α interagissent facilement avec les atomes environnants et perdent rapidement leur énergie, même lorsqu'ils se propagent dans l'air ambiant. Quelques centimètres d'air suffisent alors à les arrêter.

Comme applications principales de la radioactivité α , citons les détecteurs de fumée par ionisation¹ et les microgénérateurs d'énergie (pour les satellites par exemple). Un exemple de microgénérateur d'énergie utilisant la radioactivité α est représenté sur la figure 8.1.

Réactions β . Un noyau peut être instable parce qu'il est trop riche en neutrons. Un de ses neutrons peut alors se transformer spontanément en proton suivant la réaction :



où $\bar{\nu}_e$ est une particule neutre, très légère, appelée *antineutrino électronique*.

1. Dans ce type de détecteur, un matériau radioactif émetteur de particules α ionise l'air entre deux électrodes, créant ainsi un courant électrique mesurable. Lorsque des particules de fumée pénètrent dans le détecteur, elles absorbent une partie des particules α , induisant une baisse du courant qui déclenche l'alarme. Du fait de la nocivité des produits radioactifs utilisés, ces détecteurs sont désormais interdits en France.

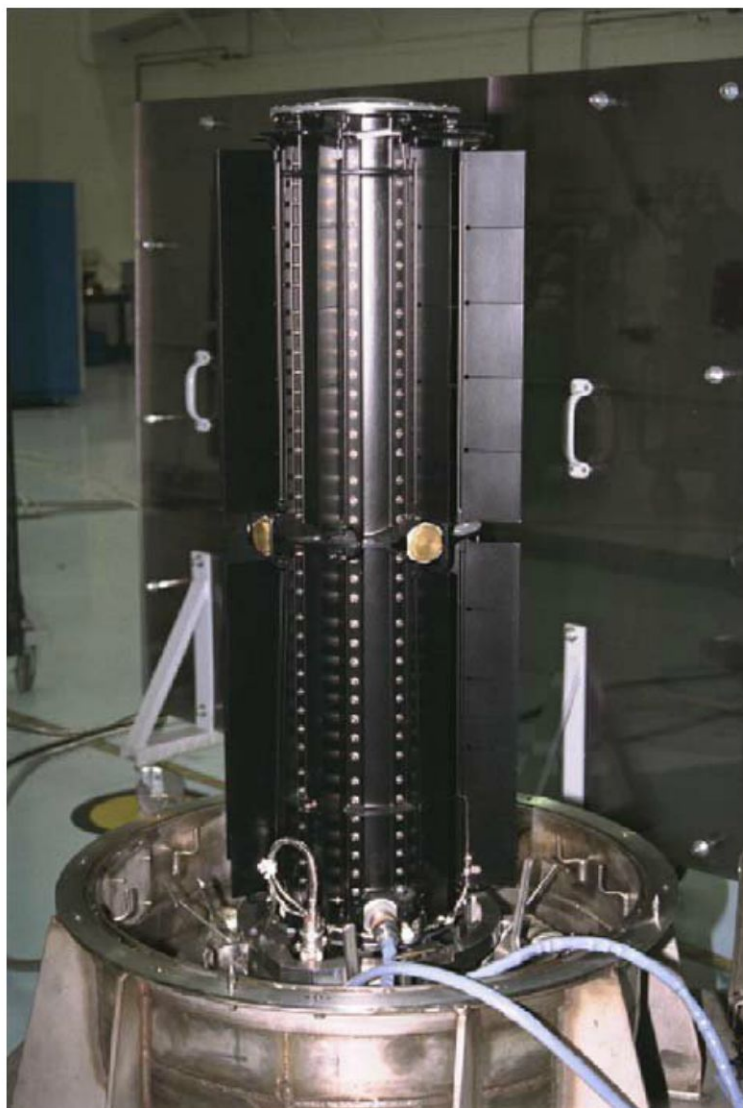


Figure 8.1 – Photo du microgénérateur d’énergie à radioactivité α utilisé dans la sonde spatiale Cassini.

(source libre : <http://en.wikipedia.org/wiki/Cassini-Huygens>)

La désintégration β^- d’un noyau A_ZX s’écrit donc :

$${}^A_ZX \xrightarrow{\beta^-} {}^A_{Z+1}Y + {}^0_{-1}e^- + {}^0_0\bar{\nu}_e \quad (8.28)$$

où la particule β^- n’est rien d’autre que l’électron produit lors de la transformation d’un neutron en proton.

Par exemple :

$${}^{90}_{38}Sr \longrightarrow {}^{90}_{39}Y + {}^0_{-1}e^- + {}^0_0\bar{\nu}_e \quad (+0,546 \text{ MeV}) \quad (8.29)$$

8.2. De multiples réactions nucléaires

Les particules β^- étant chargées, elles interagissent assez fortement avec la matière qu'elles traversent (mais moins que les particules α). Quelques millimètres d'aluminium suffisent à les arrêter. Elles sont généralement ultra-relativistes.

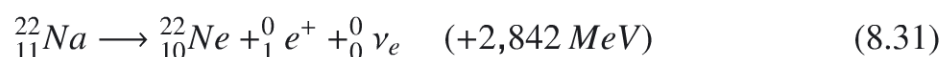
Les applications principales de la radioactivité β^- concernent le traitement de certains cancers des yeux et des os, ainsi que certains tests de qualité dans la fabrication du papier.

Un noyau peut également être instable parce qu'il contient trop de protons. Il y a conversion d'un des protons du noyau en 1 neutron, 1 positron et 1 neutrino électronique :



Les positrons e^+ possèdent les mêmes propriétés physiques que l'électron excepté la charge qui est opposée.

Par exemple :



La tomographie par émission de positrons (PET) est l'une des applications médicales principales de la radioactivité β^+ . Elle consiste à réaliser une image à trois dimensions d'un organe où l'on a préalablement fixé des noyaux radioactifs « traceurs », émetteurs de positrons. On ne détecte en fait pas directement les positrons, mais les rayons γ émis lors de la désintégration des positrons.

Réaction γ . Chaque désintégration α ou β d'un noyau instable donne généralement naissance à un noyau fils dans un *état excité métastable*. Celui-ci peut alors se désexciter en émettant un ou plusieurs photons γ :



où l'énergie libérée Q est l'énergie d'excitation E^* du noyau initial. Celle-ci se partage en énergie cinétique du noyau désexcité ${}_Z^AX$ et en énergie $h\nu$ du photon γ émis.

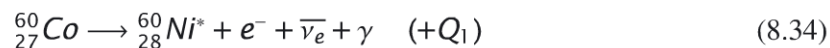
Un calcul rapide (conservation de l'énergie et de la quantité de mouvement totales) montre alors que :

$$h\nu = E^* \frac{1 + \frac{E^*}{2Mc^2}}{1 + \frac{E^*}{Mc^2}} \quad (8.33)$$

dans le référentiel où le noyau initial ${}_Z^AX^*$ est au repos. Le photon emporte donc quasiment toute l'énergie libérée E^* si celle-ci est suffisamment faible par rapport à l'énergie de masse Mc^2 , où M est la masse du noyau X .

Exemple

Émission de deux photons gamma lors de la désintégration d'un noyau de Cobalt en un noyau de Nickel :



avec $Q_1 = 1,17 \text{ MeV}$ et $Q_2 = 1,33 \text{ MeV}$.

Si l'énergie du photon émis est assez faible ($< 100 \text{ keV}$), il peut être converti en un électron par effet photoélectrique au sein même de l'atome (on reviendra sur l'effet photoélectrique dans la seconde partie de cet ouvrage). Le transfert d'énergie peut également être direct (sans faire appel à une réaction γ) et conduit à l'émission d'un *électron de conversion interne*.

Les rayons γ sont, comme les neutrons rapides, très pénétrants et il faut généralement plusieurs centimètres de plomb pour les arrêter.

8.2.3 Fission, fusion

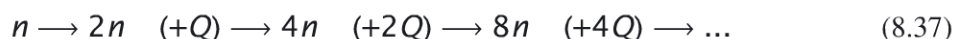
Fission. Les noyaux très lourds peuvent spontanément, ou artificiellement, se scinder en plusieurs noyaux dont l'énergie de liaison par nucléon est plus élevée (c'est-à-dire qu'ils sont plus stables). Par exemple, les réactions α sont un cas particulier de *fission* « naturelle ». On peut aussi *provoquer* la fission d'un noyau en le percutant avec des particules de haute énergie, des neutrons par exemple.

Certains noyaux très lourds ont aussi la particularité d'émettre et d'absorber facilement des neutrons, ce qui peut conduire, si certaines conditions sont vérifiées, à des *réactions en chaîne*. Ces réactions donnent lieu à une libération considérable d'énergie.

Exemple



avec $Q = 192,9 \text{ MeV}$. Chaque neutron produit peut alors initier une autre réaction de fission, donnant ainsi naissance à une réaction en chaîne du type :



Remarques

1. Il faut généralement que les neutrons initiaux aient une énergie cinétique supérieure à une valeur *seuil* (5,3 MeV dans le cas de l'uranium-235);
2. La réaction en chaîne ne se réalise que si les neutrons produits rencontrent suffisamment de noyaux d'uranium sur leur parcours. Il y a donc une *masse critique* d'uranium-235 pour que la réaction se produise (quelques dizaines de kg en pratique).

En dehors de la réalisation d'armes nucléaires, la production industrielle d'énergie nucléaire est aujourd'hui la principale application des réactions en chaîne de fission nucléaire. Environ 12 % de l'énergie produite en 2013 dans le monde était d'origine nucléaire. Ce chiffre monte à 75 % en France.

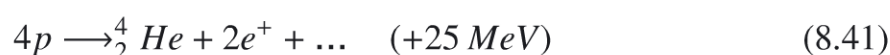
La fission complète de 1 kg d'uranium 235 libère une énergie de $7,9 \cdot 10^{13}$ J, à comparer aux $4,5 \cdot 10^7$ J libérés par la combustion totale de 1 kg de pétrole. Cette différence de *concentration* d'énergie (nucléaire ou chimique) explique en partie pourquoi il ne faut qu'une trentaine de tonnes d'uranium enrichi pour produire 1000 MW d'énergie électrique par an alors qu'il faut 1,5 millions de tonnes de pétrole ou 2,3 millions de tonnes de charbon.

Fusion. Deux noyaux légers peuvent, dans certains cas, se rapprocher suffisamment pour que l'attraction due à l'interaction forte entre nucléons dépasse la répulsion coulombienne des protons. Il en résulte un noyau plus lourd et plus stable. Le processus est fortement exoénergétique mais nécessite des conditions de température et de pression extrêmement sévères.

Ces conditions sont remplies au cœur des étoiles où se produisent des réactions de *fusion* de l'hydrogène en hélium 3 et 4 :



dont l'équation bilan est finalement :



La recherche d'une *fusion contrôlée* de l'hydrogène est actuellement très active (projet mondial ITER de fusion par confinement magnétique, projets américains et français NIF et LMJ utilisant des lasers produisant des impulsions de très grande énergie). Une maîtrise industrielle de ce type de fusion permettrait en effet de réaliser des centrales fournissant beaucoup plus d'énergie que les centrales utilisant la fission d'atomes (10 fois plus environ), et créant beaucoup moins de déchets radioactifs.

8.3 RADIOACTIVITÉ ET DOSIMÉTRIE

8.3.1 Loi de décroissance exponentielle

La probabilité par unité de temps qu'un noyau instable se désintègre se révèle être une constante, appelée *constante radioactive* λ (unité : s^{-1}). L'inverse de λ est appelée *constante de temps* de la désintégration, et est notée $\tau = 1/\lambda$.

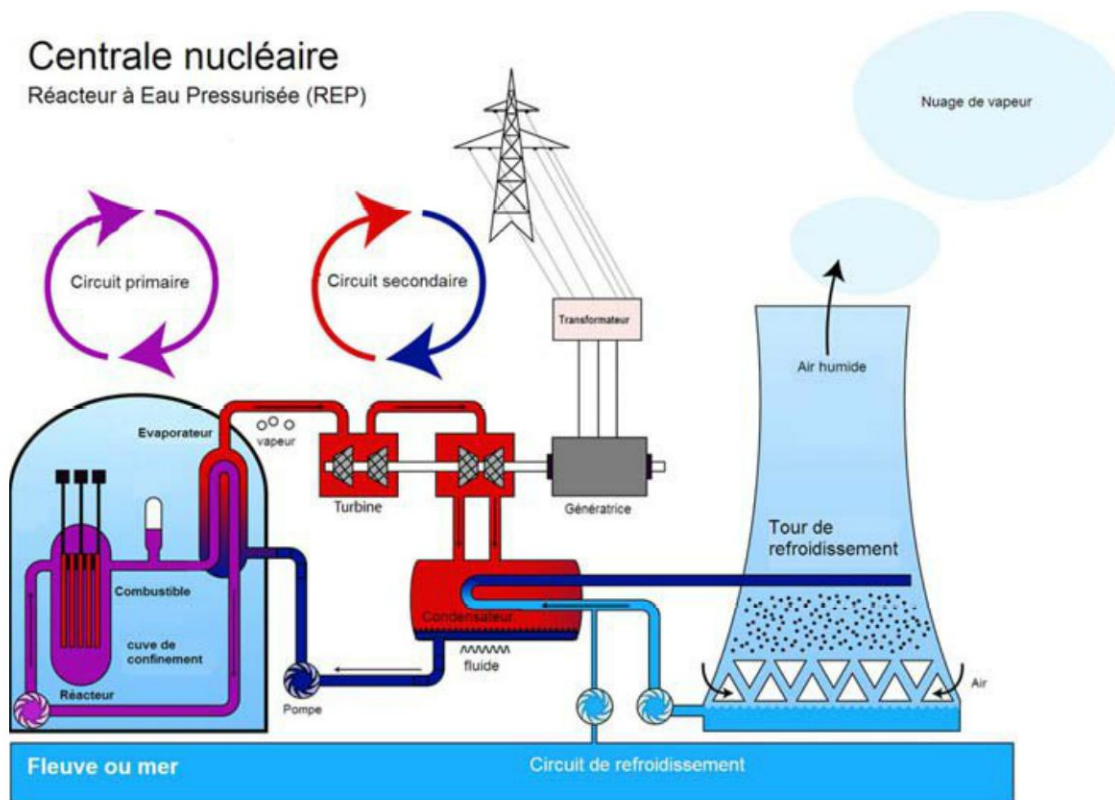


Figure 8.2 – Schéma de principe d'une centrale nucléaire produisant de l'électricité. Ce réacteur est de type à eau pressurisée (REP). L'eau du circuit primaire est sous haute pression (150 bars environ). Elle est utilisée comme fluide caloporteur pour la vaporisation de l'eau du circuit secondaire, permettant de créer de l'électricité à l'aide de turbines et de transformateurs. L'eau primaire est également utilisée comme modérateur pour réfrigérer le cœur du réacteur. Le refroidissement entraîne alors la création d'importants nuages de vapeur. Les REP sont les plus répandus des réacteurs.
(source libre : http://fr.wikipedia.org/wiki/Centrale_nucleaire)

Le caractère *probabiliste* du processus fait que l'on ne peut pas connaître l'instant où *un* noyau donné va se désintégrer (cette propriété est d'origine *quantique* comme on le verra dans la seconde partie de cet ouvrage). En revanche, on peut statistiquement prédire le comportement d'un grand nombre de noyaux radioactifs. Par exemple, sur l'histogramme de la figure 8.3, on a représenté le nombre de désintégrations radioactives d'un ensemble initial de N_0 noyaux entre deux instants successifs séparés d'une durée τ . La probabilité de désintégration entre les instants t et $t + \tau$ étant égale à λt , la courbe approchant l'histogramme est une courbe exponentiellement décroissante.

En effet, soit $N(t)$ le nombre de noyaux radioactifs, la loi des grands nombres nous indique que la variation $dN(t)$ de noyaux radioactifs pendant le temps dt s'écrit :

$$dN(t) = -\lambda N(t) dt \quad (8.42)$$

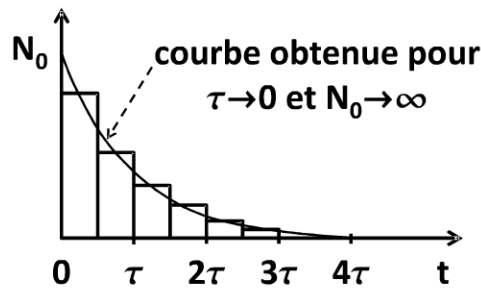


Figure 8.3 - Statistique des instants de désintégration de noyaux radioactifs.

qui s'intègre en :

$$N(t) = N_0 e^{-\lambda t} \quad (8.43)$$

où N_0 est le nombre de noyaux radioactifs à l'instant initial $t = 0$. Cette loi de décroissance exponentielle porte le nom de *loi de Soddy*.

Pour caractériser cette décroissance radioactive on utilise le temps de demi-vie T , appelé aussi période radioactive, qui est défini comme le temps au bout duquel le nombre de noyaux initialement présents a été divisé par deux :

$$N(T) = N_0/2 \quad (8.44)$$

ce qui conduit aux relations :

$$T = \frac{\ln(2)}{\lambda} = \tau \ln(2) \quad (8.45)$$

et :

$$N(t) = \frac{N_0}{2^{t/T}} \quad (8.46)$$

Remarques

1. *Elimination biologique* : Après avoir été ingérés ou inhalés, les éléments (radioactifs ou non) sont progressivement éliminés par les corps biologiques. Cette décroissance est également exponentielle, avec une *constante de temps biologique* τ_b longue et différente de τ . Pendant le temps dt , les deux types d'élimination des éléments radioactifs s'ajoutent :

$$dN(t) = -\frac{N(t) dt}{\tau} - \frac{N(t) dt}{\tau_b} \quad (8.47)$$

conduisant à la loi de décroissance exponentielle suivante :

$$N(t) = N_0 e^{-t/\tau_{eff}} \quad (8.48)$$

avec la *constante de temps effective* :

$$\tau_{eff} = \left(\frac{1}{\tau} + \frac{1}{\tau_b} \right)^{-1} \quad (8.49)$$

2. Ce type de décroissance multiple se retrouve également dans le cas où un élément radioactif peut se désintégrer de plusieurs façons différentes. Si chaque processus a une constante de temps τ_i , la constante de temps effective sera égale à $\tau_{eff} = 1/(\sum_i 1/\tau_i)$. L'inverse de chaque constante de temps étant égale à la probabilité par unité de temps du processus correspondant (constantes radioactives λ_i), les rapports $\lambda_i/\lambda_{eff} = \tau_{eff}/\tau_i$ de ces probabilités sont appelés *rapports de branchement*. Leur somme vaut bien évidemment 1.

8.3.2 Activité d'une source radioactive

L'*activité* d'une source radioactive est, par définition, le nombre de désintégrations par unité de temps de cette source :

$$A(t) = \lambda N(t) = A_0 e^{-t/\tau} = \frac{A_0}{2^{t/T}} \quad (8.50)$$

avec $A_0 = \lambda N_0$.

Si une source contient plusieurs éléments radioactifs, l'activité de cette source est la somme des activités de ses constituants.

L'unité est le *Becquerel* (Bq), qui correspond à *une* désintégration par seconde. D'autres unités sont parfois utilisées, comme le *Curie* (Ci) qui correspond à 37 GBq. Très souvent, on utilise les *activités massiques* (en Bq.kg⁻¹).

Exemples

lait $\approx 80 \text{ Bq.kg}^{-1}$

un être humain de 70 kg $\approx 10^4 \text{ Bq}$

le combustible d'un réacteur nucléaire $\approx 10^{19} \text{ Bq}$



8.1 Dose radioactive : effets de la radioactivité

Les particules émises (α , n , p , e^- , e^+ , γ ...) lors des réactions nucléaires pouvant ioniser la matière qu'elles traversent, elles sont couramment dénommées *rayonnements ionisants*. On définit la *dose reçue* (ou *dose absorbée*) « D » comme l'énergie déposée dans une unité de masse de matière. L'unité est le *Gray* : $1 \text{ Gy} = 1 \text{ J.kg}^{-1}$

L'activité et la dose reçue sont des *grandeurs objectives*, c'est-à-dire pouvant être mesurées avec des appareils (compteur Geiger, calorimètre...). Elles ne donnent que peu d'indications sur les *effets biologiques* produits lors de l'absorption de rayonnements ionisants. Ces effets peuvent aller de l'*altération* à la *destruction* des cellules du vivant. L'altération du patrimoine génétique (l'ADN) d'une cellule conduit parfois, lorsque le gène altéré est dominant, à l'apparition de cancers.

La nature des rayonnements ionisants est particulièrement importante : à dose reçue égale, les dégâts sont bien plus importants dans le cas d'une exposition à des particules massives (particules alpha, neutrons...) plutôt qu'avec des particules peu ou non massives (électrons, positrons, rayons gamma...). On introduit

alors un *facteur de pondération radiologique* W_R pour chaque type de radiation : 1 pour les rayons X et gamma ; 1 pour les électrons et positrons ; 5 pour les protons ; de 5 à 20 pour les neutrons en fonction de leur énergie ; 20 pour les particules alpha et les noyaux lourds. La *dose équivalente* « H » est alors définie comme la somme des doses reçues D_R pondérées par les facteurs W_R correspondant aux types de rayonnements ionisants considérés :

$$H = \sum_R W_R D_R \quad (8.51)$$

L'unité est le *Sievert* (Sv) : $1 \text{ Sv} = \text{J.kg}^{-1}$

Exemples

Vol Paris-New-York² = 0,02 mSv ; radiographie des poumons = 0,3 mSv ; radioactivité naturelle du sol en France = de 1 à 2 mSv/an ; scanner du corps entier = 150 mSv.

L'effet de cette dose équivalente n'est pas le même suivant l'organe touché. Pour rendre compte des effets différents selon la nature des tissus biologiques irradiés, on introduit un *facteur de pondération tissulaire* W_T (par exemple, $W_{\text{estomac}} = W_{\text{prostate}} = 0,12$; $W_{\text{foie}} = W_{\text{thyroïde}} = 0,05$; $W_{\text{peau}} = W_{\text{os}} = 0,03$). La *dose efficace* « E_D » est alors la somme pondérée des doses équivalentes H_T reçues par tous les organes et tissus irradiés d'un corps biologique :

$$E_D = \sum_T W_T H_T \quad (8.52)$$

$$= \sum_T \sum_R W_T W_R D_{RT} \quad (8.53)$$

où D_{RT} est la dose de rayonnement ionisant R absorbée par le tissu T . L'unité est le Sv.

Remarque

On utilise souvent les *débits de dose* (absorbée, équivalente, efficace), c'est-à-dire des doses par unité de temps. En France, les limites fixées par la réglementation sont actuellement de 20 mSv/an pour les travailleurs du nucléaire et de 1 mSv/an pour le reste de la population.

Les effets des rayonnements ionisants sont de deux types : *déterministes* (brûlures, syndrome d'irradiation aigüe) ou *stochastiques* (développement aléatoire de cancers). Si les effets dus aux fortes irradiations sont relativement bien compris, les effets stochastiques dus aux faibles irradiations sont encore très mal connus et sont régulièrement réévalués. On estime que l'excès de cancers dûs à la radioactivité est de l'ordre de 5 % par Sv absorbé. Pour se prémunir des irradiations, on peut *s'éloigner* des sources radioactives (division des doses par

2. S'élever en altitude diminue l'épaisseur de la couche d'atmosphère protectrice qui absorbe les rayonnements ionisants issus du soleil et, plus généralement, de l'espace.

un facteur d^2 si l'on multiplie la distance à la source par d) et/ou utiliser des *écrans protecteurs* dont l'épaisseur est adaptée au type de rayonnement ionisant. La loi d'absorption de ces rayonnements est en effet *exponentiellement décroissante dans la matière* :

$$I(x) = I_0 e^{-\mu x} \quad (8.54)$$

où I est l'intensité du rayonnement, x l'épaisseur de matière traversée, et μ le coefficient d'absorption linéique du matériau pour le rayonnement ionisant considéré. Ainsi, par exemple, pour réduire de moitié l'intensité de rayonnements γ , il faut typiquement : 1 cm de plomb ; ou 2,5 cm d'acier ; ou 6 cm de béton ; ou 18 cm d'eau ; ou 29 cm de bois ; ou 150 m d'air...

LES QUATRE INTERACTIONS FONDAMENTALES, LA RELATIVITÉ GÉNÉRALE

9

La cohésion de la matière et les interactions entre particules élémentaires sont assurées par diverses forces, qui, selon la conception actuelle de la physique, peuvent être décrites au moyen de quatre interactions fondamentales seulement : ce sont la gravitation, l'électromagnétisme, et les interactions faible et forte. Ces quatre forces fondamentales sont très différentes, tant par leur *intensité* que par leur *portée* ou par la nature des *charges* concernées.

Bien que la recherche d'une description unifiée des interactions fondamentales soit l'un des objectifs majeurs de la recherche en physique, la gravitation reste aujourd'hui encore à part, et seules les interactions électromagnétique, faible et forte sont pour l'instant unifiées au sein d'un modèle théorique appelé *modèle standard*.

Le but de ce chapitre est d'introduire plusieurs de ces notions, en insistant plus particulièrement sur la théorie de la *relativité générale*, prolongement naturel des idées et principes de la théorie de la relativité restreinte.

9.1 INTERACTION GRAVITATIONNELLE ET NOTION DE RELATIVITÉ GÉNÉRALE

Dans les chapitres précédents, nous avons vu que, dans le cadre de la théorie de la relativité restreinte, les équations de la dynamique et de l'électromagnétisme satisfaisaient aux mêmes lois (de transformation de Lorentz). On peut alors se demander s'il en est de même des trois autres interactions fondamentales, à commencer par la plus commune de toutes à notre échelle : l'interaction gravitationnelle.

Il est cependant facile de montrer que les équations de la théorie classique de la gravitation, c'est-à-dire celle de Newton, ne sont pas en accord avec les transformations de Lorentz et le principe de relativité. En effet, en gravitation newtonienne, le champ de gravitation créé par une masse s'adapte instantanément dans tout l'espace à toute modification de cette masse ou de son mouvement, violant ainsi le principe de relativité et le fait qu'aucun objet ou information ne peut se déplacer plus vite que la vitesse de la lumière. Il est par ailleurs évident, au simple vu de la formule de la force gravitationnelle entre deux masses rappelée ci-dessous (équation 9.1), que

toute loi de ce type (ne faisant intervenir que la *distance* entre les objets) ne peut pas être invariante par transformation de Lorentz. Il en est par exemple de même des lois de l'électrostatique et de la magnétostatique qui ne tiennent pas compte du temps requis pour la propagation d'une modification de la distribution des charges ou de leurs vitesses.

Par souci de cohérence, il apparaît donc nécessaire de construire une théorie *relativiste* de la gravitation, qui serait l'analogue pour la gravité des équations de Maxwell pour l'électromagnétisme.

Une telle théorie remet alors inévitablement en question le principe de relativité restreinte puisque celui-ci ne considère que des référentiels galiléens (en translation uniforme, à vitesse constante, les uns par rapport aux autres) et ignore les référentiels accélérés, en particulier sous l'effet de forces gravitationnelles. On peut évidemment, comme on l'a fait au chapitre 7, parler de *dynamique relativiste* de particules soumises à des forces (et donc à des accélérations) dans le cadre de la relativité restreinte, mais un examen plus précis des effets dus à ces accélérations montre que la relativité restreinte n'est valable que lorsque ces effets sont négligeables. Plus précisément, si a est l'accélération subie par un objet sur une distance d , il faut corriger les résultats de relativité restreinte par des quantités proportionnelles à ad/c^2 .

L'établissement d'un principe de relativité *généralisé* (aux référentiels accélérés) semble donc également nécessaire.

C'est le *principe d'équivalence*, que nous allons voir dans cette partie, qui s'est révélé être la clef pour réaliser ces deux objectifs. Celui-ci a conduit non seulement à la théorie de la *relativité générale* mais également à la refonte complète de notre vision de l'espace-temps, devenu variable dynamique de la théorie et non plus cadre rigide dans lequel les objets évolueraient.

9.1.1 Force de gravitation, forces d'inertie

Interaction gravitationnelle. L'interaction gravitationnelle est responsable de l'attraction réciproque des corps massifs entre eux. C'est la plus faible des quatre interactions fondamentales au niveau microscopique mais, grâce à son action à longue distance (sa *portée* est infinie) et à son caractère cumulatif (contrairement, par exemple, aux forces électromagnétiques qui tendent à se neutraliser du fait de l'existence de charges électriques opposées), elle est la plus intense des quatre interactions aux grandes échelles de distance, dans le domaine de la cosmologie par exemple.

Loi de Newton. Selon la loi de Newton, deux masses ponctuelles m et m' distantes de r s'attirent selon la loi :

$$\mathbf{F} = -G \frac{mm'}{r^2} \mathbf{e}_r \quad (9.1)$$

9.1. Interaction gravitationnelle et notion de relativité générale

où $G = 6,67 \cdot 10^{-11} \text{ m}^3 \text{kg}^{-1} \text{s}^{-2}$ est une constante appelée *constante gravitationnelle* et où \mathbf{e}_r est un vecteur unitaire reliant les deux points où sont situées les masses.

D'une façon générale, le lien entre la densité de matière ρ et le potentiel gravitationnel $\phi(r)$ créé par cette matière au point r est donné par l'*équation de Poisson* :

$$\Delta\phi = 4\pi G\rho(r) \quad (9.2)$$

dont la résolution permet d'obtenir $\phi(r)$ ainsi que la force gravitationnelle agissant sur une masse test ponctuelle m :

$$\mathbf{F} = m\mathbf{g} \quad (9.3)$$

où :

$$\mathbf{g} = -\nabla\phi \quad (9.4)$$

est le *champ de gravitation* au point r .

La force dans (9.1) décroît en $1/r^2$ mais ne s'annule pas totalement à grande distance : sa portée est infinie.

Un autre type de force : les forces inertielles. D'autres forces que celles issues des quatre interactions fondamentales peuvent également se faire sentir : ce sont les *forces d'inertie*, appelées parfois *forces fictives* puisqu'elles résultent du mouvement et non d'une interaction physique. Ce sont, par exemple, les forces centrifuge et de Coriolis dans un référentiel en rotation, ou bien la force d'écrasement que l'on ressent dans un véhicule lors d'une forte accélération.

Plus généralement, ce sont les forces qui agissent sur un objet massif lorsque celui-ci appartient à (ou est observé à partir d') un référentiel non-galiléen. C'est pour cela que l'on appelle parfois *référentiels inertiels* les référentiels galiléens : ce sont ceux où aucune force d'inertie n'apparaît.

L'existence de tels référentiels inertiels dans l'univers est sujette à caution car tout chose (masse ou énergie) tend à créer des effets non-inertiels autour d'elle. L'univers dans sa globalité semble être cependant un bon candidat (par exemple en considérant le référentiel où le rayonnement du fond diffus cosmologique, présent partout dans l'univers, est isotrope), même si cela ne manque pas de soulever des questions quant à l'existence d'un tel référentiel universel, peu en accord avec le principe même de relativité...

En pratique, on ne se pose pas ce genre de questions et l'on suppose qu'il existe bien au moins *un* référentiel inertiel, que l'on définit expérimentalement par l'absence de toute force d'inertie sur les objets immobiles dans ce référentiel. Comme il est dit plus bas, ce référentiel n'est inertiel que localement, c'est-à-dire dans une région limitée de l'espace.

9.1.2 Deux masses différentes et pourtant égales !

Deux masses interviennent en physique :

- la *masse inertielle* m_i qui mesure comment un corps s'oppose à un changement de mouvement, et qui intervient dans le Principe Fondamental de la Dynamique : $\mathbf{F} = m_i \mathbf{a}$.
- la *masse gravitationnelle* m_g qui mesure comment un corps réagit dans un champ gravitationnel et qui intervient dans la loi de Newton (9.1) $\mathbf{F} = -G \frac{m_g m'_g}{r^2} \mathbf{e}_r$.

Il n'y a pas de raison, *a priori*, pour que ces deux masses soient égales. Ainsi le mouvement d'un corps dans un champ de pesanteur \mathbf{g} doit, en toute rigueur, s'écrire :

$$m_i \mathbf{a} = m_g \mathbf{g} \quad (9.5)$$

Les expériences de Galilée, d'Eötvös et de bien d'autres (utilisant essentiellement la chute libre de masses bien calibrées ou des balances à torsion) semblent néanmoins montrer que ces masses sont proportionnelles, et même égales :

$$m_g = m_i \quad (9.6)$$

à une impressionnante précision relative inférieure à 10^{-12} ... qui pourrait être améliorée d'un facteur 1000 dans les prochaines années grâce aux projets spatiaux STEP et MICROSCOPE.

Par conséquent, tous les corps tombent (dans un champ de pesanteur uniforme) à la même vitesse dans le vide :

$$\boxed{\mathbf{a} = \mathbf{g} \quad \forall m} \quad (9.7)$$

C'est cette égalité expérimentale qu'Einstein postule comme principe premier, et qui sert de base à la relativité générale.

9.1.3 Un principe fondateur : le principe d'équivalence

On peut l'énoncer comme ceci :

Il y a équivalence complète entre un champ de gravitation uniforme et une accélération uniforme.

Ainsi, un observateur ne peut pas déterminer, par aucune expérience interne que ce soit, si le référentiel qu'il occupe est dans un champ de gravitation uniforme ou s'il est dans un référentiel uniformément accéléré par rapport à un référentiel inertiel. C'est l'extension aux mouvements uniformément accélérés du principe de relativité énoncé précédemment pour les seuls mouvements uniformes.

9.1. Interaction gravitationnelle et notion de relativité générale

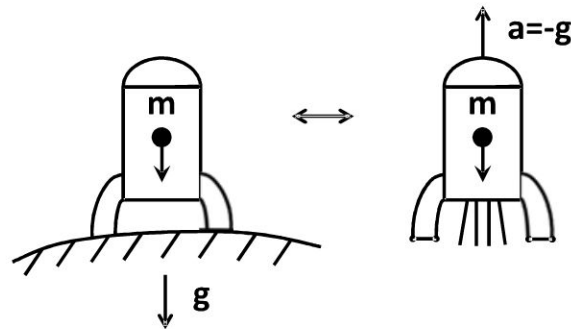


Figure 9.1 – Illustration du principe d'équivalence : chute d'un corps dans une fusée, accélérée ou non.

Considérons, pour fixer les idées, l'exemple d'une fusée dans laquelle un observateur lâche un objet de masse m (voir la figure 9.1) qui tombe verticalement vers le sol. Sans connaissance de l'environnement extérieur à la fusée (c'est-à-dire sans avoir de hublot ou de système permettant d'apprécier, de l'intérieur, le mouvement de la fusée par rapport à un référentiel quelconque), il peut tirer les conclusions suivantes :

1. la fusée est immobile dans un champ de gravitation (posée à la surface d'une planète par exemple) et l'objet lâché est attiré par la gravité ;
2. ou bien la fusée est accélérée par rapport à un référentiel inertiel et l'objet lâché va dans le sens opposé à ce mouvement accéléré.

Dans le premier cas, la chute est due à la gravité (g), alors que dans le second cas elle est due à une accélération (a). Si $a = g$, les deux mouvements sont identiques et l'on ne peut pas déterminer, par une expérience interne à la fusée, laquelle des deux propositions précédentes est correcte.

Une conséquence directe est que la lumière est défléchie par un champ de gravitation. En effet, puisqu'un rayon lumineux suit une ligne droite dans un référentiel inertiel, le principe d'équivalence implique qu'il subira une accélération, et donc qu'il sera courbé, lorsqu'il traversera un champ gravitationnel.

Ce phénomène est illustré sur la figure 9.2 où le référentiel $\mathcal{R}_{fusée}$ est le référentiel d'une fusée en chute libre dans le champ de gravitation terrestre. Celui-ci étant un référentiel inertiel, la lumière s'y propage en ligne droite. Dans le référentiel terrestre \mathcal{R}_{Terre} , en revanche, la fusée, et par conséquent la lumière, ont une trajectoire courbe (parabolique, en fait, dans un champ de gravitation uniforme).

Les phénomènes de *lentille* et de *mirage gravitationnels* sont directement reliés à cet effet de courbure des rayons lumineux par un champ de gravitation. En effet, lorsqu'un objet très massif (le cœur d'un amas de galaxies par exemple) est situé entre un observateur et une source lumineuse, il courbe les rayons lumineux de celle-ci et déforme ainsi l'image reçue par un observateur (figures 9.3 et 9.4). Ces phénomènes ont été observés dans de multiples circonstances depuis leur découverte en 1979.

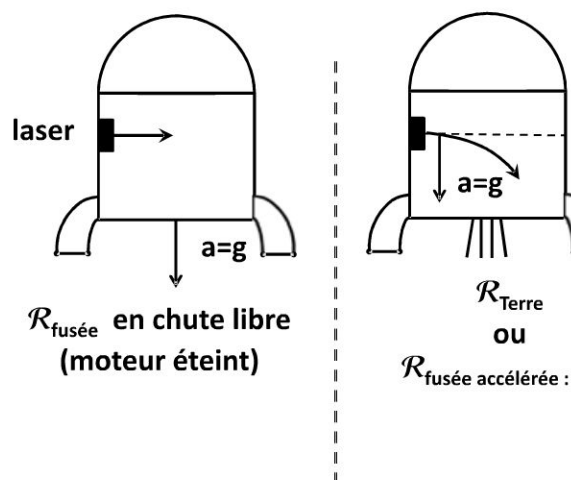


Figure 9.2 – Courbure d'un rayon lumineux dans un champ de gravitation. Le référentiel d'une fusée en chute libre dans le champ de gravitation terrestre étant un référentiel inertiel, la lumière s'y propage en ligne droite. Dans le référentiel terrestre, en revanche, la fusée, et par conséquent la lumière, ont une trajectoire courbe (parabolique dans un champ de gravitation uniforme).

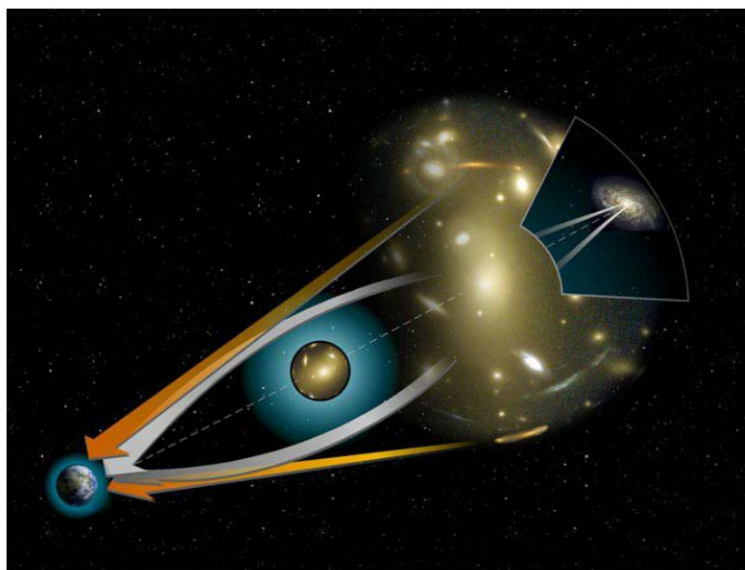


Figure 9.3 – Schéma de principe de la déviation des rayons lumineux par un objet très massif (au centre). La lumière émise par une galaxie est courbée par le champ gravitationnel créé par cet objet massif. Il en résulte un effet de mirage gravitationnel, sous forme d'arcs lumineux ou d'images multiples observés sur Terre.

(source libre : http://fr.wikipedia.org/wiki/Lentille_gravitationnelle)

Il est important de noter que le principe d'équivalence, c'est-à-dire l'équivalence entre forces inertielle et gravitationnelle, n'est vrai que *localement*, c'est-à-dire dans une région réduite par rapport aux dimensions caractéristiques de variation du champ gravitationnel. Il existe en effet deux différences essentielles entre forces inertielle et

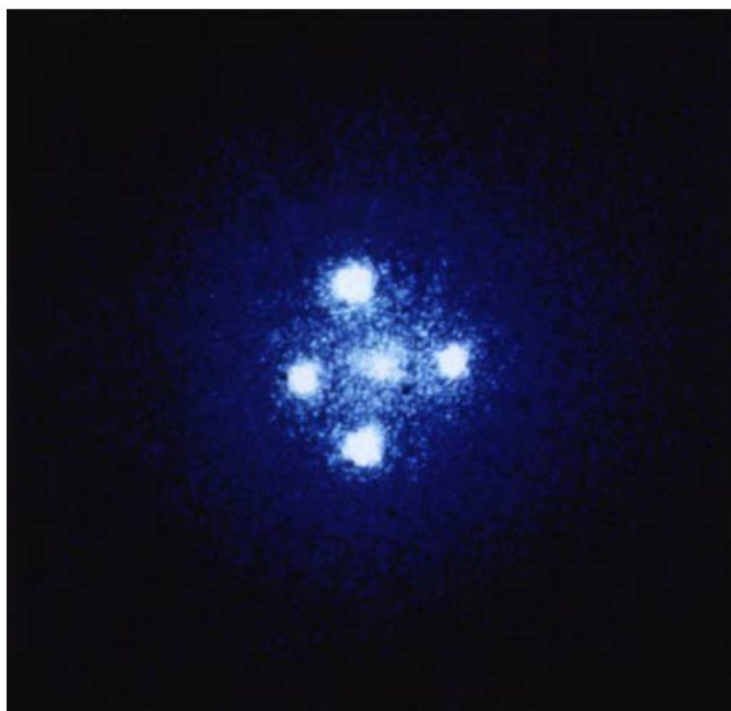


Figure 9.4 – Image célèbre appelée « croix d'Einstein » où quatre images d'un même quasar lointain sont formées par effet de lentille gravitationnelle autour d'une galaxie très massive, située au centre de l'image.

(source libre : http://fr.wikipedia.org/wiki/Croix_d'Einstein)

gravitationnelle : 1) une force gravitationnelle n'est *jamais* uniforme, et 2) sa décroissance particulière à l'infini fait qu'elle ne peut pas être compensée par un changement de référentiel global. Par exemple, si la force centrifuge créée par une rotation est bien capable de reproduire à l'identique une force gravitationnelle en une zone très localisée de l'espace, cette force d'inertie diverge à l'infini, contrairement à la force gravitationnelle qui décroît vers zéro¹.

Une force de gravitation n'étant pas uniforme, ses variations, appelées *forces de marée*, tendent à déformer tout objet étendu.

9.1.4 Effets de la gravité sur les mesures de longueurs et de durées

D'un point de vue expérimental, le principe d'équivalence traduit le fait que le référentiel d'un objet en chute libre dans un champ de gravitation est un *référentiel inertiel*, dans lequel il n'y a pas de champ de gravitation (d'où l'effet *d'apesanteur* qu'expérimentent les astronautes, cosmonautes ou spationautes).

1. La force gravitationnelle varie comme l'inverse du carré de la distance aux objets massifs qui la créent, alors que la force centrifuge croît linéairement avec la distance au centre de rotation.

Dit autrement, pour ressentir les effets d'un champ de gravitation dans son environnement, un observateur ne doit pas être en chute libre dans ce champ de gravitation. Au contraire, un observateur *immobile* à une distance r d'une masse M verra les objets autour de lui tomber avec l'accélération GM/r^2 (si ceux-ci sont en chute « libre » bien sûr).

Cette interprétation du principe d'équivalence permet alors de comprendre pourquoi, et comment, la gravité agit sur les mesures de longueur et de durée.

Comparons pour cela les résultats de mesure (de longueur et de durée) d'un observateur O' , en chute libre dans le champ de gravitation créé par une masse M , avec les résultats d'observateurs O_r , immobiles dans ce champ de gravitation, à différentes distances r de la masse M (figure 9.5).

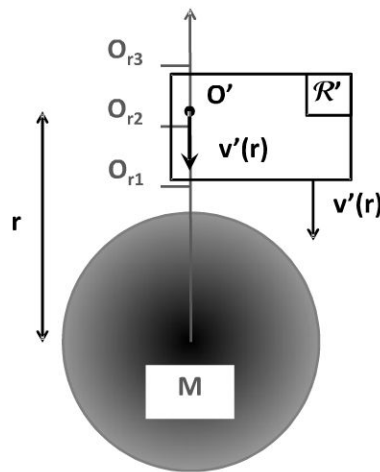


Figure 9.5 – Différents observateurs dans le champ de gravitation créé par la masse M : O' (référentiel propre \mathcal{R}') est en chute libre alors que O_{r1} , O_{r2} et O_{r3} sont des observateurs immobiles dans le référentiel où la masse M est immobile. La vitesse de \mathcal{R}' par rapport à \mathcal{R} augmente lorsque O' se rapproche de la masse M .

Le référentiel \mathcal{R}' de l'observateur O' étant un référentiel inertiel, l'observateur O' y mesure des longueurs et durées propres, par exemple L_0 et Δt_0 . À cause de la gravité, ce référentiel \mathcal{R}' « tombe » sur la masse M avec une vitesse qui augmente d'autant plus que la distance r est faible. Par exemple, partant d'une vitesse initiale nulle à une distance infinie de M , la loi de conservation de l'énergie totale (entre l'infini et la distance r) nous indique² que le référentiel \mathcal{R}' sera animé, à la distance r de M , d'une

2. La vitesse de O' par rapport à O_r est la solution de l'équation de conservation de l'énergie totale (entre l'infini et la distance r) :

$$0 = -G \frac{Mm}{r} + (\gamma_r - 1) mc^2 \simeq -G \frac{Mm}{r} + \frac{1}{2} mv^2 \quad (9.8)$$

c'est-à-dire $v \simeq \sqrt{2GM/r}$.

Conformément à ce qui a été dit dans l'introduction de la partie 9.1, cette formule n'est valable que tant que l'accélération GM/r^2 subie par l'observateur O' est suffisamment faible, c'est-à-dire tant que

9.1. Interaction gravitationnelle et notion de relativité générale

vitesse d'environ $\sqrt{2GM/r}$ par rapport à l'observateur immobile O_r . Inversement, l'observateur en chute libre verra les observateurs O_r aller de plus en plus vite par rapport à lui (jusqu'à ce que O' s'écrase à la surface de la masse M !).

Cette vitesse relative non-nulle entre observateurs conduit alors inévitablement à des différences dans les résultats de mesure de longueurs et de durées dans un référentiel donné. Par exemple, dans le référentiel \mathcal{R}' , lorsque l'observateur O' mesure des longueurs et durées propres L_0 et Δt_0 , l'observateur O_r mesure des longueurs plus courtes $L = L_0/\gamma_r$ et des durées plus longues $\Delta t = \gamma_r \Delta t_0$, d'un facteur :

$$\gamma_r = \frac{1}{\sqrt{1 - \frac{2GM}{rc^2}}} \simeq 1 + \frac{GM}{rc^2} \quad (9.9)$$

qui augmente lorsque la distance r diminue. Le même raisonnement s'applique bien évidemment aux mesures effectuées dans le référentiel \mathcal{R} où la masse M est immobile. Dans ce cas, c'est l'observateur O' , en mouvement à la vitesse $\sqrt{2GM/r}$ par rapport à \mathcal{R} , qui mesure des distances plus faibles et des durées plus longues, avec le même facteur γ_r .

Ainsi, dans le cas où la masse M est la masse terrestre, plus les observateurs sont proches du centre de la terre, plus les effets de la gravité terrestre sur les mesures de longueur et de durée sont importants.

9.1.5 Application : décalage vers le rouge d'origine gravitationnelle

L'exemple du paragraphe précédent nous montre que la présence d'un champ de gravitation modifie la perception qu'un observateur a des durées. Par extension, tout phénomène périodique verra sa fréquence modifiée dans un champ de gravitation. Pour les ondes électromagnétiques, le lien entre ce décalage spectral et le champ de gravitation peut s'obtenir aisément à l'aide du calcul habituel de l'effet Doppler comme on l'indique ci-dessous.

Considérons, par exemple, le cas d'une source S émettant une onde lumineuse de fréquence ν (sous forme d'un flash de courte durée) verticalement, vers le haut, à partir de la surface d'une planète. Quelle fréquence ν' va capter un détecteur D situé à une hauteur h au-dessus du point d'émission ?

Pour répondre à cette question, utilisons le fait 1) que la lumière se déplace à la vitesse c dans tout référentiel inertiel ; et 2) qu'un référentiel en chute libre est un référentiel inertiel.

$GM/rc^2 \ll 1$. Au-delà, l'effet de la gravité ne peut plus être décrit dans le cadre de la relativité restreinte et il faut considérer la théorie de la relativité générale, que nous aborderons plus loin.

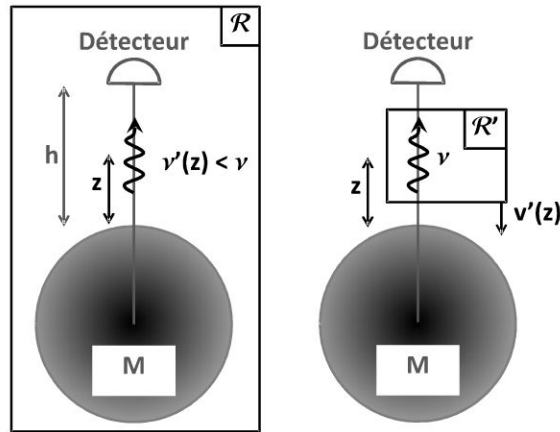


Figure 9.6 – Décalage vers le rouge d'origine gravitationnelle d'une onde lumineuse de fréquence ν . Dans le référentiel \mathcal{R} où la masse M et le détecteur sont au repos, le signal lumineux parcourt la distance h à la vitesse c . Dans le référentiel \mathcal{R}' , en chute libre par rapport à la masse M , le signal lumineux se déplace encore à la vitesse c , mais la vitesse relative entre \mathcal{R} et \mathcal{R}' crée un effet Doppler tel que $\nu'(z) < \nu$, où z est la distance verticale parcourue par le signal depuis le sol. La fréquence au niveau du détecteur est alors $\nu'(h) \simeq (1 - gh/c^2)\nu$.

Dans le référentiel en chute libre, l'onde lumineuse parcourt donc la distance h à la vitesse c et atteint le détecteur au bout du temps h/c (figure 9.6). Pendant ce temps, le référentiel en chute libre a acquis la vitesse $v = gh/c$ par rapport à la surface terrestre (si l'on suppose, pour simplifier, que l'accélération gravitationnelle est uniforme et égale à g). Dans le référentiel en chute libre, la fréquence de l'onde lumineuse n'a pas changé et est restée égale à ν . Par contre, ce référentiel inertiel s'éloigne à la vitesse $v = gh/c$ du détecteur D au moment de la détection de l'onde lumineuse. Il en résulte un effet Doppler vers le rouge (diminution de la fréquence, voir le chapitre 5) :

$$\nu' = \sqrt{\frac{1 - v/c}{1 + v/c}} \nu \simeq (1 - v/c) \nu = (1 - gh/c^2) \nu \quad (9.10)$$

où la formule approchée n'est valable que dans le cas d'un faible champ de gravitation. La variation relative de fréquence est alors égale à :

$$\nu' - \nu \simeq -\frac{gh}{c^2} \nu \quad (9.11)$$

qu'on peut écrire sous la forme :

$$\boxed{\frac{\Delta \nu}{\nu} = \frac{\Delta \phi}{c^2}} \quad (9.12)$$

9.1. Interaction gravitationnelle et notion de relativité générale

où $\Delta\phi$ est la *variation* du potentiel gravitationnel sur la distance h . Par conséquent, un observateur situé à une hauteur h au-dessus de la surface d'une planète mesurera un décalage vers le rouge des ondes émises par la planète.

Puisqu'une fréquence est l'inverse d'un temps ($\frac{\Delta t}{t} = -\frac{\Delta\nu}{\nu} \simeq \frac{gh}{c^2}$), une autre interprétation de ce résultat est la suivante : *le rythme d'une horloge située dans un champ gravitationnel augmente d'autant plus que le champ (c'est-à-dire l'accélération) est fort.*

Cet effet a été testé à de nombreuses reprises, lors de l'émission de signaux électromagnétiques par les sondes explorant le système solaire par exemple, ou en comparant le temps indiqué par deux horloges atomiques situées à des hauteurs différentes. Celles-ci sont maintenant si précises que l'effet a pu être mis en évidence sur une différence d'altitude de 30 cm (variation relative de 10^{-17}) !

L'effet de décalage vers le rouge d'origine gravitationnelle est ainsi la principale cause de dérive temporelle relativiste du système GPS (voir le chapitre 6). En effet, les deux effets relativistes principaux sont :

1. la dilatation des temps due à la vitesse orbitale des satellites ($\simeq 3,8$ km/s) :

$$\frac{\delta t}{t} \simeq 1 - \gamma \simeq -8.10^{-11} \quad (9.13)$$

conduisant à une dérive de $-7 \mu\text{s}$ par jour (soit une erreur de positionnement d'environ 2 km) ;

2. le redshift gravitationnel (satellites orbitant à une altitude moyenne $h_{\text{sat}} = 20\,200$ km) :

$$\frac{\delta t}{t} \simeq \frac{GM_{\text{terre}}}{c^2} \left(\frac{1}{R_{\text{terre}}} - \frac{1}{R_{\text{terre}} + h_{\text{sat}}} \right) \simeq +5,4.10^{-10} \quad (9.14)$$

conduisant à une dérive de $+46 \mu\text{s}$ par jour (soit une erreur de positionnement d'environ 14 km).

La somme de ces deux effets relativistes conduit alors à une erreur systématique (si elle n'est pas corrigée) de $+39 \mu\text{s}$ par jour, c'est-à-dire à une erreur de positionnement d'environ 12 km. Il est à souligner que l'effet relativiste gravitationnel est bien plus important que l'effet de relativité restreinte.

9.1.6 Le principe de relativité générale

Selon le principe d'équivalence, la gravitation n'existe pas en soi (puisque on peut la supprimer en se plaçant dans un référentiel en chute libre) mais n'est qu'une manifestation de notre appartenance à un certain référentiel.

Ceci généralise ce que l'on sait déjà sur le mouvement uniforme, à vitesse constante, qui *est comme rien* puisqu'on ne peut pas savoir, par des expériences internes au référentiel de l'objet en mouvement, s'il est véritablement en mouvement ou au repos.

Le principe d'équivalence va plus loin et pose que la gravitation, et d'une façon générale tout mouvement (accéléré ou non), *est comme rien* puisqu'on peut toujours choisir, localement, un référentiel dans lequel n'est plus ressentie d'accélération. Le **principe de relativité générale** d'Einstein étend alors le principe de relativité restreinte à tout type de mouvement :

Les lois de la physique doivent avoir la même forme dans tous les référentiels, quel que soit leur état de mouvement.

Du point de vue mathématique, ce principe se traduit par le fait que la forme des équations de la physique doit rester inchangée dans une transformation *quelconque* du système de coordonnées (et non plus seulement dans une transformation de Lorentz). Cette propriété mathématique porte le nom de *principe de covariance généralisée*.

9.1.7 Relativité générale et géométrie

On a vu aux paragraphes 9.1.4 et 9.1.5 que les longueurs et les durées sont modifiées par la gravité, c'est-à-dire que la *géométrie* (au sens étymologique du terme) de l'espace-temps est modifiée par un champ de gravitation. Or, dans tous les référentiels inertiels (c'est-à-dire ceux en chute libre dans un champ de gravitation), on doit retrouver les lois usuelles de la physique (c'est-à-dire la dynamique relativiste *restreinte*), et en particulier le principe d'inertie : « un corps soumis à aucune force se déplace en ligne droite » (et donc reste au même endroit s'il n'a pas de vitesse initiale). Le problème est alors le suivant : peut-on garder ce principe d'inertie dans tous les référentiels ? Comment le généraliser aux référentiels accélérés ? Ceci oblige à redéfinir ce que l'on entend par « ligne droite » car, dans un référentiel accéléré, les trajectoires des objets lâchés sont courbes ! (paraboliques en fait, si l'accélération est uniforme).

Einstein suggère alors de considérer une autre géométrie de l'espace-temps, non-euclidienne, dans laquelle les trajectoires de ces objets seraient des « lignes droites », c'est-à-dire les lignes de plus court chemin entre deux points. En mathématiques, de telles lignes s'appellent des *géodésiques* et leur forme dépend de la géométrie choisie. Par exemple, les géodésiques de l'espace euclidien usuel sont les lignes droites, alors que les géodésiques d'une sphère sont ses grands cercles, comme le savent très bien les navigateurs.

9.1. Interaction gravitationnelle et notion de relativité générale

Ainsi, la gravité (et toute accélération, de quelque origine que ce soit) affecte la géométrie de l'espace-temps en changeant la forme des géodésiques et, par conséquent, la forme des trajectoires suivies par les objets et particules (et en particulier les rayons lumineux). L'espace-temps n'est alors plus plat, on dit qu'il est courbe. D'une façon générale, toute forme d'énergie (non nécessairement gravitationnelle) contribue à changer la courbure de l'espace-temps.

Dans le cas de faibles champs de gravitation, la géométrie euclidienne habituelle reste valide, au moins localement, c'est-à-dire sur de faibles distances. Les résultats de relativité restreinte ne sont, eux, valides que tant que les effets des champs gravitationnels peuvent être négligés.

Le principe de covariance généralisée nous précise que les lois de la physique doivent s'exprimer à l'aide d'objets géométriques définis indépendamment des systèmes de coordonnées (des référentiels) choisis. Ces objets, dont les quadri-vecteurs de la relativité restreinte en sont un exemple, portent le nom de « tenseur » et l'équation fondamentale de la relativité générale est l'équation tensorielle suivante :

$$[\mathbf{T}] = \frac{8\pi G}{c^4} [\mathbf{G}] \quad (9.15)$$

où $[\mathbf{T}]$ est le tenseur impulsion-énergie de l'objet (matière et/ou champ électromagnétique) en mouvement dans l'espace-temps, et $[\mathbf{G}]$ le tenseur associé à la géométrie courbe de l'espace-temps créée par la distribution de masse. Cette équation n'exprime rien d'autre que l'équivalence entre géométrie et matière-énergie :

Energie et Matière = Géométrie

Il est à remarquer que l'espace-temps courbe de la relativité générale n'est pas localement bien défini car un corps en mouvement dans un champ de gravitation modifie lui-même ce champ de gravitation (à cause de sa propre masse par exemple !) et donc l'espace-temps dans lequel il évolue. Réciproquement, un champ de gravitation modifie la masse d'un objet, et l'on ne peut pas, en toute rigueur, définir sa masse indépendamment du champ de gravité... Ces effets surprenants sont cependant généralement si faibles que l'on n'en tient pas compte.

Notons également que, de même que l'accélération d'une charge crée des ondes électromagnétiques, l'accélération d'une masse doit créer des ondes, appelées ondes gravitationnelles, se propageant à la vitesse de la lumière (figure 9.7). De très faible intensité, les ondes gravitationnelles n'ont à ce jour été observées qu'indirectement, mais il existe de nombreux projets terrestres et spatiaux pour les détecter directement. Certains sont déjà opérationnels comme les détecteurs à barre de Weber (en géométrie

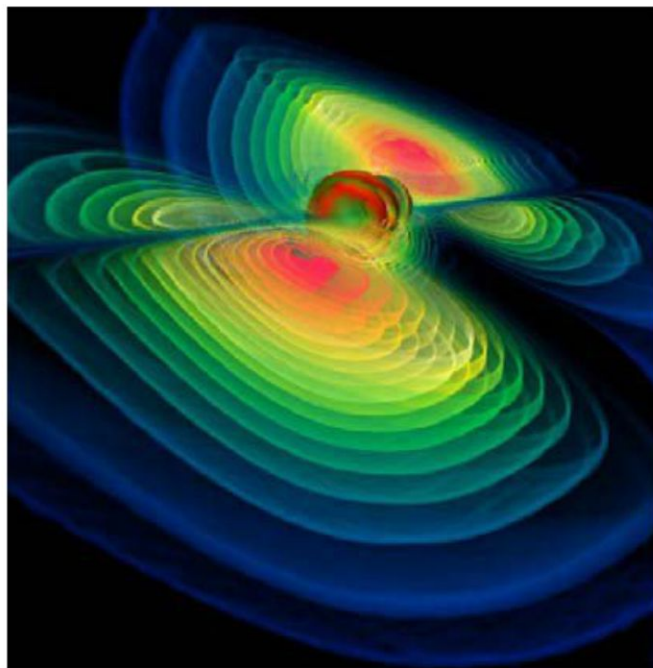


Figure 9.7 – Simulation numérique d'ondes gravitationnelles émises lors de la collision de deux trous noirs.

(source : Werner Benger, werner@cct.lsu.edu : <http://www.sfb.tpi.uni-jena.de/Projects/B7.shtml>)

cylindrique ou sphérique³) ou les gigantesques interféromètres de Michelson LIGO, VIRGO, GEO et TAMA (Figure 9.8).

9.2 INTERACTIONS ÉLECTROMAGNÉTIQUE, FORTE ET FAIBLE

Chaque interaction fondamentale peut être caractérisée par essentiellement trois paramètres :

1. le type de **charge** soumise à l'interaction (ainsi que son caractère cumulatif ou neutralisant) ;
2. son **intensité relative**, caractérisée par un nombre pur appelé *constante de couplage* ;
3. sa **portée**.

3. Les barres de Weber sont d'énormes barres de métal très rigides, isolées des vibrations externes et refroidies par cryogénie pour diminuer le bruit (interne) d'agitation thermique. Leur fonctionnement est très proche des antennes radio usuelles : la distorsion de l'espace-temps créée par le passage d'une onde gravitationnelle peut exciter la barre à sa fréquence de résonance et induire une déformation suffisamment importante pour être détectée par des capteurs piézoélectriques.

9.2. Interactions électromagnétique, forte et faible



Shématique de Virgo

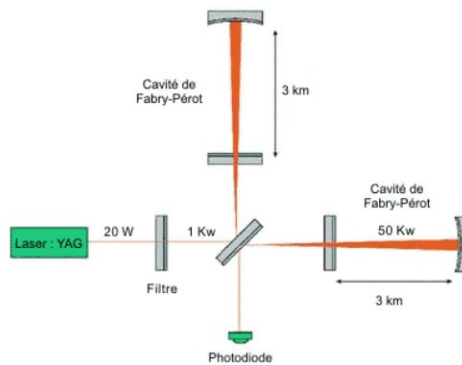


Figure 9.8 – Photo et schéma de principe de l’interféromètre VIRGO, situé près de Pise en Italie. Les deux bras de l’interféromètre de Michelson font 3 km.

(source : <http://lappweb.in2p3.fr/virgo/> et http://www.ego-gw.it/virgodescription/pag_4.html, contact : ego@ego-gw.it)

Le *modèle standard* conduit à envisager toute interaction entre deux particules comme un échange entre elles d’une autre particule, dite médiatrice de l’interaction (voir la figure 9.9). Le lien entre la portée d’une interaction et la masse m de la particule médiatrice de l’interaction s’exprime souvent sous la forme :

$$\text{portée} \simeq \frac{\hbar}{mc} \quad (9.16)$$

où \hbar/mc est la longueur d’onde Compton λ_{Compton} .

Cette relation peut se comprendre de la façon suivante : lors d’une interaction entre deux particules élémentaires de matière, la particule médiatrice est produite de manière « virtuelle » car son émission viole la loi de conservation de l’énergie relativiste (en effet, une particule élémentaire au repos de masse m_0 ne peut pas, selon la relativité restreinte, se transformer en deux particules élémentaires dont l’une aurait encore la masse m_0). C’est la mécanique quantique, dont l’étude se fera dans

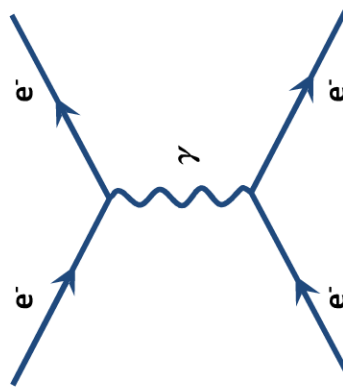


Figure 9.9 – Diagramme de Feynman décrivant l'interaction entre deux électrons par l'intermédiaire d'un photon γ (la particule médiatrice de l'interaction électromagnétique). Le temps s'écoule de la gauche vers la droite sur ce genre de diagramme.

la seconde partie de cet ouvrage, qui permet de résoudre le problème en autorisant la violation de la conservation de l'énergie pendant un court instant δt . En effet, la quatrième relation d'indétermination de Heisenberg indique qu'il peut y avoir un écart d'énergie δE à condition que la condition suivante soit vérifiée (chapitre 13.3) :

$$\delta E \cdot \delta t \simeq \hbar \quad (9.17)$$

Si la particule médiatrice de l'interaction a une masse m , cet écart en énergie doit être au moins égal à mc^2 , ce qui limite la durée pendant laquelle a lieu l'échange de cette particule virtuelle :

$$\delta t \lesssim \frac{\hbar}{mc^2} \quad (9.18)$$

et, par conséquent, la durée de vie de celle-ci. Pendant cet intervalle de temps, la particule virtuelle ne peut alors parcourir une distance qu'au plus égal à \hbar/mc , ce qui fournit un ordre de grandeur de la portée de l'interaction. La relation (9.16) indique qu'une portée infinie correspond à l'échange de particules de masse nulle. C'est le cas de l'électromagnétisme et de la gravitation, pour lesquelles ces particules se nomment photons et gravitons (bien que ces derniers soient hypothétiques car ils n'ont pas encore été détectés).

Nous venons de détailler l'interaction gravitationnelle, passons donc maintenant rapidement en revue certaines caractéristiques des trois autres interactions fondamentales : électromagnétique, forte et faible.

9.2.1 Interaction électromagnétique

Les particules concernées sont toutes les particules chargées électriquement. La loi de Coulomb donne l'expression de la force entre deux particules de charges q et q' :

$$\mathbf{F} = \frac{1}{4\pi\epsilon_0} \frac{qq'}{r^2} \mathbf{e}_r \quad (9.19)$$

où \mathbf{e}_r est le vecteur unitaire joignant les centres des charges concernées.

Il existe deux types de charge, positive et négative, ce qui implique l'existence de forces attractives ou répulsives. La charge électromagnétique est quantifiée à l'échelle microscopique (exemple : proton de charge $+e$, électron de charge $-e$, quark de charge $-e/3$ ou $+2e/3$).

L'interaction électromagnétique est au cœur des processus chimiques et biologiques. A l'échelle atomique (10^{-10} m), c'est l'interaction dominante. Pour caractériser cette intensité, on peut considérer par exemple le potentiel électrostatique entre deux charges identiques (protons ou électrons) distantes de $\lambda_{Compton}$: $e^2/(4\pi\epsilon_0\hbar/mc)$. La constante de couplage de l'interaction électromagnétique est alors définie comme le quotient de cette énergie potentielle par l'énergie de masse au repos d'une des charges :

$$\alpha_{e.m.} = \frac{e^2}{4\pi\epsilon_0\hbar/mc} / mc^2 = \frac{e^2}{4\pi\epsilon_0\hbar c} \simeq \frac{1}{137} \sim 10^{-2} \quad (9.20)$$

Cette constante est appelée *constante de structure fine*.

La portée de l'interaction électromagnétique est infinie à l'instar de la gravitation. A grande échelle (c'est-à-dire pour les objets ou zone macroscopiques), les forces peuvent être colossales mais très souvent se neutralisent, du fait de la présence de charges opposées. La force résiduelle est appelée force de van der Waals. Bien que très faible, cette force résiduelle joue un rôle important dans de nombreux domaines.

Les particules médiatrices de l'interaction sont les photons, de masse nulle (leur portée est bien infinie).

9.2.2 Interaction forte

C'est la force responsable de la cohésion des noyaux atomiques. Les particules de matière concernées sont les *quarks* représentant la moitié des constituants élémentaires de la matière (faite de « leptons » et de quarks, et des « antiparticules » associées, mais l'interaction forte n'agit que sur les quarks et antiquarks, pas sur les leptons (électrons, neutrinos...) ni sur les photons). Dans le cadre du modèle standard, il existe six quarks, dont sont constitués les mésons (assemblage d'un quark et d'un antiquark) et les baryons (assemblages de trois quarks), qui sont les deux sous-familles des « hadrons ».

Les deux quarks les plus présents dans la matière environnante sont les quarks « u » (pour up) et « d » (pour down). La charge électrique de ces quarks est différente : $+2e/3$ pour le quark u, et $-e/3$ pour le quark d. Ainsi, le baryon « uud » porte bien une charge $+e$ (c'est le proton), et le baryon « udd » est bien neutre électriquement (c'est le neutron).

La charge de l'interaction forte n'est pas la charge électrique, c'est la charge de couleur. Il existe en fait trois charges de « couleur » : rouge, bleu, vert (ces couleurs n'ont évidemment rien à voir avec les couleurs de la lumière). Chacun des six quarks existe en trois couleurs (et trois anti-couleurs pour les six anti-quarks). Comme en électromagnétisme, les charges de couleur identiques se repoussent et les charges de couleur opposées s'attirent. Les hadrons sont des assemblages de quarks de couleur « blanche » (équivalent à la neutralité électrique). Comme exemples remarquables, on peut citer les nucléons (protons et neutrons) qui sont des assemblages de trois quarks (c'est-à-dire des baryons) de couleurs différentes, et les pions qui résultent de l'assemblage d'un quark et d'un antiquark (ce sont donc des mésons).

L'interaction forte est la plus intense des quatre interactions fondamentales au niveau microscopique. Comme pour l'électromagnétisme, son intensité peut s'exprimer sous la forme d'une constante de couplage qui est le quotient de l'interaction forte entre deux nucléons distants de $\lambda_{Compton}$ par l'énergie de masse au repos d'un nucléon :

$$\alpha_{Forte} = \frac{g_F^2}{\hbar c} \approx 15 \quad (9.21)$$

où g_F est une constante intervenant dans l'expression du potentiel de Yukawa $V(r) = g_F^2 \exp[-r/\lambda_{Compton}]/r$ décrivant l'interaction forte entre nucléons.

Sa portée est très courte : les quarks n'existent que sous forme liée. L'interaction forte est très intense en dessous de 10^{-15} m, mais rapidement négligeable au-delà. Il en résulte un *confinement* des quarks dans les hadrons.

Les particules médiatrices de l'interaction forte entre quarks sont les « gluons ». Il existe huit gluons différents. Ils sont « colorés », c'est-à-dire qu'ils portent une charge de couleur comme les quarks. Ils sont donc sensibles à l'interaction qu'ils véhiculent, contrairement aux photons qui véhiculent l'interaction électromagnétique entre charges électriques mais ne portent pas eux-mêmes de charge électrique. Ainsi, à l'instar de l'interaction gravitationnelle en relativité générale, la théorie de l'interaction forte entre quarks (la *chromodynamique quantique*) est non-linéaire, ce qui complique sérieusement son étude.

Comme en électromagnétisme, il existe une force résiduelle : c'est la force nucléaire entre hadrons qui résulte de l'échange de pions entre les nucléons. Il existe trois types de pions médiateurs : Π^+ , Π^- , et Π^0 .

9.2. Interactions électromagnétique, forte et faible

Remarque

Pour contrebalancer et dépasser la répulsion électrique entre protons au sein des noyaux atomiques, la force nucléaire nécessaire est extrêmement forte. Il y a tout de même une répulsion à très faible distance (traduisant le fait que les nucléons ne s'interpénètrent pas). Il existe cependant des étoiles massives où cette répulsion entre nucléons est contrebalancée par l'interaction gravitationnelle, ce sont les étoiles à neutrons ou étoiles à quarks.

9.2.3 Interaction faible

Seules certaines particules sont concernées : celles de spin 1/2 (section 15.4) (par exemple, les nucléons). C'est une force très faible par rapport aux autres forces (forte et électromagnétique) qui ne peut engendrer de systèmes liés. Sa constante de couplage est de l'ordre de 10^{-5} :

$$\alpha_{faible} \simeq 10^{-5} \quad (9.22)$$

Elle change en fait plus volontiers le type de la matière que son mouvement : elle peut ainsi provoquer un changement d'identité, une transmutation des particules.

Elle est responsable des radioactivités β (transmutation de neutrons en protons) et joue un rôle majeur dans les processus de fusion nucléaire (chapitre 8). Les temps de désintégration associés sont extrêmement longs (par exemple, la durée de vie d'un neutron en dehors d'un noyau est de 15 minutes).

Les particules médiatrices de l'interaction faible sont les « bosons intermédiaires » W^+ , W^- et Z^0 dont la masse est très élevée (traduisant une portée d'interaction très faible). La prédiction de l'existence (et de la masse) de ces particules avant confirmation expérimentale a été l'un des grands succès de la théorie électrofaible, c'est-à-dire de la théorie unificatrice des interactions faible et électromagnétique.

L'interaction faible possède une propriété surprenante : contrairement aux autres interactions fondamentales, elle n'est pas invariante par symétrie miroir (c'est-à-dire par inversion spatiale des coordonnées) : elle ne conserve pas la « parité » comme illustré sur la figure 9.10.

9.2.4 Variation des constantes de couplage avec l'énergie

Du fait du phénomène d'écrantage par les autres particules élémentaires (virtuelles ou non), les constantes de couplage des interactions fondamentales varient avec l'échelle spatiale à laquelle on observe les particules, c'est-à-dire avec l'énergie de ces particules.

En électromagnétisme le phénomène d'écran à longue distance est dû à la polarisation du milieu par les charges électriques. Inversement, aux courtes distances, la constante de couplage électromagnétique croît.

Il en est de même pour la constante de couplage faible grâce au « champ de Higgs », responsable supposé de l'acquisition de masse des bosons intermédiaires.

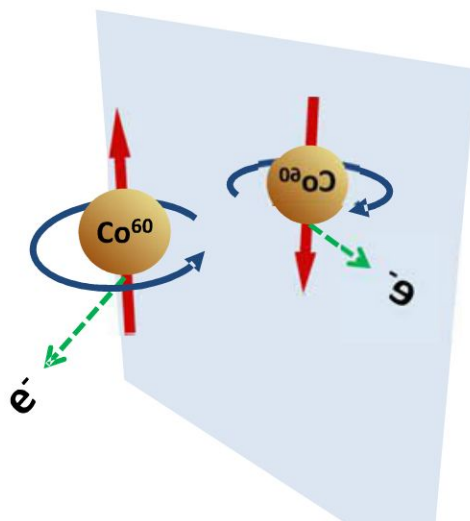


Figure 9.10 – Exemple de violation de parité dans la désintégration β d'un noyau de cobalt 60. A gauche, le phénomène tel qu'il a été observé expérimentalement : l'électron est émis préférentiellement dans la direction opposée à celle du spin du noyau (symbolisé par la flèche). A droite, l'image dans un miroir parallèle à la direction du spin de ce même phénomène : comme le spin est lié à une rotation des charges, cette rotation, et donc le spin, est inversée dans l'image, alors que la direction préférentielle d'émission des électrons est inchangée. Cette situation « miroir » n'est donc pas celle qui est observée expérimentalement : on en conclut que l'interaction responsable de cette désintégration n'est pas invariante par parité.

Inversement, aux courtes distances, les quarks sont quasi-libres (on parle de « liberté asymptotique ») par effet d'anti-écran des gluons.

Les différentes constantes de couplage semblent converger vers une même valeur aux hautes énergies : les interactions faible et électromagnétique s'unifient vers 10^{-18} m (10^3 GeV) ; les interactions électrofaible et forte s'unifient vers 10^{-31} m (10^{16} GeV). Il semble que les quatre interactions pourraient s'unifier vers 10^{-34} m (10^{19} GeV). De grands projets expérimentaux sont en cours (le Large Hadron Collider (LHC) au CERN de Genève par exemple) pour étudier ces régimes de très haute énergie.

9.3 CONCLUSION : THÉORIES D'UNIFICATION

Il n'aurait pas été juste de clore ce chapitre sans évoquer les tentatives d'unification des quatre interactions fondamentales, qui occupent actuellement plusieurs dizaines de milliers de scientifiques à travers le monde. L'unification des théories scientifiques est une démarche historique constante. Ainsi en est-il de l'électromagnétisme qui résulta de l'union, par Maxwell (et d'autres), de l'électricité et du magnétisme au sein d'une même théorie. De même, la relativité générale d'Einstein peut être vue

comme l'aboutissement d'un processus unificateur de la gravitation et du mouvement qui commença il y a très longtemps.

La théorie décrivant dans un même cadre interactions faible et électromagnétique est la théorie dite électrofaible. Développée par Glashow, Salam et Weinberg, elle fut confirmée par la découverte en 1983 des bosons intermédiaires qui étaient prédits par cette théorie.

La théorie décrivant dans un même cadre interactions électrofaible et forte est le *modèle standard*. Développée par de nombreux physiciens, dont S. Glashow, S. Weinberg, A. Salam, M. Gell-Mann et G. Zweig, c'est une très belle construction intellectuelle, basée sur la prise en considération de groupes de symétrie. Les données expérimentales actuelles, et en particulier la découverte en 2012 du Boson de Higgs, sont en bon accord avec les prédictions de cette théorie, qui reste néanmoins incomplète car elle contient des paramètres libres que seules les expériences peuvent fixer.

Il existe plusieurs tentatives de « grande unification » du modèle standard avec l'interaction gravitationnelle : théories des « cordes » et des « supercordes » par exemple. L'unification est rendue très complexe par l'absence de charge gravitationnelle négative, par l'égalité mystérieuse entre masse inertielle et masse gravitationnelle (c'est-à-dire le principe d'équivalence), par la forte non-linéarité de la relativité générale, et par le lien entre énergie et géométrie de l'espace-temps. Du fait des énergies gigantesques qu'il faut atteindre, les expériences actuelles ne permettent pas de tester directement ces théories, mais la recherche de tests indirects est très active.

Il faut noter que la théorie électrofaible et le modèle standard ont pour cadre conceptuel commun une description à la fois relativiste et quantique des phénomènes, ce qui fournit une transition naturelle vers la seconde partie de cet ouvrage, même si celle-ci n'ambitionne que de présenter les rudiments de la physique quantique, et non ses développements ultimes.

Partie 2

Mécanique quantique

LE MONDE QUANTIQUE

Avec les progrès technologiques des dernières décennies, la planète Terre nous est maintenant accessible dans son ensemble : on peut par exemple obtenir à l'aide de son smartphone ou son ordinateur des vues très détaillées de n'importe quel point de sa surface prises depuis un satellite. Et pourtant la différence entre l'échelle des objets nous environnant (typiquement le décimètre) et l'échelle Terrestre (typiquement 10 000 km) est énorme : huit ordres de grandeur !

Effectuons maintenant le même saut de huit ordres de grandeur, mais vers l'infiniment petit au lieu de l'infiniment grand : nous atteignons le nanomètre ($10^{-9}m$), et nous entrons dans le monde des atomes, des molécules, des mailles cristallines. C'est l'échelle privilégiée du *monde quantique*, dont la description fait l'objet de la deuxième partie de ce livre.

Pour l'infiniment petit, les technologies ont fait aussi de grands progrès dans les dernières décennies, et l'on dispose maintenant de l'équivalent microscopique des satellites d'observation Terrestre : *on peut « voir » les objets et les phénomènes quantiques individuels, et agir sur eux de manière individualisée* et non plus seulement avoir accès à des phénomènes résultant de l'accumulation d'un nombre gigantesque de phénomènes microscopiques comme c'était le cas auparavant. On assiste actuellement à l'essor rapide des *nanosciences* et des *nanotechnologies* qui explorent et exploitent ces nouvelles possibilités.

Des particules extrêmement petites comme l'électron, le proton et d'autres particules élémentaires sont relativement aisées à visualiser individuellement lorsqu'elles sont produites dans les accélérateurs de particules en raison de leur très grande énergie. La figure II.1 donne un exemple d'« événement » individuel, en l'occurrence le résultat d'une collision frontale entre deux protons de très grande énergie.

Ce n'est que plus récemment qu'on a pu voir des objets microscopiques individuels d'énergie bien moindre. La figure II.2 montre à titre d'exemple une photographie de l'expérience de D. Wineland, prix Nobel 2012, où l'on voit trois ions qui sont piégés par un champ électromagnétique approprié en des points bien précis de l'espace. On peut ici préciser ce que « voir » un objet quantique veut dire : la figure montre l'image, enregistrée par une caméra, de la lumière émise par ces trois ions bien individualisés lorsqu'on les a excités par un laser. Les lois de l'optique interdisent que les taches lumineuses observées sur la figure aient une dimension inférieure à la longueur d'onde de la lumière utilisée, en l'occurrence le micromètre. Il faut donc imaginer que

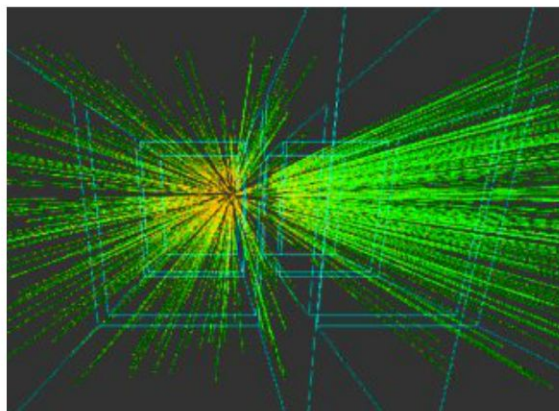


Figure II.1 – Image d'une collision proton-proton unique reconstruite par le détecteur ATLAS du Large Hadron Collider du CERN à Genève, montrant les trajectoires de toutes les particules produites par la collision.

(source : CERN)

chaque ion occupe une zone mille fois plus petite que la tache lumineuse créée par sa présence. « Voir » veut donc dire disposer d'un témoin capable de suivre au cours du temps la présence d'un objet microscopique individuel. Il en va de même pour les trajectoires montrées sur la figure II.1, qui ont été déterminées par ordinateur à partir des signaux électriques provenant des différents détecteurs entourant le point de collision.

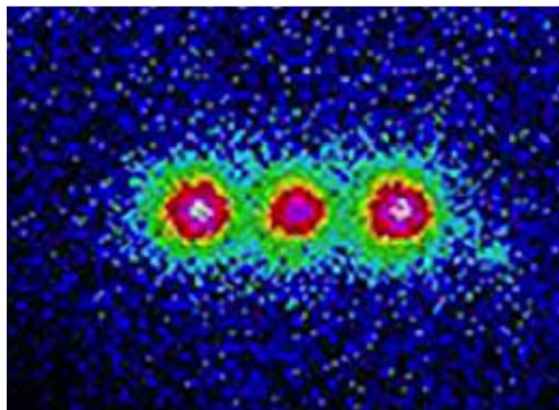


Figure II.2 – Image de la lumière de fluorescence émise par trois ions de Béryllium excités par laser.

(source : NIST, david.wineland@nist.gov : http://www.nist.gov/public_affairs/factsheet/quantum_infosci.cfm)

On peut aussi voir, dans le sens que nous venons de préciser, des particules aussi fragiles que des photons, qui sont d'ordinaire très faciles à détruire. La figure II.3 montre un résultat de l'expérience de S. Haroche, prix Nobel 2012, qui retrace l'« histoire » de quelques photons individualisés piégés dans une cavité formée de miroirs

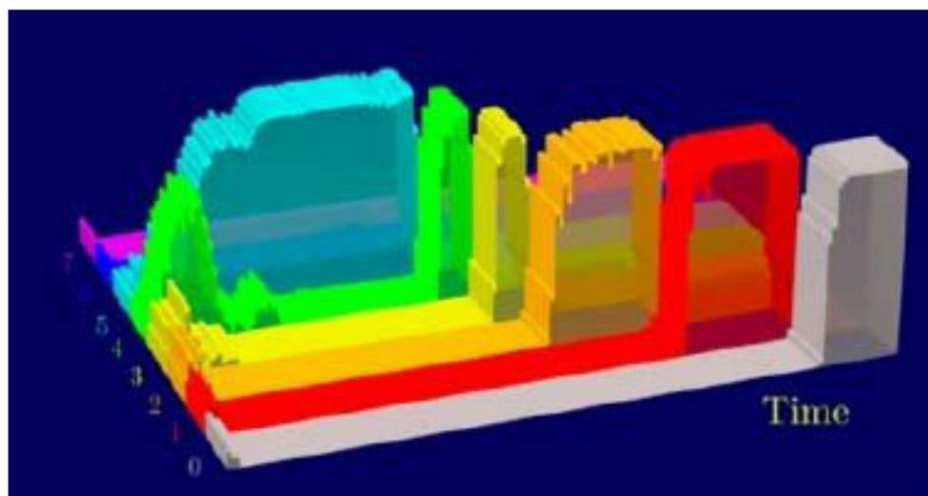


Figure II.3 – Les différentes traces montrent l'évolution en fonction du temps de la présence d'un nombre donné de photons piégés dans une cavité (de 0 à 5 du premier plan au dernier plan) et mesurés de manière non destructive. Aux temps courts, une phase initiale d'excitation permet de peupler la cavité avec 1, 2, 3, 4 puis finalement 5 photons qui subsistent dans la cavité pendant plusieurs millisecondes. Les inévitables « fuites » de la cavité éliminent ensuite un par un les photons qu'elle contient, jusqu'à son vidage complet (trace au premier plan). La « perte » de chacun de ces photons est un événement soudain précisément enregistré.

(source : <http://www.nobelprize.org>)

de très grande réflectivité et dont on peut connaître le nombre à tout instant sans les détruire.

D'autres techniques sont l'analogue exact du satellite d'observation Terrestre, car elles permettent de « voir » la surface d'un solide à l'échelle de l'atome, et la présence éventuelle d'atomes différents « adsorbés » sur cette surface, c'est-à-dire collés sur cette surface par les forces à courte portée créées par les atomes du cristal. C'est ce que l'on observe sur la figure II.4 qui montre sur quelques nanomètres la surface d'un cristal d'oxyde de Cérium observée par microscope à force atomique, avec ses atomes régulièrement disposés, sur laquelle on voit quelques atomes supplémentaires adsorbés.

On peut en fait non seulement observer le monde quantique mais agir sur lui de manière individualisée, par exemple rectifier la position d'un ion particulier dans un piège, remettre un photon dans la cavité dès qu'on sait qu'un s'est échappé, ou encore modifier la topographie d'une surface cristalline, comme on peut le voir sur l'exemple spectaculaire de la figure II.5.

On peut aussi façonner atome par atome des édifices cristallins ayant des propriétés particulièrement intéressantes pour des applications diverses. On peut par exemple créer des cristaux semi-conducteurs par *épitaxie*, c'est-à-dire par dépôt successif des couches cristallines individuelles, en modifiant de manière contrôlée le type d'atomes

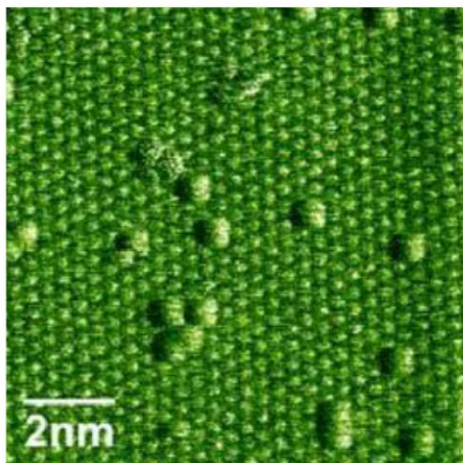


Figure II.4 – Surface d'oxyde de Cérium observée par microscope à force atomique.

(source : <http://spie.org/x48127.xml>, Christoph Richter *et al.*, Nanoworld, Erlangen)

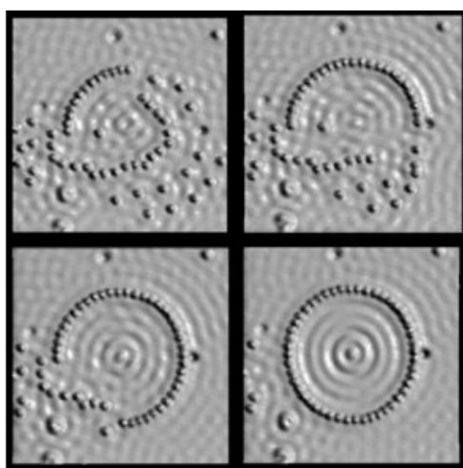


Figure II.5 – Surface de cuivre sur laquelle l'expérimentateur construit atome par atome un cercle d'atomes de fer adsorbés.

(« corral quantique », source : Commie, Lutz, Eigler IBM)

déposé. On crée ainsi un monocristal de taille macroscopique dont la composition est contrôlée au niveau de l'atome, comme on le voit sur la figure II.6 qui montre une coupe transverse d'un cristal ainsi fabriqué. Ce type de matériau a en particulier des propriétés optiques intéressantes. Il est à la base de la fabrication de « lasers à puits quantique » qui sont largement utilisés dans les réseaux actuels de télécommunication. On voit donc que les nanotechnologies ne sont pas de simples amusements de laboratoires, mais débouchent sur des applications essentielles pour notre monde actuel, en microélectronique pour construire des ordinateurs de plus en plus compacts et performants et dans les technologies de l'information.

Dans les chapitres qui suivent, nous allons décrire progressivement les caractéristiques essentielles du monde quantique et introduire les lois qui le régissent, connues

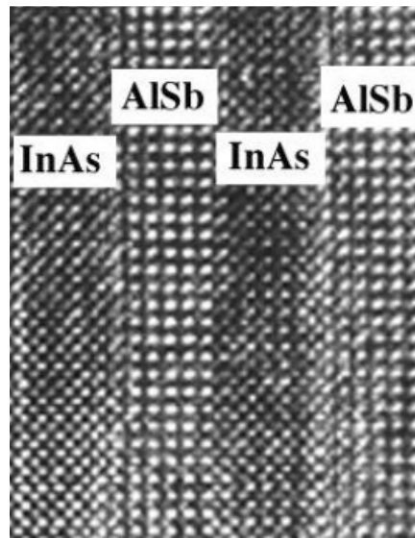


Figure II.6 – Coupe d’une hétérostructure cristalline InAs/AlSb fabriquée par épitaxie par jet moléculaire. Les atomes formant la structure sont clairement visibles et ont été déposés couche par couche. L’épaisseur de chaque couche d’InAs ou d’AlSb est de 1,8 nm. On remarque que la structure régulière du cristal se maintient même lorsque le cristal change de composition.

(source : Institut Max Planck de la physique des microstructures, Halle)

sous le nom de *mécanique quantique*. Celles-ci vont nous sembler souvent étranges, car elles diffèrent de celles qui sont tirées de notre expérience quotidienne et bâties sur des observations effectuées à notre échelle humaine, 10^8 fois plus grande. L’objet de ce livre est d’en saisir les aspects essentiels, de se forger peu à peu une « intuition quantique », sans toutefois entrer dans les détails de la description d’objets comme les atomes, les molécules, les solides ou la lumière, sujets pour lesquels le lecteur se réfèrera à des manuels plus détaillés et spécialisés, dont nous mentionnons certains dans la bibliographie à la fin de ce manuel.

LE PHOTON : UNE INTRODUCTION À LA PHYSIQUE QUANTIQUE

10



10.1 Le photon : retour à la conception corpusculaire de la lumière ?

Comme nous l'avons indiqué au chapitre 2, dans la deuxième moitié du XIX^e siècle, Maxwell avait établi sur des bases solides la nature ondulatoire de la lumière, en montrant qu'elle était une onde de champ électrique et magnétique obéissant à quatre équations, les équations de Maxwell. De nombreux physiciens avaient approfondi cette analyse, en particulier Hertz et Lorentz, et montré que les phénomènes lumineux connus trouvaient une explication satisfaisante dans ce cadre. Hertz avait expérimentalement montré qu'il existait aussi des ondes électromagnétiques de plus basse fréquence, les ondes dites hertziennes, qui avaient des propriétés analogues à celles de la lumière. À la fin du XIX^e siècle, les physiciens pouvaient être satisfaits : ils avaient résolu pour l'essentiel le problème de la lumière.

Il revient au même Hertz d'avoir introduit une première lézarde dans cet édifice, par sa découverte fortuite du phénomène dit « effet photoélectrique », que nous décrirons en détail dans les paragraphes 10.1.1 et 10.2.2, et qui ne trouvait pas d'explication simple dans l'approche électromagnétique. De même, la répartition spectrale de la lumière émise par un corps chauffé (appelé « rayonnement du corps noir »), étudiée expérimentalement à la fin du XIX^e siècle, ne semblait pas pouvoir être expliquée de manière satisfaisante en termes d'ondes de Maxwell. C'est Max Planck qui, à l'orée du XX^e siècle proposa une explication à ce dernier phénomène, et montra qu'on trouvait la bonne répartition spectrale à condition de supposer que les échanges d'énergie entre la matière et la lumière se faisaient par blocs insécables, ou « quanta », $h\nu$ proportionnels à la fréquence ν du rayonnement, introduisant ainsi sa fameuse constante h , la constante de Planck, caractéristique des phénomènes quantiques.

Reprenant et approfondissant cette hypothèse révolutionnaire, Einstein en 1905, en même temps qu'il établissait les bases de la relativité restreinte, émettait l'hypothèse que « l'énergie d'un rayon lumineux émis par une source ponctuelle n'est pas continûment distribuée dans un volume de plus en plus grand, mais consiste en un nombre fini de quanta d'énergie qui sont des points localisés dans l'espace, se déplacent sans se diviser, et qui peuvent être produits et absorbés en unités complètes ». Il était alors en mesure d'expliquer simplement l'effet photoélectrique découvert par Hertz : il s'agissait donc du retour, mais sur des bases bien différentes, de l'hypothèse corpusculaire chère à Newton.

Mais il fallait bien évidemment rendre compatible l'approche ondulatoire de Maxwell, qui expliquait l'ensemble des phénomènes d'interférence et de diffraction observés jusqu'alors, et l'approche corpusculaire d'Einstein qui expliquait l'effet photoélectrique et le rayonnement du corps noir. C'est ce qui a été accompli au XX^e siècle par un certain nombre de physiciens au premier rang desquels on peut citer Dirac et Glauber. À partir des concepts quantiques généraux, c'est-à-dire du formalisme de la mécanique quantique élaboré dans les années 1920, ils mirent au point une théorie, appelée « électrodynamique quantique », ou « optique quantique », qui explique par un formalisme unique l'ensemble des phénomènes lumineux connus jusqu'à présent. Il n'est pas question de la présenter dans ce manuel, dont le but est d'introduire un certain nombre de concepts simples permettant une première approche de ces phénomènes.

10.1 LE PHOTON : UN FAIT EXPÉRIMENTAL

De nombreuses expériences simples permettent de démontrer la nature corpusculaire de la lumière. Prenons l'exemple d'une mesure réalisée quotidiennement par les physiciens contemporains : la mesure de la distance entre la Terre et la Lune (figure 10.1). Grâce à des réflecteurs posés sur la Lune par les différentes missions spatiales, qu'elles soient américaines ou russes, il est possible d'obtenir cette distance en mesurant le temps que met une impulsion lumineuse émise depuis un observatoire Terrestre pour faire l'aller-retour. Les ordres de grandeur de cette expérience sont tout à fait remarquables. Des impulsions lumineuses d'environ une centaine de picosecondes et d'énergie 200 mJ sont émises depuis la Terre, mais du fait de la diffraction, qui élargit le faisceau à l'aller et au retour, et de la petite taille des réflecteurs lunaires, seule une toute petite partie de l'énergie revient dans le détecteur environ 2 s plus tard. Si on fait un bilan énergétique complet on trouve qu'il y a un rapport 10^{20} entre la puissance envoyée et la puissance reçue ! Expérimentalement, on observe qu'à ces puissances extrêmement faibles le détecteur ne se comporte plus de façon linéaire. Il produit un signal qui n'est pas celui qu'on observerait sur l'impulsion initiale divisé par 10^{20} : la plupart du temps le signal est nul, mais pour environ une impulsion sur cinquante on observe l'apparition d'une brève « bouffée », que nous appellerons « clic » et qui est d'intensité constante : l'arrivée d'un photon a été enregistrée par le détecteur. L'arrivée de ces photons est totalement irrégulière : c'est une illustration directe de la *nature probabiliste de la théorie quantique* que nous allons décrire dans ce chapitre. L'énergie moyenne reçue par impulsion étant environ 50 fois plus faible que celle d'un photon unique, celui-ci étant insécable, au lieu de détecter 1 cinquantième de photon à chaque tir on détecte un photon arrivant au hasard, en moyenne une fois tous les 50 tirs.

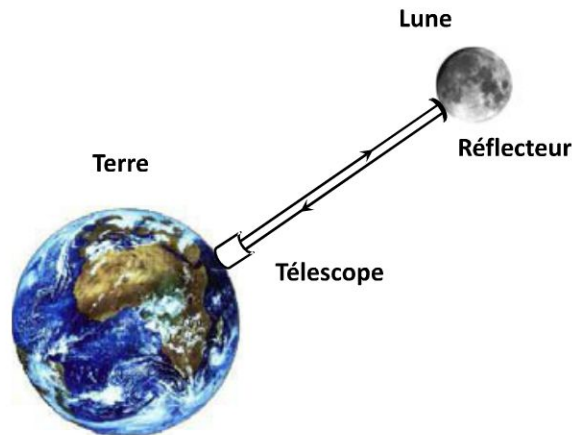


Figure 10.1 – Laser Terre-Lune : on mesure le temps d’aller-retour d’une brève impulsion laser, émise depuis un télescope, et réfléchi par un dispositif posé sur la Lune. Le détecteur Terrestre reçoit en moyenne un photon toutes les cinquante impulsions lumineuses.

10.1.1 L’effet photoélectrique

Revenons à l’effet photoélectrique mentionné dans l’encart 10.1 : Hertz découvre en 1887 que de la lumière incidente sur un conducteur électrique peut lui arracher des électrons. En 1902, Lenard démontre expérimentalement que c’est le nombre d’électrons émis qui augmente linéairement avec la puissance du rayonnement incident et non l’énergie de ces électrons. Par contre, cette dernière dépend linéairement de la fréquence du rayonnement.

En 1905, Einstein postule l’existence de quanta de lumière, qui seront appelés plus tard « photons », dont l’énergie E est proportionnelle à la fréquence ν de la lumière considérée :

$$E = h\nu = \hbar\omega \quad (10.1)$$

où $\omega = 2\pi\nu$ est la pulsation de la lumière, $h \approx 6,6 \cdot 10^{-34}$ J.s est la constante de Planck et $\hbar = \frac{h}{2\pi} \approx 10^{-34}$ J.s est appelée « constante de Planck réduite ».

L’explication de l’effet photoélectrique est alors très simple si l’on suppose que l’interaction se fait quanta par quanta avec conservation de l’énergie pour chaque processus élémentaire : un photon incident sur la plaque métallique éjecte un électron, et l’énergie E_{electron} de l’électron émis est égale à l’énergie du photon incident diminuée de l’énergie d’extraction $W_{\text{extraction}}$ nécessaire pour arracher l’électron à la plaque.

$$E_{\text{electron}} = h\nu - W_{\text{extraction}} \quad (10.2)$$

Il en résulte qu’il y a un seuil d’énergie, donc de fréquence, pour que le processus soit possible : l’énergie $h\nu$ du photon incident doit être supérieure à l’énergie d’extraction $W_{\text{extraction}}$. Enfin le nombre d’électrons arrachés est, dans les conditions idéales, égal au nombre de photons incidents.

10.1.2 La lumière, c'est combien de photons ?

Le spectre de la lumière visible se situe autour de la longueur d'onde de 600 nm, et la fréquence associée est d'environ $\nu = 500$ THz. L'énergie d'un photon visible est donc de l'ordre de $E = h\nu \approx 4 \cdot 10^{-19}$ J. On remarque que c'est de l'ordre de l'électron-Volt ($1,6 \cdot 10^{-19}$ J). Le photon a donc une énergie très petite, qui est néanmoins détectable par des détecteurs suffisamment sensibles. Pour se rendre compte de l'énergie du photon on peut utiliser l'unité traditionnelle du nombre de particules individuelles dans le monde macroscopique : la mole. Une mole de photons contient une énergie de $6 \cdot 10^{23} \times 4 \cdot 10^{-19} \approx 200$ kJ. Une ampoule électrique qui émet une puissance lumineuse d'une dizaine de Watt émet 0,00005 mol de photons par seconde, autrement dit avec une mole de photons on peut éclairer une pièce pendant environ 5 heures 30 ! Ainsi, avec une simple pile on peut alimenter une lampe pendant des dizaines d'heures (les moles de réactif dans la pile étant directement transformées en moles de photons)¹.

On peut aussi prendre l'exemple du pointeur laser qui délivre une puissance lumineuse d'1 mW. En divisant cette puissance par $h\nu \approx 4 \cdot 10^{-19}$ J, on obtient le flux numérique de photons, qui est d'environ $2,5 \cdot 10^{15}$ photons par seconde, tous de même fréquence et se propageant dans la même direction.

10.1.3 La détection des photons

La détection des photons est assurée par divers dispositifs :

- les **bolomètres**, ou **Wattmètres**, détectent la puissance lumineuse par l'effet d'échauffement produit par la lumière. La plupart fonctionnent avec des flux de photons importants, correspondant à des puissances lumineuses macroscopiques, où la « granularité » de la lumière n'est pas visible. Certains, refroidis à la température de l'Hélium liquide, sont suffisamment sensibles pour mesurer l'échauffement produit par un seul photon !
- les **photodiodes** sont basées sur l'effet photoélectrique, où, comme nous venons de le voir, des électrons sont arrachés à la matière sous l'effet des photons incidents. Le courant d'électrons ainsi créé est proportionnel à la puissance lumineuse P . Lorsqu'on atténue fortement la lumière pour atteindre des puissances de l'ordre de 10^{-15} W et qu'on utilise des photodiodes suffisamment sensibles, on constate comme dans l'exemple de la mesure de la distance Terre-Lune que le courant mesuré ne diminue pas uniformément : il apparaît comme formé de brèves « bouffées » de courant ou « clics », de formes identiques, correspondant à l'arrivée aléatoire des photons, arrivées de plus en plus rares lorsqu'on diminue l'éclairement. C'est le recouvrement de ces « bouffées » lorsqu'on augmente la puissance incidente qui

1. Pour plus de détails voir J. M. Courty et N. Treps, « La Lumière, c'est combien de photons ? », *Dossier Pour la Science* **53**, octobre-novembre 2006.

crée le signal macroscopique continu. Il existe aujourd'hui des photodiodes ayant des efficacités quantiques supérieures à 90%, autrement dit qui ont 90% de chance de détecter un photon incident.

- L'œil humain est un remarquable détecteur, car il est sensible à quelques photons : en effet, connaissant la taille de la pupille, la luminance de sources (comme une nuit de pleine Lune, ou l'étoile la plus faible visible à l'œil) et le temps de rémanence de l'œil, on trouve qu'une dizaine de photons suffit à former une image visible.

10.2 PROPRIÉTÉS DU PHOTON

10.2.1 Énergie et quantité de mouvement

Chaque photon contenu dans un faisceau de lumière monochromatique décrit par une onde plane de fréquence ν et de vecteur d'onde \mathbf{k} a une énergie E et une quantité de mouvement \mathbf{p} données par :

$$E = h\nu \quad ; \quad \|\mathbf{p}\| = h/\lambda \quad (10.3)$$

La direction du vecteur \mathbf{p} est celle de la direction de propagation de l'onde, et h est la constante de Planck. On peut aussi écrire la relation plus symétrique :

$$E = \hbar\omega \quad ; \quad \mathbf{p} = \hbar\mathbf{k} \quad (10.4)$$

À partir de la relation relativiste (7.55) du chapitre 7 reliant E , \mathbf{p} et la masse m de la particule : $E^2 - \mathbf{p}^2 c^2 = m^2 c^4$ et de la relation $\omega^2 = c^2 \mathbf{k}^2$, on déduit que $m = 0$. *Le photon est donc une particule de masse nulle* et qui se déplace dans le vide à vitesse constante, toujours égale à c . Elle est donc à la fois quantique et relativiste.

10.2.2 Processus élémentaires d'émission et d'absorption de photons

a) Propriété fondamentale

L'interaction entre la lumière et un objet tel qu'un atome, une molécule, un solide, *se fait par processus élémentaires mettant en jeu un seul photon*. Il y a conservation de l'énergie et de la quantité de mouvement dans chaque processus élémentaire. Si E et \mathbf{p} sont l'énergie et la quantité de mouvement de l'objet avant le processus, et E' et \mathbf{p}' l'énergie et de la quantité de mouvement de l'objet après le processus, on peut écrire, pour un processus d'absorption de la lumière par l'objet (donc d'annihilation du photon) :

$$E + h\nu = E' \quad ; \quad \mathbf{p} + \hbar\mathbf{k} = \mathbf{p}' \quad (10.5)$$

et pour un processus d'émission de la lumière par l'objet, donc de création d'un photon (figure 10.2) :

$$E = E' + h\nu \quad ; \quad \mathbf{p} = \mathbf{p}' + \hbar\mathbf{k} \quad (10.6)$$

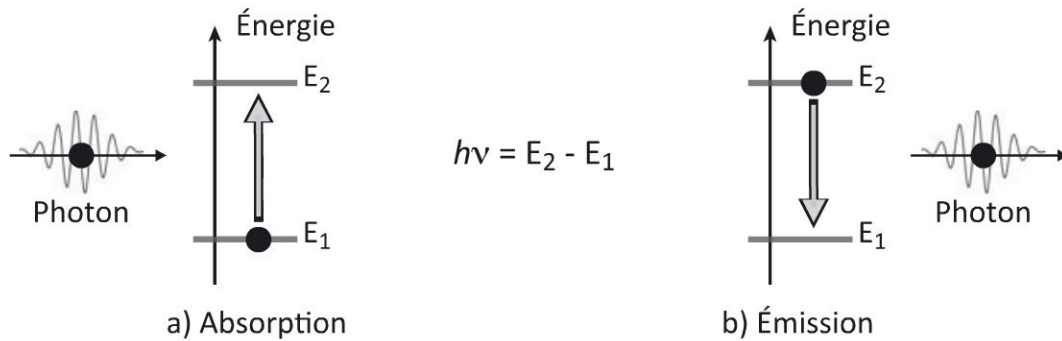


Figure 10.2 - Diagramme montrant les deux niveaux de plus basse énergie E_1 et E_2 d'un atome. a) Excitation de l'atome par absorption d'un photon incident ; b) Désexcitation de l'atome par émission d'un photon.

b) Exemples

Comme on le verra par la suite de manière plus détaillée, *l'énergie E d'un objet quantique ne peut pas prendre n'importe quelle valeur*. Elle ne peut prendre qu'un certain nombre de valeurs bien précises E_1, E_2, \dots , correspondant à des configurations, ou états, différents. On dit alors que l'énergie de ces systèmes ou objets est quantifiée. C'est le cas des atomes ou des molécules (figure 10.2). Elle peut aussi appartenir à des intervalles d'énergie $E_{\min} < E < E_{\max}$. On parle alors de « bandes » d'énergie, ou de « continuums » (figure 10.3). C'est le cas des solides. Lorsque le système a la plus basse énergie possible, on dit qu'il est dans l'*état fondamental*.

Donnons quelques exemples de ces processus d'émission et d'absorption de lumière :

- **Diode électroluminescente (LED).** Le système formé par le contact entre deux semi-conducteurs constitue une diode, système quantique qui présente des bandes d'énergie : une bande supérieure $E_{c,\min} < E < E_{c,\max}$, appelée *bande de conduction*, et une bande inférieure $E_{v,\min} < E < E_{v,\max}$, appelée *bande de valence* (avec $E_{c,\min} > E_{v,\max}$). Un courant électrique I (en ampères) appliqué aux bornes de cette diode permet d'apporter des électrons dans la bande supérieure. Il peut y avoir alors émission d'un photon, et le système se trouve à la fin dans la bande de valence de plus basse énergie. La fréquence du photon émis est alors telle que $h\nu > E_{c,\min} - E_{v,\max}$. Dans les conditions idéales, chaque électron formant le courant électrique d'alimentation fournit un photon : le courant numérique de photons est

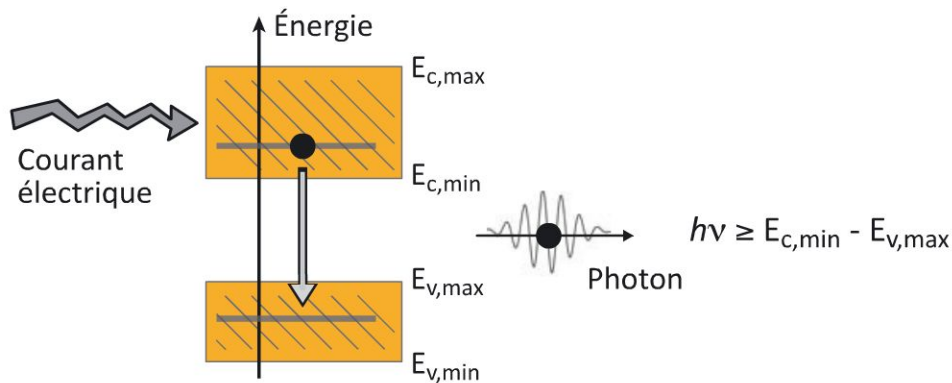


Figure 10.3 – Diagramme montrant les deux bandes d'énergie $[E_{c,min}, E_{c,max}]$ et $[E_{v,min}, E_{v,max}]$ accessibles à un électron dans une diode électroluminescente (LED). Un courant électrique appliqué sur la diode apporte les électrons à une énergie située dans la bande d'énergie supérieure. La diode se désexcite vers un niveau d'énergie situé dans la bande inférieure avec émission d'un photon.

égal au courant numérique électrique. La puissance lumineuse émise vaut alors :

$$P = \frac{h\nu I}{q} \quad (10.7)$$

avec $q = 1,6 \cdot 10^{-19}$ C. Si ν correspond à la lumière jaune de longueur d'onde $0,6 \mu\text{m}$, un courant de 1 mA dans la LED produit une puissance lumineuse de 2,5 mW, proche de celle du pointeur laser.

- **Réaction photochimique.** On considère deux molécules A et B qui, mises en contact forment, par réaction chimique, une autre molécule AB, système quantique qui possède des niveaux d'énergie définis (on dit aussi « discrets »). Dans certains cas, la réaction exothermique produit la molécule dans un niveau d'énergie E_1 supérieur à l'énergie E_0 de l'état fondamental, et il y a émission d'un photon lorsque la molécule AB transite vers l'état fondamental : la réaction chimique produit de la lumière. La fréquence du photon émis vaut $(E_1 - E_0)/h$. Une mole de réactif A donne alors naissance à une *mole de photons*, soit 200 kJ pour la lumière jaune. Si la réaction est rapide, voire explosive, et épuise les réactifs en une seconde, on a un flash de lumière très intense, de puissance 200 kW. Si la réaction est lente et dure une heure, il y a émission pendant ce temps de lumière de puissance 56 W. C'est ce qui se produit dans les « bâtons lumineux » qu'on trouve dans les fêtes foraines.
- **Effet photoélectrique.** Nous l'avons déjà abordé dans le paragraphe d'introduction : un électron dans un métal possède un état fondamental d'énergie E_0 et une bande d'énergie, ou continuum dont l'énergie inférieure vaut E_{min} (figure 10.4). Le premier état correspond à l'électron confiné dans le métal et libre de s'y déplacer. Le continuum correspond à l'électron éjecté du métal. Lorsqu'on éclaire

le métal, il y a un processus d'absorption, où le photon du faisceau est absorbé et communique son énergie à l'électron. Si sa fréquence est trop faible, telle que $h\nu < E_{\min} - E_0 = W_{\text{extraction}}$, le processus ne peut avoir lieu. Si au contraire $h\nu > E_{\min} - E_0 = W_{\text{extraction}}$, l'électron peut se trouver à l'issue du processus dans un état du continuum, donc être éjecté du métal, et on peut le recueillir sur une électrode. La lumière produit donc dans ce cas un courant d'électrons. C'est le processus inverse de celui qui se produit dans une diode électroluminescente : le courant numérique d'électrons y est (dans les conditions idéales) égal au courant numérique des photons incidents. Pour la lumière jaune, 2,5 mW de lumière produisent 1 mA de courant électrique, ou de manière équivalente, 1 mW de lumière produit 0,4 mA de courant.

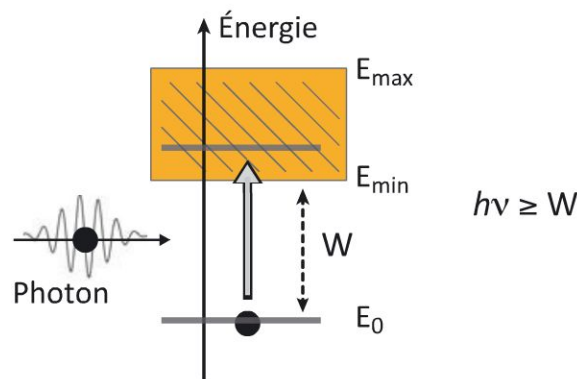


Figure 10.4 - Diagramme montrant le niveau de plus basse énergie E_0 et la bande d'énergie $[E_{\min}, E_{\max}]$ accessibles à un électron dans un métal. Sous l'effet d'un photon incident, l'électron initialement dans l'état de plus basse énergie passe dans la bande d'énergie supérieure, qui correspond à un électron éjecté hors du métal avec une certaine énergie cinétique (effet photoélectrique).

- **Excitation et modification de la vitesse d'un atome.** Un atome possède des niveaux d'énergie discrets $E_1 < E_2 < \dots$. S'il se trouve dans son état fondamental de plus basse énergie, il peut effectuer un saut vers le niveau excité d'énergie E_2 en absorbant un photon de fréquence ν telle que :

$$\nu = \frac{E_2 - E_1}{h} \quad (10.8)$$

C'est la *règle de Bohr*. Il y a aussi conservation de la quantité de mouvement au cours du processus. Si M est la masse totale de l'atome, \mathbf{v}_i sa vitesse initiale, et \mathbf{v}_f sa vitesse après absorption du photon de vecteur d'onde \mathbf{k} , on a donc :

$$M\mathbf{v}_i + \hbar\mathbf{k} = M\mathbf{v}_f \quad (10.9)$$

L'atome change donc de vitesse sous l'effet du rayonnement, d'une quantité qui vaut $\hbar k/M$, de l'ordre du cm.s^{-1} pour un atome de masse $M = 10^{-25}$ kg. La variation est faible, mais non négligeable. Une fois excité, l'atome revient dans son état

10.3. Temps et lieu d'arrivée du photon

de plus basse énergie en un temps de l'ordre de 10^{-8} s. Il est alors prêt à absorber un nouveau photon qui va provoquer un nouveau changement de vitesse identique, de l'ordre du cm.s^{-1} . Le taux de changement de vitesse, c'est-à-dire l'accélération ou la décélération appliquée par un faisceau laser à l'atome est donc de $1 \text{ cm.s}^{-1}/10^{-8} \text{ s} = 10^6 \text{ m.s}^{-2}$, c'est-à-dire 10^5 fois l'accélération de la pesanteur Terrestre !

Ainsi, si on utilise des faisceaux convergents de fréquence bien ajustée (figure 10.5), on peut amener la vitesse de l'atome à une valeur pratiquement nulle, correspondant à des températures de l'ordre de 10^{-6} K .

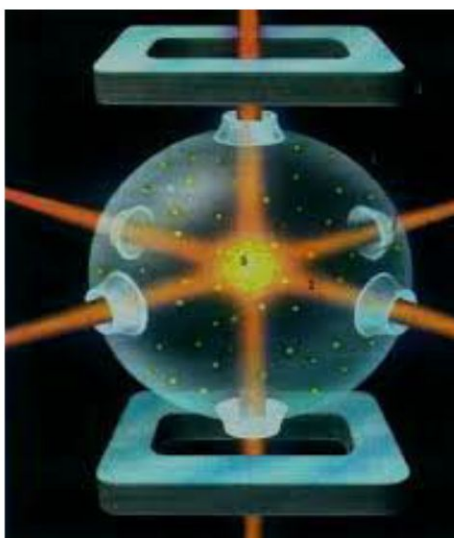


Figure 10.5 – Les atomes contenus dans une cellule et soumis à six faisceaux laser convergents en présence d'un champ magnétique inhomogène sont piégés et maintenus à très basse température au centre du dispositif.

(source : Institut Non Linéaire de Nice)

Remarque

Il existe en fait d'autres processus élémentaires que ceux que nous venons d'évoquer dans cette partie, qui mettent en jeu non pas un mais plusieurs photons. On les appelle *processus multiphotoniques*. Par exemple, il peut y avoir transition d'un atome de son état fondamental à un état excité par absorption simultanée de deux photons. Ces processus sont la plupart du temps beaucoup moins probables que ceux à un seul photon, ou bien nécessitent de très grands éclaircissements lumineux.

10.3 TEMPS ET LIEU D'ARRIVÉE DU PHOTON

10.3.1 Temps d'arrivée du photon dans un faisceau lumineux stationnaire

On constate expérimentalement que les photons contenus dans un faisceau lumineux stationnaire, de puissance constante, arrivent au hasard, et que l'instant exact d'arrivée d'un photon ne peut être prédit, même en contrôlant parfaitement la source. Le

nombre de photons N arrivant dans un intervalle de temps T donné est une quantité intrinsèquement fluctuante : *les phénomènes quantiques individuels sont de nature statistique*.

Considérons un phénomène stationnaire, où la probabilité de détecter un photon ne dépend pas du moment où on fait la détection, mais de la durée de cette détection. C'est une situation typique d'une source lumineuse continue composée d'un grand nombre d'émetteurs aléatoires indépendants. Ainsi, pour un détecteur faisant de multiples mesures, chacune pendant un temps Δt , on appelle X_i la variable aléatoire associée au nombre de photons détectés lors de la i^{e} mesure. Le nombre moyen de photons mesurés \bar{N} pendant une mesure est alors égal à la *moyenne* des mesures effectuées (pour un grand nombre de mesures). On écrit, en anticipant sur les notations de la mécanique quantique, $\bar{N} = \langle X \rangle$ (où l'indice i a été omis car la moyenne est indépendante de la mesure considérée). On peut montrer ² que les probabilités de mesure suivent une *loi de Poisson* (figure 10.6) :

$$\text{Proba}(X_i = n) = e^{-\bar{N}} \frac{\bar{N}^n}{n!} \quad (10.10)$$

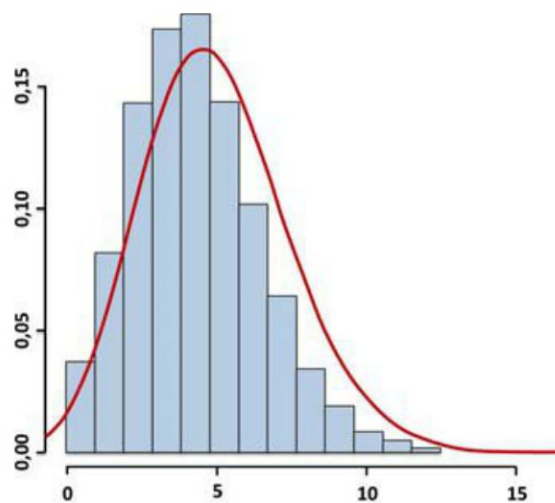


Figure 10.6 – Histogramme du nombre de photons mesurés à l'aide d'un compteur de photons sur une source lumineuse très faible pendant un intervalle de temps donné. La valeur moyenne du nombre de photons détectés est de 5, et l'écart-type autour de cette valeur est de $\sqrt{5}$.

Considérons maintenant la *variance* du nombre de photons mesurés, qui représente les fluctuations d'une mesure à l'autre. On a :

$$\text{Var}(X) = \Delta X^2 = \langle X^2 \rangle - \langle X \rangle^2$$

2. John Taylor, *Incertitudes et analyses des erreurs dans les mesures physiques*, Dunod 2000.

Dans le cas d'une loi de Poisson, on montre que la variance est égale à la valeur moyenne $\Delta X^2 = \bar{N}$. Cela signifie que si on effectue une mesure unique on a 68 % de chance qu'elle soit dans l'intervalle $[\bar{N} - \sqrt{\bar{N}}, \bar{N} + \sqrt{\bar{N}}]$ et 95 % de chance qu'elle soit dans l'intervalle $[\bar{N} - 2\sqrt{\bar{N}}, \bar{N} + 2\sqrt{\bar{N}}]$.

On voit que naturellement le nombre de photons par unité de temps fluctue. Cela a-t-il une influence sur des mesures macroscopiques ? Si on considère un faisceau laser de quelques milliwatts, cela correspond à environ 10^{16} photons par seconde. La variance associée à une mesure d'une seconde est donc 10^8 photons, soit une fluctuation relative de 10^{-8} , toute petite. Prenons par contre le cas d'un appareil photo numérique, et faisons le calcul dans l'autre sens, quel est le nombre minimal de photons par pixel permettant une image de qualité ? Si on considère qu'une fluctuation de 1 % est raisonnable, cela correspond à 10 000 photons en moyenne par pixel. Pour 10 millions de pixels, il faut donc environ 10^{11} photons pour former une image correcte. C'est ce nombre qui fixe la limite du temps d'exposition pour les prises de nuit.

10.3.2 Lieu d'arrivée du photon : éclairage uniforme

Certains appareils photographiques numériques CCD (*Charge Coupled Device*) sont extrêmement sensibles, à tel point qu'ils peuvent enregistrer l'arrivée d'un photon unique sur un de leurs pixels. On peut ainsi savoir où arrive le photon sur la surface sensible du détecteur.

Considérons tout d'abord le cas où le CCD est soumis à l'éclairage uniforme d'une onde lumineuse plane : si elle est suffisamment intense, chaque pixel fournira un signal identique, et la photo montrera une teinte plate (c'est-à-dire uniforme). Si l'on baisse l'intensité incidente tout en maintenant le caractère uniforme de l'éclairage, on constate que les photons, maintenant détectés un par un, arrivent en des points aléatoires de la surface du CCD. On retrouve sur l'aspect spatial le caractère probabiliste de la détection qu'on avait vu précédemment sur l'aspect temporel. Si on mesure le nombre de photons arrivant sur différentes zones de même surface du CCD pendant le temps d'ouverture de l'appareil photo, on obtiendra une qualité fluctuante de zone en zone, avec ici aussi une statistique de Poisson et une variance égale à la racine carrée de la valeur moyenne.

10.3.3 Lieu d'arrivée du photon : cas de la lame semi-réfléchissante

Considérons un système composé d'une source lumineuse et d'une lame séparatrice (L) (figure 10.7) qui divise le faisceau incident en deux faisceaux de même puissance. Nous avons déjà rencontré une telle lame dans le chapitre 2 au cœur de l'interféromètre de Michelson.

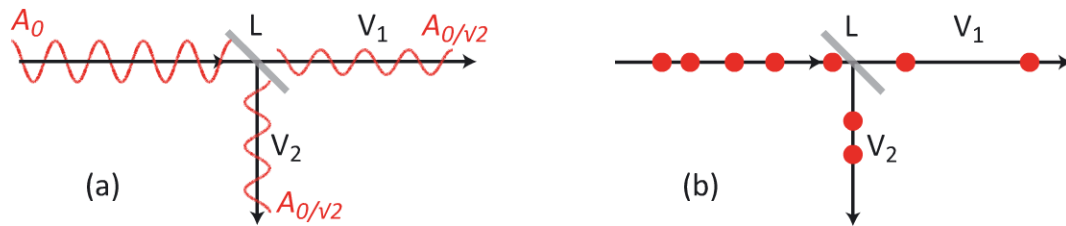


Figure 10.7 – a) Une onde lumineuse arrivant sur la lame semi-réfléchissante se divise en deux ondes d’amplitudes égales b) Les photons arrivant sur la lame ne peuvent pas être coupés en deux : ils se répartissent aléatoirement sur les deux faisceaux de sortie.

Dans le point de vue de Maxwell, si $A = A_0 \exp(i(kz - \omega t))$ représente le champ électrique en notation complexe avant la lame, les champs électriques sur les deux faisceaux issus de la lame s’écrivent respectivement $A_1 = \frac{A_0}{\sqrt{2}} \exp(i(kz - \omega t))$ et $A_2 = \frac{A_0}{\sqrt{2}} \exp(i(kz - \omega t))$. De cette façon, les puissances lumineuses $P = \alpha |A|^2$ (où α est une constante de proportionnalité réelle positive) des faisceaux de sortie sont égales, et la somme des puissances sur les deux voies de sortie est égale à la puissance incidente $P_1 + P_2 = P$.

Qu’en est-il quand on baisse cette puissance et que les photons arrivent un par un sur la lame ? Par définition, ceux-ci ne peuvent se couper en deux. Ils doivent donc être soit transmis, soit réfléchis par la lame. On est amené à adopter dans ce cas aussi une description probabiliste du comportement du photon au passage de la lame, très différente de la description déterministe de l’approche ondulatoire : on ne peut savoir si un photon donné incident sur la lame va être réfléchi ou transmis. Les probabilités de transmission et de réflexion valent :

$$\text{Proba}_{D_1}(1) = \text{Proba}_{D_2}(1) = \frac{1}{2}$$

où l’on a adopté la notation $\text{Proba}_D(n)$ pour exprimer la probabilité que le détecteur D détecte n photons. Ces probabilités sont parfaitement déterminées, et si on se place à la limite d’un grand nombre de photons, on va évidemment retrouver la loi classique.

On a ici un exemple de tirage au sort caractéristique des phénomènes quantiques. Le phénomène fondamentalement aléatoire qui se produit à la traversée d’une lame semi-réfléchissante est d’ailleurs utilisé pratiquement dans des dispositifs commerciaux pour produire des suites de nombres réellement aléatoires, données utiles pour le calcul numérique et les jeux de hasard sur Internet, et qu’il est très difficile d’obtenir par d’autres moyens.

10.4 INTERFÉRENCES LUMINEUSES ET PHOTONS

Comment concilier un phénomène aussi intrinsèquement ondulatoire qu'une interférence avec la granularité introduite par la présence du photon ? C'est ce que nous allons analyser plus en détail dans ce paragraphe.

Considérons deux ondes lumineuses planes, de vecteurs d'onde \vec{k}_1 et \vec{k}_2 , se propageant dans des directions différentes du plan xOz , inclinées d'un angle θ de part et d'autre de l'axe Oz . Elles se superposent donc dans une partie de l'espace (figure 10.8). D'un point de vue expérimental, on peut par exemple superposer deux faisceaux lasers, ou bien envoyer un faisceau préalablement élargi sur un *biprisme de Fresnel* (figure 10.8b), qui est formé de deux prismes de verre identiques P_1 et P_2 accolés par leurs bases et qui dévient les parties supérieures et inférieures du faisceau lumineux incident dans des directions symétriques par rapport à l'axe de propagation initial.

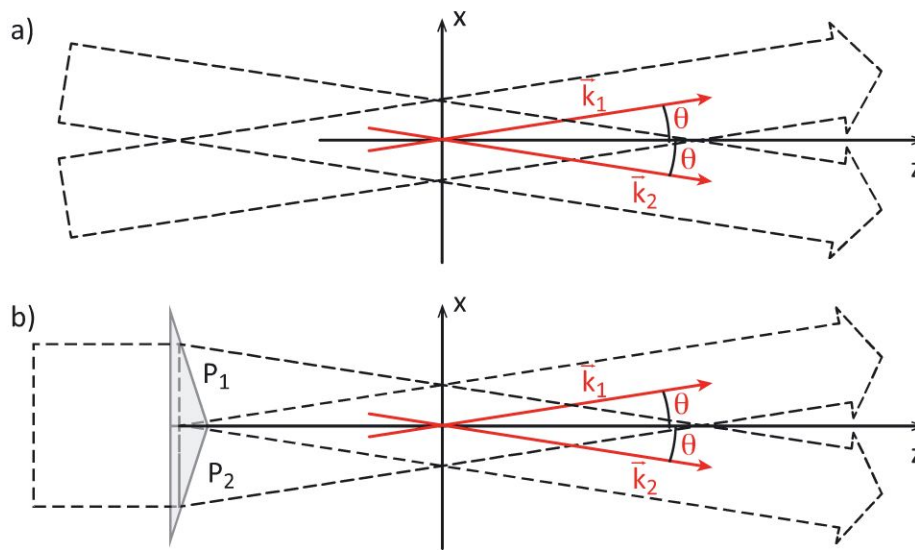


Figure 10.8 – a) Superposition de deux ondes planes ayant des directions de propagation, donc des vecteurs d'onde \vec{k}_1 et \vec{k}_2 , symétriques par rapport à Oz . b) Production de deux ondes planes symétriques à l'aide d'un faisceau lumineux élargi se réfractant dans deux prismes de verre P_1 et P_2 symétriques.

10.4.1 Interférence entre deux ondes planes : approche classique

Envisageons tout d'abord la lumière comme une onde, que nous supposons de pulsation ω et de vecteur d'onde k . Les vecteurs d'onde des deux ondes planes ont pour composantes cartésiennes $(-k \sin \theta, 0, k \cos \theta)$ et $(k \sin \theta, 0, k \cos \theta)$. Leurs amplitudes

complexes, introduites à la section 10.3.3, sont égales dans ce cas à :

$$\begin{aligned} A_1(\mathbf{r}, t) &= A_0 e^{i(kz \cos \theta - kx \sin \theta - \omega t)} \\ A_2(\mathbf{r}, t) &= A_0 e^{i(kz \cos \theta + kx \sin \theta - \omega t)} \end{aligned}$$

Dans la région de recouvrement, l'onde résultante a une amplitude qui est la somme des amplitudes de ces deux ondes et qui est donc égale à :

$$A(\mathbf{r}, t) = A_1 + A_2 = 2A_0 e^{i(kz \cos \theta - \omega t)} \cos(kx \sin \theta) \quad (10.11)$$

La puissance lumineuse dans la région de recouvrement est proportionnelle au module carré de l'amplitude complexe de cette onde et vaut donc :

$$P(\mathbf{r}) = 4\alpha A_0^2 \cos^2(kx \sin \theta) \quad (10.12)$$

La *puissance lumineuse n'est plus uniforme* : elle dépend de l'abscisse x du point d'observation : si on insère dans cette région un capteur CCD perpendiculaire à la direction de propagation Oz on va donc enregistrer une image comportant des bandes parallèles à Oy alternativement éclairées et sombres, appelées *franges d'interférences*, avec des points où $P = 0$, donc non éclairées. En ces points « de la lumière plus de la lumière donne de l'obscurité », car les deux ondes y sont exactement égales en valeur absolue et opposées en signe : ce phénomène d'*interférence destructive* est caractéristique de la nature ondulatoire de la lumière. Nous l'avons déjà rencontré dans une configuration différente, celle de l'interféromètre de Michelson (section 2.3.3)

10.4.2 Interférences photon par photon

Utilisons maintenant de la lumière très atténuée (figure 10.9). À très court temps de pose, le nombre moyen de photons incidents est très faible, et les photons arrivent sur le détecteur CCD en des points aléatoires, sans que l'on sache où va arriver un photon donné. On retrouve l'aspect statistique des phénomènes quantiques individuels déjà signalé.

En revanche, lorsqu'on accumule les arrivées de photons sur le détecteur, on voit apparaître progressivement une modulation dans la densité des points d'arrivée, qui ressemble de plus en plus à la variation de la puissance lumineuse de l'onde déterminée dans le paragraphe précédent : les franges d'interférence apparaissent. On constate qu'en fin de compte la valeur moyenne du nombre de photons mesurés en un point donné varie comme l'éclairement ou comme la puissance lumineuse P en ce point.

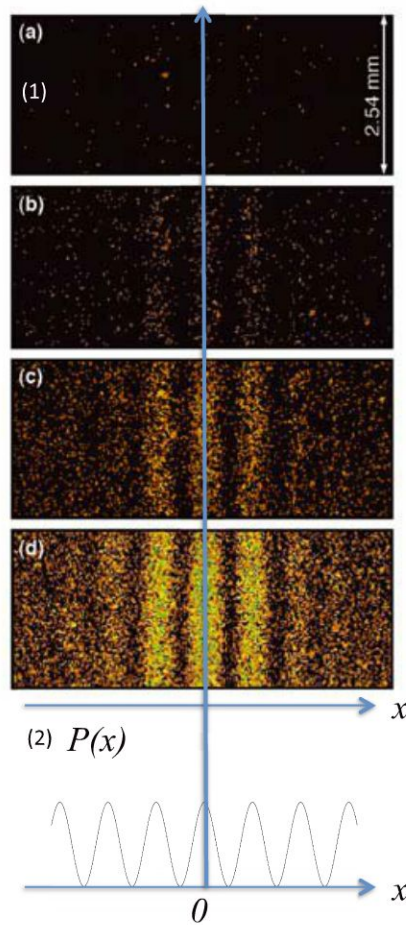


Figure 10.9 - (1) Signal détecté par le CCD pour une puissance lumineuse très faible et pour des temps de pose de plus en plus longs (de (a) à (d) : 10, 100, 500 et 2000 s ; (2) Puissance $P(x)$ de l'onde électromagnétique lumineuse en fonction de la position x : la probabilité de détecter le photon est proportionnelle à la puissance lumineuse.
(source : ENS Cachan, http://www.physique.ens-cachan.fr/old/franges_photon/interference.htm)

On est donc amené à admettre que la probabilité $Proba(\mathbf{r})$ de trouver le photon au point \mathbf{r} est proportionnelle à la puissance lumineuse P en ce point, elle-même proportionnelle au carré du module de l'onde complexe en ce point³ :

$$Proba(\mathbf{r}) \propto P \propto |A(\mathbf{r}, t)|^2$$

L'onde « guide » en quelque sorte le photon. Comme le carré du module de l'amplitude $A(\mathbf{r}, t)$ donne la probabilité de présence, à une constante multiplicative près, on qualifie souvent $A(\mathbf{r}, t)$ d'« amplitude de probabilité ».

3. Le symbole « \propto » veut dire « proportionnel à ».

10.4.3 Discussion

Prenons l'exemple de l'expérience du biprisme éclairé par un faisceau lumineux très atténué dans lequel les photons se présentent de manière isolée à l'entrée du dispositif et considérons les trois situations suivantes :

- 1) On bloque la lumière issue du prisme P_2 . L'onde plane A_1 passe seule. L'éclairement sur le CCD est uniforme et il n'y a pas de phénomène d'interférence. La probabilité $Proba_1$ obtenue dans ce cas ne dépend pas de la position,
- 2) On bloque la lumière issue du prisme P_1 . L'onde plane A_2 passe seule, il n'y a pas d'interférence et la probabilité $Proba_2$ est uniforme,
- 3) On laisse passer toute la lumière : les deux ondes passent, se superposent et interfèrent. La probabilité $Proba_{1+2}$ n'est plus uniforme.

On a donc :

$$Proba_{1+2} \neq Proba_1 + Proba_2 \quad (10.13)$$

Supposons un instant que le photon soit une particule classique à laquelle on peut attribuer une trajectoire bien définie : il passe donc soit par le prisme supérieur, soit par le prisme inférieur, et la probabilité de trouver le photon en un point donné du CCD doit être la somme de celles des deux configurations exclusives 1 et 2. Cette hypothèse implique donc que $Proba_{1+2} = Proba_1 + Proba_2$.

Comme cette égalité n'est pas satisfaite par le photon, on en conclut que celui-ci n'est pas une particule classique dont on peut connaître à tout instant la position. Pour décrire ses propriétés, il faut à la fois considérer la lumière comme une onde dont on ajoute les amplitudes, et comme une particule dont la probabilité de présence est liée au carré du module de l'amplitude de cette onde. On constate en particulier qu'à l'emplacement des franges noires, où il y a interférence parfaitement destructive pour l'onde, on peut dire avec certitude que le photon ne sera jamais détecté en ces points. De la même manière que de la lumière plus de la lumière peut donner une absence de lumière, on peut dire que pour le photon « de la probabilité plus de la probabilité peut donner une probabilité nulle ».

10.5 BILAN

Ce chapitre nous a permis de voir que des expériences finalement assez courantes nécessitent une description de la lumière allant au-delà de la physique classique pour être interprétées. Il faut admettre que la lumière a aussi un aspect « granulaire » et aléatoire. La dualité onde corpuscule pour la lumière s'exprime par les relations $E = \hbar\omega$, $\mathbf{p} = \hbar\mathbf{k}$. Nous avons vu que les échanges d'énergie au niveau microscopique se faisaient photon par photon, en respectant la relation $h\nu = E_2 - E_1$. Les phénomènes quantiques isolés ont un caractère intrinsèquement aléatoire, et l'on ne

peut faire que des prédictions de caractère statistique. Plus précisément la probabilité de détection du photon à t et en \mathbf{r} est proportionnelle au module carré de l'amplitude complexe $|A(\mathbf{r}, t)|^2$ du champ électrique de l'onde lumineuse. Lorsqu'on augmente l'éclairement donc qu'on passe à des flux de photons très importants, la granularité de la lumière, et son caractère aléatoire associé, n'est plus observable et on retrouve une onde « classique » dont l'amplitude varie continûment.



10.2 lame semi-réfléchissante, interférences et dualité onde-corpuscule

L'interféromètre de Mach-Zehnder est une variante de l'interféromètre de Michelson que nous avons déjà étudié dans le chapitre 2 (figure 10.10a). Une lame semi-réfléchissante L_1 sépare une onde incidente monochromatique de longueur d'onde λ en deux faisceaux distincts, qui sont ensuite superposés à nouveau sur une autre lame semi-réfléchissante L_2 grâce aux miroirs M_1 et M_2 . L'intensité de la lumière issue de cette recombinaison varie en fonction de la différence de longueur $\delta = L_1 M_1 L_2 - L_1 M_2 L_2$ entre les trajets des deux ondes : elle présente des maxima quand les deux ondes sont en phase après L_2 ($\delta = p\lambda$, avec p entier). Elle s'annule quand les deux ondes sont en opposition de phase lorsqu'elles se recombinent ($\delta = (p + 1/2)\lambda$).

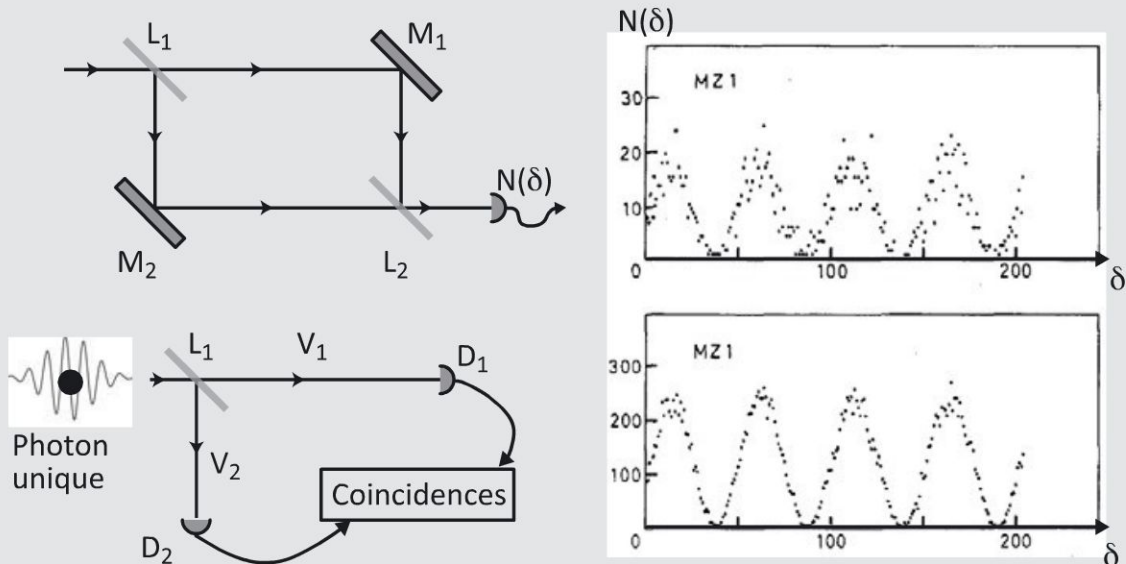


Figure 10.10 - a) Interféromètre de Mach-Zehnder ; b) L'absence de comptage en coïncidence sur les voies V_1 et V_2 montre que le photon est passé par une seule des deux voies V_1 et V_2 ; c) Signal interférométrique détecté par le compteur de photons pour deux temps de pose différents (en haut 10 s, et en bas 2000 s).

(Figure reproduite avec permission depuis A. Aspect, P. Grangier, G. Roger, *Europhysics Letters*, vol. 1, p. 173 1986)

L'expérience⁴ que nous allons décrire, réalisée à l'Institut d'Optique d'Orsay, permet de bien mettre en évidence l'étrangeté de la dualité onde-corpuscule. Elle consiste à envoyer de manière répétitive un photon unique à l'entrée de l'interféromètre de Mach-Zehnder et à réaliser successivement deux séries de mesures :

- La première (figure 10.10b) consiste à détecter d'éventuelles coïncidences entre les arrivées des photons sur deux compteurs de photons D_1 et D_2 insérés sur les deux voies de sortie V_1 et V_2 de la lame semi-réfléchissante L_1 . L'expérience n'en détecte aucune aux incertitudes expérimentales près, ce qui veut bien dire que le photon incident isolé est passé soit sur la voie V_1 soit sur la voie V_2 , mais jamais sur les deux à la fois : comme indiqué plus haut, on ne peut pas couper en deux un photon.
- Dans la deuxième série, on retire les compteurs de photons D_1 et D_2 , et on aligne l'interféromètre de Mach-Zehnder en recombinaison les deux faisceaux V_1 et V_2 sur la lame L_2 . On met un compteur de photons sur le faisceau de sortie, qui enregistre le signal $N(\delta)$ résultant de l'accumulation pendant un temps T donné des comptages pour une différence de trajet δ donnée. On change ensuite la valeur de δ et on répète l'expérience. La figure 10.10c donne le nombre de photons $N(\delta)$ en fonction de δ pour des valeurs de plus en plus grandes de T . On constate qu'on voit apparaître peu à peu la variation sinusoïdale du signal en fonction de δ , démontrant comme dans le cas du biprisme le caractère à la fois ondulatoire et statistique des phénomènes liés à ces photons uniques. Mais la situation est nouvelle car on est maintenant sûr, grâce à l'expérience précédente, que le photon n'est passé que par un seul des deux chemins possibles !

Cette expérience illustre donc la propriété quantique surprenante, voire paradoxale, que le phénomène d'interférence se produit, non pas lorsque deux photons se « retrouvent » après avoir pris des chemins différents, mais bien lorsqu'un seul photon « a le choix » entre deux chemins possibles. Le photon passe en quelque sorte « à la fois » par les deux chemins, de manière à être sensible à δ et par un seul de ces chemins, pour ne pas donner lieu à comptage en coïncidence. On a ici une illustration du phénomène spécifiquement quantique d'*interférence de chemin*, où ce sont les différentes « histoires » possibles de la particule qui donnent lieu à interférence avec addition des amplitudes de probabilité correspondantes.

4. A. Aspect, P. Grangier, G. Roger *Europhysics Letters* vol.1, p. 173 1986.

POLARISATION DE LA LUMIÈRE : ASPECTS CLASSIQUES ET QUANTIQUES

11

11.1 DESCRIPTION CLASSIQUE DE LA POLARISATION

Avant d'aborder la description quantique d'un faisceau lumineux nous allons préciser certaines propriétés classiques de la lumière liées au caractère vectoriel du champ électrique, qui se traduit par le phénomène de *polarisation de la lumière*.

11.1.1 Onde électromagnétique et champ électrique transverse

Considérons une onde plane lumineuse de pulsation ω se propageant dans la direction z . Cette onde consiste en un vecteur champ électrique oscillant, dont Maxwell a montré qu'il oscille dans une direction transverse à la direction de propagation de la lumière, donc dans le plan xOy . Tout vecteur de ce plan peut se décomposer sur une base de deux vecteurs unitaires \mathbf{e}_H parallèle à Ox et \mathbf{e}_V parallèle à Oy (figure 11.1).

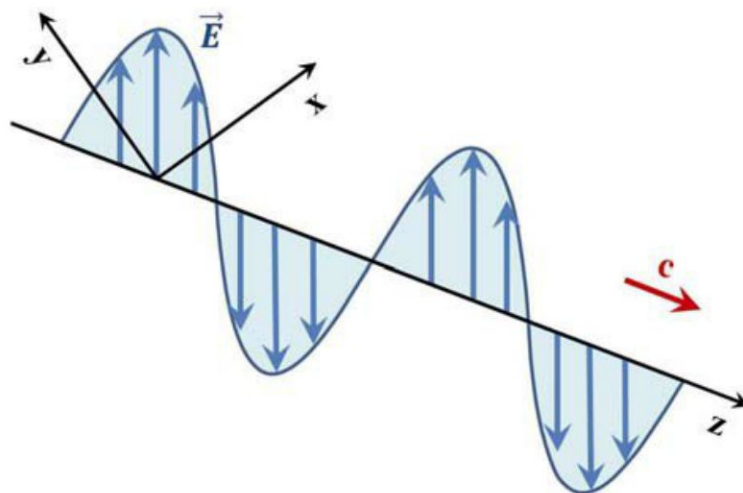


Figure 11.1- Onde plane lumineuse : oscillations des vecteurs champ électrique \vec{E} et champ magnétique \vec{B} .

On peut donc, en toute généralité, décomposer le vecteur champ électrique sur cette base et écrire, en utilisant comme précédemment la notation complexe pour les

champs oscillants :

$$\mathbf{E}(z,t) = A_H e^{i(kz + \phi_H - \omega t)} \mathbf{e}_H + A_V e^{i(kz + \phi_V - \omega t)} \mathbf{e}_V \quad (11.1)$$

$$= e^{i(kz - \omega t)} \begin{pmatrix} A_H e^{i\phi_H} \\ A_V e^{i\phi_V} \end{pmatrix}_{H,V} \quad (11.2)$$

où A_H et A_V sont les amplitudes de l'onde selon les directions horizontales et verticales. On appelle polarisation de la lumière la trajectoire décrite par l'extrémité du vecteur champ électrique en un plan transverse donné ($z = \text{constante}$ dans notre cas).

On s'intéressera en particulier dans ce cours au cas de la *polarisation linéaire*, correspondant au cas où $\phi_H = \phi_V = \phi$. Un vecteur unitaire à coefficients réels suffit dans ce cas pour décrire le champ :

$$\mathbf{E}(z,t) = A e^{i(kz + \phi - \omega t)} \mathbf{e}_p \quad (11.3)$$

avec :

$$\begin{aligned} \mathbf{e}_p &= \alpha \mathbf{e}_H + \beta \mathbf{e}_V \\ \alpha &= A_H/A, \quad \beta = A_V/A, \quad \alpha^2 + \beta^2 = 1 \\ A &= \sqrt{A_H^2 + A_V^2} \end{aligned} \quad (11.4)$$

La trajectoire décrite par l'extrémité du vecteur champ électrique lors de son oscillation à la fréquence optique ω est dans ce cas un segment de droite, de direction définie par le vecteur \mathbf{e}_p . Dans le cas particulier où $A_V = 0$ on parle de polarisation horizontale, et verticale pour le cas $A_H = 0$. La base dans laquelle on décrit le champ électrique n'est pas unique, et n'importe quelle rotation de la base $\{\mathbf{e}_H, \mathbf{e}_V\}$ est une base possible. On utilisera souvent la base tournée à 45° , définie par :

$$\begin{aligned} \mathbf{e}_{45} &= \frac{1}{\sqrt{2}} \mathbf{e}_H + \frac{1}{\sqrt{2}} \mathbf{e}_V \\ \mathbf{e}_{-45} &= -\frac{1}{\sqrt{2}} \mathbf{e}_H + \frac{1}{\sqrt{2}} \mathbf{e}_V \end{aligned} \quad (11.5)$$

qui vérifie bien $\mathbf{e}_{45} \cdot \mathbf{e}_{-45} = 0$ (vecteurs orthogonaux) et $\mathbf{e}_{45} \cdot \mathbf{e}_{45} = \mathbf{e}_{-45} \cdot \mathbf{e}_{-45} = 1$ (vecteurs unitaires). La nouvelle base est bien orthonormée.

De façon plus générale, à partir de tout vecteur $\mathbf{e}_\theta = \cos \theta \mathbf{e}_H + \sin \theta \mathbf{e}_V$ on peut définir une base en lui adjoignant le vecteur $\mathbf{e}'_\theta = -\sin \theta \mathbf{e}_H + \cos \theta \mathbf{e}_V$. L'ensemble $\{\mathbf{e}_\theta, \mathbf{e}'_\theta\}$ forme alors une base orthonormale directe du plan transverse. Ainsi, un vecteur champ électrique quelconque comme celui défini à l'équation (11.3) peut se décomposer dans cette nouvelle base :

$$\mathbf{E}(z,t) = (\mathbf{E}(z,t) \cdot \mathbf{e}_\theta) \mathbf{e}_\theta + (\mathbf{E}(z,t) \cdot \mathbf{e}'_\theta) \mathbf{e}'_\theta \quad (11.6)$$

$$= \begin{pmatrix} \mathbf{E}(z,t) \cdot \mathbf{e}_\theta \\ \mathbf{E}(z,t) \cdot \mathbf{e}'_\theta \end{pmatrix}_{\{\mathbf{e}_\theta, \mathbf{e}'_\theta\}} \quad (11.7)$$

11.1. Description classique de la polarisation

Mentionnons qu'il est aussi possible d'avoir des phases ϕ_H et ϕ_V différentes dans l'expression (11.1) du champ. L'extrémité du vecteur champ décrit dans ce cas *une ellipse*, et l'on parle de polarisation elliptique (circulaire si $A_H = A_V$). Les coefficients α et β de la décomposition du vecteur polarisation \mathbf{e}_p sont dans ce cas complexes, avec la restriction $|\alpha|^2 + |\beta|^2 = 1$. Par souci de simplicité, nous nous restreindrons la plupart du temps dans ce chapitre au cas où ces coefficients sont réels et la polarisation linéaire.

11.1.2 Passage à travers un polariseur : loi de Malus

Il existe dans la nature de nombreux matériaux dont les propriétés optiques dépendent de la polarisation, appelés biréfringents (indice dépendant de la polarisation) et dichroïques (absorption dépendant de la polarisation). Ils permettent de sélectionner ou de manipuler la polarisation de la lumière, et nous en rencontrerons de nombreux exemples tout au long de ce cours. Considérons ici les polariseurs linéaires, qui sont des milieux absorbant la lumière oscillant selon une direction de polarisation et transmettant la lumière oscillant selon la direction orthogonale. Un cas particulier est le film polariseur formé de longues molécules alignées. Si on place un tel film sur le trajet d'un faisceau lumineux, alors seule la composante de polarisation dans la direction de l'axe passant (ou axe propre) du polariseur est transmise (figure 11.2).

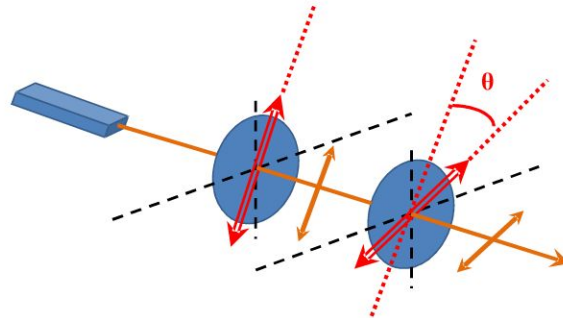


Figure 11.2 - Illustration de la loi de Malus : une onde lumineuse polarisée rectilignement (après passage d'une onde lumineuse quelconque dans un polariseur linéaire par exemple) traverse un polariseur (jouant le rôle d'analyseur) dont l'axe passant fait un angle θ avec la polarisation de l'onde. En sortie, l'onde lumineuse est encore polarisée rectilignement, mais sa direction de polarisation est celle de l'analyseur, et son amplitude est réduite car multipliée par un facteur $\cos\theta$.

Pour calculer le champ transmis, il faut alors l'exprimer dans la base de polarisation définie par le polariseur.

Considérons un polariseur « \mathbf{e}_θ », dont l'axe propre est selon la direction \mathbf{e}_θ et une onde lumineuse $\mathbf{E}_{\text{in}}(z, t)$ incidente sur ce polariseur. L'expression du champ dans la base $\{\mathbf{e}_\theta, \mathbf{e}'_\theta\}$ est donnée par l'équation (11.6). Le champ à la sortie du polariseur est

donc simplement :

$$\mathbf{E}_{\text{out}}(z, t) = (\mathbf{E}_{\text{in}}(z, t) \cdot \mathbf{e}_\theta) \mathbf{e}_\theta \quad (11.8)$$

Dans le cas où le champ incident est polarisé horizontalement :

$$\begin{aligned} \mathbf{E}_{\text{in}}(z, t) &= A_H e^{i(kz + \phi - \omega t)} \mathbf{e}_H \\ \mathbf{E}_{\text{out}}(z, t) &= A_H e^{i(kz + \phi - \omega t)} \cos \theta \mathbf{e}_\theta \end{aligned} \quad (11.9)$$

La puissance lumineuse est proportionnelle au carré du champ, donc en prenant le carré du module, on en déduit directement l'expression de la puissance de sortie en fonction de l'angle de rotation du polariseur et de la puissance d'entrée (figure 11.3) :

$$P_{\text{out}} = \cos^2 \theta P_{\text{in}} \quad (11.10)$$

C'est ce qu'on appelle la *loi de Malus*, qui décrit la puissance transmise par un polariseur linéaire en fonction de l'angle entre l'axe du polariseur et la polarisation de la lumière incidente.

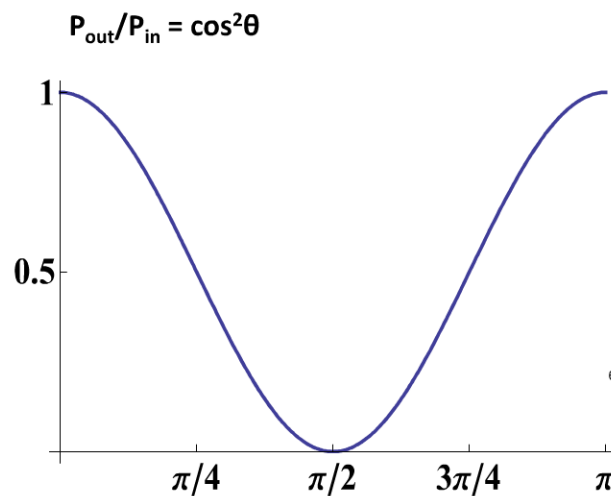


Figure 11.3 – Loi de Malus : le rapport entre puissances transmise et incidente est égale à $\cos^2 \theta$, où θ est l'angle entre la polarisation de l'onde incidente et l'axe passant du polariseur.

Mentionnons aussi dans cette section le cube séparateur de polarisation (figure 11.4). Cet élément transmet une direction de polarisation et réfléchit à 90° la polarisation orthogonale.

L'avantage par rapport au polariseur est que la lumière n'est pas absorbée, et que les deux directions de polarisation correspondant aux axes propres du cube sont disponibles. Ainsi, si on note 1 et 2 les deux sorties de ce cube, alignées de telle sorte qu'elles correspondent respectivement aux polarisation horizontale et verticale, et qu'on injecte de la lumière polarisée telle que décrite dans l'équation (11.1), les deux

11.1. Description classique de la polarisation

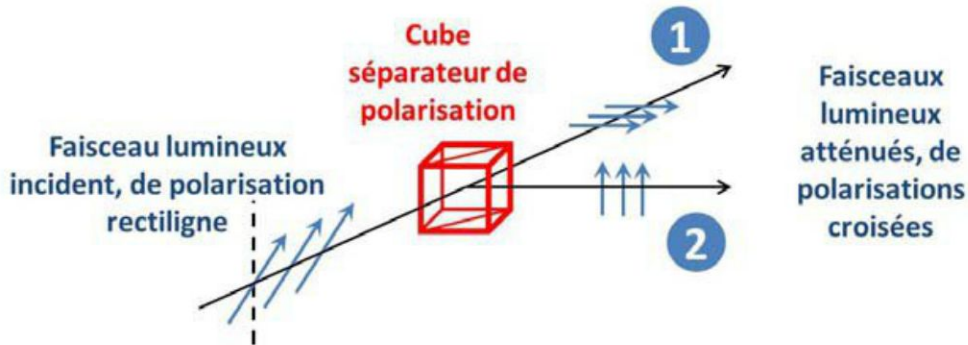


Figure 11.4 – Schéma de principe d'un cube séparateur de polarisation.

champs à la sortie s'écrivent :

$$\begin{aligned} \mathbf{E}_1(z, t) &= A_H e^{i(kz + \phi_H - \omega t)} \mathbf{e}_H \\ \mathbf{E}_2(z, t) &= A_V e^{i(kz + \phi_V - \omega t)} \mathbf{e}_V \end{aligned}$$

Dans le cas particulier d'un faisceau incident polarisé linéairement selon la direction \mathbf{e}_p (équation 11.3), les deux faisceaux à la sortie s'écrivent :

$$\begin{aligned} \mathbf{E}_1(z, t) &= A e^{i(kz + \phi - \omega t)} (\mathbf{e}_p \cdot \mathbf{e}_H) \mathbf{e}_H \\ \mathbf{E}_2(z, t) &= A e^{i(kz + \phi - \omega t)} (\mathbf{e}_p \cdot \mathbf{e}_V) \mathbf{e}_V \end{aligned} \quad (11.11)$$

11.1.3 Association de polariseurs

Nous venons de voir que pour décrire la propagation d'une onde électromagnétique il faut travailler dans un espace de dimension 2, alors que ce qu'on mesure, l'intensité, est un scalaire (c'est-à-dire un nombre). Cela mène parfois à des comportements extrêmement non intuitifs. Nous allons en voir ici un exemple.

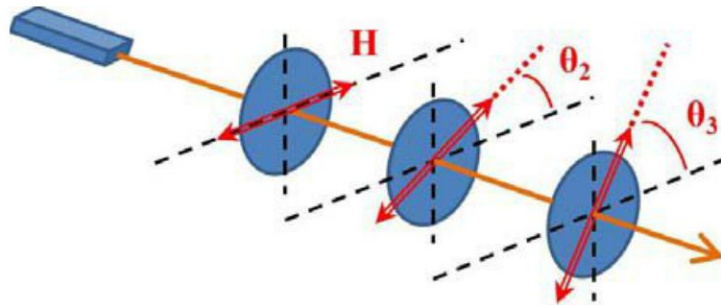


Figure 11.5 – Association de trois polariseurs dont les axes passants font, par rapport à l'axe horizontal, un angle de 0° , θ_2 et θ_3 .

Considérons le dispositif de la figure 11.5, où de la lumière est incidente sur un jeu de trois polariseurs successifs, « H », « θ_2 », « θ_3 ». Décrivons comme précédemment

le champ lumineux incident de manière générale : $\mathbf{E}_{in}(z, t) = A_H \exp(-i(\omega t - kz + \phi_H))\mathbf{e}_H + A_V \exp(-i(\omega t - kz + \phi_V))\mathbf{e}_V$. Le premier polariseur « H » étant aligné selon la direction horizontale, après ce polariseur la lumière peut être décrite par :

$$\mathbf{E}_1(z, t) = A_H e^{i(kz + \phi_H - \omega t)} \mathbf{e}_H.$$

Supposons tout d'abord que le deuxième polariseur n'est pas présent (ou qu'il est aligné avec le premier). En ce cas, la transmission par le 3^e polariseur se décrit comme dans le paragraphe précédent, équation (11.10). La puissance transmise vaut donc $P_3 = \cos^2 \theta_3 P_1$, où θ_3 est l'angle que fait le troisième polariseur avec l'horizontale. En particulier, lorsque les deux polariseurs, 1 et 3, sont croisés (i.e. que leurs axes sont orthogonaux), aucune intensité n'est transmise, les polariseurs absorbent toute la lumière.

Partons de cette configuration, où le premier polariseur est horizontal, et le troisième vertical ($\theta_3 = \pi/2$). Toute la lumière est absorbée par le dispositif. Ajoutons le polariseur 2. Le champ après le premier polariseur est inchangé, le champ après le polariseur « θ_2 » est donc :

$$\mathbf{E}_2(z, t) = A_H e^{i(kz + \phi_H - \omega t)} \cos \theta_2 \mathbf{e}_{\theta_2}.$$

Le champ incident sur le troisième polariseur n'est donc plus polarisé orthogonalement à l'axe de ce polariseur. On peut calculer le champ transmis, il vient :

$$\begin{aligned} \mathbf{E}_3(z, t) &= A_H e^{i(kz + \phi_H - \omega t)} \cos \theta_2 \cos(\pi/2 - \theta_2) \mathbf{e}_V \\ P_3 &= \cos^2 \theta_2 \cos^2(\pi/2 - \theta_2) P_1. \end{aligned} \quad (11.12)$$

Ainsi, en particulier quand $\theta_2 = \pi/4$, $I_3 = I_1/4$. On augmente la puissance transmise par le dispositif en rajoutant un élément absorbant ! Ce comportement tout à fait étonnant est directement lié à l'aspect vectoriel de la polarisation.

11.2 DESCRIPTION QUANTIQUE

Nous allons, dans cette section, donner une description qui permette de prendre en compte à la fois les aspects ondulatoires et corpusculaires de la lumière dans un même formalisme. La lumière se prête très bien à une description quantique directe, car le caractère ondulatoire est déjà décrit par les équations de Maxwell, il ne reste donc qu'à les rendre compatibles avec l'aspect probabiliste évoqué précédemment. Le formalisme développé ici fournira une solide base de départ pour aborder la description de tous les autres systèmes physiques, et en particulier l'électron et l'atome. La description quantique de la lumière est d'autant plus simple que les photons n'ont pas de

masse et n'interagissent pas entre eux, ils sont donc parfaitement isolés de leur environnement et sont naturellement des systèmes quantiques très purs. Ainsi, c'est avec de la lumière qu'ont été réalisées pour la première fois la plupart des expériences de pensée imaginées par les fondateurs de la mécanique quantique.

11.2.1 Effet d'un polariseur sur un photon

Nous avons déjà effectué une première description quantique de la lumière au chapitre précédent en introduisant des probabilités au niveau de la détection et du passage dans une lame semi-réfléchissante et en reliant la probabilité de détection au carré du champ électrique. Il nous faut affiner cette approche pour y inclure les propriétés de polarisation. Nous considérerons ici uniquement des ondes planes, où la probabilité de détection ne dépend pas du point considéré pour faire apparaître uniquement les propriétés liées à la polarisation. Nous avons vu que celles-ci se manifestent dans un espace vectoriel de dimension deux. Nous verrons, dans le chapitre suivant, que toute la richesse de la théorie quantique apparaît dès que l'espace des états possibles est de dimension supérieure à un, et on pourra alors très simplement généraliser ce que nous faisons ici.

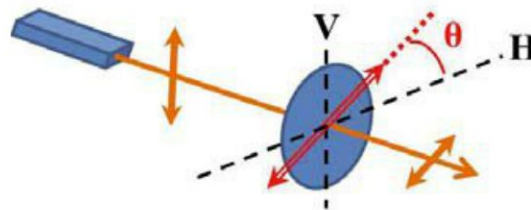


Figure 11.6 – Passage d'un photon unique, issue d'une onde lumineuse polarisée verticalement, à travers un polariseur dont l'axe passant fait un angle θ avec l'horizontal. En sortie, l'onde lumineuse est encore polarisée rectilignement, mais sa direction de polarisation est celle de l'analyseur, et son amplitude est réduite car multipliée par un facteur $\cos\theta$. Cette réduction d'amplitude se traduit par une diminution de la *probabilité de passage* du photon en $\cos^2\theta$.

Plaçons-nous dans le cadre de la figure 11.6, où un photon unique de lumière polarisée verticalement est incident sur un polariseur « θ ». Comme pour le cas de la lame séparatrice, il n'est pas possible de couper le photon en entités d'énergie plus faible : *il va donc se produire un phénomène probabiliste au passage du polariseur*, avec des probabilités non nulles de transmission ou d'absorption du photon, la probabilité devant être telle que la moyenne sur un grand nombre de détections donne la valeur classique, c'est-à-dire la loi de Malus. Ainsi, la probabilité de détecter un photon après le polariseur doit avoir pour valeur :

$$\text{Proba}_{\theta}(1) = \cos^2 \theta. \quad (11.13)$$

De manière plus générale, si la lumière incidente est polarisée linéairement selon la direction \mathbf{e}_p , on a pour les photons composant cette lumière :

$$\text{Proba}_\theta(1) = (\mathbf{e}_p \cdot \mathbf{e}_\theta)^2. \quad (11.14)$$

On voit que les probabilités de passage du photon à travers le polariseur sont données directement par le produit scalaire entre le vecteur polarisation incident et le vecteur représentant la direction de l'axe du polariseur. Ces deux vecteurs étant unitaires, on obtient bien un nombre compris entre 0 et 1, 1 n'étant obtenu que lorsqu'ils ont la même direction (mais pas forcément le même sens).

On peut maintenant décrire les probabilités liées au passage à travers un cube séparateur de polarisation. Considérons simplement le cas où le photon incident est polarisé linéairement selon la direction \mathbf{e}_p , et reprenons en termes de probabilité les résultats des équations (11.11). Il vient directement que :

$$\begin{aligned} \text{Proba}_{D_1}(1) &= |\mathbf{e}_p \cdot \mathbf{e}_H|^2 \\ \text{Proba}_{D_2}(1) &= |\mathbf{e}_p \cdot \mathbf{e}_V|^2 \end{aligned} \quad (11.15)$$

Comme \mathbf{e}_p est normalisé, on constate que $\text{Proba}_{D_1}(1) + \text{Proba}_{D_2}(1) = 1$, qui illustre qu'un seul photon est incident sur le cube et que ce dernier n'absorbe pas la lumière, il ne fait que l'aiguiller vers deux directions différentes en fonction de sa polarisation.

11.2.2 État quantique de polarisation

Pour décrire de manière commode les propriétés du photon, il faut faire appel à une notion complètement nouvelle, mais extrêmement puissante, celle d'**état quantique**, qui va nous permettre de décrire de manière unifiée les probabilités des résultats des différentes mesures que nous venons de trouver.

La mécanique quantique postule que les objets mathématiques qui décrivent les photons *appartiennent à un espace vectoriel* : on les appelle *vecteurs d'état* du photon. Cette propriété, valable pour tous les objets quantiques et pas seulement pour les photons, est à la base de beaucoup de propriétés étranges, mais confirmées par l'expérience, des objets quantiques. Le paragraphe précédent nous a montré que, pour décrire les propriétés de polarisation d'un photon unique, l'outil important est le vecteur unitaire de polarisation. Nous allons donc nous inspirer de l'espace des vecteurs unitaires de polarisation pour décrire le *vecteur d'état* du photon. L'intérêt de commencer par la polarisation du photon est de tirer profit de l'analogie directe qui existe entre vecteur de polarisation et vecteur d'état. Nous verrons dans les chapitres suivants que la notion de vecteur d'état s'étend à des systèmes plus complexes et à des espaces vectoriels de dimensions plus importantes, où cette analogie directe n'existe plus, mais où les propriétés de base que nous allons introduire subsistent.

Nous postulons que dans le cas où l'on considère de la lumière se propageant selon la direction z , monochromatique et de pulsation ω , l'état quantique du photon unique associé est parfaitement défini par son vecteur unitaire de polarisation. Nous introduisons les notations :

Notation	État quantique	Vecteur unitaire équivalent
$ H\rangle$	Photon polarisé horizontalement	$\mathbf{e}_H = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$
$ V\rangle$	Photon polarisé verticalement	$\mathbf{e}_V = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$

Le vecteur d'état quantique associé au photon est noté par le symbole $|\rangle$ qui contient un nom ou label formé de lettres ou de chiffres caractérisant l'état du photon. Il porte le nom usuel de *ket*, attribué par l'inventeur du vecteur d'état, Paul Dirac. La structure de l'espace mathématique des états quantiques est celle d'un *espace vectoriel*, en l'occurrence l'espace des vecteurs polarisation. Ainsi pour un photon polarisé selon une direction \mathbf{e}_p quelconque, l'analogie entre vecteur d'état et vecteur polarisation nous permet d'écrire le vecteur d'état $|\psi_p\rangle$ qui le caractérise sous la forme :

$$|\psi_p\rangle = \alpha |H\rangle + \beta |V\rangle \quad \text{où} \quad \mathbf{e}_p = \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \quad \text{avec} \quad \alpha^2 + \beta^2 = 1. \quad (11.16)$$

Comme nous l'avons vu en début de section, les polarisations linéaires ne représentent qu'une fraction des états de polarisation accessibles. Il suffit de considérer que les coefficients α et β sont à valeurs complexes pour engendrer tout l'espace des polarisations possibles. L'état quantique le plus général pour un photon unique est donc :

$$|\psi\rangle = \alpha |H\rangle + \beta |V\rangle \quad \text{où} \quad (\alpha, \beta) \in \mathbb{C}^2 \quad \text{et} \quad |\alpha|^2 + |\beta|^2 = 1. \quad (11.17)$$

11.2.3 Principe de superposition : espace de Hilbert des états de polarisation du photon

La structure d'espace vectoriel que nous venons d'introduire implique un *principe de superposition*, selon lequel *toute combinaison linéaire d'états d'un photon est encore un état possible pour le photon*. Cette propriété d'apparence mathématique et technique a en fait de profondes conséquences physiques.

L'espace vectoriel des états quantiques de polarisation du photon est muni d'un produit scalaire *hermitien* (hermitien car il concerne des vecteurs à coefficients complexes), qui est l'analogue du produit scalaire usuel des vecteurs polarisation. Le produit scalaire hermitien de deux vecteurs d'états $|\psi\rangle$ et $|\psi'\rangle$ de composantes respectives

(α, β) et (α', β') est noté $\langle \psi | \psi' \rangle$ et a pour expression :

$$\langle \psi | \psi' \rangle = \alpha^* \alpha' + \beta^* \beta' \quad (11.18)$$

où $*$ dénote la conjugaison complexe. On a en particulier :

$$\langle H | V \rangle = 0, \quad \langle H | H \rangle = \langle V | V \rangle = \langle \psi_p | \psi_p \rangle = \langle \psi | \psi \rangle = 1 \quad (11.19)$$

La base $(|H\rangle, |V\rangle)$ est donc orthonormée.

On appelle *espace de Hilbert* un espace vectoriel à coefficients complexes muni d'un produit scalaire hermitien. Tous les systèmes quantiques isolés, aussi complexes soient-ils, peuvent être décrits par des vecteurs d'état appartenant à des espaces de Hilbert de dimensions très grandes, parfois infinies. Ils obéissent donc tous au principe de superposition que nous avons énoncé plus haut.

11.2.4 Traversée d'un polariseur par un photon

Le formalisme que nous venons d'introduire permet de décrire simplement les phénomènes de nature probabiliste qui se produisent lorsqu'un photon est incident sur un polariseur, à travers les postulats suivants :

- À tout polariseur « θ », d'axe dirigé selon l'angle θ , est associé un état de polarisation du photon spécifique, celui qui passe avec une probabilité unité à travers ce polariseur, c'est-à-dire celui dont la direction de polarisation est parallèle à l'axe du polariseur. Son vecteur d'état est $|\theta\rangle = \cos \theta |H\rangle + \sin \theta |V\rangle$.
- Lorsqu'un photon décrit par un vecteur d'état $|\psi\rangle$ est incident sur un polariseur « θ », la probabilité $Proba_\theta(|\psi\rangle)$ pour qu'il passe à travers le polariseur vaut :

$$Proba_\theta(|\psi\rangle) = |\langle \theta | \psi \rangle|^2 \quad (11.20)$$

- Dans le sous-ensemble des cas où le photon a été effectivement transmis, l'état conditionnel du photon à l'issue de la traversée du polariseur est $|\theta\rangle$, indépendamment de l'état du photon incident sur le polariseur.

Il est facile de voir que les deux règles précédentes permettent de retrouver les probabilités (11.14) et (11.15) des paragraphes précédents, et le fait qu'une fois le polariseur traversé, le photon se retrouve « intact » à sa sortie, c'est-à-dire dans l'état d'un photon transmis à coup sûr.

11.2.5 Association de polariseurs

Il est maintenant possible de calculer ce qu'il advient d'un photon après son passage à travers un jeu de plusieurs polariseurs. Considérons un photon unique $|\psi_p\rangle$ polarisé

selon \mathbf{e}_p incident sur deux polariseurs successifs « θ » et « θ' ». La probabilité pour le photon de passer à travers le premier polariseur est donnée par l'équation précédente (11.20). Ce photon passé à travers ce polariseur est maintenant polarisé selon l'axe passant du polariseur, nous pouvons le noter $|\theta\rangle$. La probabilité que ce photon passe à travers le second polariseur est donc $|\langle\theta'|\theta\rangle|^2 = |\mathbf{e}_{\theta'}^* \cdot \mathbf{e}_\theta|^2$. Finalement, la probabilité de détecter le photon après les deux polariseurs est le produit des probabilités successives, soit :

$$\text{Proba}_{\mathbf{e}_{\theta}, \mathbf{e}_{\theta'}}(1) = |\mathbf{e}_{\theta'}^* \cdot \mathbf{e}_\theta|^2 |\mathbf{e}_\theta^* \cdot \mathbf{e}_p|^2 \quad (11.21)$$

Cette expression permet de retrouver les résultats du paragraphe 10.1.3, mais cette fois-ci concernant l'arrivée de photons *isolés* sur les polariseurs.



11.1 Nature de l'incertitude quantique

Nous venons de voir que les photons polarisés à 45° , noté $|45^\circ\rangle$, ont 50 % de chance d'être transmis à travers un polariseur « H », et 50 % de chance d'être absorbés. Considérons maintenant la lumière émise par le Soleil : le processus d'émission des photons à la surface du Soleil ne privilégiant aucune direction, elle est formée de photons polarisés rectilignement dans une direction θ aléatoire. Lorsqu'un photon solaire arrive sur le polariseur « H », il a une probabilité $\cos^2 \theta$ d'être transmis. L'angle θ étant aléatoire, il faut moyenner cette probabilité sur tous les angles θ possibles, ce qui donne une probabilité moyenne de transmission de 0,5. Les photons solaires ont donc, comme les photons polarisés à 45° , 50 % de chances d'être transmis à travers un polariseur « H », et 50 % de chance d'être absorbés.

Correspondent-ils pour autant au même état de la lumière ? Non, car les photons polarisés à 45° ont une probabilité unité d'être transmis par un polariseur « 45° », alors que les photons de la lumière solaire ont encore 50 % de chances d'être transmis à travers ce polariseur, 50 % d'être absorbés.

L'incertitude de polarisation des photons venant du Soleil vient du fait qu'on ne maîtrise pas leur processus d'émission à la surface du Soleil, elle est liée à *une préparation incomplète, probabiliste* du système : on dit qu'il y a *mélange statistique* des états du photon. En revanche, le caractère aléatoire de la transmission d'un photon $|45^\circ\rangle$ à travers un polariseur est fondamental car il n'est pas dû au fait qu'un ou plusieurs paramètres du système ne sont pas maîtrisés : on ne pourra jamais connaître avec certitude si un photon donné sera transmis ou absorbé. En effet, le physicien John Bell a montré que si on fait l'hypothèse inverse et qu'on suppose que l'incertitude quantique vient aussi du fait qu'on ne maîtrise pas les paramètres hypothétiques, pour l'instant inconnus (d'où leur nom habituel de « variables cachées »), qui gouvernent la transmission ou l'absorption d'un photon par le polariseur, on en déduit une inégalité (connue sous le nom d'*inégalité de Bell*) que doivent vérifier les résultats de mesure. De nombreuses expériences ont montré que cette inégalité est systématiquement violée, et qu'il faut donc abandonner l'hypothèse des « variables cachées » à la base du raisonnement de J. Bell. Notons qu'il existe plusieurs types de variables cachées, et que seules les théories à variables cachées « locales » (c'est-à-dire définies localement, au moment de l'interaction photon-polariseur par exemple) sont concernées par

l'inégalité de Bell. La violation de cette inégalité n'exclut donc pas les théories à variables cachées non-locales, comme la théorie de de Broglie-Bohm¹.

L'incertitude quantique est donc intrinsèque et irréductible : il faut se résoudre à admettre une certaine limitation dans la prédictibilité des phénomènes physiques individuels dans le monde quantique. Fort heureusement, l'incertitude sur la prédiction des phénomènes est de plus en plus faible lorsqu'on accumule un grand nombre de ces phénomènes, car la *loi des grands nombres* fait converger la fréquence des événements avec la probabilité donnée par les règles quantiques.

1. Voir à ce sujet F. Laloë, *Comprenons-nous vraiment la mécanique quantique ?*, EDP Sciences, 2011.

SYSTÈMES À PLUSIEURS ÉTATS : NOTION D'ÉTATS QUANTIQUES

12

12.1 ÉTATS QUANTIQUES

Nous avons introduit au chapitre précédent le concept de vecteur d'état, avec sa notation sous forme de *ket* introduite par Dirac. Ce vecteur d'état appartient à un espace vectoriel à coefficients complexes et permet de prédire tous les résultats de nature probabiliste concernant les mesures faites sur le système considéré. La puissance de la mécanique quantique réside dans la grande généralité de cette approche : ce concept peut être étendu à d'autres particules (électrons, protons) et à des systèmes composites, simples ou complexes, en faisant la liste des différents états possibles. Par exemple, pour une particule de masse non nulle comme l'électron, on introduira des états permettant de faire des prédictions sur sa position et sa quantité de mouvement, son énergie, etc. Cette approche est très efficace pour des systèmes isolés dont la liste des états possibles est bien contrainte. Dès que le système est en interaction avec le monde extérieur, cette liste peut devenir gigantesque et il faudra utiliser d'autres outils et procéder à des approximations, mais cela n'entre pas dans le cadre de ce cours introductif.

Nous allons, dans cette partie, considérer uniquement les grandeurs caractéristiques d'un système qui sont susceptibles de prendre seulement un ensemble dénombrable de valeurs, car elles sont plus simples mathématiquement, et nous aborderons les grandeurs qui prennent un nombre continu de valeurs, comme la position d'une particule, dans le chapitre suivant. Dans le présent chapitre, pour une propriété donnée du système (polarisation, énergie, etc.) nous aurons une liste dénombrable de configurations possibles, que nous appellerons des *états quantiques*, en étendant ce que nous avons vu en dimension deux, pour la polarisation, à une dimension quelconque pour toute grandeur physique dénombrable.

La notion de *dimension de l'espace des états* est essentielle en mécanique quantique. Nous verrons qu'elle est mathématiquement très bien définie, mais on peut en donner une approche plus expérimentale. Pour l'exemple de la polarisation, nous avons vu que cet espace est de dimension deux. Qu'est-ce que cela signifie ? Simplement, deux états de polarisations suffisent à décrire tous les états, car un ket $|\psi\rangle$ quelconque se décompose de façon unique sur une base à deux dimensions. En quoi

des états peuvent-ils former une base ? Pour l'exemple de la polarisation, on a vu que si un photon est V (resp. H), et qu'on utilise un polariseur « H » (resp. « V »), on est sûr de ne mesurer aucun photon. Ces deux états peuvent donc être parfaitement discriminés expérimentalement : ils sont dits orthogonaux. On dira que ce sont *les états propres* de l'appareil de mesure constitué par le polariseur « H » car ils donnent une valeur sûre à la mesure de polarisation : la dimension de l'espace des états est donc le nombre minimum d'états orthogonaux ou exclusifs à partir desquels il est possible de décrire tout état quantique.



12.1 Un autre système à deux états : le spin $\frac{1}{2}$. Expérience de Stern et Gerlach

L'une des expériences les plus emblématiques mettant en évidence un système à deux états est celle d'Otto Stern et Walther Gerlach, réalisée en 1922. L'idée était de mesurer le moment magnétique des atomes d'argent, lié à la rotation des charges dans l'atome, donc au moment cinétique de l'atome. Pour cela, ils ont produit dans une enceinte à vide un jet d'atomes d'argent. La source des atomes d'argent (un four chauffé et percé d'un petit trou) produit des atomes dont le moment magnétique est orienté dans des directions aléatoires, de la même manière que le Soleil produit des photons de polarisation aléatoire. Ces atomes passent ensuite dans un champ magnétique variant selon z (figure 12.1). En présence de ce champ magnétique non uniforme, ils subissent une force proportionnelle au gradient de champ magnétique et à la projection de leur moment magnétique selon la direction du gradient. Ils sont donc déviés par le dispositif, et cette déviation est mesurable en observant les points où s'accumulent les atomes sur l'écran. Le dispositif permet donc de connaître la projection du moment magnétique de l'atome selon z . Cette quantité peut prendre *a priori* des valeurs continues car l'orientation du moment magnétique par rapport à Oz n'est pas fixée par le dispositif.

Lors de cette expérience, Stern et Gerlach n'ont pas observé la distribution continue attendue, mais seulement deux points d'impact symétriques par rapport à la déviation nulle, donc deux valeurs opposées de la projection du moment magnétique. *Il n'y a donc que deux valeurs possibles $\pm\mu_B$ de la projection du moment magnétique de l'atome sur un axe donné*, et ce quelle que soit l'orientation choisie pour cet axe : c'est ce qu'on appelle la *quantification spatiale*. L'appareil de Stern et Gerlach est donc pour les atomes l'exact analogue du polariseur pour les photons.

Pour en rendre compte théoriquement on introduit une nouvelle observable quantique, le spin, permettant de décrire cette propriété. On lui associe deux vecteurs de base $|+\rangle_z$ et $|-\rangle_z$ décrivant respectivement un atome d'argent dévié à coup sûr vers la tache supérieure et inférieure, donc de projection du moment magnétique sur Oz égal respectivement à $+\mu_B$ et $-\mu_B$. L'espace de dimension deux associé au spin a exactement la même structure que celui de la polarisation et suit donc les règles exposées précédemment pour le photon.

On montre, en outre, que les mesures de la projection du moment magnétique selon les directions z , x ou y obéissent aux mêmes règles que les mesures de

polarisation du photon selon les directions $\{H,V\}$, $\{D,\bar{D}\}$, ou circulaire gauche et circulaire droite, et on peut donc, par exemple, décrire l'état de l'atome dont la projection du moment magnétique de spin selon x vaut $+\mu_b$ par le vecteur d'état $|+\rangle_x = \frac{1}{\sqrt{2}}(|+\rangle_z + |-\rangle_z)$.

Le spin est la première quantité que nous rencontrons qui est de nature purement quantique, car elle n'a pas d'équivalent classique. De manière générale, pour décrire toute particule élémentaire il est nécessaire de donner son spin. Les électrons ou les neutrons ont, comme l'atome d'argent, un spin dit « $1/2$ », qui correspond à un espace de dimension deux, mais d'autres particules, qu'elles soient élémentaires ou composites, peuvent avoir des spins plus élevés, correspondant à un nombre plus grand de valeurs quantifiées possibles pour la mesure de la projection du moment magnétique sur un axe. Le spin d'une particule est une propriété qui rentre dans le cadre de la description quantique du moment cinétique, étude qui va au-delà du cadre de ce livre.

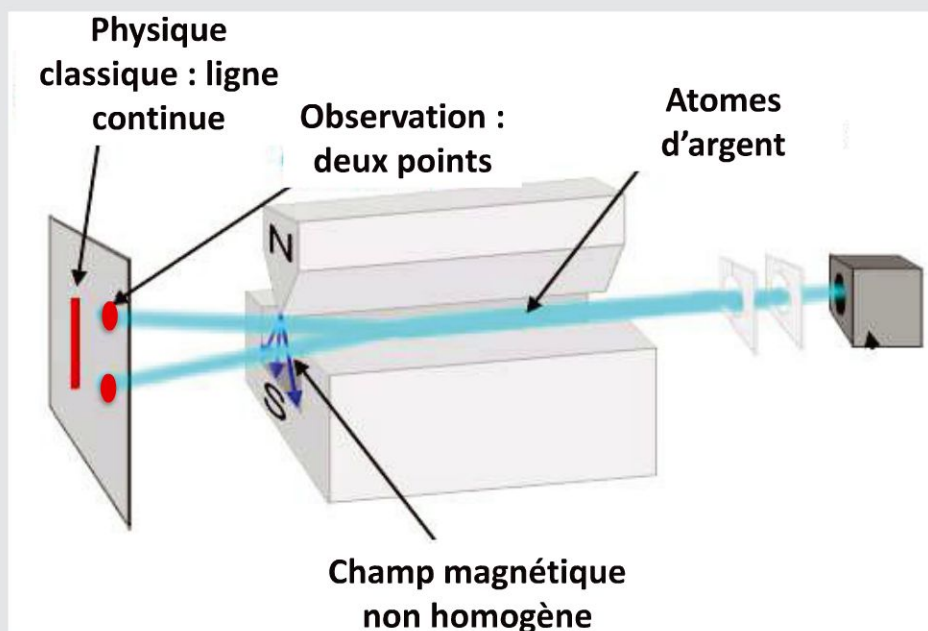


Figure 12.1 – Schéma de principe de l'expérience de Stern et Gerlach : des atomes d'argent sont envoyés, à l'intérieur d'une enceinte à vide, dans l'entrefer d'un aimant où règne un champ magnétique variable le long de l'axe vertical, où ils subissent une force proportionnelle à la projection de leur moment magnétique sur cet axe. En physique classique, cette projection prend un ensemble continu de valeurs. L'expérience montre qu'en sortie, les atomes se regroupent en deux faisceaux, chacun correspondant à une valeur, quantifiée, de cette projection.

(source libre : http://fr.wikipedia.org/wiki/Expérience_de_Stern_et_Gerlach)

12.2 MANIPULATION ET MESURE D'ÉTATS QUANTIQUES

12.2.1 Le cadre mathématique

Prenons pour l'instant le problème d'un point de vue abstrait, en suivant la procédure engagée dans le chapitre précédent. Supposons que nous soyons en présence d'un système isolé dont l'espace des états possibles est de dimension N . Rappelons que $N = 2$ pour la polarisation d'un photon ou le moment magnétique d'un atome d'argent (encart 12.1). Nous verrons en exemple un cas où $N = 3$: les neutrinos. Un autre exemple est constitué par une cavité optique piégeant la lumière, dans laquelle les états possibles pour la lumière sont les états à nombre de photons donné, de zéro à quelques unités.

Supposons que nous soyons capable d'identifier N états quantiques formant une base de cet espace. Cette base est notée $\{|H\rangle, |V\rangle\}$ dans le cas de la polarisation du photon, et nous la notons $\{|\psi_n\rangle\}_{n=1\dots N}$ dans le cas général. Dans le cadre de la polarisation, l'état le plus général est combinaison linéaire des états $|H\rangle$ et $|V\rangle$. Ici encore, si notre système est dans un état inconnu $|\psi\rangle$, il est possible de le décomposer sur les états de base :

$$|\psi\rangle = \sum_{n=1}^N c_n |\psi_n\rangle \quad \text{où} \quad \begin{cases} \{c_1, \dots, c_N\} \in \mathbb{C}^N \\ |c_1|^2 + \dots + |c_N|^2 = 1 \end{cases} \quad (12.1)$$

Les coefficients c_i sont les coefficients complexes de la décomposition unique de $|\psi\rangle$ dans la base des états. L'ensemble des états possibles forme un espace vectoriel à coefficients complexes, aussi appelé espace de Hilbert, muni du produit scalaire hermitien :

Si $|\psi_a\rangle$ et $|\psi_b\rangle$ sont deux vecteurs de l'espace des états :

$$\begin{aligned} |\psi_a\rangle &= c_1 |\psi_1\rangle + c_2 |\psi_2\rangle + \dots + c_N |\psi_N\rangle \\ |\psi_b\rangle &= d_1 |\psi_1\rangle + d_2 |\psi_2\rangle + \dots + d_N |\psi_N\rangle \end{aligned} \quad (12.2)$$

on définit le produit scalaire hermitien :

$$\langle \psi_b | \psi_a \rangle = c_1 d_1^* + c_2 d_2^* + \dots + c_N d_N^*$$

Pour un état physique, on impose aussi la règle de normalisation :

$$|\langle \psi | \psi \rangle|^2 = 1.$$

12.2.2 Le principe de superposition

La description précédente, inspirée de celle des états de polarisation, nous a amené à poser que tout vecteur d'état se décompose de façon unique dans la base des vecteurs

d'états. La réciproque est vraie : toute combinaison linéaire normée des vecteurs de base est un état quantique possible.

Si $|\psi_a\rangle$ et $|\psi_b\rangle$ sont deux états quantiques possibles, alors
 $|\psi\rangle = \alpha |\psi_a\rangle + \beta |\psi_b\rangle$ où $|\langle\psi|\psi\rangle|^2 = 1$
 est un état quantique possible.

Cette notion de superposition d'états est propre à la mécanique quantique, et elle reflète le caractère ondulatoire de cette théorie. En mécanique des particules classiques, cette notion n'existe pas. On la retrouve par contre pour les ondes classiques, qu'elles soient lumineuses, sonores ou autres, où elle est à l'origine des phénomènes d'interférences. La théorie que nous sommes en train de développer contient bien les ingrédients pour décrire les interférences observées expérimentalement, aussi bien pour la lumière que pour la matière.

12.2.3 Introduction à la mesure dans le monde quantique

Lorsqu'on a décrit les états possibles pour un système (sa polarisation, son énergie), cette liste d'états correspond à la liste des résultats possibles lors de la mesure de la quantité physique considérée. Pour la polarisation, nous avons vu que la liste des états possibles dépend du choix de l'appareil de mesure : horizontal (H) et vertical (V), ou diagonal (D) et anti-diagonal (\bar{D}). D'ailleurs, dans notre description, la notion même de l'existence d'une base d'états quantiques est indissociable de l'existence d'un appareil de mesure capable de discriminer entre ces différents états. C'est un cube polariseur et deux détecteurs dans le cadre de la polarisation. C'est un compteur de photons pour l'énergie du champ électromagnétique. Nous avons vu que la probabilité de détection d'un photon ayant une direction de polarisation donnée était égale au carré du module du produit scalaire entre l'état du photon et le vecteur polarisation. En particulier, quand le photon est décrit dans la même base que l'appareil de mesure, ce sont directement les modules carrés des coefficients qui donnent les probabilités.

Supposons donc que nous avons un appareil de mesure capable de discriminer les états de la base $\{|\psi_n\rangle\}_{n=1\dots N}$, c'est-à-dire qui donne la réponse certaine numérotée par n (que nous appellerons réponse n) si et seulement si on soumet à son examen l'état quantique $|\psi_n\rangle$. Nous postulons que la probabilité que la mesure sur un état physique $|\psi\rangle$ donne la réponse n est :

$$\text{Proba}_{|\psi\rangle}(|\psi_n\rangle) = |\langle\psi_n|\psi\rangle|^2 = |c_n|^2 \quad (12.3)$$

d'après la définition (12.1) et la règle du produit scalaire (12.2). On retrouve bien que le résultat de la mesure est certain si on envoie sur le détecteur un état $|\psi_{n'}\rangle$: la probabilité vaut 1 si $n' = n$ et 0 si $n' \neq n$.

Cette règle, dite règle de Born, est l'extension de la règle de projection sur un polariseur du chapitre précédent. Elle peut se généraliser : supposons qu'il existe un appareil de mesure capable de déterminer si le système est dans l'état $|\chi\rangle$ et uniquement dans cet état-là. Alors, la probabilité de détection sur ce détecteur à partir de l'état $|\psi\rangle$, autrement dit la probabilité de détecter le système $|\psi\rangle$ dans l'état $|\chi\rangle$ est donnée par :

$$\text{Proba}_{|\psi\rangle}(|\chi\rangle) = |\langle\chi|\psi\rangle|^2. \quad (12.4)$$

Nous avons vu ici l'un des principes fondamentaux de la mécanique quantique, qui est que *le résultat d'une mesure unique est aléatoire*. Seules sont bien déterminées les probabilités de mesure, et donc les moyennes pour un grand nombre de mesures reproduisant à l'identique la même expérience. Attention, il est très important de ne pas confondre une série de mesures où on refait à chaque fois l'expérience complète sur des systèmes toujours préparés dans le même état quantique, comme nous venons de le considérer ici, avec plusieurs mesures successives effectuées toujours sur le même système comme nous allons le voir maintenant.

12.2.4 Et si la mesure ne détruit pas l'état ?

Nous avons pour l'instant décrit les probabilités d'apparition d'un « clic » sur le détecteur de polarisation qui signale que le détecteur a observé et détruit l'état quantique en question. Cependant, il est possible d'imaginer, et de réaliser expérimentalement, des détecteurs permettant de mesurer une grandeur sur un système tout en ne le détruisant pas. Que devient alors le système après la détection ? En physique classique, la réponse est simple, du moins en théorie, puisque, l'interaction entre le système et le détecteur étant bien connue, on peut connaître exactement l'effet de la mesure sur le système et la prendre en compte pour son évolution future. Ce n'est plus le cas en physique quantique, à cause du caractère probabiliste des interactions au niveau microscopique.

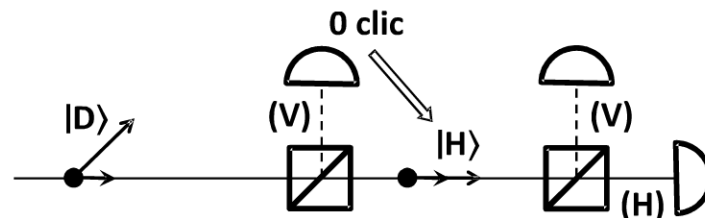


Figure 12.2 – Mesures successives d'un photon polarisé diagonalement par deux cubes séparateurs de polarisation. Si le premier détecteur ne détecte rien, cela signifie que le photon est transmis. Il peut alors être décrit par le ket $|H\rangle$, et on est sûr de le détecter sur la voie H du deuxième cube. Ainsi, la mesure non destructive (absence de clic) nous donne une nouvelle description de l'état du photon, conditionnelle à cette mesure.

Supposons que le détecteur a caractérisé l'état du système en donnant la réponse n . Il semble raisonnable d'imposer que le même détecteur redonnera le même résultat n si on répète la mesure sur le même système déjà mesuré. Utilisons une fois encore l'exemple de la polarisation du photon. Un détecteur partiellement non destructif est décrit en figure 12.2. Un cube séparateur de polarisation est placé sur le chemin d'un photon, et un seul compteur de photons est placé sur la voie réfléchie (polarisation verticale), la voie transmise est laissée libre. Ainsi, si le détecteur ne détecte rien, on sait que le photon a été transmis par le cube, et on a donc détecté un photon polarisé horizontalement. Que se passe-t-il maintenant si on place un deuxième cube polariseur après le premier ? En optique classique, toute la lumière transmise par le premier cube est transmise par le deuxième (même polarisation). Comme la physique quantique doit refléter, aux grands nombres de photons, les résultats classiques, le photon transmis par le premier cube est forcément transmis par le deuxième, avec une probabilité un. L'état du photon après le cube est donc $|H\rangle$.

Ce résultat est tout à fait remarquable. Partant d'un photon quelconque $|\psi\rangle$, l'absence de « clic » sur un détecteur *change* l'état du photon en $|H\rangle$. Cette modification drastique et irréversible de l'état quantique lors d'une mesure non destructive a beaucoup perturbé les scientifiques et les épistémologues. En fait, il faut bien comprendre qu'un détecteur non destructif est un système complexe avec beaucoup de degrés de liberté, sa description complète en est donc extrêmement difficile. Effectuer une mesure à l'aide d'un détecteur, même idéal, implique des interactions entre le système et le détecteur qui modifient fatalement l'état quantique du système.

Il faut aussi insister sur le fait que l'état après la mesure est *conditionné* par le résultat précis de cette mesure : ce n'est pas l'état du photon résultant de l'interaction avec l'appareil de mesure que l'on trouve de manière certaine après le premier séparateur de polarisation. C'est l'état dans lequel se trouve le photon dans le sous-ensemble de cas (dont on ne peut déterminer avec certitude quand ils vont se produire) où la mesure a donné le résultat 0.

12.3 APPLICATION À LA CRYPTOGRAPHIE

Nous allons illustrer les propriétés fondamentalement nouvelles de la mesure au niveau quantique par un exemple où le caractère fondamental de l'incertitude quantique est mis à profit pour « cacher » de l'information. Il s'agit de la *cryptographie quantique*.

12.3.1 Carte de crédit quantique

Commençons par un exemple simple conceptuellement mais éloigné de la réalité expérimentale, qui consiste à concevoir une carte de crédit dont on est sûr qu'elle ne peut être copiée, donc à sécurité absolue.

Elle est formée de n boîtes contenant chacune un photon polarisé $|H\rangle$, $|V\rangle$, $|D\rangle$ ou $|\bar{D}\rangle$, et donnée par la banque à son client. Par exemple, pour $n = 4$, la carte contient les photons ($|D\rangle$, $|H\rangle$, $|H\rangle$, $|\bar{D}\rangle$). Le lecteur de carte de la banque est formé d'une série de n polariseurs choisis parmi H, V, D, \bar{D} dans un ordre connu d'elle seule, par exemple \bar{D}, H, V, D , avec laquelle elle va mesurer de manière non-destructive le contenu de la carte lorsque le client voudra faire un retrait. Elle trouvera donc 0,1,0,0 et attribuera le retrait au compte repéré par ce numéro. Sa mesure aura pour effet de transmettre les photons dans les états mesurés, donc ne changera pas le contenu de la carte, car la banque a pris soin de mettre dans la carte des photons dont la mesure donne un résultat sans incertitude.

Supposons maintenant qu'un malfaiteur ait la possibilité de mesurer le contenu de la carte, afin de la reproduire et d'utiliser la copie pour effectuer des retraits. Il faut donc qu'il mesure la polarisation des photons, mais il ne connaît pas la série des polariseurs utilisés par la banque. Il utilisera par exemple les polariseurs H, V, D, D . Il trouvera bien l'état du deuxième et quatrième photon, mais obtiendra un résultat aléatoire pour le premier et troisième photon. Il n'aura donc l'état exact de la carte qu'avec une probabilité de 25 %. En outre, après la mesure, l'état de la carte aura changé car le premier photon sera maintenant $|H\rangle$ ou $|V\rangle$, et le troisième $|D\rangle$ ou $|\bar{D}\rangle$, et la lecture pirate de la carte sera décelée.

Ce schéma est très loin d'une possible réalisation pratique (la « boîte à photons » qui les conserve pendant un temps infini n'existe pas encore actuellement), mais il illustre bien les possibilités ouvertes par les mesures sur deux différentes bases, c'est-à-dire par le principe de superposition quantique, et par le fait que la mesure affecte de manière fondamentalement incontrôlable l'état d'un système quantique.

12.3.2 Théorème de non-clonage

Notre malfaiteur est peut être très fort en physique : que se passe-t-il s'il dispose d'une machine qui permet de dupliquer un état quantique quelconque sans avoir à le mesurer ? Il pourra dans ce cas fabriquer une nouvelle carte de crédit identique à l'originale, un clone en quelque sorte, et faire des retraits sans être décelé.

Fort heureusement, les lois de la mécanique quantique, et plus précisément le principe de superposition, interdisent l'existence d'une telle machine : c'est le *théorème de non-clonage*, dont la démonstration est extrêmement simple.

Supposons en effet que la machine produise deux photons dans l'état X , soit l'état $|X, X\rangle$, à partir d'un photon $|X\rangle$, et deux photons dans l'état Y , soit l'état $|Y, Y\rangle$, à

partir d'un photon $|Y\rangle$. Le principe de superposition implique que si on lui injecte l'état $|45\rangle = (|X\rangle + |Y\rangle)/\sqrt{2}$ la machine produira l'état $(|X,X\rangle + |Y,Y\rangle)/\sqrt{2}$. Il est facile de voir que cet état n'est pas le clone de $|45\rangle$. En effet, le clone de $|45\rangle$ est l'état :

$$\begin{aligned} |45,45\rangle &= ((|X\rangle + |Y\rangle)/\sqrt{2}, (|X\rangle + |Y\rangle)/\sqrt{2}) \\ &= (|X,X\rangle + |Y,Y\rangle + |X,Y\rangle + |Y,X\rangle)/2 \end{aligned}$$

La carte de crédit quantique est donc bien incopiable, non pas pour des raisons de difficulté technique comme c'est le cas actuellement, et qui peuvent être remises en question par d'éventuels progrès de la technique, mais pour des raisons liées aux fondements de la mécanique quantique.

12.3.3 Quelques notions de cryptographie

L'objet de la cryptographie est de transmettre de manière parfaitement confidentielle un message entre deux opérateurs. Ce message peut être représenté par un code binaire. À titre d'exemple supposons que nous souhaitons transmettre le message, certes un peu abscon, 011000. Comme il est impossible de s'assurer d'une parfaite confidentialité des lignes de communications, l'idée est qu'Alice code le message, et que Bob soit le seul à pouvoir le décoder. Deux stratégies principales sont à l'œuvre :

- *cryptographie à clé publique* : c'est la stratégie mise en œuvre dans les transmissions sécurisées sur Internet par exemple. Elle repose sur des principes mathématiques fondamentaux. Bob choisit deux nombres premiers n_1 et n_2 , et les multiplie $n = n_1 \times n_2$. Il envoie de manière publique ce produit à Alice. Alice utilise une procédure mathématique¹ pour coder le message avec le nombre n . Cette procédure est telle qu'il est impossible de retrouver le message d'origine sans connaître les facteurs n_1 et n_2 . La sécurité repose alors sur la difficulté de factoriser le nombre n , l'algorithme le plus efficace, connu pour factoriser les grands nombres en ses facteurs premiers, étant exponentiellement complexe. Donc la sécurité ne dépend que du temps de calcul nécessaire pour retrouver la clé de déchiffrement, qui croît exponentiellement avec sa longueur ;
- *cryptographie à clé privée* : dans ce cas, on suppose qu'Alice et Bob sont les seuls à partager une clé, qui est idéalement un nombre de la même taille que le message à transmettre. Supposons par exemple que ce nombre est 110011. Alice fait alors l'addition bit à bit :

$$\begin{array}{r} 0\ 1\ 1\ 0\ 0\ 0 \\ \& 1\ 1\ 0\ 0\ 1\ 1 \\ \hline = 1\ 0\ 1\ 0\ 1\ 1 \end{array}$$

1. R. L. Rivest, A. Shamir, and L. Adleman, *Com. of the ACM* **21**, 120 1978, R. L. Rivest, A. Shamir, and L. Adleman, « Cryptographic communications system and method » (1983), *US Patent* 4,405,829.

Elle envoie le résultat à Bob. Ce dernier, avec sa clé peut faire l'opération inverse et retrouver le message. La sécurité ici repose donc sur la possession de clés secrètes, et le problème est donc transféré au partage confidentiel de ces clés.

Il existe de nombreux protocoles classiques, dont le protocole RSA universellement utilisé de nos jours dans les transmissions sécurisées d'informations sur le web et qui est basé sur la difficulté de factoriser de très grands nombres. Sa sécurité n'est pas absolue, car rien n'interdit qu'une méthode plus rapide de factorisation soit trouvée un jour et rende ce code obsolète.

12.3.4 Cryptographie quantique

La cryptographie quantique n'a pas les défauts précédents : comme la carte de crédit quantique, elle procure un moyen parfaitement sûr - au moins dans son principe - de transférer ces clés secrètes.

Nous n'entrerons pas dans ce paragraphe dans le détail de la procédure, décrite dans de nombreux ouvrages et abordée dans l'encart 12.2. Elle est basée, comme pour la carte de crédit, sur l'utilisation des deux bases (H,V) et (45,-45) pour mesurer la polarisation des photons échangés. Tout espion qui a connaissance des messages échangés qui permettent de construire la clé est immédiatement détecté car les règles de la mécanique quantique impliquent, comme dans le paragraphe précédent, qu'il modifie nécessairement ces messages, et donc que sa présence peut être détectée. La sécurité du système n'est ainsi pas une conséquence, comme c'est le cas pour toutes les autres méthodes de cryptage, de la complexité de la procédure. Elle est assurée par les lois même de la mécanique quantique.

La cryptographie quantique est maintenant sortie des laboratoires de recherche : on peut trouver dans le commerce des appareils de cryptage quantique fonctionnant sur des distances de plusieurs kilomètres. On a ainsi sécurisé en 2007, par un dispositif de cryptographie quantique, la transmission des données électorales du canton de Genève.



12.2 Le protocole de cryptographie quantique BB84

En 1984, Bennett et Brassard² ont proposé un protocole d'échange de clés secrètes basé sur l'échange de photons uniques polarisés, que nous allons rapidement expliciter ici. Ce protocole se décompose en trois étapes : génération d'une clé, transmission, réconciliation.

1. Génération d'une clé secrète

Pour que le protocole fonctionne, Alice doit disposer d'une source parfaite de nombres aléatoires. Pour ce faire elle utilise un des principes de base de la

2. C. H. Bennett and G. Brassard, « Proceedings of the IEEE International Conference on Computers », *Systems and Signal Processing*, n° 175, Bangalore, India 1984.

mécanique quantique, qui dit que toute mesure est intrinsèquement aléatoire. En envoyant alors un photon unique sur une lame séparatrice, on a 50 % de chances pour le mesurer transmis et 50 % de chances de le mesurer réfléchi. On associe le zéro à la première situation et le un à la deuxième, et on obtient alors une suite parfaitement aléatoire de zéros et de uns comme décrit en figure 12.3.

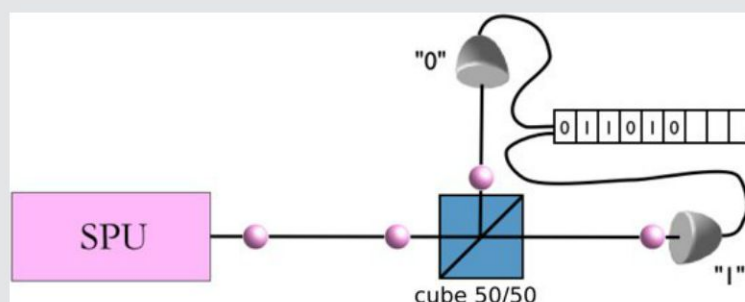


Figure 12.3 – Protocole de génération de nombres parfaitement aléatoires : une source de photons uniques (SPU) envoie des photons sur une lame séparatrice, telle que le photon a exactement 50 % de chances d'être transmis ou réfléchi. Une chaîne aléatoire de 0 et de 1 est ainsi générée à partir de ces photons : 0 si le photon est mesuré transmis, 1 s'il est mesuré réfléchi. La nature aléatoire du phénomène est intrinsèque puisque assurée par les lois de la mécanique quantique. La réalisation expérimentale n'est néanmoins pas parfaite, et un biais dans la lame séparatrice ou une efficacité moindre des deux détecteurs peut dégrader l'aspect aléatoire du phénomène. En pratique, donc, ce type de générateur nécessite une calibration extrêmement précise.

(source : S. Tanzilli, LPMC, Université de Nice)

2. Transmission de la clé

Alice code la clé sur des photons uniques polarisés. Par exemple, elle choisit d'attribuer la valeur 0 au photon polarisé horizontalement : $|H\rangle$, et la valeur 1 à celui polarisé verticalement : $|V\rangle$. Elle envoie donc une succession de photons à Bob, qui fait une mesure sur la base $\{H, V\}$ et peut donc reconstruire le message. Cependant, cette technique ne suffit pas à assurer la sécurité de la transmission. En effet, supposons qu'un espion, traditionnellement appelé Ève (de l'anglais *eavesdropping*), cherche à intercepter le message. Si Ève choisit de faire une mesure dans la base $\{H, V\}$, elle va détruire le photon mais pourra connaître exactement ce qui a été envoyé par Alice. Elle peut donc envoyer le même photon à Bob, passant ainsi inaperçue.

L'aléatoire en mécanique quantique arrive quand la base de mesure n'est pas celle où est construit l'état. Ainsi, si Ève faisait la mesure sur la base $\{D, \bar{D}\}$ elle ne pourrait connaître parfaitement l'état. Pour assurer la sécurité de la transmission, Alice doit donc choisir de manière parfaitement aléatoire la base sur laquelle elle encode le message, pour qu'Ève ne sache pas sur quelle base elle doit mesurer. En pratique, elle choisit entre la base $\{H, V\}$, où $|H\rangle$ est associé à 0 et $|V\rangle$ est associé à 1, et la base $\{D, \bar{D}\}$, où $|D\rangle$ est associé à 0 et $|\bar{D}\rangle$ est associé à 1. Ce choix se faisant là encore en utilisant un générateur de nombres aléatoires quantique ! Ainsi, Ève ne sait pas sur quelle base elle doit faire la

mesure. Si elle se trompe, le photon renvoyé sera différent du photon qu'elle a reçu, et sa présence pourra être détectée. Cependant, Bob non plus ne sait pas sur quelle base faire sa mesure. Comment alors s'assurer de la transmission du message ?

3. Réconciliation

Le protocole complet est décrit sommairement sur la figure 12.4. Alice choisit aléatoirement la base sur laquelle est codé le message. Bob, à la mesure, choisit également aléatoirement, car il n'a aucune connaissance des choix faits par Alice. Il a donc une chance sur deux d'avoir choisi la bonne base. Après l'échange complet effectué, Bob indique donc à Alice, de manière publique, quelles bases il a utilisé pour faire ses mesures, et Alice lui indique dans quels cas il a bien choisi, et donc Alice et Bob ne conservent que les nombres correspondants au choix communs de bases. Alice et Bob partagent alors une clé aléatoire de taille en moyenne moitié de celle de départ.

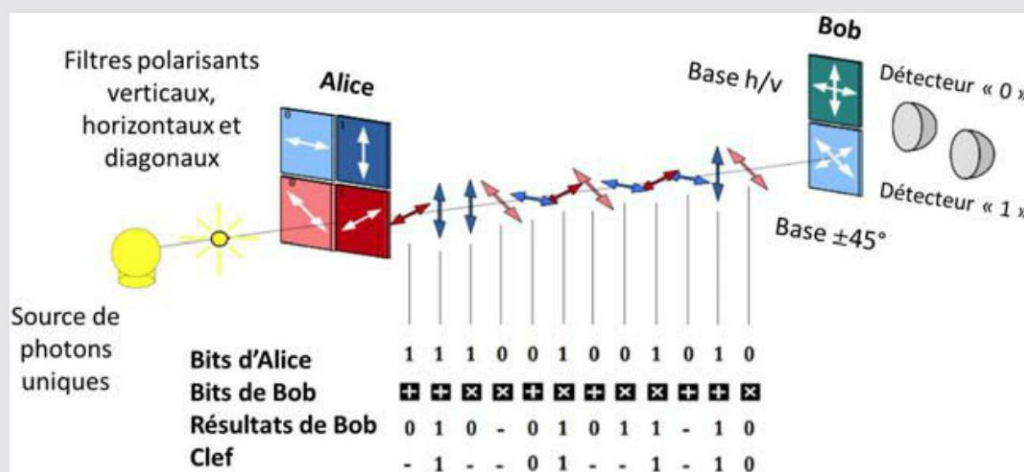


Figure 12.4 – Protocole de cryptographie BB84 : Alice code ses qbits aléatoirement sur l'une des deux bases complémentaires de polarisation. Bob choisit une base de lecture, qu'il communique après coup à Alice. Seuls les bits d'information correspondant aux bases communes sont conservés.

Quel est l'intérêt de cette procédure complexe ? Comme nous l'avons vu, si un espion, Ève, tente d'intercepter le message, elle ne sait pas quelle est la base propre de mesure. En essayant d'observer, elle va donc introduire des erreurs, en moyenne pour un bit sur deux. Alice et Bob vont donc tenter de déceler la présence d'un espion en examinant les erreurs de transmission. Pour cela, ils choisissent un sous-ensemble des bits échangés, qu'ils rendent publics (ils ne font plus alors partie de la clé !) et comparent. Si ces bits sont identiques, cela signifie que personne n'a intercepté le message, le reste de la clé est donc sûr. Sinon, en fonction des différences ils peuvent évaluer la probabilité de la présence d'un espion, et garder ou non la clé échangée.

Est-ce que cela fonctionne vraiment dans la pratique ? Oui. Mais c'est un peu plus compliqué. Tout d'abord, le protocole précédent se fait au prix d'une perte importante des bits échangés, et surtout permet de détecter un espion mais ne dit

pas comment s'en affranchir. Les protocoles modernes permettent, en évaluant le taux d'erreur, de connaître la quantité d'informations maximale recueillie par l'espion. Avec des protocoles d'augmentation de la confidentialité, il peuvent alors construire une clé secrète avec les bits leur restant, au prix là encore d'une perte importante, mais même en présence d'espion ils peuvent échanger une clé aléatoire.

De nombreuses expériences ont récemment mis en œuvre ce protocole en utilisant du codage d'information sur divers paramètres quantiques et non uniquement sur la polarisation du photon. Des appareils de cryptographie quantique sont même commercialisés³ et ont déjà servi pour des opérations de démonstration. Le marché reste néanmoins pour l'instant assez petit, mais qui sait, si un jour les physiciens développent un ordinateur permettant de factoriser les grands nombres en un temps court, alors pourront-ils vendre beaucoup plus efficacement leurs protocoles de cryptographie quantique ?

12.4 ÉVOLUTION DES ÉTATS QUANTIQUES

Jusqu'à présent, nous avons décrit les systèmes quantiques à un instant donné, en expliquant comment l'utilisation d'un espace de Hilbert, inspiré de ce qui se passe pour la polarisation de la lumière, permet de rendre compte des mesures faites à cet instant. Néanmoins, rien n'a encore été dit sur l'évolution de ces états, ce que nous allons faire ici. Appuyons-nous une fois encore sur l'analogie avec l'optique et une des règles de base de la dualité onde-particule : $E = \hbar\omega$, et considérons une onde lumineuse monochromatique de pulsation ω . Nous avons vu qu'elle peut s'écrire sous la forme complexe $A \exp(-i(\omega t - kz))$. Le facteur temporel est ici $e^{-i\omega t}$. On peut donc écrire, de manière empirique, que l'évolution de l'onde associée au photon est le facteur multiplicatif $e^{-iEt/\hbar}$. C'est cette expression que nous allons généraliser ci-dessous.

12.4.1 Représentation matricielle de l'énergie et équation de Schrödinger

a) Évolution temporelle

Soit un état quantique décrit à $t = 0$ par le ket $|\psi_E\rangle$, « état propre » de l'énergie de valeur propre E . Ceci signifie que si on effectue une mesure idéale de l'énergie sur ce système on est sûr à 100 % de trouver la valeur E .

3. Plusieurs entreprises proposent ou développent des solutions, On peut citer ID quantique (Suisse), <http://idquantique.com>, MagiQ Technol. (USA), <http://www.magiqtech.com/>, SeQureNet (France), <http://www.sequren.net/>, Auréa Technol. (France), <http://aureatechnology.fr/>, NuCrypt (USA), <http://www.nucrypt.net/>, QuTools (Allemagne), <http://www.qutools.com/>, QuintessenceLabs (Australie), <http://qlabsusa.com/>.

Nous posons comme principe qu'à tout instant t ultérieur, l'état du système s'écrit :

$$|\psi_E(t)\rangle = e^{-iEt/\hbar} |\psi_E\rangle. \quad (12.5)$$

Ce postulat est compatible avec l'évolution temporelle en $e^{-i\omega t} = e^{-iEt/\hbar}$ de l'onde lumineuse complexe introduite dans les chapitres précédents.

Supposons que nous connaissions tous les états $\{|\psi_n\rangle\}_{n=1\dots N}$ qui donnent une valeur bien définie E_n à une mesure d'énergie (nous verrons en section 13.2 comment on peut trouver ces états dans un cas particulier). À $t = 0$ le système se trouve dans un état quelconque $|\psi(t=0)\rangle$, c'est-à-dire qu'il n'est pas nécessairement un état dont l'énergie est bien définie (ce qui peut paraître surprenant si l'on garde en tête une image classique du système, une particule par exemple). Comme les états propres de l'énergie forment une base, il existe donc des coefficients c_n tels que :

$$|\psi(t=0)\rangle = \sum_n c_n |\psi_n\rangle. \quad (12.6)$$

L'espace des états quantiques étant vectoriel, tout comme d'ailleurs l'espace des ondes monochromatiques en optique, l'évolution temporelle peut se décomposer sur celle des états propres de l'énergie, et on a donc de façon générale :

$$|\psi(t)\rangle = \sum_n c_n e^{-iE_n t/\hbar} |\psi_n\rangle. \quad (12.7)$$

Ainsi, chaque composante $c_n |\psi_n\rangle$ d'énergie bien définie E_n évolue avec le coefficient complexe $e^{-iE_n t/\hbar}$. Insistons sur le fait que les coefficients $e^{-iE_n t/\hbar}$ sont à l'intérieur de la somme précédente et ne peuvent donc pas être mis en facteur de celle-ci (à moins que toutes les énergies E_n soient égales).

b) Équation d'évolution associée

Maintenant que l'on a pu décrire de manière générale l'évolution d'un état quantique, il est possible de dériver l'équation différentielle associée. Reprenons tout d'abord ce qui se passe pour un état propre de l'énergie, et dérivons par rapport au temps l'équation (12.5), il vient :

$$i\hbar \frac{\partial |\psi_E\rangle}{\partial t} = E |\psi_E\rangle. \quad (12.8)$$

L'évolution d'un état quantique, état propre de l'énergie, est directement gouvernée par l'énergie du système. Ce fait peut être mis en relation avec ce qui se passe en mécanique classique, où par exemple la vitesse d'une particule est directement reliée à son énergie totale.

Cette équation se généralise comme dans le paragraphe précédent. Ainsi, en dérivant l'équation (12.7) on trouve :

$$i\hbar \frac{\partial |\psi\rangle}{\partial t} = \sum_n c_n E_n |\psi_n\rangle. \quad (12.9)$$

C'est cette équation qui régit l'évolution de tout vecteur d'état lorsque l'on connaît les propriétés d'énergie du système.

En fait, l'écriture précédente reste encore spécifique, car elle n'est utile que si l'on connaît les états propres de l'énergie. La représentation matricielle de l'énergie va nous permettre de mettre à jour une version plus générale de l'équation d'évolution : l'équation de Schrödinger.

c) Représentation matricielle et équation de Schrödinger

La structure de l'espace de Hilbert est une structure naturellement vectorielle. Ainsi, si on considère que l'on connaît une base du système physique $\{|\psi_n\rangle\}_{n=1\dots N}$, on sait que tout état quantique peut s'écrire de la forme : $|\psi\rangle = \sum_n c_n |\psi_n\rangle$. Connaissant cette base, la donnée d'un vecteur d'état est donc complètement équivalente à la donnée des coefficients complexes c_n . On peut donc formellement l'écrire sous la forme d'un vecteur colonne de dimension N :

$$|\psi_a\rangle = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_N \end{pmatrix} \quad (12.10)$$

On revient ainsi à la notation d'origine, celle de la polarisation, mais généralisée à N dimensions. Cette notation permet de réaliser facilement toutes les opérations de la mécanique quantique. Par exemple, le produit scalaire hermitien, que nous avons introduit en section 12.2, s'écrit comme un produit de matrices lignes et colonnes si on introduit la notation *duale* : $\langle\psi| = (c_1^*, c_2^*, \dots, c_N^*)$. Pour deux vecteurs d'état $|\psi_a\rangle$ et $|\psi_b\rangle$ il vient :

$$\langle\psi_b | \psi_a\rangle = (d_1^*, d_2^*, \dots, d_N^*) \cdot \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_N \end{pmatrix} = \sum_n d_n^* c_n \quad (12.11)$$

On retrouve bien la formule du produit scalaire.

Cette représentation trouve tout son intérêt pour exprimer de façon générale l'équation d'évolution (12.9). Le terme de droite de l'équation peut s'écrire comme le produit d'une matrice et d'un vecteur. En effet, introduisons la matrice :

$$\hat{H} = \begin{pmatrix} E_1 & 0 & \dots & 0 \\ 0 & E_1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & E_N \end{pmatrix} \quad (12.12)$$

qui est une matrice diagonale, dont les coefficients diagonaux sont les énergies propres du système. Il vient, en multipliant cette matrice par le vecteur d'état :

$$\hat{H}|\psi\rangle = \begin{pmatrix} E_1 & 0 & \dots & 0 \\ 0 & E_1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & E_N \end{pmatrix} \cdot \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_N \end{pmatrix} = \begin{pmatrix} E_1 c_1 \\ E_2 c_2 \\ \vdots \\ E_N c_N \end{pmatrix} = \sum_n c_n E_n |\psi_n\rangle. \quad (12.13)$$

Cette équation, que l'on peut écrire sous la forme condensée suivante :

$$i\hbar \frac{\partial |\psi\rangle}{\partial t} = \hat{H} |\psi\rangle. \quad (12.14)$$

est l'équation centrale de la mécanique quantique : elle est appelée « équation de Schrödinger ».

Son interprétation complète dépasse le cadre de ce cours, mais nous pouvons néanmoins en dire quelques mots. La matrice \hat{H} s'appelle le Hamiltonien, et représente l'énergie du système. On peut montrer à partir des expressions (12.7) et (12.9) que l'équation précédente est indépendante de la *base choisie*. Cela signifie que si on exprime le vecteur d'état sur une base qui n'est pas celle des états propres de l'énergie, on peut encore utiliser cette équation à condition bien entendu d'exprimer le Hamiltonien dans cette même base.

12.4.2 Oscillation de neutrinos

L'une des illustrations les plus marquantes, ces dernières années, de la mécanique quantique en dimension finie est celle de l'étude des neutrinos émis par le Soleil et détectés sur Terre. Comme nous l'avons vu dans la partie relativité de cet ouvrage, une réaction de fusion élémentaire se produisant dans le Soleil met en jeu quatre protons qui, après une réaction en chaîne exothermique produisent un noyau d'hélium, deux positrons et deux neutrinos électroniques. Le neutrino, particule introduite pour satisfaire la conservation de l'impulsion lors de réactions nucléaires (section 7.4.4), est longtemps resté un mystère pour les scientifiques. En effet, la théorie ne donne pas

d'indication précise sur sa masse, mais en fournit simplement une limite supérieure : si elle est différente de zéro elle est forcément faible. Expérimentalement, le neutrino est incroyablement difficile à détecter, réagissant très peu avec la matière : sur Terre, il arrive environ 50 milliards de neutrinos par cm^2 et par seconde, sans que nous ne nous en apercevions ! Les scientifiques ont donc cherché dans les années 1960 à mesurer ces neutrinos, en utilisant leur interaction, faible mais non nulle, avec des protons. Pour cela, ils ont utilisé des réservoirs contenant de gigantesques quantités de liquide, souvent à base d'eau. On citera en particulier l'expérience *Homestake* aux États-Unis, réalisée dans un réservoir contenant 40 m^3 de perchloroéthylène, liquide riche en chlore, élément interagissant plus facilement que les autres avec les neutrinos. Ou encore l'observatoire Super-Kamiokande au Japon, construit dans une mine, et constitué d'un réservoir contenant environ 50 000 tonnes d'eau pure, entouré de plus de 11 000 photomultiplicateurs (figure 12.5). On trouvera un descriptif de l'ensemble des expériences ayant servi à cette aventure scientifique à l'adresse Internet http://en.wikipedia.org/wiki/Solar_neutrino_problem.

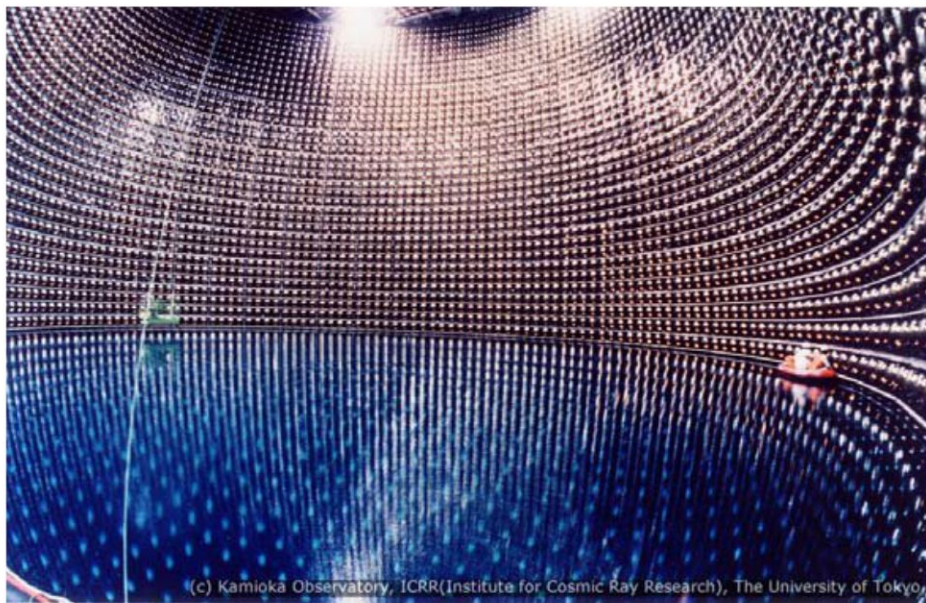


Figure 12.5 – Observatoire Super-Kamiokande au Japon. Le réservoir central contient 50 tonnes d'eau dans lesquels baignent plusieurs milliers de photomultiplicateurs (les demi-sphères brillantes sur la figure).

(Source : <http://www-sk.icrr.u-tokyo.ac.jp/sk/gallery/index-e.html>)

Pourquoi répéter de si nombreuses et coûteuses expériences ? La raison est que dès les premières mesures effectuées par l'expérience *Homestake*, les scientifiques ont remarqué que seul un tiers des neutrinos prévus par la théorie était détecté. Ce déficit étant largement supérieur à l'incertitude sur la valeur attendue, il est donc devenu une énigme que de nombreux physiciens ont cherché à résoudre. Ce n'est que dans les années 1980 que la réponse fut claire, mais dès 1968 un physicien italien, Bruno

Pontecorvo, proposa une explication basée sur la masse des neutrinos, explication étayée ensuite dans le modèle standard, et confirmée expérimentalement en 1998 (Raymond Davis Jr. et Masatoshi Koshiba reçurent le prix Nobel de physique 2002 pour la détection des neutrinos solaires).

Nous avons vu dans ce cours tous les éléments permettant de donner une explication simple à ce phénomène. Le concept est directement ancré dans le problème de l'évolution temporelle des états quantiques d'un système physique. L'idée des physiciens fut de dire que si on ne détectait qu'un tiers des neutrinos électroniques prévus par la théorie, c'était que, pendant leur trajet depuis le Soleil, ils avaient changé de nature, pour devenir des neutrinos muonique ou tauique. Plusieurs hypothèses permettent d'arriver à cette conclusion. Tout d'abord, on suppose que les trois types de neutrinos sont trois états quantiques d'un même système physique, premier système de dimension trois que nous rencontrons. Ensuite, il est nécessaire de supposer que ce ne sont pas des états propres du Hamiltonien \hat{H} , correspondant à la mesure de l'énergie du système. En effet, si le neutrino électronique était un état propre de \hat{H} alors il resterait, lors de son évolution temporelle, un neutrino électronique, comme nous venons de le voir dans la section précédente.

Ainsi, nous sommes en présence de deux bases de dimension trois pour le même système physique. Tout d'abord la base des états du neutrinos intervenant dans les réactions nucléaires, que nous noterons :

$$|\nu_e\rangle, \quad |\nu_\mu\rangle, \quad |\nu_\tau\rangle \quad \text{neutrinos intervenants dans les réactions nucléaires.} \quad (12.15)$$

D'un autre côté, il existe la base des états propres de l'énergie, que nous noterons :

$$|\nu_1\rangle, \quad |\nu_2\rangle, \quad |\nu_3\rangle, \quad \text{états d'énergies respectives } E_1, \quad E_2, \quad E_3. \quad (12.16)$$

Chaque état d'un neutrino peut donc se décomposer dans la base des états propres de l'énergie :

$$\begin{aligned} |\nu_e\rangle &= \alpha_1 |\nu_1\rangle + \alpha_2 |\nu_2\rangle + \alpha_3 |\nu_3\rangle \\ |\nu_\mu\rangle &= \beta_1 |\nu_1\rangle + \beta_2 |\nu_2\rangle + \beta_3 |\nu_3\rangle \\ |\nu_\tau\rangle &= \gamma_1 |\nu_1\rangle + \gamma_2 |\nu_2\rangle + \gamma_3 |\nu_3\rangle \end{aligned} \quad (12.17)$$

avec des coefficients α_i, β_i et γ_i *a priori* complexes.

Regardons maintenant l'évolution temporelle d'un neutrino émis par le Soleil. Au moment de l'émission il s'agit d'un neutrino électronique, son état quantique s'écrit donc $|\psi(t=0)\rangle = |\nu_e\rangle$. D'après le principe de l'évolution temporelle, on a donc pour tout instant ultérieur :

$$|\psi(t)\rangle = \alpha_1 e^{-iE_1 t/\hbar} |\nu_1\rangle + \alpha_2 e^{-iE_2 t/\hbar} |\nu_2\rangle + \alpha_3 e^{-iE_3 t/\hbar} |\nu_3\rangle \quad (12.18)$$

Au temps T , de l'ordre de 8 mn, auquel le neutrino arrive sur Terre, la probabilité de détecter un neutrino électronique s'écrit :

$$\begin{aligned}
 P_{|\nu_e\rangle}(T) &= |\langle \nu_e | \psi(T) \rangle|^2 \\
 &= |(\alpha_1^* \langle \nu_1 | + \alpha_2^* \langle \nu_2 | + \alpha_3^* \langle \nu_3 |) \\
 &\quad (\alpha_1 e^{-iE_1 T/\hbar} |\nu_1\rangle + \alpha_2 e^{-iE_2 T/\hbar} |\nu_2\rangle + \alpha_3 e^{-iE_3 T/\hbar} |\nu_3\rangle)|^2 \\
 &= |\alpha_1|^2 e^{-iE_1 T/\hbar} + |\alpha_2|^2 e^{-iE_2 T/\hbar} + |\alpha_3|^2 e^{-iE_3 T/\hbar} \quad (12.19)
 \end{aligned}$$

En développant cette dernière expression, et en regroupant les termes exponentiels sous forme de cosinus, on trouve finalement :

$$\begin{aligned}
 P_{|\nu_e\rangle}(T) &= |\alpha_1|^4 + |\alpha_2|^4 + |\alpha_3|^4 \\
 &\quad + 2|\alpha_1|^2 |\alpha_2|^2 \cos((E_1 - E_2)T/\hbar) \\
 &\quad + 2|\alpha_2|^2 |\alpha_3|^2 \cos((E_2 - E_3)T/\hbar) \\
 &\quad + 2|\alpha_3|^2 |\alpha_1|^2 \cos((E_3 - E_1)T/\hbar) \quad (12.20)
 \end{aligned}$$

On voit que cette probabilité évolue en fonction du temps et oscille au rythme des différences d'énergie, avec une amplitude dépendant des poids dans la décomposition. Expérimentalement, le fait qu'on n'observe qu'un tiers des neutrinos à l'arrivée, donc pour $T \approx 8$ mn, donne des renseignements à la fois sur les coefficients de la décomposition et sur les énergies propres. L'hypothèse de départ est ainsi confirmée : les neutrinos électronique, muonique et tauique ne sont pas des états propres de l'énergie. De plus, on peut en déduire des bornes inférieures sur les différences entre les énergies des états propres. Enfin, cette expérience a aussi permis de confirmer que les neutrinos ont une masse différente de zéro.

Cette analyse a été brillamment confirmée en 2001 à l'observatoire de Sudbury au Canada. Dans cette expérience, 1000 tonnes d'eau lourde (formée de molécules d'eau dont l'hydrogène est remplacé par du deutérium) et quelques milliers de tonnes d'eau légère furent utilisées pour mesurer à la fois le nombre total de neutrinos et le nombre de neutrinos électroniques.

ONDES DE MATIÈRE : UNE INTRODUCTION À LA FONCTION D'ONDE

13

À l'instar de la lumière et de son double comportement ondulatoire et particulaire, et aussi étrange que cela puisse paraître, la matière, dont on a de bonnes raisons de penser qu'elle est formée de particules ponctuelles, peut également présenter un aspect « onde ». Un électron, un neutron, un atome, ou même une grosse molécule peut ainsi se comporter comme une particule bien localisée, voire ponctuelle, mais aussi comme une onde répartie dans l'espace, et donner lieu à des phénomènes typiquement ondulatoires (interférences, diffraction, ondes évanescentes, etc.).

L'aspect ondulatoire de la matière permet d'expliquer un très grand nombre de phénomènes, primordiaux comme la stabilité des atomes et de la matière et la conduction électrique, ou plus exotiques, comme la condensation de Bose-Einstein et la supraconductivité. Les applications sont multiples (diodes à effet tunnel, transistor, lasers, microscopie électronique, nanolithographie, horloges atomiques...). L'efficacité prédictive de la théorie quantique de la matière, dont nous allons présenter les bases dans ce chapitre, s'accompagne de questionnements profonds qu'une telle dualité ne manque pas de soulever : comment en effet ces deux aspects onde et particule sont-ils reliés ? De quel type sont ces ondes de matière ? À quelles équations obéissent-elles ? Les réponses que la mécanique quantique apporte à ces questions font appel aux notions que nous avons introduites dans le chapitre précédent, mais aussi à des notions nouvelles, comme la fonction d'onde ou les inégalités de Heisenberg, que ce chapitre et le suivant vont s'attacher à présenter, d'abord de façon simple (le présent chapitre) puis de façon plus précise et détaillée (chapitre 14).

13.1 ONDE DE MATIÈRE

13.1.1 L'idée de Louis de Broglie

Louis de Broglie a l'intuition en 1924 que, de même qu'à l'onde lumineuse est associée une particule, le photon, de même à toute particule matérielle doit être associé un phénomène ondulatoire, dont il ne précise pas la nature exacte. Il utilise « dans l'autre sens » les relations qui lient les propriétés caractéristiques de l'onde lumineuse à celles du photon (équation 10.4).

Nous postulons donc qu'à toute particule de masse m , d'énergie E et de quantité de mouvement \mathbf{p} est associée une onde de pulsation ω_{dB} et de vecteur d'onde \mathbf{k}_{dB} donnés par :

$$\omega_{dB} = \frac{E}{\hbar} \quad ; \quad \mathbf{k}_{dB} = \frac{\mathbf{p}}{\hbar} \quad (13.1)$$

La constante commune qui permet de passer des propriétés ondulatoires (ω, \mathbf{k}) aux propriétés corpusculaires (E, \mathbf{p}) est la constante de Planck réduite $\hbar = h/2\pi \simeq 10^{-34} \text{ J.s}$.

On déduit de la deuxième relation (13.1) l'expression de la longueur d'onde de l'onde de matière en fonction de la vitesse v de la particule, appelée relation de de Broglie :

$$\lambda_{dB} = \frac{h}{mv} \quad (13.2)$$

La matière étant décrite comme un phénomène ondulatoire, il y a donc possibilité d'interférences pour les particules, et en particulier d'interférences destructives : ainsi, de même que « de la lumière plus de la lumière peut donner de l'obscurité » (section 10.4), « de la matière plus de la matière peut donner une absence de matière » !

13.1.2 Ordres de grandeur, descriptions classique et quantique

La dualité onde-corpuscule défie généralement notre intuition car nous avons du mal à la percevoir à notre échelle, d'autant plus pour la matière dont les attributs classiques habituels (solidité, présence de bords...) semblent inconciliables avec une description ondulatoire. Il est donc impératif de comprendre quelles sont les conditions d'observation de cet aspect « onde » pour la matière. La relation de de Broglie se révèle alors particulièrement utile pour déterminer si, dans tel ou tel contexte, l'aspect ondulatoire d'un objet physique est prédominant par rapport à son aspect corpusculaire.

De façon intuitive, on comprend aisément qu'un objet matériel se comportera d'autant plus comme un corpuscule que l'onde de de Broglie qui lui est associée sera bien localisée, c'est-à-dire qu'elle occupera une zone très réduite de l'espace (de façon à pouvoir attribuer à l'objet une position et une trajectoire bien définies). Cette situation est en fait très similaire à ce qui se passe pour la lumière lorsque l'on cherche les conditions pour lesquelles elle apparaît sous forme d'un rayon lumineux. C'est la fameuse *approximation de l'optique géométrique*, valable lorsque la longueur d'onde de la lumière est petite par rapport à la taille caractéristique des objets avec lesquels elle interagit. Dans le cas contraire apparaît le phénomène de *diffraction*, d'autant plus prononcé que le milieu parcouru par la lumière présente des variations (de taille,

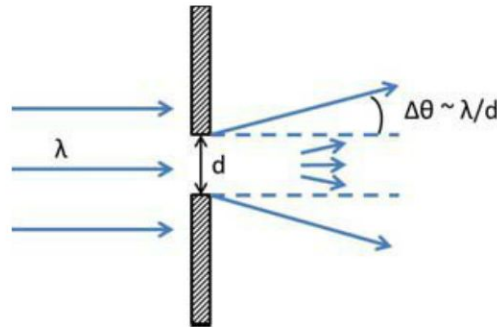


Figure 13.1 – Schéma montrant la diffraction d'un faisceau lumineux parallèle de longueur d'onde λ par une fente de largeur d . La dispersion angulaire, en sortie de la fente, est de l'ordre de $\Delta\theta \approx \lambda/d$. Lorsque $\lambda \ll d$, le faisceau de sortie (en pointillés) est formé de rayons parallèles : c'est l'approximation de l'optique géométrique.

d'indice optique...) importantes sur une distance de l'ordre de quelques longueurs d'onde.

Par exemple, lorsque l'on fait passer un faisceau lumineux parallèle à travers une fente de largeur d (figure 13.1), on observe que la lumière en sortie présente une dispersion angulaire $\Delta\theta$ de l'ordre de λ/d . On constate alors que ce phénomène de diffraction est d'autant plus marqué que λ/d est grand. Dans le cas contraire ($\lambda \ll d$), on retrouve bien les conditions de l'optique géométrique : la lumière en sortie de la fente se présente sous la forme d'un faisceau lumineux collimaté de largeur d .

Dans le cas des particules de matière, cette situation revient à comparer la longueur d'onde de de Broglie λ_{dB} à la taille caractéristique d des objets avec lesquels elle interagit (ou, de façon plus rigoureuse, avec la longueur typique de variation du milieu dans lequel elle évolue). Si $\lambda_{dB} \ll d$, on retrouve la condition de l'optique géométrique, c'est-à-dire qu'une description *classique* du mouvement de la particule sera suffisante, à des corrections de l'ordre de λ_{dB}/d près. Dans le cas contraire (λ_{dB} supérieure à, ou de l'ordre de, d), l'aspect onde de la particule prédominera et l'on ne pourra pas se contenter d'utiliser la mécanique classique pour décrire, même grossièrement, les phénomènes (ondulatoires) observés : il faudra nécessairement utiliser la mécanique quantique (voir, par exemple, sur les figures 13.2, 13.3 et 13.4, les phénomènes de diffraction et d'interférences de neutrons).

Écrite sous la forme $mvd \gg \hbar$, la condition $\lambda_{dB} \ll d$ montre que la constante de Planck joue un rôle similaire à la vitesse de la lumière en relativité restreinte : en effet, comme on l'a vu dans la première partie de ce livre, une description non-relativiste d'un phénomène (comparaison de mesures de durées ou de longueurs, effet Doppler, dynamique et collisions de particules...) est suffisante lorsque les vitesses en jeu sont négligeables par rapport à la vitesse de la lumière.

Par exemple, est-il nécessaire d'utiliser la mécanique quantique pour décrire le mouvement (très lent !) d'un escargot ? Non, bien sûr ! La longueur d'onde de de

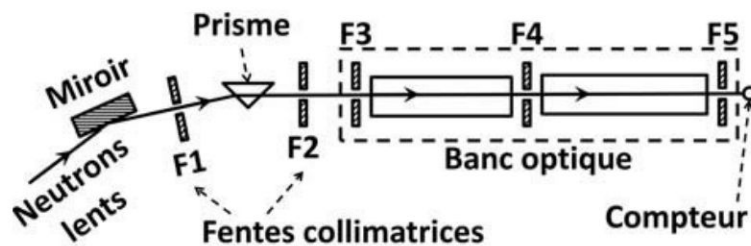


Figure 13.2 – Schéma de l'appareil utilisé par le groupe de A. Zeilinger en 1988 pour étudier l'aspect ondulatoire d'un faisceau de neutrons. La partie située à gauche de la fente F4 permet d'obtenir un faisceau parallèle et monocinétique de neutrons. Lorsque la zone diffractante F4 correspond à une fente unique, on observe, après la fente F5 en zone de sortie de l'appareil, la figure de diffraction 13.3. Lorsque l'on dispose deux fentes en F4, on observe la figure d'interférences 13.4.

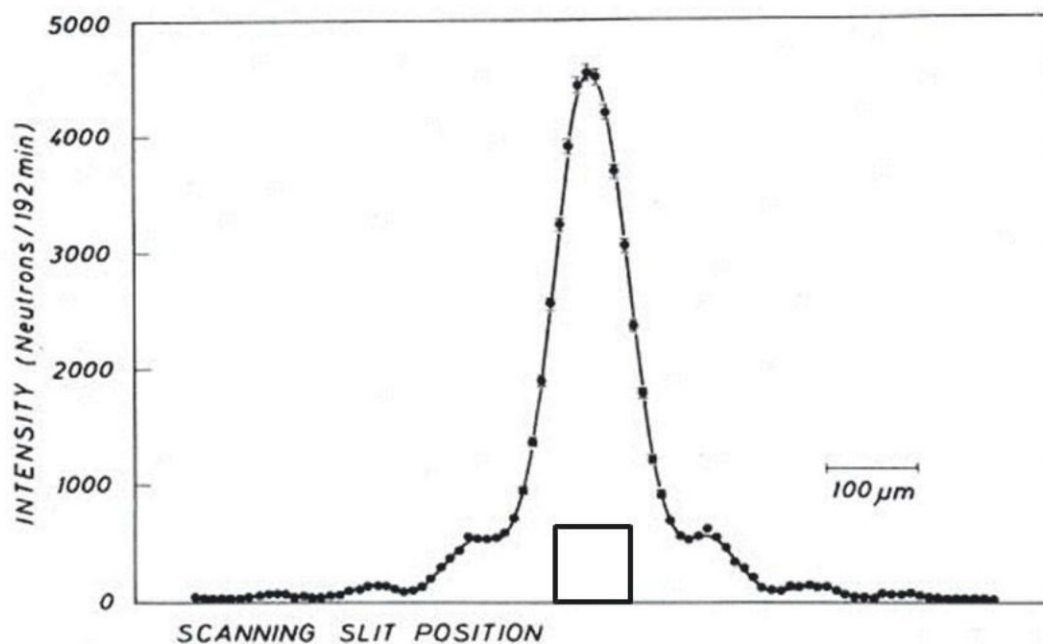


Figure 13.3 – Diffraction d'un faisceau de neutrons de $\lambda_{dB} = 19,3 \text{ \AA}$ par une fente de $90 \text{ }\mu\text{m}$ de largeur (figure 13.2 pour la configuration expérimentale). Comme attendu, la dimension de la tache centrale de diffraction est supérieure à la largeur de la fente. On observe également, de chaque côté de la tache centrale, des franges typiques de diffraction par une fente. Les points sont expérimentaux mais la courbe continue correspond à la prévision théorique. Le carré noir correspond à la largeur de la fente, c'est-à-dire à l'image géométrique du faisceau. (source : A. Zeilinger, R. Gähler, C.G. Shull, W. Treimer, and W. Mampe, *Reviews of Modern Physics*, vol. 60, 1988)

Broglie associée est en effet de $h/mv \approx 7.10^{-29} \text{ m}$ (pour une vitesse de 1 mm.s^{-1} et une masse de 10 g), évidemment très inférieure aux diverses longueurs caractéristiques du problème (taille de l'escargot, des herbes ou des obstacles qu'il rencontre

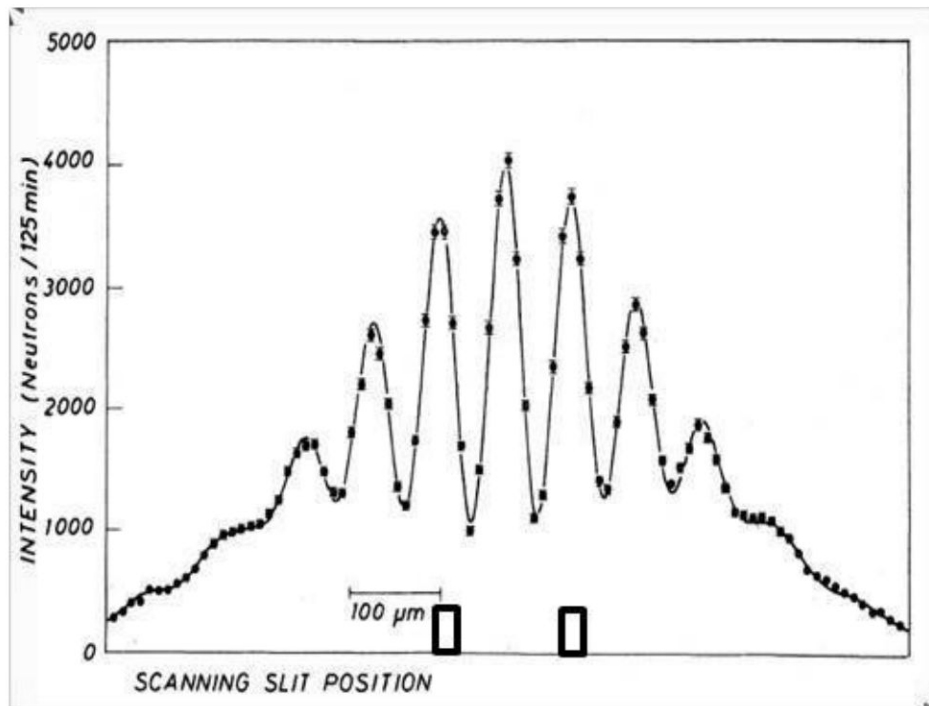


Figure 13.4 - Interférences neutroniques. La configuration expérimentale est la même qu'en 13.2, excepté qu'il y a maintenant deux fentes de largeur $22\ \mu\text{m}$ (espacées de $104\ \mu\text{m}$) en zone F4. Les carrés noirs correspondent aux deux fentes. (source : A. Zeilinger, R. Gähler, C.G. Shull, W. Treimer, and W. Mampe, *Reviews of Modern Physics*, vol. 60, 1988)

sur son chemin). Une description quantique du problème n'apporterait que des corrections relatives de l'ordre de 10^{-26} aux résultats classiques !

Au contraire, dans le cas de l'électron d'un atome d'hydrogène (masse $\approx 10^{-30}\ \text{kg}$, énergie typique $\approx 10\ \text{eV}$, taille caractéristique $\approx 0,1\ \text{nm}$), on obtient $\lambda_{dB} \approx 0,5\ \text{nm}$ qui est du même ordre de grandeur que la longueur caractéristique du problème. On en conclut qu'une description classique de l'électron d'un atome d'hydrogène est totalement insuffisante et qu'une description quantique est nécessaire dans ce cas.

La relation (13.2) indique que la longueur d'onde de de Broglie est d'autant plus grande que la masse de la particule est faible et que sa vitesse est faible. On sait par ailleurs que, pour réaliser des interférences (lumineuses ou acoustiques par exemple), il faut contrôler les trajets parcourus par les ondes qui interfèrent à une fraction de longueur d'onde près. Pour une vitesse de $1\ \text{m.s}^{-1}$ par exemple, la longueur d'onde de de Broglie vaut $6.10^{-25}\ \text{m}$ pour un petit objet macroscopique comme une poussière de $1\ \mu\text{g}$, et on ne pourra donc pas observer d'interférences avec ces objets, *a fortiori* lorsqu'ils sont encore plus massifs. Par contre, toujours pour une vitesse d' $1\ \text{m.s}^{-1}$, cette longueur d'onde vaut $0,4\ \mu\text{m}$ pour un neutron et $0,7\ \text{mm}$ pour un électron. Ainsi, si la source de particules les produit avec une vitesse suffisamment bien définie (donc

que l'onde associée est suffisamment monochromatique et de direction donnée), on pourra espérer observer de telles interférences.

La relation (13.2) montre aussi que si $v \rightarrow 0$, alors $\lambda_{dB} \rightarrow \infty$. Il semble donc que pour observer des interférences de matière, il suffit d'utiliser des particules au repos, même macroscopiques. C'est sans compter sur l'*agitation thermique* qui empêche tout objet d'être exactement au repos, sauf à une température T nulle. On montre en thermodynamique que la vitesse moyenne d'agitation thermique v_{therm} d'une particule de masse m vaut :

$$v_{therm} \simeq \sqrt{\frac{k_B T}{m}} \quad (13.3)$$

où $k_B = 1,38 \cdot 10^{-23} \text{ J.K}^{-1}$ est la constante de Boltzmann. On en déduit une *longueur d'onde de de Broglie thermique* :

$$\lambda_{dB,therm} = \frac{h}{\sqrt{mk_B T}} \quad (13.4)$$

qui vaut à température ambiante ($T = 300\text{K}$) $3 \cdot 10^{-19} \text{ m}$ pour la poussière ; $0,05 \text{ nm}$ pour un atome de sodium de masse atomique 23 ; $0,2\text{nm}$ pour le neutron et 11 nm pour l'électron. On aura donc une longueur d'onde de l'ordre du nanomètre à température ambiante avec des électrons, à 10 K avec des neutrons, à 1 K avec des atomes, mais toujours pas avec des particules macroscopiques, les plus basses températures obtenues à l'heure actuelle sur de tels objets étant de l'ordre du mK.

Une remarque pour finir cette partie. Selon l'argument précédent, on pourrait penser qu'il est nécessaire d'utiliser une description quantique des molécules de dihydrogène du vide interstellaire au vu de la très faible température du milieu (3 K)... Mais il n'en est rien car la densité du gaz d'hydrogène est si faible que la distance intermoléculaire est de l'ordre du centimètre, c'est-à-dire bien supérieure à la longueur d'onde de de Broglie $\lambda_{dB,therm}$ de l'hydrogène à cette température (qui est de l'ordre du nanomètre). Une description classique du vide interstellaire est donc suffisante, à de petites corrections quantiques près.

13.1.3 Expériences

De très nombreuses expériences d'interférence et de diffraction ont été réalisées avec des ondes de matière, confirmant l'hypothèse de de Broglie. Elles utilisent des dispositifs analogues à ceux qui ont été développés pour observer les interférences lumineuses, comme les trous d'Young, l'interféromètre de Michelson (chapitre 2) ou le biprisme de Fresnel (chapitre 10). De nombreuses expériences utilisent la diffraction par un réseau, c'est-à-dire par une structure périodique dans l'espace : dans cette configuration (figure 13.5) les différentes ondes diffractées par les éléments périodiques de la structure interfèrent constructivement uniquement dans des directions

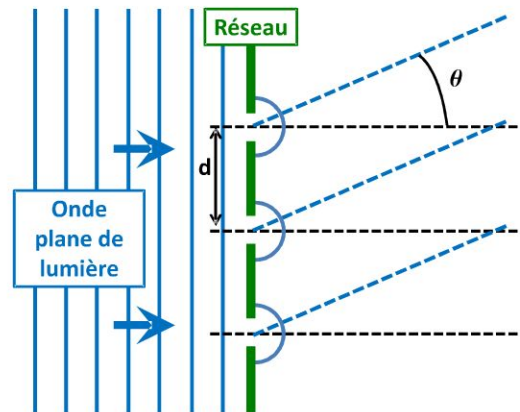


Figure 13.5 – Diffraction d'un faisceau lumineux par un réseau périodique de fentes. Le champ transmis, somme des champs diffractés par chacune des fentes du réseau périodique de période d , se renforce par interférence constructive uniquement dans des directions $\theta = \theta_n$, obéissant à la condition de Bragg $\sin \theta_n = n\lambda/d$ (n entier) où la différence de phase entre les faisceaux diffractés est un multiple de 2π . Si l'on place un écran à une grande distance du réseau, on observera une figure de diffraction similaire à celle de la figure 13.6.

bien définies. Plus précisément, une onde monochromatique incidente de longueur d'onde λ donne naissance à une onde diffusée dont l'intensité est maximale selon un ensemble discret de directions $\theta = \theta_n$ telles que $\sin \theta_n = n\lambda/d$, où n est un nombre entier et d la période spatiale de la structure diffractante (figure 13.6). Les irisations que l'on observe par exemple à la surface d'un disque, vinyle ou DVD, sont dues à la diffraction de la lumière sur les sillons gravés sur la surface de celui-ci dans des directions dépendant de la longueur d'onde.

La première expérience à avoir confirmé l'hypothèse de de Broglie a été réalisée par Davisson et Germer en 1924. Ils utilisèrent un faisceau d'électrons interagissant avec le réseau des atomes régulièrement disposés dans un monocristal de Nickel et observèrent que les électrons étaient diffusés par le cristal dans les directions prédites par la formule ci-dessus avec $\lambda = \lambda_{dB}$. Depuis, de nombreuses expériences ont confirmé la réalité des ondes de matière. De telles expériences de diffraction et d'interférence de matière sont maintenant régulièrement utilisées pour sonder les propriétés de solides ou de surfaces variées.

Les interférences de matière sont observées non seulement avec des particules élémentaires comme les électrons ou les neutrons, mais aussi avec des objets complexes et complexes, comme les atomes ou les molécules. Jusqu'à quel niveau de complexité un objet microscopique donne-t-il lieu à interférence ? En 2006, une équipe à Vienne a observé des interférences avec la molécule de $C_{60}F_{48}$, de masse 1 600 fois celle d'un nucléon, et de dimension de l'ordre de la dizaine de nm (figure 13.7). Il faut donc se rendre à l'évidence expérimentale : des molécules complexes, avec de très

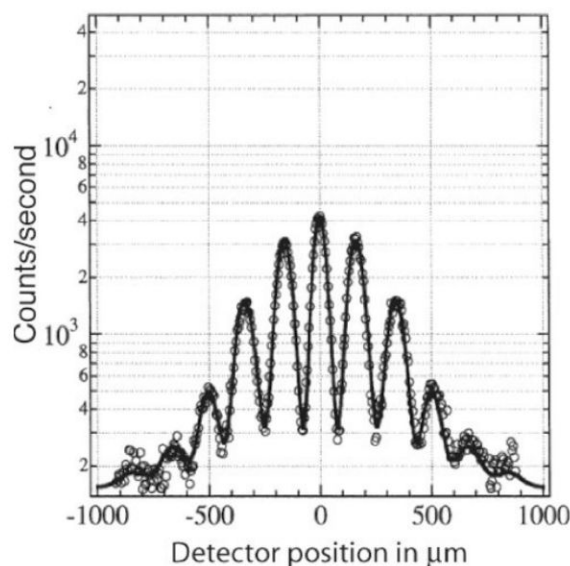


Figure 13.6 – Diffraction d'un faisceau de molécules de sodium après leur passage à travers un réseau de fentes de $100\text{ }\mu\text{m}$ de large. Le dispositif expérimental est sensiblement le même que pour la diffraction neutronique : une zone de préparation, une zone de collimation (pour créer un faisceau de molécules), une zone diffractante (le réseau), et une zone de détection. Les petits cercles représentent le nombre de molécules détectées à différentes distances du centre du faisceau. La ligne continue correspond au résultat théorique de diffraction de l'onde de matière associée. (équipe de D. Pritchard du MIT à Cambridge, USA).

(source : http://www.learner.org/courses/physics/visual/visual.htmlshortname=atom_waves)

nombreux degrés de liberté, ont l'étrange propriété de « passer » simultanément par des chemins différents pour interférer constructivement ou destructivement en sortie de l'interféromètre.

On peut alors imaginer et réaliser toute une gamme d'appareils (polariseurs, miroirs, cavités, fibres, hologrammes...) pour reproduire, avec les ondes de matière, ce que l'on sait faire avec les autres types d'onde, en particulier avec les ondes lumineuses. Au domaine classique de la physique constitué par l'« optique lumineuse » ou *optique photonique* s'est récemment ajouté un domaine en rapide essor : l'*optique atomique* (le mot atomique devant être compris au sens large, c'est-à-dire décrivant tout type de matière). La longueur d'onde de de Broglie dépendant de la masse m , les interféromètres atomiques sont très sensibles aux effets inertiels et gravitationnels dont ils forment des capteurs extrêmement performants.

13.1.4 Nature statistique de l'interférence de matière

Lorsque des particules matérielles pénètrent une à une dans un interféromètre, on observe en sortie des impacts qui se répartissent spatialement de manière aléatoire, exactement comme ce qui se passe pour les photons de la figure 10.9. Ce n'est que si

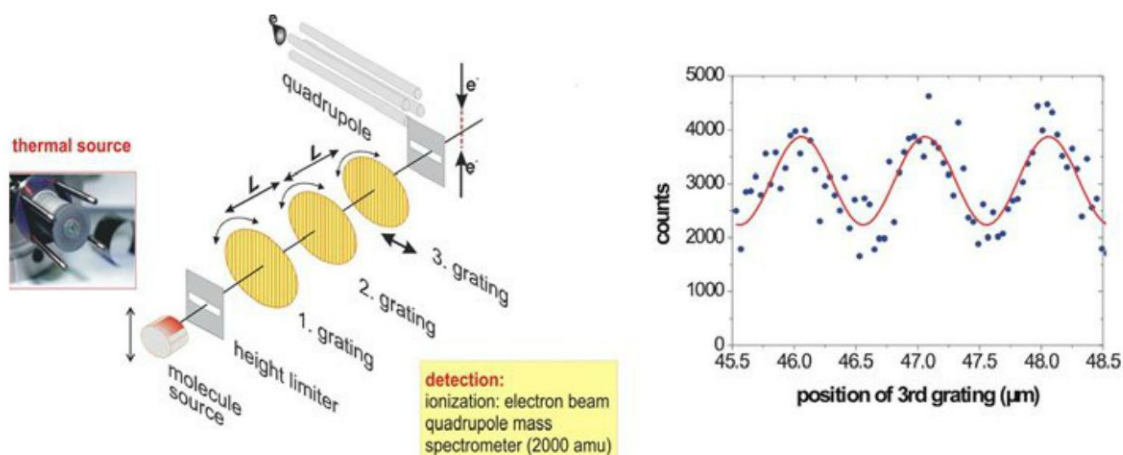


Figure 13.7 – Principe de l'interféromètre à ondes de matière utilisé pour réaliser des interférences de grosses molécules à Vienne en Autriche (molécules de fullerène $C_{60}F_{40}$). Après traversée des deux grilles préparatrices et de la grille diffractante, chacune des molécules est détectée en un endroit précis (par un signal de fluorescence par exemple). Progressivement, on observe que le lieu de ces impacts forme un interférogramme, comme celui de la figure de droite.

on accumule un très grand nombre de particules que l'on voit se dessiner de plus en plus nettement la figure d'interférence : comme pour la lumière, on peut considérer que l'onde de matière « guide » en quelque sorte la particule.

L'expérience de N. Tonomura en 1989 constitue l'une des plus belles illustrations de cet aspect statistique des interférences à ondes de matière (figure 13.8). Des électrons sont en effet envoyés, un par un, sur un biprisme électronique (un filament chargé entre deux électrodes planes) qui dévie différemment les trajectoires des électrons selon qu'ils passent à gauche ou à droite du fil. Ce biprisme sert en fait de « séparatrice électronique », c'est-à-dire qu'il joue le rôle des fentes dans l'expérience de Young ou celui du biprisme de Fresnel pour les interférences lumineuses (section 10.4). Sur le détecteur, en sortie, apparaissent alors des impacts dont la position semble aléatoire. Progressivement, cependant, on observe (figure 13.8) que ces impacts se répartissent de façon ordonnée, en bandes périodiquement espacées... ce sont les franges d'interférences électroniques !

Cette expérience est l'exact analogue de celle que nous avons décrite pour les photons au chapitre 10.

Plus récemment, en 2012, l'équipe de M. Arndt a pu visualiser cette construction progressive des franges d'interférences pour des molécules (voir le film, d'où sont tirées les images de la figure 13.9, de cette construction à l'adresse Internet : www.quantumnano.at/far-field-mode.3953.html). On peut ainsi aujourd'hui « voir » *en temps réel* se construire, atome par atome ou molécule par molécule, des franges d'interférences de matière...

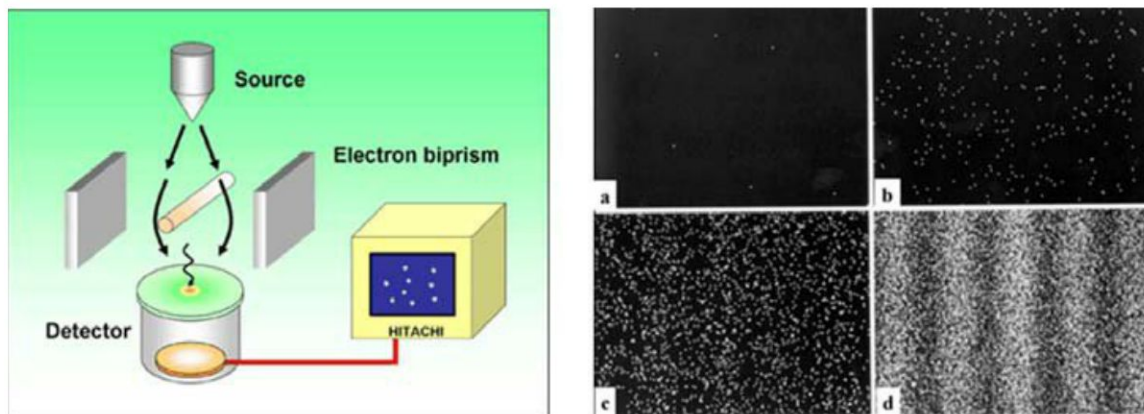


Figure 13.8 – *Figure de gauche* : principe de l'expérience d'interférences électroniques « électron par électron » de N. Tonomura en 1989. C'est un biprisme électronique qui joue le rôle du biprisme optique du chapitre 10. *Figure de droite* : construction progressive des interférences électroniques. Le nombre d'électrons détectés est de 8 (a), 100 (b), 3000 (c), 70000 (d).

(Source : <http://www.aps.org/units/fed/newsletters/fall2010/zollman.cfm>)

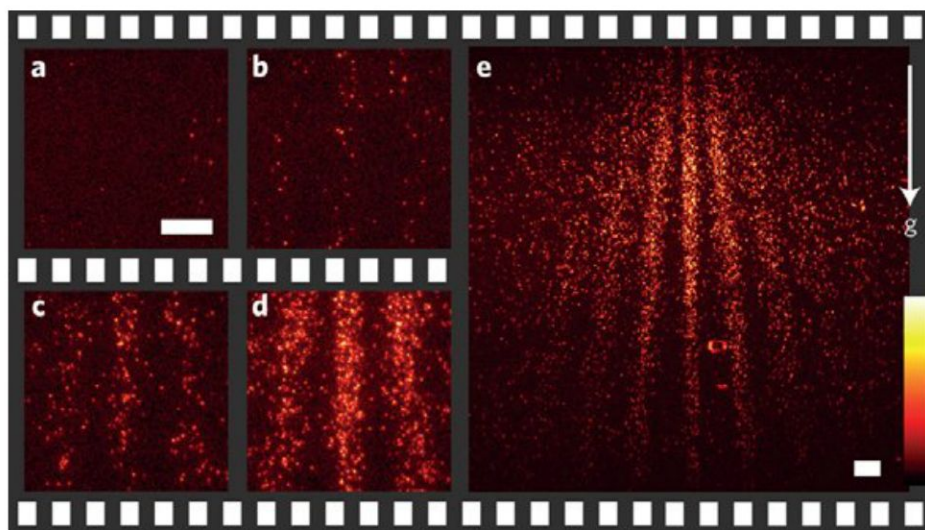


Figure 13.9 – La dualité onde-corpuscule visible en temps réel! Construction progressive (de « a » à « e ») d'interférences de molécules de phthalocyanine ($C_{32}H_{18}N_8$, masse égale à plus de 500 fois la masse d'un neutron) après passage à travers une grille de taille nanométrique (dispositif expérimental similaire à celui de la figure 13.7). Ces images sont tirées du film réalisé par l'équipe de M. Arndt en 2012 (film disponible à l'adresse Internet : www.quantumnano.at/far-field-mode.3953.html).

13.1.5 Introduction de la fonction d'onde

Comment décrire mathématiquement l'onde de matière introduite par de Broglie ? C'est Schrödinger qui réalisa cette avancée capitale, avec l'introduction en 1926 de la *fonction d'onde* $\psi(x,t)$, fonction complexe de l'espace et du temps, et de l'équation qui porte son nom. En fait, Schrödinger introduisit cette onde sans expliciter

son lien avec les grandeurs observables, tout comme Fresnel introduisit une fonction d'onde lumineuse sans connaître sa signification physique exacte, bien avant qu'on sache qu'il s'agissait d'un champ électromagnétique. Il s'avère en effet que l'onde $\psi(x,t)$ associée à tout objet physique n'est pas du même type que l'onde d'un champ classique, comme le champ électromagnétique ou le champ de pression de l'onde acoustique.

Il revient à Born d'avoir donné le premier, en 1926, l'interprétation de la fonction d'onde en termes de probabilité : son module au carré $|\psi(x,t)|^2$ donne la densité de probabilité de trouver à l'instant t l'objet physique à la position x . Plus précisément, la probabilité P de trouver la particule à l'instant t au point de coordonnée x telle que $x_1 \leq x \leq x_2$ a pour expression (on ne considère ici qu'une seule dimension d'espace pour simplifier, mais la généralisation à trois dimensions est immédiate) :

$$P = \int_{x_1}^{x_2} dx |\psi(x,t)|^2 \quad (13.5)$$

Il en résulte la condition de normalisation :

$$\int_{-\infty}^{\infty} dx |\psi(x,t)|^2 = 1 \quad (13.6)$$

qui n'exprime rien d'autre que le fait que la particule se trouve nécessairement quelque part dans l'espace, la somme des probabilités devant donc être égale à un.

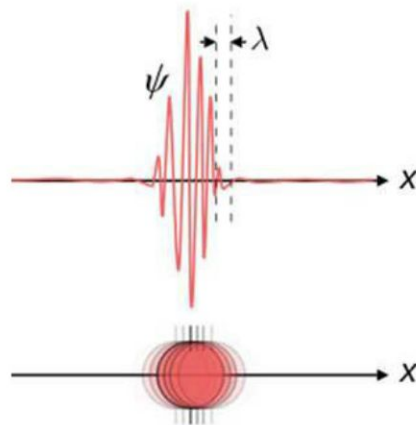


Figure 13.10 – Dessin illustrant la dualité onde corpuscule et l'interprétation probabiliste de la fonction d'onde $\psi(x)$. L'onde complexe ψ est bien localisée dans l'espace (en termes techniques, c'est un « paquet d'ondes »). La longueur d'onde de Broglie λ correspond à sa période spatiale moyenne. Sous l'onde ψ sont représentées quelques occurrences différentes (disques rouges) de la particule décrite par une telle fonction d'onde. L'opacité des disques correspond à la densité de probabilité de présence $|\psi(x)|^2$.

(Source : http://en.wikipedia.org/wiki/Wave_particle_duality (mai 2012))

La fonction complexe $\psi(x,t)$ est ainsi une « *onde de probabilité* », et le lien onde-particule est de nature probabiliste. Cette interprétation permet en particulier d'expliquer les expériences décrites dans la section précédente. Elle est l'exact analogue de ce que l'on a vu dans le chapitre 10 pour le photon : la probabilité de le trouver en un point donné est proportionnelle au module carré de l'onde lumineuse complexe. Il y a donc une certaine parenté entre l'onde lumineuse complexe $A(x,t)$, qui nous renseigne sur la probabilité de présence d'un photon unique et la fonction d'onde $\psi(x,t)$ qui nous renseigne sur la probabilité de présence de la particule considérée.

La forte analogie entre mécanique quantique du photon et mécanique quantique de la matière et de l'électron, déjà contenue dans les relations introduites par de Broglie, ne doit toutefois pas être poussée trop loin. Elle est plus ténue dès lors qu'on s'intéresse à un faisceau lumineux contenant beaucoup de photons... De plus, le photon n'est pas une particule comme les autres : sa masse est nulle et sa vitesse est toujours c . Il en résulte que les équations d'évolution de l'onde lumineuse et de l'onde de matière sont différentes. Enfin, comme son énergie de masse mc^2 est nulle, le photon est une particule qui peut être facilement créée ou détruite, alors que ce n'est pas le cas pour l'électron et toute autre particule matérielle.

Dans le chapitre 12, nous avons indiqué dans les postulats généraux qu'à tout système physique était associé un vecteur d'état appartenant à un espace de Hilbert. C'est évidemment aussi le cas pour une particule matérielle. Il nous faut donc faire le lien entre la fonction d'onde d'un système matériel $\psi(x,t)$ et son vecteur d'état $|\Psi(t)\rangle$. C'est l'objet de l'encart 13.1. Même si ces deux objets sont reliés, il faut prendre garde à ne pas confondre le vecteur d'état $|\Psi(t)\rangle$, vecteur dans l'espace abstrait des états de la particule (espace de Hilbert), avec la fonction d'onde $\psi(x,t)$, fonction complexe de la variable x , dont l'ensemble des valeurs constitue les « coordonnées généralisées » du vecteur d'état sur la base continue $\{|x\rangle\}$ définie dans l'encart.

13.2 QUANTIFICATION DE L'ÉNERGIE POUR UNE PARTICULE CONFINÉE

Une propriété capitale des systèmes ondulatoires est l'existence de *cavités résonantes* pour ces ondes. Nous avons vu, par exemple, les cavités optiques pour la lumière, formées de miroirs qui se font face, dans le chapitre 10, et l'exemple des cordes vibrantes et des tuyaux d'orgue pour l'acoustique est bien connu : ces dispositifs ont pour point commun de n'entretenir que des ondes dont *la fréquence ne prend qu'un nombre discret de valeurs*. En effet, il n'y a interférence constructive entre toutes les ondes lors de leurs réflexions multiples sur les parois de la cavité que pour des fréquences bien déterminées. Nous allons voir que c'est aussi le cas pour les ondes de matière, et grâce à la relation $E = h\nu$, la quantification des fréquences entraînera pour les systèmes quantiques la quantification de l'énergie.

13.2.1 Boîte quantique à une dimension

Pour simplifier l'analyse, nous supposons que la particule, de masse m , ne peut se déplacer que le long de l'axe Ox , et nous nous plaçons à un instant t donné. Entre les points d'abscisse 0 et L , elle n'est soumise à aucune force, mais rencontre en ces deux points des « murs » infranchissables (figure 13.11). Nous verrons dans le paragraphe 13.2.4 comment réaliser en pratique de tels dispositifs.

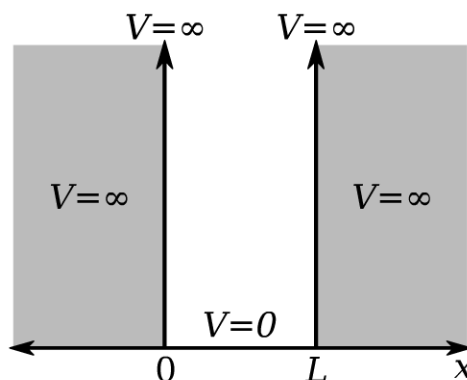


Figure 13.11 - Puits de potentiel infini. Le potentiel est nul à l'intérieur du puits (entre $x = 0$ et $x = L$) et infini en dehors.

(Source : http://fr.wikipedia.org/wiki/Particule_dans_une_boîte)

Classiquement, la particule est libre entre $x = 0$ et $x = L$, donc sa vitesse v est constante. La particule inverse le signe de sa vitesse lorsqu'elle rencontre l'un des deux murs, sur lesquels elle « rebondit » élastiquement. Son énergie est uniquement cinétique $E = mv^2/2$. E peut donc prendre n'importe quelle valeur positive, et sa plus petite valeur possible est 0 (particule au repos en un point quelconque entre 0 et L).

Décrivons maintenant la particule par sa fonction d'onde quantique. La probabilité de trouver la particule en dehors de l'intervalle $[0, L]$ est nulle, quelle que soit son énergie cinétique, d'où :

$$\psi(x < 0) = \psi(x > L) = 0 \quad (13.7)$$

Nous postulons que la fonction d'onde $\psi(x)$ est continue, on a donc :

$$\psi(x = 0) = \psi(x = L) = 0 \quad (13.8)$$

Nous cherchons ici une fonction d'onde décrivant une particule de vitesse, donc de longueur d'onde de de Broglie, donnée. La condition précédente impose une annulation de cette onde en deux points : on est dans une situation analogue à celle de la corde vibrante retenue en ses deux extrémités (condition dite d'onde stationnaire). Poursuivant l'analogie, on considère une solution sinusoïdale du type $\sin(2\pi x/\lambda_{dB})$ ou $\cos(2\pi x/\lambda_{dB})$, où la longueur d'onde de de Broglie λ_{dB} apparaît naturellement.

Comme elle doit s'annuler en $x = 0$, c'est nécessairement un sinus, et comme elle doit s'annuler en $x = L$, on doit avoir : $\sin(2\pi L/\lambda_{dB}) = 0$, donc :

$$n \frac{\lambda_{dB}}{2} = L \quad n = 1, 2, \dots \quad (13.9)$$

c'est-à-dire qu'il ne peut y avoir qu'un nombre entier de demi-longueurs d'onde entre 0 et L .

Grâce à (13.2) on en déduit l'expression de la vitesse v , puis de l'énergie $E = mv^2/2$ de la particule :

$$v = n \frac{h}{2mL} \quad ; \quad E = E_n = n^2 \frac{h^2}{8mL^2} \quad (13.10)$$

À la différence du système classique équivalent, *l'énergie du système quantique est donc quantifiée*. On appelle « niveaux d'énergie » ces différentes valeurs. À chacun d'entre eux est associée une fonction d'onde de forme bien déterminée, donc un vecteur d'état $|\Psi_n\rangle$ bien caractérisé (figure 13.12). On appelle « état fondamental » l'état quantique de plus basse énergie, et « états excités » les autres. L'état fondamental est celui dans lequel se trouve généralement le système en l'absence d'excitation extérieure.

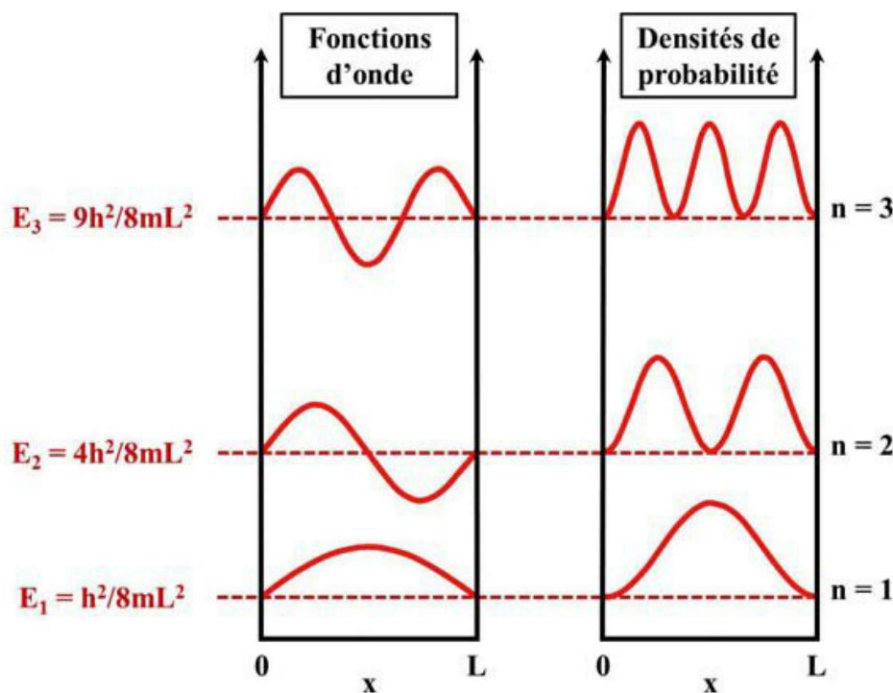


Figure 13.12 – Premiers niveaux d'énergie, fonctions d'onde associées, et densités de probabilité pour un puits de potentiel carré infini de largeur L .

(source libre : http://commons.wikimedia.org/wiki/File:Particle_in_a_box_wavefunctions.png)

13.2. Quantification de l'énergie pour une particule confinée

Pour un électron et pour de petites valeurs de n , les énergies E_n sont extrêmement faibles dès lors que L est supérieur au μm . Elles sont par contre de l'ordre de l'électronvolt si L est de l'ordre du nanomètre. Dans les « boîtes quantiques » de taille de l'ordre du nanomètre, l'électron a donc une énergie comparable à celle d'un atome, d'où le nom d'« atome artificiel » donné à ces boîtes quantiques. Leurs niveaux d'énergie sont ajustables en variant la dimension de la boîte lors de la fabrication.

Remarquons pour finir que la plus basse énergie possible de la particule est $E_1 = h^2/8mL^2$, quantité strictement positive, donc supérieure à l'état de plus basse énergie $E = 0$ de la particule classique. Nous aurons plus loin une explication qualitative de cette propriété.

La quantification de l'énergie est une des propriétés fondamentales des objets quantiques : on retrouve une dépendance de l'énergie en fonction d'un ou plusieurs nombres entiers dans tous les systèmes quantiques dans lesquelles la particule se trouve confinée, d'une manière ou d'une autre, dans une région finie de l'espace, comme les atomes ou les molécules. En particulier, l'atome d'hydrogène a des niveaux d'énergie donnés par la relation $E_n = -13,6/n^2 \text{ eV}$ où n est entier, expression que l'on peut retrouver en imposant à l'onde de matière (électronique) de comporter un nombre entier de longueurs d'onde de de Broglie sur une orbite circulaire de l'électron autour du noyau.

13.2.2 Application : système à deux états, qubit

Si l'on se restreint à des états quantiques peu excités de la boîte qui se décomposent uniquement sur l'état fondamental $n = 1$ et le premier niveau excité $n = 2$, l'état le plus général dans ce contexte s'écrit :

$$|\Psi\rangle = \alpha_1|\Psi_1\rangle + \alpha_2|\Psi_2\rangle \quad (13.11)$$

C'est un système à deux états proche de celui d'un photon polarisé, mais localisé dans une boîte de très petite taille. Les deux états sont des « bits » sur lesquels on peut faire du calcul, comme sur les bits 0 et 1 d'un ordinateur. Mais ici le système est quantique, et il s'agit de « qubit », sur lequel on peut faire des opérations quantiques différentes de celles d'un ordinateur classique. C'est le domaine de l'information quantique, en plein développement à l'heure actuelle car on a montré théoriquement que le calcul quantique pouvait réaliser certaines opérations beaucoup plus vite qu'un ordinateur classique (le calcul d'une opération logique sur une superposition de N états permettant le calcul *parallèle* de 2^N données en même temps). Il s'agit maintenant de passer de la théorie à la pratique !

À l'heure actuelle, en 2012, la réalisation pratique de tels qubits se fait au moyen d'ions piégés, de boîtes quantiques, de circuits supraconducteurs... En mars 2011,

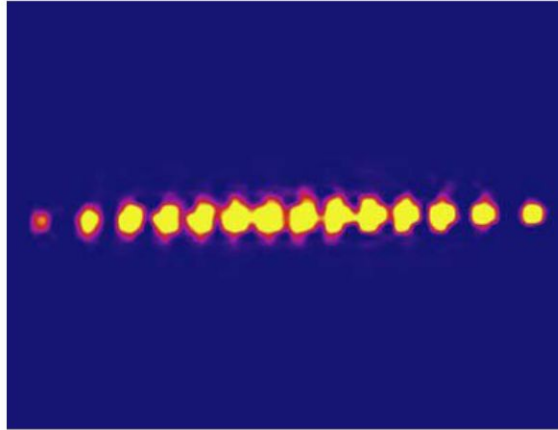


Figure 13.13 – « Record » de fabrication, en 2012, de 14 qubits, réalisés avec des ions piégés, dans le laboratoire de Rainer Blatt en Autriche.

(source : « 14-Qubit Entanglement : Creation and Coherence », T. Monz, P. Schindler, J. T. Barreiro, M. Chwalla, D. Nigg, W. A. Coish, M. Harlander, W. Hänsel, M. Hennrich, R. Blatt, *Phys. Rev. Lett.* **106**, 130506, 2011)

le record du nombre de qubits dans une architecture viable pour former un (futur) ordinateur quantique était de 14 qubits (figure 13.13).

13.2.3 Interaction avec la lumière

Lorsqu'un électron est piégé dans une boîte quantique, les règles d'absorption et d'émission de photons (chapitre 10) impliquent que les seules longueurs d'onde électromagnétiques émises ou absorbées par l'électron sont celles qui correspondent au passage d'un niveau d'énergie à un autre (règle de Bohr) :

$$|E_{n'} - E_n| = h\nu_{nn'} \implies \lambda_{n \rightarrow n'} = \left| \frac{hc}{E_{n'} - E_n} \right| = \frac{8mcL^2/h}{|n'^2 - n^2|} \quad (13.12)$$

Par exemple, pour la transition entre le premier état excité $n = 2$ et l'état fondamental $n' = 1$, on obtient : $\lambda_{21} \simeq 10^{-6} \cdot L^2$ où L est exprimée en nanomètres. Les nanotechnologies permettent de fabriquer des « boîtes quantiques » ayant des tailles de l'ordre de l'intervalle entre atomes dans un cristal, donc de l'ordre du nanomètre. En réglant la taille de ces boîtes, on peut déterminer à volonté la couleur qu'elles absorberont ou émettront.

13.2.4 Les boîtes quantiques en pratique...

Ce contrôle très précis de la couleur (et des autres propriétés électroniques) des boîtes quantiques intéresse beaucoup l'ingénierie laser. En effet, l'utilisation de boîtes quantiques comme milieu actif pour produire des lasers aux propriétés bien contrôlées (longueur d'onde, intensité, stabilité) est aujourd'hui un domaine en plein essor. Les

13.2. Quantification de l'énergie pour une particule confinée

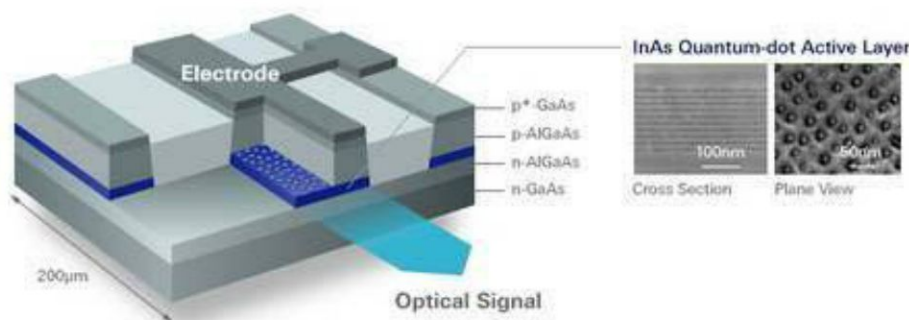


Figure 13.14 – Exemple de laser à boîtes quantiques. Au sein d'un sandwich de semi-conducteurs (arséniure de gallium et d'aluminium), les nanocristaux, semi-conducteurs eux aussi (InAs, arséniure d'indium), forment une couche de boîtes quantiques (Quantum-dot Active Layer) qui émet une lumière laser (Optical Signal). À droite, deux images des nanocristaux en microscopie électronique, en coupe (*Cross Section*) et vus de dessus (*Plane View*). « Quantum Dot Laser » de gallium indium phosphide.

(source : http://www.futura-sciences.com/fr/news/t/technologie-1/d/un-laser-quantique-miniature-pour-les-picoprojecteurs_20616)

applications en médecine (scalpel laser, tomographie optique) et dans les nouvelles technologies de l'information et de la communication sont sans cesse plus nombreuses. Elles sont en particulier au cœur des lasers qui produisent les impulsions lumineuses qui transportent l'information dans le réseau Internet.

D'un point de vue pratique, une boîte quantique n'est rien d'autre qu'un petit cristal de matériau semi-conducteur au sein duquel un ou plusieurs électrons sont confinés (à une, deux ou trois dimensions d'espace). Les bords du puits de potentiel ressenti par ces électrons correspondent donc (à peu près) aux bords physiques de ce cristal. Par conséquent, on peut moduler à volonté les propriétés de ces boîtes quantiques (entre autres la couleur qu'elles émettent ou absorbent) simplement en réglant la taille du cristal.

Comment alors créer ces petits cristaux ? Et, surtout, comment les créer de telle sorte qu'ils aient tous la même taille ?

Il existe en fait plusieurs méthodes pour créer des boîtes quantiques aujourd'hui. La plus utilisée est celle basée sur la croissance (chimique) de nanocristaux en solution, dont on contrôle la température et la concentration (méthode dite « colloïde »). Formées d'alliages binaires (séléniure ou sulfure de cadmium, arséniure ou phosphure d'indium) ou ternaires (sélénio-sulfure de cadmium), les boîtes quantiques qui en résultent ont une taille (bien contrôlée) de l'ordre de quelques nanomètres, et peuvent être produites par paquets de plusieurs millions en même temps.

D'autres méthodes de production, comme la lithographie ou la croissance par épitaxie, sont aujourd'hui également exploitées.

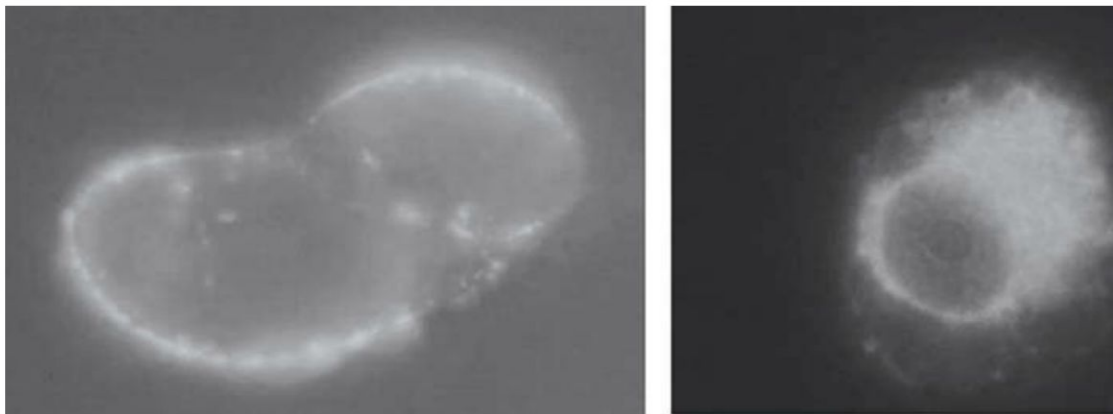


Figure 13.15 – Image de fluorescence de boîtes quantiques injectées en très grand nombre dans une cellule (d'une taille de quelques dizaines de micromètres). Quinze minutes après le début de l'expérience, les boîtes quantiques sont dans la membrane cellulaire (image de gauche). Au bout de quatre heures, elles ont pénétré la cellule et se sont distribuées dans le fluide cellulaire (image de droite, où la zone centrale plus sombre est le noyau cellulaire). L'utilisation de boîtes quantiques pour introduire un complexe actif dans une cellule tout en suivant sa progression est de plus en plus étudiée.

(source : www.washington.edu/news/2008/06/26/gene-silencer-quantum-dots-reduce-protein-production/)

Les propriétés particulières des boîtes quantiques en font d'excellentes candidates pour de multiples applications : en micro-électronique et en information quantique, par exemple, mais aussi, et surtout, dans le domaine de la biologie cellulaire où elles permettent d'avoir accès, *in vivo* et avec une résolution inégalée, aux processus inter-cellulaires ou au traçage en temps réel de cellules (figure 13.15).

Bien évidemment, l'introduction de nanocristaux semi-conducteurs dans un corps vivant n'est pas sans risque et de nombreuses équipes tentent actuellement de déterminer leur toxicité, à court et plus long terme.



13.1 Lien entre fonction d'onde et vecteur d'état

On a dit au chapitre 12 qu'à tout système physique était associé un vecteur d'état $|\psi(t)\rangle$. Le lien entre la fonction d'onde d'un système matériel $\psi(\mathbf{r}, t)$ et son vecteur d'état $|\psi(t)\rangle$ est l'objet de cet encart.

Pour le photon polarisé, on avait introduit le vecteur d'état $|\mathcal{P}\rangle = \alpha|H\rangle + \beta|V\rangle$, combinaison linéaire des états $|H\rangle$ et $|V\rangle$, qui sont les deux seuls états quantiques du photon qui ont un comportement certain lors de la traversée d'un polariseur « X », c'est-à-dire lors de la mesure de la polarisation du photon selon Ox : $|H\rangle$ passe à coup sûr, et $|V\rangle$ est absorbé à coup sûr. $|\mathcal{P}\rangle$ est donc défini par la donnée des nombres complexes α et β (avec la condition de normalisation $|\alpha|^2 + |\beta|^2 = 1$), coordonnées du vecteur $|\mathcal{P}\rangle$ sur la base $\{|H\rangle, |V\rangle\}$. Si le photon incident est décrit par un tel état, son passage à travers le polariseur X est aléatoire, $|\alpha|^2$ étant la probabilité de transmission et $|\beta|^2$ la probabilité d'absorption (loi de Malus).

13.2. Quantification de l'énergie pour une particule confinée

On se place maintenant dans une configuration analogue à celle d'un photon décrit par une onde plane se propageant dans la direction Oz , en l'occurrence une particule de masse m se déplaçant le long de l'axe Oz dont on veut connaître la position selon la direction Ox . Pour cela, on utilise un *détecteur de particules* D_n qui fournit un « clic » si la particule passe entre les points de coordonnées x_n et x_{n+1} (figure 13.16). On appelle $|x_n\rangle$ le vecteur qui décrit l'état de la particule qui arrive à coup sûr sur ce détecteur (il a pu être créé par exemple par un écran absorbant qui comporte un trou entre les abscisses x_n et x_{n+1}). On peut ensuite aligner le long de l'axe Ox N détecteurs analogues D_n ($n = 1, \dots, N$) qui « couvrent » tout l'axe, et permettent de savoir où se trouve la particule (figure 13.16). Par analogie avec le cas du photon polarisé rappelé dans le paragraphe précédent, un état quantique quelconque de la particule est alors décrit par le vecteur d'état :

$$|\Psi\rangle = \sum_{n=1}^N \psi_n |x_n\rangle \quad (13.13)$$

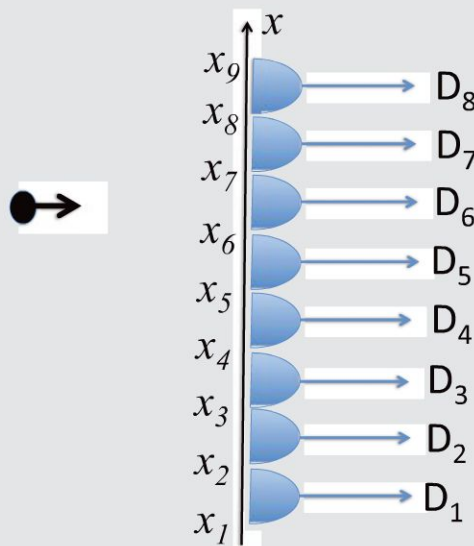


Figure 13.16 - Ensemble de détecteurs permettant de connaître la position le long de l'axe Ox d'une particule se déplaçant avec une vitesse parallèle à Oz

où les ψ_n sont les coordonnées du vecteur $|\Psi\rangle$ sur la base $|x_n\rangle$. $|\Psi\rangle$ est donc défini par la donnée de la liste des nombres complexes ψ_n (avec la condition de normalisation $\sum_n |\psi_n|^2 = 1$). Si la particule incidente est dans l'état $|\Psi\rangle$, elle sera détectée aléatoirement par l'un de ces détecteurs, $|\psi_n|^2$ étant la probabilité de détection par le détecteur D_n . Précisons que nous nous sommes placés ici à un instant fixé t , et qu'il faut donc écrire en toute rigueur :

$$|\Psi(t)\rangle = \sum_{n=1}^N \psi_n(t) |x_n\rangle \quad (13.14)$$

Si le détecteur D_n ne détruit pas la particule (c'est par exemple un trou dans un écran suivi d'un détecteur de grande dimension), le postulat P2 implique que

dans le cas où la particule a franchi le trou, son état conditionnel après traversée de l'écran est l'état $|x_n\rangle$, indépendant de l'état incident.

Pour connaître le plus précisément possible la position de la particule, on diminue la taille $x_{n+1} - x_n$ des détecteurs. Lorsque celle-ci tend vers zéro (et le nombre de détecteurs vers l'infini), on *passse à la limite continue* en remplaçant par une intégrale la somme dans (13.14), et l'on écrit :

$$|\Psi(t)\rangle = \int_{-\infty}^{\infty} dx \psi(x,t) |x\rangle \quad (13.15)$$

$|x\rangle$ étant l'état de la particule qui est détectée à coup sûr par un détecteur de taille infinitésimale dx situé en x . $|\Psi\rangle$ est donc défini par la donnée de la *fonction complexe* $\psi(x,t)$, *appelée fonction d'onde*. Si la particule incidente est dans l'état $|\Psi\rangle$ elle sera détectée aléatoirement par l'un de ces détecteurs, $|\psi(x,t)|^2 dx$ étant la probabilité de détection de la particule à l'instant t et entre les positions x et $x + dx$.

13.3 INÉGALITÉS DE HEISENBERG

Nous avons vu que la fonction d'onde ne contient qu'une information de type probabiliste sur les résultats de mesure de position que l'on peut faire sur le système. Sauf cas très particulier, la valeur de la position mesurée ne sera donc pas certaine, et si l'on fait des mesures répétées sur le système quantique toujours préparé dans le même état quantique, on enregistrera des résultats fluctuants. Nous allons maintenant préciser les contraintes qui existent sur ces fluctuations.

13.3.1 Écart-type

Prenons comme exemple la mesure de la position le long de l'axe Ox . La valeur trouvée x est une grandeur aléatoire dont la loi de probabilité est $P(x) = |\psi(x,t)|^2$. La valeur moyenne (ou espérance mathématique) $\langle x \rangle$ et l'écart-type Δx des valeurs mesurées sont donnés par les formules habituelles du calcul des probabilités :

$$\langle x \rangle = \int x P(x) dx \quad ; \quad \langle x^2 \rangle = \int x^2 P(x) dx$$

$$\Delta x = \sqrt{\langle x^2 \rangle - \langle x \rangle^2}$$

L'écart-type, dans le cas d'une distribution de probabilité formée d'une seule « bosse », donne une idée de la dispersion des résultats de mesure autour de la valeur moyenne, donc de l'incertitude sur la mesure. C'est alors, en ordre de grandeur, la « largeur » de la bosse.

13.3.2 Inégalité de Heisenberg position-impulsion

Il est aussi possible, comme nous le verrons plus en détail dans le chapitre suivant, de faire une mesure de la vitesse de la particule, et en particulier de sa composante v_x le long de l'axe Ox . Cette mesure donne aussi des résultats généralement incertains, avec une loi de probabilité connue qui s'exprime elle aussi en fonction de la fonction d'onde $\psi(x,t)$ du système. Cette dépendance commune de deux types de mesure par rapport à une même fonction induit des contraintes sur les caractéristiques de ces mesures. Une démonstration mathématique permet d'établir le résultat suivant :

Dans n'importe quel état quantique, les écarts-types relatifs aux mesures de position et de vitesse dans le même état obéissent à l'inégalité :

$$m\Delta v_x \Delta x \geq \frac{\hbar}{2} \quad (13.16)$$

Cet énoncé est parfois appelé, principe d'incertitude, doublement à tort car d'une part il laisse penser que cette inégalité résulte d'une imprécision expérimentale alors qu'en fait elle est de nature fondamentale, et d'autre part il ne s'agit pas d'un nouveau principe de la théorie, mais une propriété mathématique se déduisant des principes déjà énoncés. Il serait préférable de parler d'indéterminations quantiques Δx et Δv_x vérifiant, *avec certitude*, la relation (13.16).

L'inégalité de Heisenberg (13.16) implique qu'un état dans lequel la position de la particule est parfaitement définie ($\Delta x = 0$) présente une dispersion infinie sur la mesure de sa vitesse ($\Delta v_x = \infty$), et qu'un état dans lequel la vitesse de la particule est parfaitement définie ($\Delta v_x = 0$) présente une dispersion infinie sur la mesure de sa position ($\Delta x = \infty$). Elle exclut tout état pour lequel on aurait $\Delta x = \Delta v_x = 0$, ce qui entraîne qu'on ne peut pas connaître avec certitude à la fois la position et la vitesse d'une particule. Elle condamne donc toute notion de *trajectoire précise de la particule* dans un état quantique.

Mentionnons qu'en outre pour des états quantiques de basse énergie, le produit $m\Delta v\Delta x$ est souvent de l'ordre de \hbar .

La relation de Heisenberg ne concerne que les indéterminations suivant la *même direction* de l'espace et au même instant. Ainsi, on a aussi, suivant les axes y et z :

$$m\Delta y \Delta v_y \geq \frac{\hbar}{2} \quad (13.17)$$

$$m\Delta z \Delta v_z \geq \frac{\hbar}{2} \quad (13.18)$$

mais il n'existe pas de condition similaire entre Δx et Δv_y par exemple, ou entre Δx et Δy , ou Δv_x et Δv_z ... ce qui autorise, par exemple, de réaliser des mesures infiniment précises de la position d'une particule suivant l'axe x et de sa vitesse suivant l'axe y .

13.3.3 Interdiction du repos total

Imaginons une particule au repos, donc arrêtée en un point donné de l'espace et de vitesse nulle : on connaît donc parfaitement à la fois sa vitesse et sa position, ce qui est interdit par l'inégalité de Heisenberg. Un tel état n'existe pas dans le cadre d'une description quantique. C'est la raison pour laquelle l'état quantique de plus basse énergie de la boîte quantique n'est pas d'énergie nulle.

On peut faire un raisonnement plus précis : dans une boîte quantique l'incertitude Δx sur la position de la particule est inférieure ou égale à la taille a de la boîte. L'inégalité de Heisenberg implique donc que l'incertitude sur la vitesse Δv_x est supérieure à $\hbar/(2ma)$. Dans un état stationnaire du système, par raison de symétrie, la vitesse moyenne $\langle v_x \rangle$ est nulle et donc $\langle v_x^2 \rangle = \Delta^2 v_x - \langle v_x \rangle^2 = \Delta^2 v_x$. Donc dans tout état stationnaire de la boîte, l'énergie cinétique moyenne $m \langle v_x^2 \rangle / 2$ est supérieure ou égale à $\hbar^2/(8ma^2)$, ce qui donne le bon ordre de grandeur de l'énergie minimale E_1 .

13.3.4 Lien avec la diffraction

L'inégalité de Heisenberg est en fait une conséquence du caractère ondulatoire de la matière quantique. On peut s'en persuader par l'exemple suivant où on la relie à un phénomène ondulatoire bien connu, celui de la diffraction introduite à la section 13.1 : une onde de longueur d'onde λ se propageant le long de l'axe Oz et incidente sur un trou de dimension d le long de Ox percé dans un écran absorbant situé dans le plan xOy (figure 13.1) donne naissance de l'autre côté de l'écran à une onde divergente s'écartant de Oz d'un angle maximal θ_m donné par :

$$\sin \theta_m \simeq \frac{\lambda}{d} \quad (13.19)$$

Ce phénomène de diffraction est commun à tous les phénomènes ondulatoires (lumière, son, vagues sur l'eau...) et existe aussi pour les ondes de matière. Lorsqu'une particule de masse m et de vitesse v le long de Oz est incidente sur l'écran décrit ci-dessus, son onde de matière diffracte à la sortie du trou selon la même loi (13.19) avec $\lambda = \lambda_{dB}$.

Considérons l'état quantique du système juste après le trou. Dans cet état, la particule possède une fonction d'onde qui est non nulle uniquement au niveau du trou, donc où la position le long de l'axe Ox de la particule est définie avec un écart-type Δx égal à d . À cause de la diffraction, la particule après l'écran dévie d'un angle aléatoire α par rapport à Oz , positif ou négatif, inférieur en module à l'angle de diffraction α_m . Elle acquiert donc une composante de vitesse selon la direction Ox qui est aléatoire et vaut $v \sin \alpha$. L'écart type sur v_x , Δv_x , est donc de l'ordre de $v \sin \alpha_m$. On a ainsi :

$$m \Delta v_x \Delta x \simeq m v \sin \alpha_m d = m v \lambda_{dB} = h \quad (13.20)$$

On retrouve ici un produit $m \Delta v_x \Delta x$ de l'ordre de \hbar .

13.3.5 Généralisation

Position et vitesse forment un couple appelé « complémentaire » par Bohr, parce qu'on ne peut les connaître parfaitement dans un même état quantique. Il y en a d'autres en mécanique quantique. Citons par exemple l'angle de rotation ϕ autour d'un axe, et le moment cinétique J calculé autour de cet axe, dont les écarts-types obéissent à la même inégalité

$$\Delta\phi\Delta J \geq \hbar \quad (13.21)$$

L'énergie E d'un système et un temps caractéristique τ de l'évolution du système obéissent également à une relation proche :

$$\Delta E\tau \gtrsim h \quad (13.22)$$

Par exemple, un état atomique excité n'est en général pas stable. Il a une durée de vie τ . Au bout d'un temps aléatoire de l'ordre de τ l'atome transite vers un état d'énergie inférieure en émettant un photon, d'une manière analogue à un noyau instable qui se désintègre (chapitre 8). On montre que l'énergie de l'état excité instable n'est pas parfaitement définie, et que l'incertitude ΔE sur sa valeur vaut :

$$\Delta E = \frac{h}{\tau} \quad (13.23)$$

Cette relation temps-énergie n'est pas de même nature que les autres inégalités de Heisenberg, la durée τ n'étant pas liée à un quelconque écart-type d'une distribution « en temps » d'une grandeur « temps du système » par exemple. Le temps t en mécanique quantique n'est pas une variable mais un *paramètre*.

Il existe donc en mécanique quantique des couples de grandeurs (x et p_x , ϕ et J , E et τ ...) qui ne sont pas mesurables sur un système donné avec une précision infinie. C'est une illustration d'une propriété quantique spécifique, appelée *principe de complémentarité* par Bohr. Ces grandeurs reflètent des aspects d'un même système qui ne peuvent pas être connus avec certitude dans le même état quantique. Ce n'est alors qu'en considérant l'ensemble de ces éléments complémentaires que l'on obtient l'image du système la plus complète possible autorisée par les lois quantiques.

ÉVOLUTION TEMPORELLE DE LA FONCTION D'ONDE : ÉQUATION DE SCHRÖDINGER D'UNE PARTICULE MATÉRIELLE

14

Le chapitre précédent nous a permis d'introduire la notion de fonction d'onde et plusieurs propriétés importantes de celle-ci à un instant t donné : sa longueur d'onde caractéristique λ_{dB} , son interprétation probabiliste, son lien avec la notion de vecteur d'état, les liens entre ses fluctuations caractéristiques Δx , Δv_x , ΔE ... Dans ce chapitre, nous allons introduire l'équation d'évolution à laquelle elle obéit. L'évolution temporelle n'a été vue, en effet, que dans le cas des vecteurs d'état $|\Psi(t)\rangle$ et non pour les fonctions d'onde $\psi(x, t)$.

Mais avant d'aborder l'équation d'évolution, il est utile de préciser les propriétés vérifiées par une fonction d'onde quelconque.

14.1 PROPRIÉTÉS DE LA FONCTION D'ONDE

Une fonction d'onde $\psi(x, t)$ (où l'on a fait apparaître la dépendance temporelle pour plus de généralité) doit obéir à certaines conditions et vérifier certaines propriétés. En voici la liste :

- deux fonctions d'onde différant par un nombre complexe $e^{i\theta}$ de module un (que l'on appelle communément un *facteur de phase*)

$$\varphi(x) = \psi(x) e^{i\theta} \quad \text{avec} \quad \theta \in \mathbb{R} \quad (14.1)$$

conduisent aux mêmes probabilités de présence. Cette propriété est en fait très générale et concerne tout type de mesure (pas nécessairement de position lorsque le coefficient θ ne dépend pas de x). Un tel facteur de phase (global) n'ayant aucun effet sur les résultats de mesure, la description d'un système physique peut s'effectuer, de façon totalement équivalente, à l'aide de $\psi(x)$ ou $\varphi(x)$. On dit qu'une fonction d'onde est définie à un facteur de phase près.

- $\psi(x, t)$ doit être continue partout.

- sa dérivée spatiale doit être continue partout, sauf en un nombre fini de points isolés (par exemple, là où l'énergie potentielle est infinie, comme dans le cas des bords du puits carré infini vu au chapitre précédent).
- la probabilité de détecter la particule entre les points x_1 et x_2 vaut $\int_{x_1}^{x_2} dx |\psi(x, t)|^2$ (interprétation de Born vue au chapitre précédent).
- $\psi(x, t)$ doit être de carré sommable (*condition de normalisation*) :

$$\int_{-\infty}^{+\infty} |\psi(x, t)|^2 dx = 1 \quad (14.2)$$

ce qui traduit simplement le fait que l'on est sûr (c'est-à-dire avec une probabilité égale à un) de détecter l'objet physique en question quelque part le long de l'axe x .

14.2 COMMENTAIRES

1. Mathématiquement, les propriétés ci-dessus ne traduisent rien d'autre que le fait que les fonctions d'onde appartiennent à un *espace vectoriel* : l'espace vectoriel de Hilbert des fonctions complexes, continues, de carré sommable (que l'on note $L^2(\mathbb{R})$ à une dimension d'espace et $L^2(\mathbb{R}^3)$ à trois dimensions d'espace). Cet espace vectoriel est appelé *espace des états* de la particule matérielle.
2. Comme pour tout type d'onde, les fonctions d'onde obéissent à un *principe de superposition* : si $\psi_1(x, t)$ et $\psi_2(x, t)$ sont deux fonctions d'onde possibles pour un système, alors toute combinaison linéaire (à coefficients complexes, normalisée à un) de ces fonctions d'onde est aussi une fonction d'onde *possible* pour ce système.
3. Le produit scalaire $\langle \Psi_1 | \Psi_2 \rangle$ de deux vecteurs d'état $|\Psi_1\rangle$ et $|\Psi_2\rangle$ s'exprime en fonction des deux fonctions d'onde associées par

$$\langle \Psi_1 | \Psi_2 \rangle = \int_{-\infty}^{\infty} dx \psi_1^*(x) \psi_2(x) \quad (14.3)$$

où $\psi_1^*(x)$ est le complexe conjugué du nombre complexe $\psi_1(x)$.

4. Si on utilise un détecteur de position « non destructif », qui n'absorbe pas la particule lorsqu'il la détecte (comme la chambre à fils des physiciens des particules), et que le détecteur D_n « clique », la particule après la mesure est dans l'état conditionnel $|x_n\rangle$ introduit dans l'encart du chapitre précédent (c'est-à-dire $|x\rangle$ à la limite continue) quel que soit l'état de la particule incidente. On parle de *réduction* ou de *projection* dans l'état $|x_n\rangle$.

14.3 ÉQUATION DE SCHRÖDINGER POUR LA FONCTION D'ONDE

Comme pour toute onde, le lien entre la variable x et le paramètre t découle directement de la relation entre la pulsation ω et le vecteur d'onde $k = \|\mathbf{k}\|$, c'est-à-dire de la relation entre l'énergie E et la quantité de mouvement \mathbf{p} . Cette relation, appelée *relation de dispersion*, est très importante pour savoir comment évolue une onde réelle, c'est-à-dire une onde composée de plusieurs ondes élémentaires (planes, par exemple).

Pour une particule libre de matière, on a simplement $E = \mathbf{p}^2/2m$ ce qui conduit à la relation de dispersion $\omega = \hbar \mathbf{k}^2/2m$. En revanche, pour un photon ou toute particule de masse nulle, on a $E = \|\mathbf{p}\|c$, et la relation de dispersion s'écrit simplement $\omega = \|\mathbf{k}\|c$.

14.3.1 Onde plane de matière

On considère une particule de masse m , de vitesse \mathbf{v} , et d'énergie E bien déterminées. On lui associe une onde dont la pulsation et le vecteur d'onde sont donnés par (13.1). On postule que la fonction d'onde décrivant un tel état quantique s'écrit :

$$\psi(\mathbf{r}, t) = \psi_0 \exp(i\mathbf{k} \cdot \mathbf{r} - i\omega t) = \psi_0 \exp\left(\frac{i}{\hbar}(\mathbf{p} \cdot \mathbf{r} - Et)\right) \quad (14.4)$$

où ψ_0 est fixé, à un facteur de phase près, par la condition de normalisation à 1. Cette forme est « plausible », car elle ressemble à l'onde plane lumineuse du chapitre 2. Remarquons toutefois que le champ électromagnétique de l'onde lumineuse est une quantité réelle, et que l'on n'utilise le champ complexe que par commodité de calcul, alors que l'onde de matière est *nécessairement* décrite par une exponentielle complexe. On peut justifier cette différence par le fait que l'énergie E d'une particule n'est définie qu'à une constante additive près. Un décalage global en énergie $E \rightarrow E + E_0$ pour la particule se traduit sur la fonction d'onde par la multiplication par le facteur de phase $e^{-iE_0 t/\hbar}$, qui ne change pas les prévisions physiques sur la particule. Il n'en est pas de même sur la forme réelle, en cosinus, de l'onde lumineuse, qui, étant proportionnelle à la force électrique, quantité mesurable, n'a pas la même invariance par changement de phase.

Remarquons que pour l'onde plane de matière (14.4) $|\psi(\mathbf{r}, t)|^2 = |\psi_0|^2$ ne dépend ni de \mathbf{r} ni de t . La probabilité de présence d'une particule de vitesse \mathbf{v} et d'énergie E bien déterminées est la même en tout point : on dit que la particule est *délocalisée*.

Il est à noter qu'une onde plane n'est pas véritablement une fonction d'onde puisqu'elle n'est pas de carré sommable (l'intégrale de son module au carré vaut l'infini). En fait seule une somme (continue) d'ondes planes peut vérifier cette condition. Il n'en reste pas moins que les ondes planes sont très utiles d'un point de vue pratique

(exactement comme en électromagnétisme). D'une part, l'équation de Schrödinger étant linéaire, si l'on sait comment évolue *une* onde plane, on sait comment évolue toute somme (continue ou discrète) d'ondes planes, c'est-à-dire finalement *toute* fonction d'onde. D'autre part, une onde plane peut souvent être utilisée comme première approximation d'une fonction d'onde décrivant une particule de vitesse \mathbf{p}/m délocalisée dans une zone étendue de l'espace.

14.3.2 Équation d'évolution pour la fonction d'onde

Remarquons que pour toute onde plane de matière (14.4) :

$$\frac{\partial}{\partial t}\psi(\mathbf{r},t) = -\frac{iE}{\hbar}\psi(\mathbf{r},t) \quad ; \quad \frac{\partial^2}{\partial x^2}\psi(\mathbf{r},t) = -\frac{m^2v_x^2}{\hbar^2}\psi(\mathbf{r},t) \quad (14.5)$$

où v_x est la composante de la vitesse selon x .

Considérons alors le cas où la particule est soumise à une force dérivant d'un potentiel (comme le potentiel gravitationnel par exemple, ou le potentiel électrique). Son énergie E est la somme de l'énergie cinétique et de l'énergie potentielle ¹ que nous noterons $V(\mathbf{r})$, donc E s'écrit :

$$E = \frac{1}{2}m(v_x^2 + v_y^2 + v_z^2) + V(\mathbf{r}) \quad (14.6)$$

En multipliant les deux membres de cette relation par $\psi(\mathbf{r},t)$ et en utilisant (14.5), on obtient une équation satisfaite par toutes les ondes planes de matière, quelles que soient leur énergie et leur vitesse :

$$i\hbar\frac{\partial}{\partial t}\psi(\mathbf{r},t) = -\frac{\hbar^2}{2m}\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\right)\psi(\mathbf{r},t) + V(\mathbf{r})\psi(\mathbf{r},t) \quad (14.7)$$

C'est une équation *linéaire* en ψ , qui est donc aussi **satisfaite** par toutes les superpositions linéaires d'ondes planes, et donc **par toutes les fonctions d'onde** de la particule. On l'appelle *équation de Schrödinger*.

Nous allons maintenant résoudre cette équation dans trois configurations particulières de potentiel : la particule libre, le puits carré infini, et la marche de potentiel.

1. Par exemple, le potentiel V s'écrit $V = +mgz$ dans le cas d'une particule de masse m dans un champ gravitationnel uniforme \vec{g} , de sens opposé au vecteur unitaire définissant l'axe z .

14.4 EXEMPLE 1 : LA PARTICULE LIBRE

14.4.1 Onde plane unique

Considérons une particule libre, de masse m , évoluant selon la direction x avec la vitesse $v > 0$. Conformément à la section précédente, elle est décrite par une onde plane :

$$\psi(x,t) = \psi_0 e^{\frac{i}{\hbar}(mvx - Et)} \quad (14.8)$$

où l'énergie E est égale à son énergie cinétique $\frac{1}{2}mv^2$ puisque $V = 0$.

Comme souligné précédemment, la particule ainsi décrite est délocalisée dans tout l'espace puisque le module au carré de $\psi(x,t)$ ne dépend ni de x ni de t : $|\psi(x,t)|^2 = |\psi_0|^2$.

14.4.2 Superposition d'ondes planes

Revenons à l'exemple de l'interférence d'électrons décrite à la section 13.1.4 (figure 13.8). Le fil électrique chargé dévie vers la droite un électron se propageant le long de Oz et passant à gauche du fil, et vers la gauche un électron passant à droite de celui-ci. On peut considérer, en première approximation, qu'en dessous du fil, l'électron est décrit par la somme de deux ondes planes, l'une correspondant à la particule de vitesse v dont la direction fait un angle $-\theta$ avec l'axe Oz :

$$\psi_1(\mathbf{r},t) = \psi_0 \exp\left(\frac{imv}{\hbar}(z \cos \theta - x \sin \theta) - iEt/\hbar\right) \quad (14.9)$$

et l'autre correspondant à la particule de vitesse v dont la direction fait un angle $+\theta$ avec l'axe Oz :

$$\psi_2(\mathbf{r},t) = \psi_0 \exp\left(\frac{imv}{\hbar}(z \cos \theta + x \sin \theta) - iEt/\hbar\right) \quad (14.10)$$

Il y a donc analogie parfaite avec le calcul d'interférences lumineuses fait dans le cas du biprisme au chapitre 10.

Dans la région de recouvrement de ces ondes, l'électron a pour fonction d'onde :

$$\psi(\mathbf{r},t) = \frac{1}{\sqrt{2}}(\psi_1 + \psi_2) = \sqrt{2}\psi_0 \exp\left(\frac{imvz \cos \theta - iEt}{\hbar}\right) \cos\left(\frac{mvx \sin \theta}{\hbar}\right) \quad (14.11)$$

La probabilité $P(\mathbf{r},t) = |\psi|^2$ de présence de l'électron au point \mathbf{r} vaut donc :

$$P(\mathbf{r},t) = 2|\psi_0|^2 \cos^2\left(\frac{mvx \sin \theta}{\hbar}\right) \quad (14.12)$$

Elle dépend maintenant de x . Un électron isolé sera donc détecté à une position aléatoire, à l'exception des points où $P(\mathbf{r}, t) = 0$, c'est-à-dire où il y a interférence totalement destructive. Si on envoie un par un de nombreux électrons, le nombre de points lumineux enregistrés sur le détecteur à l'abscisse x finira par être proportionnel à $P(\mathbf{r}, t)$ et donnera donc naissance à des franges d'interférences dans la densité d'électrons.

D'une façon générale, si l'on considère deux fonctions d'onde possibles, $\psi_1(x, t)$ et $\psi_2(x, t)$, pour un système physique, c'est le fait que l'on somme les *amplitudes* de probabilités $\psi_1(x, t) + \psi_2(x, t)$ et non les densités de probabilités $|\psi_1(x, t)|^2 + |\psi_2(x, t)|^2$, qui permet de rendre compte des phénomènes d'interférences d'ondes de matière :

$$|\psi_1 + \psi_2|^2 = |\psi_1|^2 + |\psi_2|^2 + \underbrace{(\psi_1^* \psi_2 + \psi_1 \psi_2^*)}_{\text{interférences}} \quad (14.13)$$

où ψ_1^* représente le complexe conjugué du nombre ψ_1 .

14.5 EXEMPLE 2 : LE Puits CARRÉ INFINI

On considère, comme dans la section 13.2, une particule se déplaçant le long de l'axe Ox et soumise à un potentiel $V(x)$ qui est infini pour $x < 0$, nul pour $0 \leq x \leq L$ et à nouveau infini pour $x > L$. Il s'agit d'une particule libre entre deux « murs » situés en $x = 0$ et $x = L$, qu'elle ne peut en aucune manière franchir, quelle que soit son énergie cinétique (figure 13.11). Nous allons maintenant approfondir l'approche de la section 13.2 en cherchant l'ensemble des solutions possibles de l'équation de Schrödinger pour la fonction d'onde $\psi(x, t)$.

La particule a une probabilité nulle de se trouver « de l'autre côté des murs », donc $\psi(x, t) = 0$ pour $x < 0$ et $x > L$ pour tout t . On en déduit par continuité que $\psi(0, t) = \psi(L, t) = 0$. On cherche donc une fonction complexe $\psi(x, t)$ de norme 1 ($\int |\psi(x, t)|^2 dx = 1$) satisfaisant à l'équation de Schrödinger :

$$i\hbar \frac{\partial}{\partial t} \psi(x, t) = -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} \psi(x, t) \quad (14.14)$$

pour x appartenant à l'intervalle $[0, L]$, et nulle en 0 et L .

14.5.1 Recherche de solutions particulières d'énergie donnée

Comme $E = \hbar\omega$, nous allons chercher des solutions « monochromatiques », de pulsation ω donnée, donc de la forme :

$$\psi(x, t) = \phi(x) e^{-iEt/\hbar} \quad (14.15)$$

14.5. Exemple 2 : le puits carré infini

On déduit de (14.14) que la fonction ϕ est solution de l'équation différentielle :

$$\frac{d^2}{dx^2}\phi = -\frac{2mE}{\hbar^2}\phi \quad (14.16)$$

et doit aussi être nulle en 0 et L .

Nous cherchons une solution d'énergie positive $E \geq 0$ à (14.16). Celle-ci s'écrit sous la forme d'une combinaison linéaire d'exponentielles complexes : $\alpha' e^{ikx} + \beta' e^{-ikx}$ avec $k = \sqrt{2mE/\hbar^2}$, et peut donc être mise, de façon équivalente, sous la forme : $\alpha \sin kx + \beta \cos kx$, où α et β sont deux constantes complexes. La condition aux limites $\phi(0) = 0$ impose alors que $\beta = 0$, tandis que la condition aux limites $\phi(L) = \sin kL = 0$ impose que k ne peut pas prendre n'importe quelle valeur. Plus précisément k est un des k_n obéissant à la relation :

$$k_n = \frac{n\pi}{L} \quad (14.17)$$

où n est un entier naturel non nul. La condition de normalisation impose alors que $\alpha = \sqrt{2/L}$ à un facteur de phase près, sans signification physique.

Toute solution d'énergie donnée E_n de l'équation de Schrödinger s'écrit donc (pour le puits carré infini) :

$$\psi_n(x,t) = \sqrt{\frac{2}{L}} e^{-iE_n t/\hbar} \sin\left(\frac{n\pi x}{L}\right) \quad (14.18)$$

l'énergie E_n valant :

$$E_n = n^2 \frac{\hbar^2 \pi^2}{2mL^2} \quad (14.19)$$

Comme nous l'avons déjà noté au chapitre précédent, dans un tel système, l'énergie ne peut prendre que certaines valeurs : elle est *quantifiée*.

La densité de probabilité $P_n(x)$ de présence de la particule en x dans un état n vaut $|\psi_n(x,t)|^2$:

$$P_n(x) = \frac{2}{L} \sin^2\left(\frac{n\pi x}{L}\right) \quad (14.20)$$

Elle est *indépendante du temps* : contrairement au cas classique (où la particule fait constamment des allers-retours entre les murs), la particule *quantique* « ne bouge pas », d'où le nom d'état stationnaire donné aux états quantiques d'énergie totale non-nulle déterminée. $P_n(x)$ définit une sorte de « nuage fixe », caractérisant de manière *statistique* la position de la particule.

La figure 14.1 montre les variations spatiales des premières densités de probabilités de présence $P_1(x)$, $P_2(x)$, $P_3(x)$. Ces probabilités ne sont maximales qu'en certains points du puits et s'annulent en d'autres points (appelés *nœuds de vibration*), en totale contradiction avec la description classique où l'on s'attend à détecter la particule

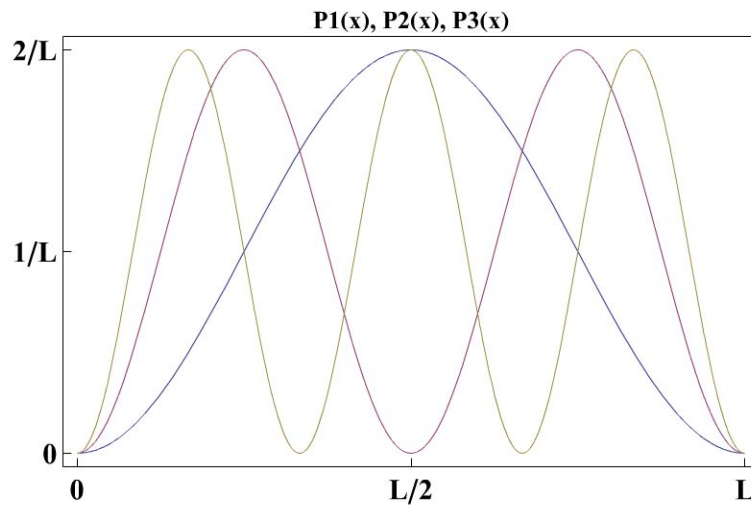


Figure 14.1 – Densités de probabilité de présence $P_1(x)$, $P_2(x)$, $P_3(x)$ du puits carré infini entre $x = 0$ et $x = L$.

en tout endroit du puits avec la même probabilité. Cependant, plus le nombre quantique n augmente, plus la distribution de probabilité $P_n(x)$ présente des nœuds de vibration de plus en plus proches les uns des autres (figure 14.2 où sont représentées $P_1(x)$, $P_{20}(x)$ et $P_{classique} = 1/L$). Pour un nombre n suffisamment grand, ces nœuds de vibration sont si proches qu'il est impossible de les distinguer expérimentalement. Dans ce cas, une mesure de position de résolution dx reviendrait à moyenner la distribution de probabilité $P_n(x)$ sur l'intervalle dx , la densité de probabilité obtenue étant uniforme et donc égale à $1/L$ comme dans le cas classique. Ceci est un exemple de ce qu'on appelle parfois le *principe de correspondance*, et qui indique que l'on s'attend à retrouver certaines propriétés classiques lorsque l'on fait tendre les nombres quantiques vers l'infini.

14.5.2 Solution générale

L'équation de Schrödinger étant linéaire, la solution générale s'obtient en superposant les solutions particulières. On peut montrer que les solutions stationnaires trouvées dans la section précédente forment une base de l'espace des solutions. La solution la plus générale de l'équation de Schrödinger s'écrit donc :

$$\psi(x,t) = \sum_n \alpha_n \psi_n(x,t) = \sum_n \alpha_n e^{-iE_n t/\hbar} \phi_n(x) \quad (14.21)$$

Elle est définie par la donnée de la liste des coefficients (α_n) .

Intéressons-nous maintenant à la dépendance temporelle d'une fonction d'onde particulière, superposition des deux premiers états d'énergie E_1 et E_2 à poids égaux

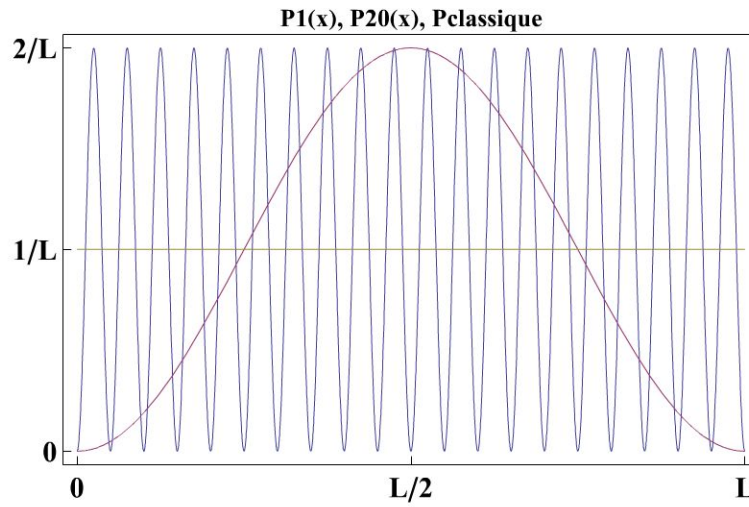


Figure 14.2 - Densités de probabilité de présence $P_1(x)$, $P_{20}(x)$, $P_{classique} = 1/L$ du puits carré infini entre $x = 0$ et $x = L$.

(c'est-à-dire $\alpha_1 = \alpha_2 = 1/\sqrt{2}$ et $\alpha_{n>2} = 0$) :

$$\psi(x,t) = \frac{1}{\sqrt{2}} \left(e^{-iE_1 t/\hbar} \phi_1(x) + e^{-iE_2 t/\hbar} \phi_2(x) \right) \quad (14.22)$$

Prenant les coefficients réels pour simplifier, on en déduit la probabilité de présence P de la particule en x :

$$P(x,t) = |\psi(x,t)|^2 = \frac{1}{2} \left[\phi_1^2(x) + \phi_2^2(x) + 2\phi_1(x)\phi_2(x) \cos\left(\frac{(E_2 - E_1)t}{\hbar}\right) \right] \quad (14.23)$$

On constate que la *probabilité de présence dépend maintenant du temps*. La particule « bouge » dans un tel état, et son mouvement présente des va-et-vient analogues à ceux du mouvement classique (figure 14.3 et l'animation <http://en.wikipedia.org/wiki/File:InfiniteSquareWellAnimation.gif>).

On note également l'importance cruciale du facteur de phase relatif $e^{-i(E_2 - E_1)t/\hbar}$ qui, contrairement au facteur de phase global (section 14.1), ne disparaît pas lorsque l'on prend le module au carré de la somme 14.22 et qui est responsable de sa variation temporelle.

14.6 GÉNÉRALISATION

En mécanique quantique, contrairement à ce qui se passe en mécanique classique, savoir comment évolue un système physique (une particule par exemple) dans un potentiel donné $V(x)$ revient à déterminer les *fonctions d'ondes solution de l'équation de Schrödinger* 14.7 (en tenant compte des éventuelles conditions de continuité aux

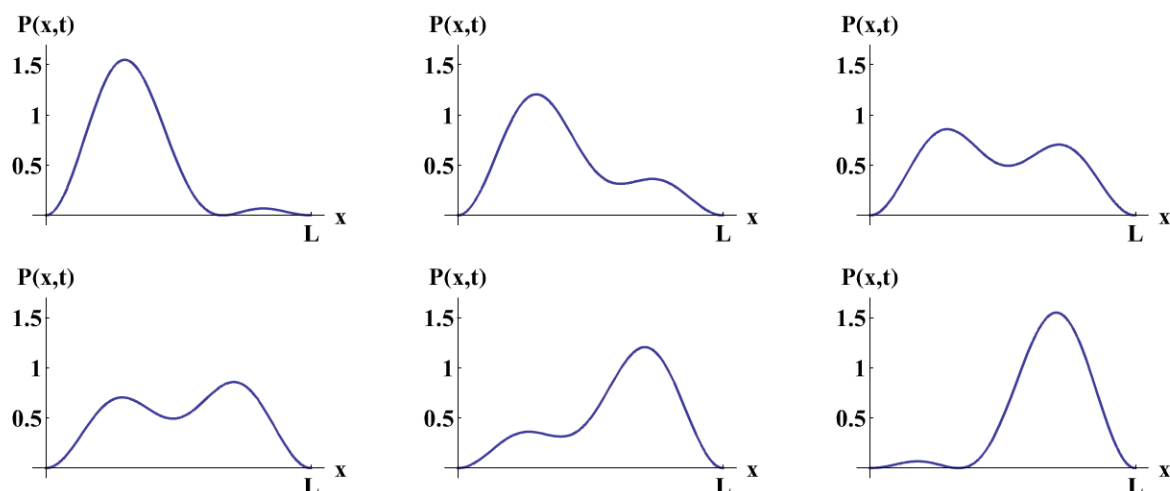


Figure 14.3 – Densités de probabilité de présence $P(x,t)$ à six instants différents. La maximum de $P(x,t)$ passe, au cours du temps, d'un bord à l'autre de la boîte.

frontières du domaine de définition du potentiel) ainsi que les *énergies possibles* du système.

Comme on peut le constater, la résolution de l'équation de Schrödinger est extrêmement simple dans le cas de la boîte quantique, c'est-à-dire du puits carré infini. Ce n'est pas surprenant car ce n'est rien d'autre qu'une version « avec frontières » (infiniment hautes) du cas de la particule libre pour lequel le potentiel est constant et uniforme (et même nul !) $V(x) = 0$.

En revanche, lorsque le potentiel dépend explicitement de la position x , l'équation de Schrödinger devient très difficile, voire impossible, à résoudre analytiquement. Ce n'est que pour une classe très restreinte de potentiels (puits parabolique, barrière rectangulaire, potentiels centraux, etc.) que l'on sait mener à bien les calculs, et il est nécessaire, dans la majorité des cas, d'utiliser l'outil numérique.

14.7 EXEMPLE 3 : MARCHE DE POTENTIEL, ONDE DE MATIÈRE ÉVANESCENTE

Nous allons maintenant nous intéresser à une situation souvent rencontrée dans les expériences, celle de la *diffusion*, ou *collision*, d'une particule, dans laquelle une particule incidente rencontre un obstacle localisé dans l'espace, sur lequel elle peut être réfléchiée, déviée, ou transmise. Loin de l'obstacle, la particule est donc libre et décrite par l'onde plane de matière $e^{i\vec{k}\vec{r}}$.

Nous traiterons ici le cas le plus simple, monodimensionnel dans la direction x , dans lequel la particule incidente venant de $x = -\infty$, d'énergie E , est libre pour toutes les valeurs négatives de x ($V(x) = 0 \forall x < 0$, région I). Elle rencontre en $x = 0$ une

14.7. Exemple 3 : marche de potentiel, onde de matière évanescence

« marche de potentiel » où l'énergie potentielle passe de 0 à une valeur constante V et conserve cette valeur pour toutes les valeurs positives de x ($V(x) = V \forall x > 0$, région II)

14.7.1 Cas d'une particule classique

Du point de vue classique, la particule n'est soumise à aucune force $F = -dV/dx$, sauf en $x = 0$ où elle subit un changement brutal d'énergie cinétique. Si son énergie cinétique initiale E est supérieure à V , elle traverse la marche et son énergie cinétique (donc sa vitesse) diminue. Si, par contre, son énergie E est inférieure à V , la particule ne pénètre pas dans la région des x positifs, mais « rebondit » sur l'obstacle : elle est réfléchiée en gardant sa vitesse initiale mais change de signe.

Expérimentalement, un potentiel en forme de marche (figure 14.4) correspond, en première approximation, au potentiel ressenti par des particules qui se heurtent à un obstacle. C'est le cas, par exemple, des électrons de conduction d'un métal lorsqu'ils atteignent la surface de ce métal. Ils restent piégés au sein du métal tant qu'ils n'ont pas l'énergie suffisante pour s'en échapper (par effet photoélectrique par exemple). La marche de potentiel correspond alors à cette énergie seuil. On peut aussi produire par épitaxie des empilements de matériaux semi-conducteurs de composition différentes (figure 9.12), qui créent pour l'électron de conduction des marches de potentiels ajustables.

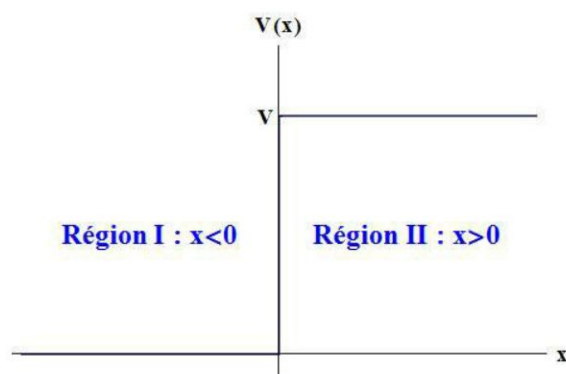


Figure 14.4 - Marche de potentiel de hauteur V .

14.7.2 Recherche de solutions d'énergie E donnée

Comme dans le cas du puits carré infini, la recherche de solutions d'énergie E positive donnée conduit aux équations :

$$x < 0 : \quad \frac{d^2}{dx^2} \phi = -\frac{2mE}{\hbar^2} \quad (14.24)$$

$$x > 0 : \quad \frac{d^2}{dx^2} \phi = -\frac{2m(E - V)}{\hbar^2} \quad (14.25)$$

Dans la région I, la solution est de la forme :

$$\phi_I(x) = Ae^{ik_I x} + Be^{-ik_I x} \quad (14.26)$$

avec $k_I = \sqrt{2mE/\hbar^2}$. Elle est la superposition de l'onde $Ae^{ik_I x}$ décrivant la particule incidente et de l'onde $Be^{-ik_I x}$ décrivant la particule réfléchie.

– Considérons tout d'abord le cas $E > V$: la solution dans la région II est alors aussi une somme d'exponentielles complexes :

$$\phi_{II}(x) = Ce^{ik_{II} x} + De^{-ik_{II} x} \quad (14.27)$$

avec $k_{II} = \sqrt{2m(E - V)/\hbar^2}$. C'est donc, en général, la somme de deux ondes se propageant vers les x positifs ou négatifs. Comme nous voulons décrire le problème de diffusion évoqué plus haut, nous allons chercher une solution où il n'y a qu'une onde transmise, se propageant vers les $x > 0$, donc avec $D = 0$.

Par analogie avec une onde lumineuse, nous appellerons coefficient de réflexion R la quantité $|B/A|^2$ et coefficient de transmission T la quantité $|C/A|^2$. En utilisant les conditions de continuité en $x = 0$ de la fonction d'onde $\phi(x)$ et de sa dérivée $\phi'(x)$, il est possible de calculer ces quantités. On trouve :

$$R = \left(\frac{1 - \sqrt{1 - V/E}}{1 + \sqrt{1 - V/E}} \right)^2 ; \quad T = 1 - R \quad (14.28)$$

On constate donc une différence importante avec le cas classique : il y a une probabilité non nulle que la particule quantique rebrousse chemin sur l'obstacle, alors qu'elle est transmise à coup sûr dans le cas classique. Remarquons que le coefficient R est proche de 1 quand $E \simeq V$ et tend vers 0 lorsque E/V tend vers l'infini. Ce phénomène porte le nom de *réflexion quantique*. Il est l'analogue, pour les ondes de matière, du phénomène bien connu de réflexion des ondes électromagnétiques lors de leur passage d'un milieu transparent à un autre d'indice optique différent.

– Considérons maintenant le cas $E < V$: la solution dans la région II est alors une somme d'exponentielles réelles :

$$\phi_{II}(x) = Ce^{-\kappa x} + De^{\kappa x} \quad (14.29)$$

avec $\kappa = \sqrt{2m(V - E)/\hbar^2}$. Lorsque $D \neq 0$ l'amplitude de l'onde de matière diverge lorsque x croît. Nous chercherons donc une solution avec $D = 0$. En utilisant les conditions de continuité en $x = 0$ de la fonction d'onde $\phi(x)$ et de sa dérivée $\phi'(x)$, il est possible de trouver une telle solution. On trouve en particulier :

$$\left| \frac{B}{A} \right|^2 = 1 ; \quad \frac{C}{A} = \frac{2ik_I}{ik_I - \kappa} \quad (14.30)$$

On remarque d'abord que le coefficient de réflexion est de module 1, ce qui implique que la particule est réfléchiée par la barrière avec certitude, comme dans le cas classique. Mais on constate aussi que $C \neq 0$: la fonction d'onde est non nulle dans la région II, classiquement interdite. Il y a donc une probabilité non-nulle de détecter la particule dans la région II, c'est-à-dire « dans » la marche ! Cette onde transmise $Ce^{-\kappa x}$ n'est pas une onde plane progressive et son amplitude décroît exponentiellement en x ; elle s'annule au bout d'une distance de quelques κ^{-1} : c'est une *onde évanescence de matière*, dont l'analogue, en électromagnétisme, correspond à l'onde évanescence existant dans un dioptré lors de la réflexion totale d'une onde électromagnétique.

Considérons alors non pas une marche, mais une barrière de potentiel de hauteur V , qui revient à la valeur $V(x) = 0$ pour $x > \ell$, avec ℓ de l'ordre de κ^{-1} : on s'attend à ce que l'onde évanescence de matière ait une valeur non négligeable en $x = \ell$ et se raccorde, en utilisant toujours les relations de continuité, à une onde progressive d'amplitude non nulle pour $x > \ell$. La particule a alors une probabilité non nulle d'être transmise au-delà d'une barrière qu'elle ne peut classiquement franchir ! C'est ce qu'on appelle l'*effet tunnel*. Ce phénomène est à la base du *microscope à effet tunnel* (figure 14.5).

14.8 MESURE DE LA VITESSE DE LA PARTICULE, NOTION D'OBSERVABLE

Jusqu'ici, les seules mesures que nous avons présentées en détail sont celles de position et d'énergie. Il existe bien évidemment d'autres grandeurs remarquables pour une particule, en premier lieu desquelles sa vitesse (ou sa quantité de mouvement).

14.8.1 État ayant une valeur certaine de la vitesse

On se place à un instant donné $t = 0$, et à une seule dimension d'espace. Un état quantique $|\Psi_v\rangle$ de vitesse v bien déterminée selon Ox a une quantité de mouvement mv donc un vecteur d'onde bien défini dans la direction Ox : $k_x = mv/\hbar$. La fonction d'onde associée est par conséquent une onde plane qui s'écrit, d'après (14.4) :

$$\psi_v(x) = \psi_0 \exp\left(i\frac{mvx}{\hbar}\right) \quad (14.31)$$

où ψ_0 est un facteur numérique sans signification physique.

Le vecteur d'état correspondant s'écrit :

$$|\Psi_v\rangle = \psi_0 \int_{-\infty}^{\infty} dx \exp\left(i\frac{mvx}{\hbar}\right) |x\rangle \quad (14.32)$$

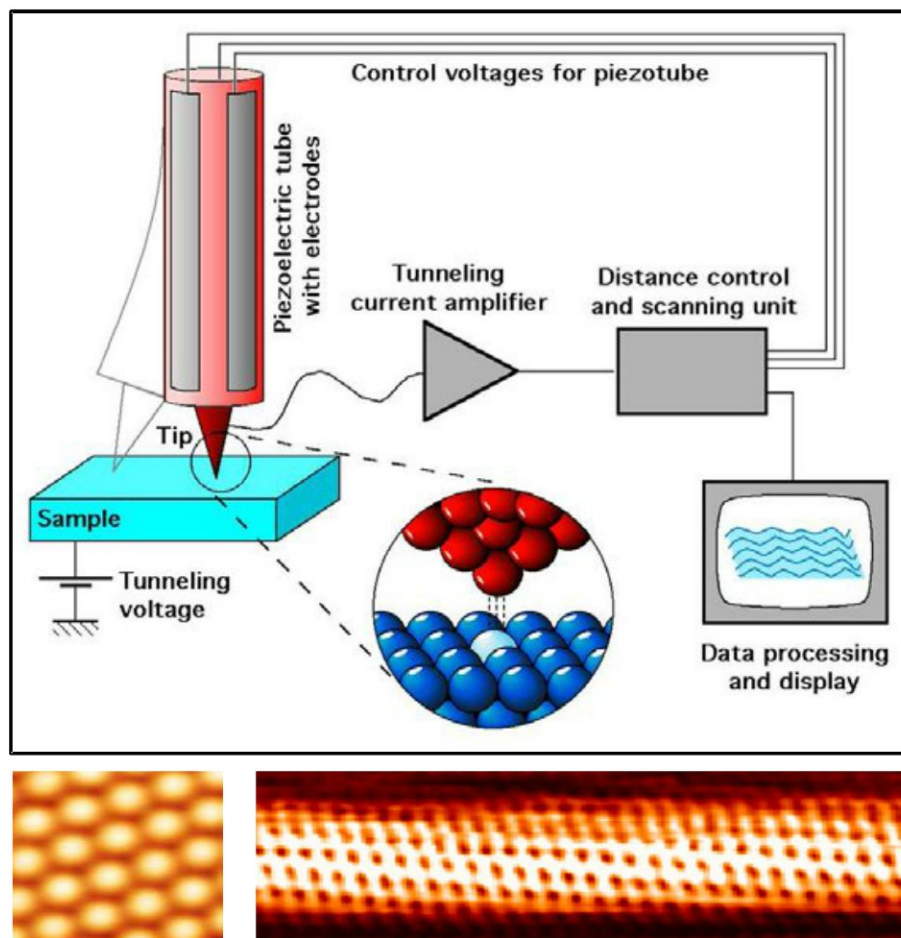


Figure 14.5 – L'une des plus importantes applications de l'effet tunnel est le *microscope à effet tunnel* dans lequel on déplace une pointe très fine près de la surface d'un échantillon conducteur. Les électrons peuvent alors passer, par effet tunnel, d'un conducteur à l'autre, c'est-à-dire de la pointe à l'échantillon, créant ainsi un courant macroscopique qui dépend de façon exponentielle, donc très sensible, de la distance pointe-échantillon. En ajustant en temps réel la hauteur de la pointe pour maintenir un courant constant et en enregistrant cette hauteur, on peut alors reconstituer le profil d'une surface. On est ainsi capable aujourd'hui de réaliser des « images » de surfaces à la précision d'« un » atome comme illustré sur les deux images expérimentales : à gauche, des atomes de silicium observés à la surface d'un cristal de carbure de silicium, et à droite, un nanotube de carbone.

(source libre : http://en.wikipedia.org/wiki/Scanning_tunneling_microscope)

14.8.2 Probabilité de mesure de la vitesse sur un état quelconque

Appliquons ici ce que l'on sait d'une mesure quantique : pour un système dans un état quelconque $|\Psi\rangle$, la mesure de la vitesse n'est certain que si $|\Psi\rangle$ coïncide avec un état $|\Psi_v\rangle$. Dans le cas général, la (densité de) probabilité $\tilde{P}(v)$ de trouver la vitesse v

14.8. Mesure de la vitesse de la particule, notion d'observable

vaut :

$$\tilde{P}(v) = |\langle \Psi_v | \Psi \rangle|^2 = \left| \int_{-\infty}^{\infty} dx \psi_v^*(x) \psi(x) \right|^2 \quad (14.33)$$

Soit, plus explicitement :

$$\tilde{P}(v) = \frac{m}{2\pi\hbar} \left| \int_{-\infty}^{\infty} dx \exp(-i\frac{mvx}{\hbar}) \psi(x) \right|^2 \quad (14.34)$$

où le facteur numérique $m/2\pi\hbar$ permet d'obtenir :

$$\int_{-\infty}^{\infty} dv \tilde{P}(v) = 1 \quad (14.35)$$

signifiant que la probabilité de trouver n'importe quelle valeur de v dans une mesure de vitesse est égale 1.

Remarquons que le résultat (14.34) est extrêmement surprenant : c'est la même quantité mathématique complexe, $\psi(x)$, qui contient les prédictions physiques concernant la position *et* la vitesse de la particule, alors qu'en physique classique on peut fixer de manière totalement indépendante position et vitesse d'un point matériel. Cette propriété implique que mesures de position et mesures de vitesse ne sont plus totalement indépendantes. Elles sont en fait liées par l'inégalité de Heisenberg que nous avons vue au chapitre 13.

La transformation mathématique qui permet de passer d'une fonction $f(x)$ à la fonction $\tilde{f}(k) = \int dx e^{-ikx} f(x)$ est bien connue des mathématiciens et des physiciens qui l'utilisent dans de nombreux domaines : il s'agit de la *transformée de Fourier*, dont la transformation inverse s'écrit $f(x) = \int dk e^{+ikx} \tilde{f}(k)/2\pi$. L'expression (14.34) nous indique donc que la probabilité de mesurer une vitesse v est proportionnelle au carré de la transformée de Fourier $\tilde{\psi}(k)$ de la fonction d'onde $\psi(x)$, avec $k = mv/\hbar$.

Il est important de noter que la connaissance des coefficients $\tilde{\psi}(k)$ est équivalente à la connaissance de la fonction d'onde $\psi(x)$. Ces *représentations* ne sont en fait que deux exemples parmi l'infinité de représentations possibles du vecteur d'état $|\psi\rangle$, et contiennent exactement la même information sur l'état du système. Ainsi, la fonction d'onde $\psi(x)$ apparaît comme une représentation particulière (appelée *représentation position*, puisqu'elle fait intervenir la variable position x) du vecteur d'état $|\psi\rangle$:

$$|\psi\rangle = \int_{-\infty}^{+\infty} \psi(x) |x\rangle dx \quad (14.36)$$

dans la base $\{|x\rangle\}$, où $|x\rangle$ est le vecteur d'état d'un système physique dont la position est détectée à coup sûr en la position x à dx près (c'est-à-dire par un détecteur de taille infinitésimale dx).

S'affranchir d'une représentation particulière en travaillant en termes de vecteurs d'état plutôt qu'en termes de fonctions d'onde est très important puisque cela permet de traiter des problèmes où la notion même de fonction d'onde n'est plus pertinente (comme dans le cas de la polarisation des photons du chapitre 11, ou celui du spin des atomes d'argent du chapitre 12).

14.8.3 Exemple de la boîte quantique

Connaissant la fonction d'onde d'un état donné du système, on en déduit par transformée de Fourier la probabilité de mesurer la vitesse. On peut le faire, par exemple, pour les trois premiers niveaux de la boîte quantique du paragraphe 14.5 (avec $k = mv/\hbar$). On trouve :

$$\tilde{P}_1(v) = \frac{2\pi L}{\hbar} \frac{\cos^2(kL/2)}{(k^2 L^2 - \pi^2)^2}; \quad \tilde{P}_2(v) = \frac{8\pi L}{\hbar} \frac{\sin^2(kL/2)}{(k^2 L^2 - 4\pi^2)^2}; \quad \tilde{P}_3(v) = \frac{18\pi L}{\hbar} \frac{\cos^2(kL/2)}{(k^2 L^2 - 9\pi^2)^2} \quad (14.37)$$

On constate, sur la figure 14.6, que dans l'état fondamental la vitesse la plus probable est 0, qui est la vitesse de l'état classique de plus basse énergie, mais que la vitesse peut prendre des valeurs positives ou négatives notablement différentes de zéro, de l'ordre de \hbar/mL . Dans les états $n = 2$ et $n = 3$ de plus grande énergie, on a deux pics centrés plus ou moins sur la vitesse de la particule classique de même énergie, mais toujours avec une dispersion notable des résultats de mesure autour de la valeur la plus probable. Pour de fortes valeurs de n , ces deux pics sont très espacés et on tend vers la situation classique où la vitesse de la particule prend seulement deux valeurs opposées (figure 14.7).

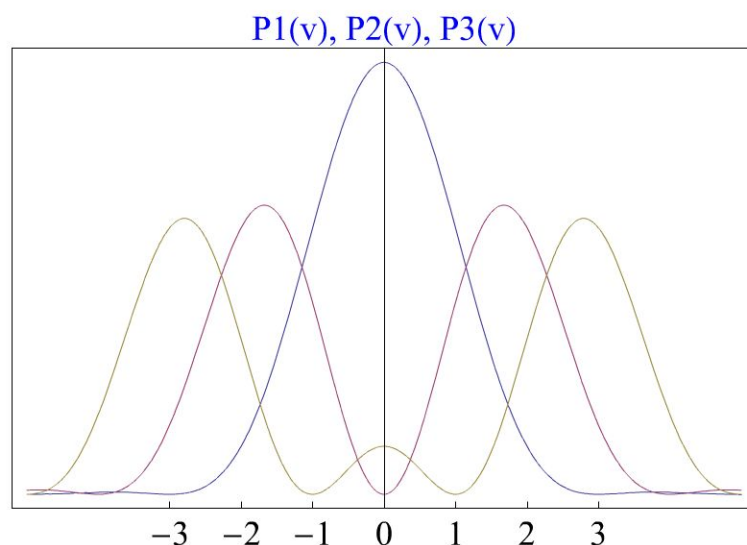


Figure 14.6 - Densités de probabilité de vitesse $P_1(v)$, $P_2(v)$, $P_3(v)$ du puits carré infini. L'axe des abscisses est en unités $\hbar/2mL$.

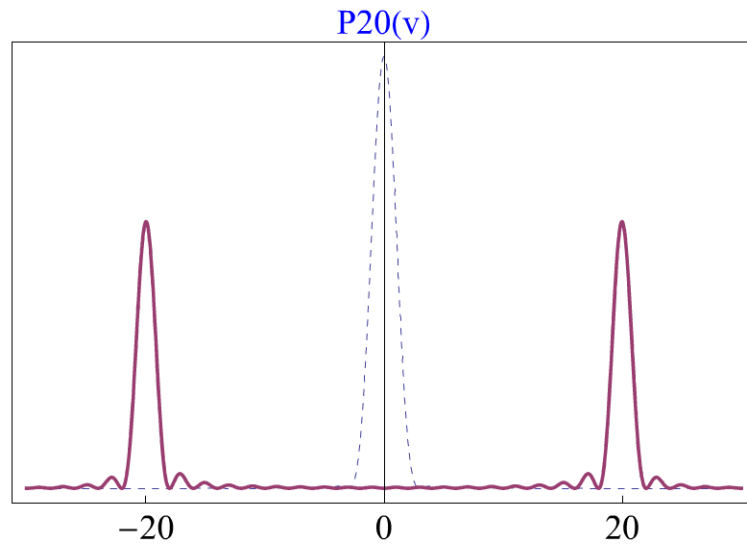


Figure 14.7 - Densité de probabilité de vitesse $P_{20}(v)$ du puits carré infini. L'axe des abscisses est en unités $\hbar/2mL$, et la distribution $P_1(v)$ figure en pointillés.

14.8.4 Valeurs moyennes et écarts-types de la quantité de mouvement

À partir de l'équation (14.34), on peut calculer les valeurs moyennes de n'importe quelle puissance de v :

$$\langle v^n \rangle = \int_{-\infty}^{\infty} v^n \tilde{P}(v) dv \quad (14.38)$$

et donc l'écart-type Δv estimant la largeur de la distribution de v :

$$\Delta v = \sqrt{\langle v^2 \rangle - \langle v \rangle^2} \quad (14.39)$$

Par exemple, dans le cas de la boîte quantique de largeur L on obtient les écarts-types :

$$\Delta x = \frac{L}{2\sqrt{3}} \sqrt{1 - \frac{6}{n^2\pi^2}} \quad \text{et} \quad \Delta p_x = n \frac{\hbar}{2L} \quad (14.40)$$

où Δx et Δp_x représentent les largeurs caractéristiques des distributions de probabilité $P_n(x)$ et $\tilde{P}_n(v)$.

Le produit $\Delta x \Delta p_x = \frac{\hbar}{2} \sqrt{\frac{n^2\pi^2}{3} - 2}$ de ces largeurs augmente avec n . On vérifie alors que sa valeur minimale, obtenue pour $n = 1$, est bien compatible avec le principe d'indétermination de Heisenberg, c'est-à-dire qu'elle est supérieure à $\hbar/2$:

$$n = 1 : \quad \Delta x \Delta p_x = \frac{\hbar}{2} \sqrt{\frac{\pi^2}{3} - 2} \simeq 0,57\hbar > \frac{\hbar}{2} \quad (14.41)$$



14.1 Notion d'observable

On remarque que la fonction d'onde $\psi_v(x)$ est solution de l'équation :

$$-i\hbar \frac{\partial}{\partial x} \psi_v(x) = mv \psi_v(x) \quad (14.42)$$

c'est-à-dire, en langage mathématique, que $\psi_v(x)$ est une *fonction propre* de l'opérateur $-i\hbar \frac{\partial}{\partial x}$, la *valeur propre* associée étant la quantité de mouvement $p = mv$.

Ce n'est pas fortuit, car on peut montrer, et c'est l'un des principes fondamentaux de la mécanique quantique, que les seuls résultats possibles d'une mesure sont en fait les valeurs propres d'opérateurs particuliers, appelés des *observables*, que l'on représente habituellement par des lettres surmontées d'un « chapeau ».

On peut ainsi définir l'opérateur quantité de mouvement suivant la direction Ox : \hat{p}_x , dont l'effet sur une fonction d'onde est donné par la relation :

$$\hat{p}_x (\psi(x, t)) = -i\hbar \frac{\partial}{\partial x} \psi(x, t) \quad (14.43)$$

L'opérateur position \hat{x} est quant à lui défini par la relation :

$$\hat{x} (\psi(x, t)) = x \cdot \psi(x, t) \quad (14.44)$$

où la quantité x désigne la variable x et non pas l'opérateur \hat{x} . L'action de l'opérateur position sur une fonction d'onde est donc très simple : elle consiste uniquement à multiplier la fonction d'onde par la variable position!²

Par exemple, si l'on considère l'expression non-relativiste de l'énergie totale $E = p_x^2/2m + V(x)$ (à une dimension), alors on peut lui associer une observable : l'opérateur Hamiltonien \hat{H} , que l'on obtient en remplaçant les variables classiques x et p_x par les opérateurs \hat{x} et \hat{p}_x dans l'expression classique de E ³ :

$$\hat{H} = \frac{\hat{p}_x^2}{2m} + V(\hat{x}) \quad (14.45)$$

L'action de \hat{H} sur une fonction d'onde quelconque redonne alors bien le membre de droite de l'équation (14.7) :

$$\hat{H} (\psi(x, t)) = -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} \psi(x, t) + V(x) \psi(x, t) \quad (14.46)$$

2. D'autres opérateurs peuvent alors être déduits de ces deux définitions (certaines grandeurs, en revanche, ne peuvent être déduites de \hat{x} et \hat{p}_x , c'est le cas des opérateurs de spin ou de polarisation du photon par exemple).

3. Le principe des opérateurs associés est en fait plus subtil et consiste à associer aux variables classiques des opérateurs correctement symétrisés. Par exemple, à la grandeur classique $x p_x$ est associé l'opérateur $(\hat{x} \hat{p}_x + \hat{p}_x \hat{x})/2$.

14.8. Mesure de la vitesse de la particule, notion d'observable

L'introduction de l'opérateur Hamiltonien permet finalement d'exprimer l'équation de Schrödinger de façon très générale, c'est-à-dire en termes de vecteur d'état :

$$i\hbar \frac{d}{dt} |\psi(t)\rangle = \hat{H} |\psi(t)\rangle \quad (14.47)$$

où \hat{H} est l'opérateur Hamiltonien du système physique décrit par le vecteur d'état $|\psi(t)\rangle$, et dont nous avons vu une expression matricielle à la section [12.4](#).

LES RÈGLES GÉNÉRALES DE LA MÉCANIQUE QUANTIQUE

15

Dans les chapitres précédents, nous nous sommes intéressés à la position du photon et à sa polarisation, puis au moment magnétique de spin d'un atome, et enfin à la position, à l'énergie et à la vitesse d'une particule massive, qu'elle soit libre ou confinée. Il est temps de conclure cet ouvrage par une revue plus générale des propriétés des objets quantiques, et d'énoncer sous forme de postulats les règles qui sont susceptibles de servir lorsqu'on traite n'importe quel système microscopique.

15.1 À LA BASE DE TOUT : LE VECTEUR D'ÉTAT

15.1.1 Énoncé

Postulat 1a

Tout système quantique, aussi complexe soit-il, est complètement décrit par la donnée d'un seul objet, le vecteur d'état, noté $|\Psi\rangle$, élément d'un espace de Hilbert.

Rappelons qu'un espace de Hilbert est un espace vectoriel à coefficients complexes muni d'un produit vectoriel complexe. Le produit scalaire complexe de deux vecteurs $|\Psi_1\rangle$ et $|\Psi_2\rangle$ est noté $\langle\Psi_1|\Psi_2\rangle$. Il est tel que :

$$\langle\Psi_2|\Psi_1\rangle = \langle\Psi_1|\Psi_2\rangle^* \quad (15.1)$$

Il existe, en outre, une contrainte et une liberté pour les vecteurs d'état décrivant un objet physique :

Postulat 1b

Pour tout vecteur d'état $|\Psi\rangle$ décrivant un objet physique : $\langle\P|\Psi\rangle = 1$.
 $|\Psi\rangle$ et $e^{i\theta}|\Psi\rangle$ décrivent le même état physique quel que soit θ .

$|\Psi\rangle$ est donc normé à 1 et est défini à un facteur de phase près.

Notons la différence importante de description entre un système classique, où il est nécessaire de connaître de nombreux paramètres différents du système si on veut le décrire complètement (position, vitesse, moment magnétique, valeur d'une onde en tout point...) et un système quantique, où le vecteur d'état d'un système physique est un objet unique qui contient *toutes* les informations mesurables sur ce système.

15.1.2 Conséquences immédiates

L'appartenance à un espace vectoriel semble une propriété technique de l'état. En fait, elle est un des aspects les plus fondamentaux de la mécanique quantique, car elle a de nombreuses conséquences. Propriété commune à tous les phénomènes ondulatoires, elle implique en particulier le « principe de superposition » :

Si $|\Psi_1\rangle$ et $|\Psi_2\rangle$ sont deux vecteurs d'état d'un même système quantique, toute combinaison linéaire $\alpha_1|\Psi_1\rangle + \alpha_2|\Psi_2\rangle$ normée à 1 de ces deux vecteurs d'état est aussi un état possible du système.

La linéarité de la mécanique quantique implique en outre le *théorème de non-clonage*, abordé au chapitre 12 :

Il n'est pas possible de dupliquer un état quantique inconnu, c'est-à-dire produire à partir de cet état deux états « copies » strictement identiques à cet état entrant.

15.1.3 Décomposition sur une base de l'espace de Hilbert

Une caractéristique importante d'un système physique est la dimension N de l'espace de Hilbert associé, c'est-à-dire le nombre de coefficients complexes qu'il faut connaître pour déterminer exactement le vecteur d'état. Celle-ci peut être finie, comme dans le cas de la polarisation du photon ou du moment magnétique de spin d'un atome (dans ces deux cas $N = 2$), ou infinie dénombrable, comme dans le cas de la boîte quantique. Elle peut même être infinie avec la puissance du continu comme dans le cas d'une particule libre. Nous allons considérer dans ce paragraphe seulement le cas dénombrable.

Il existe une infinité de bases différentes pour représenter un élément d'un espace vectoriel. Les bases les plus adaptées à la description d'un état quantique sont les bases orthonormées $\{|u_i\rangle\}$ ($i = 1, N$), telles que :

$$\langle u_i | u_j \rangle = \delta_{ij} \quad \text{avec} \quad \delta_{ii} = 1 \quad \text{et} \quad \delta_{ij} = 0 \quad \text{pour} \quad i \neq j \quad (15.2)$$

Tout vecteur d'état quantique $|\Psi\rangle$ peut se décomposer sur cette base :

$$|\Psi\rangle = \sum_{i=1}^N c_i |u_i\rangle \quad \text{avec} \quad c_i = \langle u_i | \Psi \rangle \quad \text{et} \quad \sum_{i=1}^N |c_i|^2 = 1 \quad (15.3)$$

Les coefficients $\{c_i\}_{i=1,\dots,N}$ sont les coordonnées du vecteur $|\Psi\rangle$ dans la base $\{|u_i\rangle\}$. La base orthonormée permet d'écrire simplement le produit scalaire complexe de deux vecteurs d'état $|\Psi\rangle$ et $|\Phi\rangle$, de coordonnées respectives $\{c_i\}_{i=1,\dots,N}$ et $\{d_i\}_{i=1,\dots,N}$:

$$\langle\Psi|\Phi\rangle = \sum_{i=1}^N c_i^* d_i \quad (15.4)$$

15.1.4 Cas particulier : la fonction d'onde

Considérons l'équivalent quantique d'un point matériel de masse m pouvant se déplacer uniquement le long de l'axe Ox . Nous avons vu dans le chapitre 14 que, comme on peut mesurer *a priori* sa position en n'importe quel point x de cet axe, l'espace de Hilbert associé est de dimension infinie continue. Une base possible est celle des états $\{|x\rangle\}$ ($x \in]-\infty, +\infty[$) correspondant à des états parfaitement localisés au point x , et qui vérifient :

$$\langle x|x'\rangle = 0 \quad \text{pour } x \neq x' \quad (15.5)$$

Le carré scalaire $\langle x|x\rangle$ de ces vecteurs est infini. Ils ne sont donc pas associés à des états physiques mais représentent des intermédiaires de calcul commodes.

Tout vecteur d'état quantique $|\Psi\rangle$ de la particule peut se décomposer sur cette base, non sous la forme d'une somme discrète, mais d'une intégrale :

$$|\Psi\rangle = \int_{-\infty}^{+\infty} dx \psi(x) |x\rangle \quad \text{avec} \quad \psi(x) = \langle x|\Psi\rangle \quad \text{et} \quad \int_{-\infty}^{+\infty} dx |\psi(x)|^2 = 1 \quad (15.6)$$

où $\psi(x)$ est la *fonction d'onde*, fonction complexe de carré sommable de la variable x dont l'ensemble des valeurs constitue les coordonnées du vecteur $|\Psi\rangle$ dans la base $\{|x\rangle\}$. Cette base permet d'écrire simplement le produit scalaire complexe de deux vecteurs d'état $|\Psi\rangle$ et $|\Phi\rangle$, de fonctions d'onde respectives $\psi(x)$ et $\phi(x)$:

$$\langle\Psi|\Phi\rangle = \int_{-\infty}^{+\infty} dx \psi^*(x) \phi(x) \quad (15.7)$$

Comme le vecteur d'état, la fonction d'onde est définie à un facteur de phase près. C'est une fonction continue en tout point, et sa dérivée spatiale est continue partout, sauf en un nombre fini de points isolés.

Les fonctions d'onde obéissent également à un *principe de superposition* : si $\psi_1(x)$ et $\psi_2(x)$ sont deux fonctions d'onde possibles pour un système, alors toute combinaison linéaire (à coefficients complexes, normalisée à un) de ces fonctions d'onde est aussi une fonction d'onde possible pour ce système. Cette possibilité de superposition rend possible l'existence d'*interférences destructives entre fonctions d'onde de phases opposées* analogues aux interférences destructives pour l'onde lumineuse.

On peut évidemment utiliser d'autres bases que la base $\{|x\rangle\}$, par exemple une base plus adaptée à la mesure de vitesses. Ces différentes *représentations* contiennent exactement la même information sur l'état du système. On peut facilement étendre la discussion précédente à une fonction d'onde $\psi(x, y, z)$ définie dans l'espace à trois dimensions des positions de la particule.

Précisons enfin que tous les systèmes quantiques ne sont pas descriptibles en termes de fonction d'onde : c'est le cas par exemple d'une particule matérielle ayant un spin. Bien que pratique, car elle ramène l'onde quantique à la notion plus familière qu'est le « champ » $\psi(x, t)$ ayant une valeur définie en tout point de l'espace, la notion de fonction d'onde est donc moins fondamentale que celle de vecteur d'état.

15.2 ÉVOLUTION TEMPORELLE

Au cours du temps, le vecteur d'état change en général, et il faut donc le noter $|\psi(t)\rangle$. Que ce soit en physique quantique ou classique, l'évolution dans le temps de l'état d'un système est reliée à l'expression de son énergie. Nous avons vu qu'à cette énergie, grandeur mesurable, est associée une observable, notée \hat{H} , appelée *Hamiltonien*.

15.2.1 Énoncé

Postulat 2

L'évolution du vecteur d'état $|\psi(t)\rangle$ décrivant un système physique est donnée par l'équation de Schrödinger

$$i\hbar \frac{d}{dt} |\psi(t)\rangle = \hat{H} |\psi(t)\rangle \quad (15.8)$$

où \hat{H} est l'Hamiltonien du système, qui est l'observable associée à l'énergie du système.

15.2.2 Conséquences

Précisons que cette loi d'évolution ne s'applique que pour un système quantique soumis à des contraintes fixes, comme l'effet d'une force dérivant d'un potentiel fixe. Il ne s'applique pas lorsque le système est soumis à l'interaction avec un appareil de mesure, dont l'état est par nature modifié par son interaction avec la particule quantique.

Il faut insister sur le fait que si la mécanique quantique est de nature probabiliste pour les résultats de mesure sur un système individuel à un instant donné l'évolution

15.3. La mesure, lien entre le système physique et l'observateur

du vecteur d'état est, elle, parfaitement *déterministe*. En effet, si l'on connaît le vecteur d'état du système à un instant donné $t = 0$, ainsi que l'Hamiltonien du système, alors l'équation de Schrödinger permet de déterminer le vecteur d'état du système à tout instant.

Nous avons vu au chapitre 12 que si on connaît les états $|\varphi_n\rangle$ d'énergie bien définie E_n du système, la résolution de l'équation de Schrödinger aboutit à l'expression explicite suivante du vecteur d'état au temps t :

$$|\psi(t)\rangle = \sum_n c_n e^{-iE_n t/\hbar} |\varphi_n\rangle \quad (15.9)$$

En termes de fonctions d'onde, nous avons vu que dans le cas d'un mouvement à une dimension le long de la direction x l'équation de Schrödinger d'une particule de masse m soumise à un potentiel $V(x)$ s'écrit :

$$i\hbar \frac{\partial}{\partial t} \psi(x,t) = -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} \psi(x,t) + V(x)\psi(x,t) \quad (15.10)$$

On peut aussi l'écrire sous la forme du postulat 2 :

$$i\hbar \frac{\partial}{\partial t} \psi(x,t) = \hat{H} \psi(x,t) \quad (15.11)$$

Dans cette dernière écriture, l'opérateur Hamiltonien \hat{H} est un opérateur linéaire agissant sur l'espace des fonctions :

$$\hat{H} = -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} + V(x) = -\frac{\hat{p}_x^2}{2m} + V(x) \quad (15.12)$$

où $V(x)$ est l'opérateur qui multiplie la fonction par $V(x)$, et \hat{p}_x est l'opérateur quantité de mouvement introduit à l'encart 14.1, égal à l'opérateur différentiel $-i\hbar \frac{\partial}{\partial x}$.

15.3 LA MESURE, LIEN ENTRE LE SYSTÈME PHYSIQUE ET L'OBSERVATEUR

La mesure d'une grandeur physique quelconque, que nous appellerons de manière générale A , comme la position, la vitesse, l'énergie, le moment magnétique, le courant électrique... présente des aspects spécifiques – dès lors qu'on s'intéresse au monde quantique – qui sont liés à l'aspect inévitablement aléatoire des résultats de mesure sur les objets quantiques individuels.

15.3.1 Postulats relatifs à la mesure

Postulat 3a

Il existe des états quantiques spécifiques, notés $|\phi_a\rangle$, pour lesquels la mesure de la grandeur physique A donne avec certitude le résultat réel a . L'ensemble des vecteurs $\{|\phi_a\rangle\}$ forme une base orthonormée de l'espace de Hilbert des états du système.

Dans la suite, nous supposons, pour simplifier les notations, que la liste des valeurs de a possibles forme un ensemble discret ($a_i, i = 1, 2, \dots$), et aussi qu'il n'y a qu'un seul état par valeur a_i que nous noterons $|\phi_i\rangle$ ¹.

Postulat 3b

Lorsque le système physique est dans un état décrit par un vecteur d'état $|\psi\rangle$ quelconque, la mesure ne donne en général pas un résultat certain. La probabilité $Proba(a_i, |\psi\rangle)$ d'obtenir le résultat a_i est donnée par :

$$Proba(a_i, |\psi\rangle) = |\langle\phi_i|\psi\rangle|^2$$

Il existe des cas où la mesure ne détruit pas le système mesuré. Dans ce cas on peut s'intéresser à l'état du système après la mesure

Postulat 3c

Dans les cas où la mesure a donné effectivement le résultat a_{i_0} , le système après la mesure est dans un état conditionnel (restreint au sous-ensemble des cas où on a mesuré a_{i_0}) qui est égal à l'état $|\phi_{i_0}\rangle$, et cela indépendamment de l'état du système avant la mesure.

15.3.2 Caractéristiques spécifiques de la mesure en mécanique quantique

Selon les situations, les valeurs possibles d'une mesure en mécanique quantique peuvent être très différentes :

- elles sont quantifiées et forment un ensemble discret dans le cas où A est l'énergie d'un système lié, l'état de polarisation d'un photon, la projection du moment magnétique d'un atome sur une direction donnée ;

1. On dit, dans ce cas, que les valeurs propres a_i sont *non-dégénérées*.

15.3. La mesure, lien entre le système physique et l'observateur

- elles peuvent prendre toutes les valeurs réelles dans d'autres cas, par exemple lorsque A est la position d'une particule, le champ électrique d'une onde ;
- le cas intermédiaire peut aussi se présenter, dans lequel les valeurs possibles peuvent former un (ou plusieurs) intervalles déterminés $[a_{min}, a_{max}]$ de valeurs réelles (voir les notions de bandes de conduction et de valence introduites au chapitre 10).

Le caractère intrinsèquement aléatoire des mesures effectuées sur un système quantique individuel est l'un des aspects les plus fondamentaux, mais aussi les plus inhabituels, du monde quantique : lorsque le système se trouve dans un état quantique quelconque, différent des états $|\phi_i\rangle$ introduits par le postulat (3a), le résultat d'une mesure unique de la grandeur A n'est en général pas prévisible. Il s'agit donc d'une limitation fondamentale au pouvoir prédictif de la physique. Fort heureusement, si on répète un grand nombre de fois la mesure sur un système toujours préparé de la même manière, ou si la mesure porte simultanément sur un très grand nombre de systèmes individuels, la probabilité donnée par le postulat (3b), et donc les moyennes statistiques effectuées sur ces résultats, sont parfaitement déterminées. On retrouve à ce niveau, plus « macroscopique », le déterminisme de la théorie physique.

Pour mesurer une grandeur sur un objet du monde quantique il faut franchir les huit ordres de grandeur mentionnés dans le paragraphe d'introduction au monde quantique qui nous séparent de ce monde. C'est le rôle de l'instrument de mesure, par exemple un compteur de photons ou d'électrons. Il s'agit donc d'un système macroscopique qui est susceptible de perturber fortement l'objet mesuré. Qu'un appareil de mesure perturbe l'objet mesuré est une caractéristique bien connue en physique classique : l'insertion d'un ampèremètre dans un circuit électrique modifie le courant qui y circule, ou bien l'insertion d'un thermomètre modifie la température à mesurer. Mais cette modification peut être calculée, et le résultat de mesure corrigé en conséquence.

Il n'en va pas de même au niveau quantique, où *la mesure perturbe l'objet microscopique d'une manière incontrôlable et non prévisible au niveau de la mesure individuelle*. En revanche, dès lors qu'on connaît le résultat de la mesure sans que le système ne soit détruit par cette mesure, la perturbation de l'appareil de mesure aboutit à un résultat étonnamment simple et universel, puisque, d'après le postulat (3c) l'état après la mesure est l'état $|\phi_{i_0}\rangle$, indépendamment de l'état avant la mesure. Ce postulat est appelé parfois *principe de réduction du vecteur d'état*. Insistons sur le fait que si on répète alors la mesure, on aura à coup sûr la même réponse que celle obtenue lors de la première mesure.

15.3.3 « Éléments propres » d'une mesure, observable

À toute grandeur physique A mesurable sont donc attachés la liste des résultats possibles a_i et les états associés $|\phi_i\rangle$. Il est alors commode d'introduire un opérateur linéaire, noté \hat{A} , agissant sur l'espace de Hilbert des états du système, et défini par son action sur la base des vecteurs $|\phi_i\rangle$:

$$\hat{A}|\phi_i\rangle = a_i|\phi_i\rangle \quad (15.13)$$

\hat{A} est donc une matrice diagonale sur la base $\{|\phi_i\rangle\}$. Les a_i sont les *valeurs propres* de l'opérateur \hat{A} , et les vecteurs d'état $|\phi_i\rangle$ sont les *vecteurs propres* associés. L'opérateur Hamiltonien que nous avons introduit à la section 12.4, est donc l'observable associée à la grandeur physique « énergie du système ».

L'introduction de l'observable permet en particulier de donner une expression simple à la valeur moyenne, ou espérance mathématique $E_{|\Psi\rangle}(A)$, notée en mécanique quantique $\langle\hat{A}\rangle$, de la mesure de A lorsqu'elle est effectuée un grand nombre de fois sur le système préparé à chaque fois dans un état $|\Psi\rangle$ quelconque donné. On peut montrer qu'elle est égale au produit scalaire des vecteurs d'état $\hat{A}|\Psi\rangle$ et $|\Psi\rangle$, que l'on note habituellement $\langle\Psi|\hat{A}|\Psi\rangle$.

$$\langle\hat{A}\rangle = E_{|\Psi\rangle}(A) = \sum_i a_i \text{Proba}(a_i, |\Psi\rangle) = \langle\Psi|(\hat{A}|\Psi\rangle) \quad (15.14)$$

Dans le cas de l'espace de Hilbert des fonctions d'onde, les opérateurs linéaires sont définis par leur action sur une fonction d'onde quelconque. Ils peuvent, par exemple, faire intervenir des dérivées de la fonction considérée. Par exemple, on définit l'opérateur position par :

$$\hat{x}(\psi(x,t)) = x.\psi(x,t) \quad (15.15)$$

Quant à la quantité de mouvement dans la direction x , l'opérateur associé s'écrit :

$$\hat{p}_x(\psi(x,t)) = -i\hbar \frac{\partial}{\partial x}(\psi(x,t)) \quad (15.16)$$

où la quantité x désigne la variable x et non pas l'opérateur \hat{x} . En effet, l'état donnant une valeur certaine à la quantité de mouvement p introduit en 14.1, qui a pour fonction d'onde $\psi_p(x) = \psi_0 e^{ipx/\hbar}$ est état propre de cet opérateur avec la valeur propre p :

$$-i\hbar \frac{\partial}{\partial x} \psi_p(x) = p\psi_p(x) \quad (15.17)$$

La valeur moyenne de la position de la particule, par exemple, s'écrit alors, en utilisant l'expression du produit scalaire pour la fonction d'onde :

$$E_{|\Psi\rangle}(x) = \langle\Psi|(\hat{x}|\Psi\rangle) = \int_{-\infty}^{+\infty} \psi^*(x)x.\psi(x) dx = \int_{-\infty}^{+\infty} x|\psi(x)|^2 dx \quad (15.18)$$

15.3.4 Mesures successives de grandeurs physiques

Nous avons vu en section 13.3 que l'inégalité de Heisenberg contraignait la connaissance que l'on pouvait avoir de deux variables complémentaires, comme la position et la vitesse d'une particule. Nous allons voir ici ce qu'il advient lorsque l'on essaye néanmoins de mesurer, le plus précisément possible, ces quantités. En particulier, des caractéristiques nouvelles de la mécanique quantique se manifestent lorsqu'on considère le problème de *mesures successives sur le même système*.

Nous supposons donc que la première mesure effectuée ne détruit pas le système mesuré, de sorte qu'on peut appliquer le postulat (3c), et que l'état du système n'évolue pas entre les deux mesures.

Envisageons d'abord le cas où on effectue deux fois la mesure de la grandeur A : si la première a donné le résultat a_{i_0} , le système se retrouve dans l'état $|\phi_{i_0}\rangle$, de sorte que la deuxième mesure donnera à coup sûr le même résultat. La fonction perturbatrice de la mesure ne se produit qu'à la première mesure, qui sert donc de « préparation » du système pour des mesures ultérieures qui ne sont plus entachées d'erreurs.

Considérons maintenant le cas de mesures successives de grandeurs physiques A et B différentes. La première mesure transforme l'état initial en l'état propre $|\phi_{i_0}\rangle$ associé à la valeur mesurée a_{i_0} . Il y a alors deux situations possibles :

- soit $|\phi_{i_0}\rangle$ est aussi vecteur propre de l'observable \hat{B} associée à la mesure de B et de valeur propre b_{i_0} , et la mesure de B donne à coup sûr le résultat b_{i_0} : les mesures sont dites *compatibles*, car elles ne se perturbent pas l'une l'autre ;
- soit $|\phi_{i_0}\rangle$ n'est pas vecteur propre de l'observable \hat{B} , et la mesure de B donne des résultats aléatoires : les mesures sont dites *incompatibles*. Il est facile de voir que dans ce cas, on n'obtient pas les mêmes probabilités de mesure lorsque l'on change l'ordre dans lequel ces mesures sont effectuées. On appelle *grandeurs conjuguées* de tels couples de grandeurs incompatibles.

Un exemple important est le couple position-quantité de mouvement (x, p_x) . En effet, une mesure de position d'une particule donnant la valeur x_0 transforme l'état initial en l'état $|x_0\rangle$, qui, comme nous l'avons vu au chapitre 14, n'est pas un état propre de la quantité de mouvement : la mesure de position s'accompagne donc de mesures ultérieures de la quantité de mouvement qui sont totalement aléatoires. Inversement, une mesure de quantité de mouvement donnant la valeur p_0 transforme l'état initial en l'état $|p_0\rangle$ de fonction d'onde $\psi_{p_0}(x) = \psi_0 e^{ip_0 x/\hbar}$, qui n'est pas un état propre de la position : la mesure de quantité de mouvement s'accompagne donc de mesures ultérieures de la position dont la probabilité $|\psi_{p_0}(x)|^2 = |\psi_0|^2$ est indépendante de x et qui sont totalement aléatoires. On ne peut donc pas, même avec des appareils de mesure idéaux, connaître à la fois la position d'une particule de façon infiniment précise (c'est-à-dire avec une indétermination nulle sur la position : $\Delta x = 0$) et en même temps sa quantité de mouvement de façon infiniment précise (c'est-à-dire avec une

indétermination nulle sur la quantité de mouvement : $\Delta p_x = 0$). Cette impossibilité condamne toute notion de *trajectoire* précise pour une particule en mécanique quantique, nous arrivons ainsi à la même conclusion qu'avec l'inégalité de Heisenberg.

Que se passe-t-il si on fait une mesure moins précise de position ? Si on mesure la position dans l'intervalle $[-\delta x_{\text{mesure}}/2, \delta x_{\text{mesure}}/2]$, la mesure projette le système dans l'un des états $|x_n\rangle$ définis dans l'encart 13.1 (figure 13.17). Une mesure sur cet état de la composante selon Ox de l'impulsion donne des résultats centrés sur $p_x = 0$ avec un écart-type $\delta p_{x,\text{perturb}}$ de l'ordre de $\hbar/\delta x_{\text{mesure}}$. On a alors : $\delta x_{\text{mesure}} \delta p_{x,\text{perturb}} \simeq 2\pi\hbar$

On peut montrer qu'il existe une *inégalité de Heisenberg-Ozawa* concernant l'écart type Δx_{mesure} , qui caractérise la précision d'une mesure de x , l'écart type $\Delta p_{x,\text{perturb}}$, qui caractérise la dispersion des valeurs de p_x de l'état du système perturbé par cette mesure, et les écarts types Δx et Δp_x qui caractérisent la dispersion des valeurs de x et p_x dans l'état initial du système, celles qui sont concernées par l'inégalité de Heisenberg « habituelle » (13.16). L'inégalité de Heisenberg-Ozawa s'écrit² :

$$\Delta x_{\text{mesure}} \Delta p_{x,\text{perturb}} + \Delta x \Delta p_{x,\text{perturb}} + \Delta x_{\text{mesure}} \Delta p_x \geq \hbar/2 \quad (15.19)$$

Cette relation empêche que l'on ait une mesure à la fois parfaite et non perturbatrice ($\Delta x_{\text{mesure}} = \Delta p_{x,\text{perturb}} = 0$). On peut, en revanche, avoir une mesure parfaite $\Delta x_{\text{mesure}} = 0$ avec une valeur de $\Delta p_{x,\text{perturb}}$ telle que $\Delta x \Delta p_{x,\text{perturb}} \geq \hbar/2$. Ainsi la relation (15.19) n'interdit pas que $\Delta x_{\text{mesure}} \Delta p_{x,\text{perturb}} < \hbar/2$, à condition que Δx et Δp_x soient suffisamment grands : plusieurs expériences récentes ont effectivement mis en évidence qu'on pouvait être dans ce régime, où l'effet perturbateur de la mesure est en quelque sorte « caché » dans la grande dispersion initiale des valeurs de x et de p_x . L'inégalité (15.19) ressemble à l'inégalité de Heisenberg et elle est souvent confondue avec cette dernière. Elle est cependant différente conceptuellement, car la relation (13.16) est une propriété de tout état quantique, découlant de la dualité onde-particule. Elle ne concerne pas l'effet de la mesure sur l'état et ne dépend pas du postulat 3c, alors que l'inégalité (15.19) en est une conséquence directe.

15.3.5 Une expérience de localisation spatiale d'une particule isolée

La figure 15.1 donne un exemple d'expérience de localisation d'une particule unique, en l'occurrence un atome (ou un ion) : elle consiste à exciter l'atome avec un faisceau laser se propageant le long de l'axe Ox . L'atome est alors porté dans un état excité d'énergie E_2 . Il retombe ensuite par *émission spontanée* dans son état fondamental d'énergie E_1 en émettant un photon de fréquence $\nu = (E_2 - E_1)/h$. Ce photon est émis dans une direction aléatoire. Le laser peut alors exciter à nouveau l'atome, et

2. M. Ozawa, *Physical Review* **67**, 042105, 2003

15.3. La mesure, lien entre le système physique et l'observateur

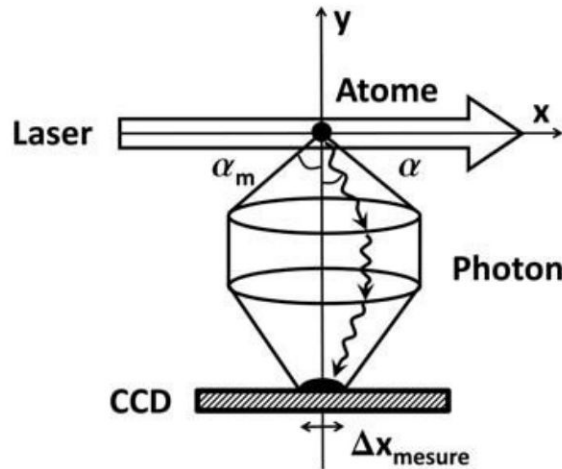


Figure 15.1 – Schéma de principe du *microscope de Heisenberg*. Le fonctionnement est détaillé dans le texte ci-contre.

le processus se répète, conduisant à l'émission de nombreux photons par la même particule isolée. L'atome est au foyer d'une lentille, qui permet de recueillir tous les photons émis dans une direction faisant un angle α avec Oy tel que $\alpha < \alpha_m$, où α_m dépend du diamètre de la lentille. Ces photons sont ensuite refocalisés par une deuxième lentille identique sur la surface d'un capteur CCD : on peut ainsi acquérir l'image de l'atome unique. La taille de cette image donne la précision Δx_{mesure} avec laquelle on peut localiser l'atome. Celle-ci est donnée par les lois de l'optique, et plus précisément celles de la diffraction :

$$\Delta x_{\text{mesure}} = \frac{\lambda}{\sin(\alpha_m)} \quad (15.20)$$

où $\sin(\alpha_m)$ est appelée *ouverture numérique* de la lentille. La précision de la localisation est donc de l'ordre de la longueur d'onde de la lumière émise, c'est-à-dire le micromètre, une valeur bien supérieure à la taille de l'atome, de l'ordre du nanomètre. Pour améliorer cette précision, il faudrait utiliser de la lumière de plus courte longueur d'onde. La figure II.2 de l'introduction donne un exemple d'image de particules isolées (en l'occurrence trois ions) enregistrée par le procédé que nous venons d'indiquer.

Cette mesure perturbe-t-elle l'ion unique ? Oui, car l'émission de chaque photon satisfait à la conservation de la quantité de mouvement totale du système « ion+photon », section 10.12.2) : l'atome recule, comme un canon, à chaque émission de photon. La variation de quantité de mouvement de l'atome $\delta \mathbf{p}$ est égale à l'opposé de la quantité de mouvement $\hbar \mathbf{k}$ du photon émis : sa vitesse est modifiée, d'une quantité faible mais non nulle, et surtout aléatoire donc imprévisible. En supposant que cette émission se fasse dans le plan xOy , on a :

$$\delta p_x = -\hbar k \sin \alpha \quad (15.21)$$

L'incertitude sur p_x induite par l'émission spontanée de photons, $\Delta p_{x,perturb}$, est donc de l'ordre de $\hbar k \sin \alpha_m$. On a alors :

$$\Delta x_{mesure} \Delta p_{x,perturb} \simeq \lambda \hbar k = 2\pi \hbar \quad (15.22)$$

en conformité avec l'inégalité de Heisenberg-Ozawa.

L'argument que nous venons de présenter est une version actualisée d'une célèbre expérience de pensée, dite du *microscope de Heisenberg*, destinée à montrer que le caractère incompressible et aléatoire de la perturbation apportée par une mesure est liée aux fondements de la théorie quantique : quantification des échanges d'énergie et de quantité de mouvement et caractère aléatoire des événements quantiques individuels.

15.4 PRINCIPE D'EXCLUSION DE PAULI

15.4.1 Notion de spin

Nous avons déjà abordé la notion de spin au chapitre 12 lors de la description de l'expérience de Stern et Gerlach. C'est une notion difficile à appréhender puisque le spin d'une particule correspond à un moment cinétique intrinsèque (c'est-à-dire n'étant pas dû au mouvement de cette particule) qui n'a pas de représentation ou d'équivalent classique.

Toute particule (élémentaire ou composite) possède un moment cinétique de spin, que l'on note généralement \vec{S} . C'est une observable (vectorielle) qui peut donc être mesurée. Les valeurs possibles de la projection de \vec{S} sur un axe donné sont en nombre fini : il y a $2S + 1$ valeurs possibles, de $-S\hbar$ à $+S\hbar$ par sauts d'une unité de \hbar . Le nombre S qui caractérise ce spin est nécessairement soit un nombre entier ($S = 0, 1, 2, \dots$), soit un nombre demi-entier ($S = 1/2, 3/2, 5/2, \dots$). Par exemple, les électrons, les protons, les neutrons et les neutrinos ont un spin $1/2$, alors que les photons ont un spin 1 et que le fameux boson de Higgs a un spin nul.

15.4.2 Particules indiscernables

Une des propriétés les plus frappantes du monde quantique est qu'à l'échelle microscopique aucune observation ne permet de distinguer deux particules ayant les mêmes propriétés physiques (comme la masse, la charge, le spin...). On parle d'*indiscernabilité*.

Par exemple, tous les électrons sont *identiques* et impossibles à distinguer par une quelconque expérience. On ne peut même pas espérer pister « un » électron particulier parmi d'autres en suivant sa trajectoire puisque celle-ci (position et vitesse à tout instant) n'est pas définie, conformément aux relations d'indétermination de Heisenberg.

La description quantique de deux particules 1 et 2 ne se fait pas en considérant l'union de deux vecteurs d'état $|\psi_1\rangle$, $|\psi_2\rangle$, mais avec un seul vecteur d'état $|\psi\rangle$ qui contient les informations sur les deux particules. Par exemple, dans le cas simple où les deux paramètres caractérisant la particule sont sa position sur l'axe Ox et la projection $m\hbar$ de son spin sur l'axe Oz , la quantité $|\langle x_1, m_1, x_2, m_2 | \psi \rangle|^2$ donne la probabilité de trouver la particule 1 en x_1 avec une projection $m_1\hbar$ et la particule 2 en x_2 avec une projection $m_2\hbar$.

Si les deux particules sont indiscernables, on ne peut pas leur attribuer de numéro, et donc cette probabilité est invariante par échange de ces numéros :

$$|\langle x_2, m_2, x_1, m_1 | \psi \rangle|^2 = |\langle x_1, m_1, x_2, m_2 | \psi \rangle|^2. \quad (15.23)$$

15.4.3 Bosons et fermions

L'indiscernabilité des particules microscopiques induit une différence majeure entre le comportement *collectif* de particules de spin entier et celui de particules de spin demi-entier.

La relation (15.23) implique que la quantité $\langle x_1, m_1, x_2, m_2 | \psi \rangle$ est invariante à un coefficient de module un près lors de l'échange des deux particules. On montre que ce coefficient est égal à $(-1)^{2S+1}$:

$$\langle x_2, m_2, x_1, m_1 | \psi \rangle = (-1)^{2S+1} \langle x_1, m_1, x_2, m_2 | \psi \rangle \quad (15.24)$$

c'est-à-dire qu'il vaut 1 pour les particules de spin entier, et -1 pour les particules de spin demi-entier.

Dit autrement, les particules de spin entier, que l'on appelle des **bosons**, sont décrites par un vecteur d'état *symétrique* sous l'échange de deux particules quelconques, alors que les particules de spin demi-entier, que l'on appelle des **fermions**, sont décrites par un vecteur d'état *anti-symétrique* sous l'échange de deux particules quelconques³.

15.4.4 Principe d'exclusion de Pauli

Considérons alors la probabilité de détecter deux fermions au même point x et avec la même projection $m\hbar$. On a :

$$\langle x_2 = x, m_2 = m, x_1 = x, m_1 = m | \psi \rangle = -\langle x_1 = x, m_1 = m, x_2 = x, m_2 = m | \psi \rangle \quad (15.25)$$

ce qui implique $\langle x, m, x, m | \psi \rangle = 0$: on ne peut pas détecter deux fermions au même endroit et avec la même valeur de m . Par contre, la démonstration ne s'applique pas à deux fermions détectés au même point, mais avec deux valeurs de m différentes. Ainsi, deux électrons (spin $S = 1/2$) dont les projections sont opposées ($m = \pm 1/2$) peuvent se trouver au même point.

3. Ce résultat porte le nom de *théorème spin-statistique*.

On montre plus généralement que *la probabilité que deux fermions soient dans le même état quantique est nulle*. Ce résultat très important est appelé « *principe d'exclusion de Pauli* » :

Deux fermions ne peuvent occuper simultanément le même état quantique.

Il joue un rôle majeur dans la physique des milieux condensés et permet d'expliquer la stabilité de la matière (des plus petits objets, comme les atomes, aux plus gros, comme les étoiles à neutrons), la classification périodique des éléments, ainsi que de nombreuses propriétés mécaniques, électriques, magnétiques, optiques et chimiques des solides.

Le principe de Pauli n'interdit pas, en revanche, que des bosons occupent simultanément le même état quantique. On peut, en fait, montrer que les bosons ont tendance à se regrouper dans le même état. Les photons d'une onde laser, les atomes d'hélium-4 superfluides, ou les atomes ultrafroids d'un gaz de Bose-Einstein en sont des exemples particulièrement frappants.

Notons, pour finir, que le principe d'exclusion de Pauli ne fait pas partie, à proprement parler, des postulats de la mécanique quantique puisqu'il peut être dérivé de la généralisation relativiste de celle-ci. Néanmoins, comme nous ne survolerons que quelques notions de *mécanique quantique relativiste* au chapitre suivant, on peut le considérer comme faisant partie des principes fondamentaux de la mécanique quantique non-relativiste.

NOTIONS DE MÉCANIQUE QUANTIQUE « ET » RELATIVISTE

16

À travers notre exploration du monde quantique, nous ne nous sommes que très rarement posé la question de la compatibilité de cette théorie avec celle, vue dans la première partie de ce livre, de la relativité restreinte.

Par exemple, lorsque nous avons introduit la notion de longueur d'onde de de Broglie, $\lambda = h/p$, d'une particule, nous n'avons pas précisé si cette expression était valable pour la version relativiste, $p = \gamma mv$, de la quantité de mouvement de cette particule (avec $\gamma = 1/\sqrt{1 - (v/c)^2}$, où v est la vitesse de la particule). Il se trouve que c'est bien le cas, et d'ailleurs, Louis de Broglie s'était directement placé dans un cadre relativiste pour concevoir l'expression qui porte aujourd'hui son nom. Il faudrait donc, en toute rigueur, considérer l'expression suivante pour λ_{dB} :

$$\lambda_{dB} = \frac{h}{mv} \sqrt{1 - (v/c)^2} \quad (16.1)$$

ou, de façon équivalente :

$$\lambda_{dB} = \frac{h}{\sqrt{2mE_c + (E_c/c)^2}} \quad (16.2)$$

en fonction de l'énergie cinétique relativiste $E_c = (\gamma - 1)mc^2$ de la particule.

En revanche, il est aisé de vérifier que l'équation d'évolution de la mécanique quantique, c'est-à-dire l'équation de Schrödinger $i\hbar \frac{\partial}{\partial t} \psi(x,t) = -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} \psi(x,t)$, n'est compatible qu'avec la relativité galiléenne, et non avec la relativité d'Einstein. Afin de concilier les deux théories, il a donc fallu construire des analogues relativistes de cette équation, appelées équations de Klein-Gordon et de Dirac, et valables respectivement pour les bosons et les fermions.

Il est bien entendu hors de question d'étudier ici ces équations et leurs solutions, qui font appel à des notions et outils mathématiques bien trop avancés par rapport au niveau visé par cet ouvrage. Il n'est cependant pas inutile de donner un aperçu de quelques-unes de leurs propriétés ou particularités. En particulier, nous verrons qu'elles soulèvent le problème fondamental de l'existence de solutions d'énergie négative (antimatière) et le fait que l'on ne peut pas, en toute rigueur, considérer l'évo-

lution d'une particule dans l'espace-temps sans prendre en compte l'ensemble des particules (réelles ou virtuelles) qui l'entourent.

16.1 ÉQUATION DE KLEIN-GORDON

On a vu au chapitre 14 que l'équation (non-relativiste) de Schrödinger s'obtenait en remplaçant, dans la définition de l'énergie classique $E = p^2/2m$, les grandeurs E et p par les opérateurs $i\hbar \frac{\partial}{\partial t}$ et $-i\hbar \frac{\partial}{\partial x}$ (principe des opérateurs associés, à une dimension d'espace).

L'équation de Klein-Gordon s'obtient de la même façon mais en considérant la définition relativiste de l'énergie $E^2 = p^2c^2 + m^2c^4$. Elle s'écrit :

$$\frac{\partial^2 \psi}{\partial (ct)^2} = \left(\frac{\partial^2}{\partial x^2} - \frac{m^2 c^2}{\hbar^2} \right) \psi \quad (16.3)$$

où l'on constate l'apparition naturelle de la longueur d'onde de Compton $\lambda_{Compton} = \hbar/mc$. Cette équation d'onde a, comme l'équation de Schrödinger, des solutions sous forme d'ondes planes (de vecteur d'onde $k = p/\hbar$), mais à chaque impulsion p peuvent maintenant correspondre deux énergies opposées :

$$E = +\sqrt{p^2c^2 + m^2c^4} \quad \text{ou} \quad E = -\sqrt{p^2c^2 + m^2c^4} \quad (16.4)$$

dont un état d'énergie négative qui aura naturellement tendance à être occupé par les particules. La conséquence principale est que l'énergie d'un système matériel ne serait pas « bornée par le bas » dans le cadre d'une telle théorie, rendant impossible la stabilité de tels systèmes.

16.2 ÉQUATION DE DIRAC, THÉORIE QUANTIQUE DES CHAMPS

L'ambiguïté de signe devant la racine carrée de l'expression précédente de l'énergie, ainsi que la dissymétrie apparente entre les rôles des variables x et t , ont conduit Paul Dirac à rechercher une équation d'onde avec uniquement des dérivées premières en t et en x , mais en prenant comme condition que le carré de cette équation devait redonner l'équation de Klein-Gordon, c'est-à-dire redonner l'expression relativiste de l'énergie.

On montre alors qu'une telle équation est nécessairement matricielle, à quatre dimensions, avec des coefficients sous forme de matrices 4×4 . Les fonctions d'onde solutions ont alors quatre composantes, et sont appelées des *spineurs* de Dirac.

16.3. Incompatibilité de la mécanique quantique et de la relativité générale

On peut également montrer que la conservation du moment cinétique total, dans le cadre de cette équation, implique l'existence de moments cinétiques (spins) demi-entiers. Le *théorème spin-statistique* fournit alors la clef pour relier ces moments cinétiques d'un genre nouveau avec les deux équations relativistes précédentes : les particules élémentaires de spin demi-entier (appelés fermions et constituant la matière) obéissent à l'équation de Dirac, alors que les particules élémentaires de spin entier (appelés bosons et véhiculant les interactions entre fermions élémentaires) obéissent à l'équation de Klein-Gordon, ou à ses généralisations¹.

Aussi belle et puissante soit elle, l'équation de Dirac ne supprime cependant pas le problème de la présence d'énergies négatives... mais Dirac en a trouvé la solution en postulant que les états d'énergie négative sont en fait tous occupés par des électrons qui, étant des fermions, ne peuvent pas occuper le même état selon le *principe de Pauli*. Le vide entourant une particule de matière est alors décrit comme une mer d'électrons *virtuels*. Ainsi, dans ce cadre, une particule d'énergie positive peut être stable puisqu'elle n'a aucun état d'énergie négative disponible à occuper ! Ce résultat signifie cependant que cela n'a plus de sens, en *physique quantique relativiste*, d'étudier l'évolution d'une seule particule... Par exemple, arracher au *vide quantique* un de ses électrons laisse, dans la mer d'électrons, un *trou*, chargé positivement et ayant la même masse qu'un électron, qu'on appelle un positron et qui est une particule d'antimatière.

On est donc, en toute rigueur, contraint de considérer l'infinité de ces particules (électrons, positrons...) qui « *vivent* » autour de la particule test que l'on veut étudier. D'où le besoin d'une théorie où les électrons et positrons sont en fait les excitations élémentaires d'un même champ de matière dans lequel évolue, ou est issue, la particule test. Ainsi, comme pour la lumière, champs et particules sont indissociables : cette approche des théories quantiques relativistes porte le nom de *théorie quantique des champs* et constitue la théorie principale sur laquelle se fonde le *modèle standard* des particules élémentaires (chapitre 9).

16.3 INCOMPATIBILITÉ DE LA MÉCANIQUE QUANTIQUE ET DE LA RELATIVITÉ GÉNÉRALE

Si la mécanique quantique s'accorde particulièrement bien avec la théorie de la relativité restreinte, il n'en est pas de même avec la théorie de la relativité générale (chapitre 9). Ces deux théories sont en effet incompatibles bien qu'elles semblent,

1. Il est à noter que toute solution de l'équation de Dirac est solution de l'équation de Klein-Gordon (mais l'inverse est faux).

chacune de leur côté, donner une description juste et précise de l'univers : vers l'infiniment grand pour la relativité générale, vers l'infiniment petit pour la mécanique quantique.

Parmi les points de divergence, on peut citer par exemple :

1. *Le hasard quantique* : la mécanique quantique est une théorie probabiliste alors que la relativité est déterministe ;
2. *La structure de l'espace-temps* : l'espace-temps de la mécanique quantique est fixe et plat alors que celui de la relativité générale est dynamique et sa courbure dépend de la répartition de la matière et de l'énergie ;
3. *Le rôle du temps* : simple paramètre en mécanique quantique, la variable « temps » a le même statut que les variables « position » en relativité ;
4. *Les relations d'Heisenberg* : les relations d'indétermination de Heisenberg empêchant de définir avec certitude les positions et vitesses des particules de matière, il n'est pas possible de déterminer la valeur du champ gravitationnel en un point ;
5. *L'énergie du vide* : elle est infinie selon la mécanique quantique, mais nulle selon la relativité².

Plusieurs théories alternatives ont été proposées au cours des dernières décennies pour tenter de concilier les grands principes de ces deux théories discordantes : théorie des cordes, gravitation quantique à boucles, géométrie non-commutative, relativité d'échelle... Mais aucune n'est, pour l'instant, suffisamment complète et consistante.

Citons, pour finir, un extrait du livre, *Rien ne va plus en physique*, de Lee Smolin paru en 2007 :

« Deux découvertes expérimentales ont été faites ces dernières décennies : d'une part, les neutrinos ont une masse et, d'autre part, l'univers est dominé par la mystérieuse matière noire et semble être en expansion accélérée. Mais nous n'avons aucune idée de la cause de la masse des neutrinos (ou de toute autre particule) et nous ne savons pas expliquer son apparition. Quant à la matière noire, elle ne s'explique avec aucune des théories physiques existantes. (...) Ces deux découvertes, la relativité et la quantique, nous ont, chacune, demandé de rompre avec la physique de Newton. Pourtant, malgré ce très grand progrès accompli au cours du siècle dernier, ces deux découvertes restent incomplètes. Chacune d'elles possède des faiblesses et des défauts, qui tendent

2. La récente découverte de l'accélération de l'expansion de l'univers a conduit certains chercheurs à introduire un nouveau type d'énergie, l'*énergie sombre*, dont serait issue la force responsable de cette accélération et qui, dans les équations de la relativité générale d'Einstein vues au chapitre 9, apparaîtrait sous la forme d'une *constante cosmologique* non-nulle. Estimée à quelques GeV/m^3 , cette énergie sombre est cependant en désaccord total avec les $10^{121} \text{ GeV}/\text{m}^3$ de l'énergie du vide quantique prédite par la théorie quantique des champs.

16.3. Incompatibilité de la mécanique quantique et de la relativité générale

à prouver l'existence d'une théorie plus fondamentale. Mais la raison la plus évidente pour laquelle chacune des deux théories est incomplète est l'existence de l'autre. Notre esprit nous incite à chercher une troisième théorie, qui unifierait toute la physique, et la raison à l'origine de cette incitation est simple. Il est évident que la nature, elle, est « unifiée ». L'univers dans lequel nous nous trouvons est interconnecté, dans le sens où tout interagit avec tout le reste. Il ne peut y avoir de solution où nous aurions deux théories de la nature, qui décrirait des phénomènes différents, comme si l'une n'avait rien à voir avec l'autre. (...) Ce problème s'appelle le problème de la « gravitation quantique ». »

La gravitation quantique reste un mystère... pour la résolution duquel vous, lecteur, aurez peut-être un rôle à jouer !

EXERCICES ET PROBLÈMES

Retrouvez les corrigés des exercices et problèmes sur la page dédiée à l'ouvrage sur dunod.com.

1 Course de vitesse « galiléenne »

Deux nageurs font une course de vitesse sur un fleuve de largeur d . Le premier effectue la traversée puis revient à son point de départ. Le second fait un aller-retour le long de la rive en parcourant la distance d dans le sens d'écoulement du fleuve, puis en revenant à son point de départ. La vitesse des nageurs par rapport à l'eau, notée v , est supérieure à celle du courant du fleuve, notée V .

- a) Exprimer le temps que met chacun des nageurs pour effectuer un aller-retour. Quel est le vainqueur de la course ?

Les nageurs sont maintenant remplacés par des photons (les photons sont les constituants élémentaires de la lumière, et plus généralement des ondes électromagnétiques, ils n'ont pas de masse et se déplacent à la vitesse c). Le courant du fleuve est remplacé par le mouvement de l'éther, cet hypothétique milieu qu'on supposait, à la fin du XIX^{ème} siècle, occuper tout l'espace, et qui est entraîné par la rotation de la Terre, à la vitesse V .

- b) Exprimer le temps que met un signal lumineux pour faire un aller-retour sur une distance d , parallèlement et perpendiculairement au mouvement de la Terre, en fonction de c , d et V .

- c) Quel est l'écart temporel δt (le retard) entre ces deux résultats ? (données pour l'application numérique : $c = 300\,000\text{ km/s}$, $d = 10\text{ cm}$, $V = 30\text{ km/s}$)

Cette course de vitesse entre photons correspond en fait au parcours de la lumière dans les deux bras d'un interféromètre de Michelson qui serait entraîné par le mouvement de la Terre. Dans un tel interféromètre, le retard δt se matérialise par un « déphasage » $\delta\phi = \omega\delta t$ entre les figures d'interférences, $\nu = \omega/2\pi$ étant la fréquence de l'onde lumineuse utilisée.

- d) Exprimer $\delta\phi$ en fonction de λ , c et δt . On rappelle que λ est la longueur d'onde de l'onde lumineuse : $\lambda = c/\nu$. Faire l'application numérique pour $\lambda = 0,5\text{ }\mu\text{m}$.

2 Une demi-vie bien relative

On considère une particule susceptible de se désintégrer, dont le temps de demi-vie est τ .

- a) Rappeler la définition de τ et dans quel référentiel cette définition est valable.

- b) On suppose maintenant que cette particule est en mouvement de vitesse constante \vec{v} par rapport à un observateur O . L'énergie de cette particule, vue depuis le référentiel de l'observateur, est E . Exprimer la vitesse de la particule en fonction de E , de la masse m de la particule et de c .
- c) Quel est le temps de demi-vie de la particule vue dans le référentiel de l'observateur ? On l'exprimera en fonction de E et m .
- d) Que peut-on dire de ce temps si la masse de la particule tend vers zéro ? Commenter sur la stabilité des particules de faible masse, et en particulier des photons.

3 Où l'on parle d'un œuf dans une bouteille...

Soit \mathcal{R}' un référentiel galiléen animé d'une vitesse $V = 0,6c$ par rapport à un autre référentiel galiléen \mathcal{R} . La vitesse V est parallèle aux axes (Ox) et (Ox') , et les horloges des deux référentiels sont ajustées de telle sorte que les trois axes, ainsi que les origines O et O' , de ces référentiels coïncident à l'instant $t = 0 = t'$. On rappelle que l'unité « année-lumière » correspond à la distance parcourue par la lumière pendant une année Terrestre.

- a) Donner l'expression du facteur de Lorentz γ en fonction de V et c , puis faire l'application numérique.
- b) Donner l'expression littérale des transformations de Lorentz permettant de passer des coordonnées du référentiel \mathcal{R} à celles du référentiel \mathcal{R}' , et inversement.
- c) Soit M_1 l'évènement de coordonnées $x_1 = 4,6$ années-lumière et $t_1 = 1$ an dans \mathcal{R} . Quelles sont ses coordonnées (x'_1, t'_1) dans \mathcal{R}' ? (Donner l'expression littérale, puis faire l'application numérique)
- d) Soit M_2 l'évènement de coordonnées $x_2 = 3,2$ années-lumière et $t_2 = 2$ ans dans \mathcal{R} . Quelles sont ses coordonnées (x'_2, t'_2) dans \mathcal{R}' ? (Donner l'expression littérale, puis faire l'application numérique)
- e) On définit les longueurs $\Delta x'$ et Δx par les relations : $\Delta x' = x'_2 - x'_1$ et $\Delta x = x_2 - x_1$. Comparer $\Delta x'$ à Δx ? Est-ce un exemple de contraction relativiste des longueurs ? (Justifier clairement)
- f) Soit un objet, immobile dans \mathcal{R}' , de longueur propre égale à L_0 . Quelle est la longueur de cet objet mesurée dans \mathcal{R} ? (Donner l'expression littérale, puis faire l'application numérique avec la valeur de $L_0 = \Delta x'$ déterminée précédemment)
- g) Reprendre la question précédente dans le cas où le référentiel \mathcal{R}' se déplace dans le sens opposé : $V = -0,6c$. Conclusion ?
- h) Donner un ordre de grandeur de la vitesse à laquelle un œuf (de poule) devrait se déplacer pour pouvoir rentrer, par effet de contraction relativiste, à travers

le goulot d'une bouteille (de champagne, par exemple). À vous de définir la géométrie du problème et d'estimer les dimensions transverses de l'œuf et du goulot ! Commenter.

4 Hors-jeu relativiste

Au football, un joueur A1 est dit hors-jeu quand un coéquipier A2 lui passe le ballon alors qu'il se situe entre le dernier défenseur D et le gardien de but G de l'équipe adverse.

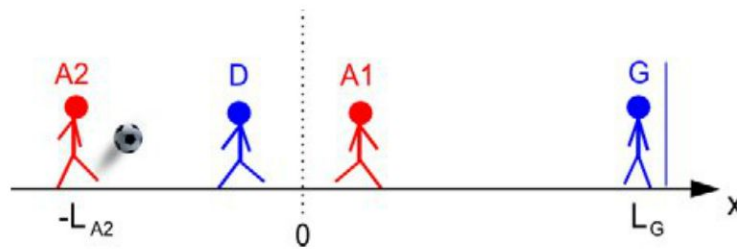


Figure 1 – Situation de hors-jeu pour A1, vu de \mathcal{R} .

Le but de cet exercice est d'étudier cette règle du football... avec une description relativiste !

Les joueurs A2 et G sont immobiles par rapport au référentiel attaché au terrain que l'on notera \mathcal{R} . A1 court (très vite !) en direction de G selon l'axe Ox avec une vitesse V par rapport à \mathcal{R} . On notera \mathcal{R}' le référentiel propre de A1. Le défenseur D court (aussi très vite !) vers A2, selon Ox avec une vitesse $-V$ par rapport à \mathcal{R} (en vue d'intercepter le ballon, mais ça c'est un autre problème !). Le référentiel propre de D est noté \mathcal{R}'' .

Initialement, les horloges des trois référentiels indiquent le même temps $t = t' = t'' = 0$, et leurs origines O , O' et O'' ainsi que leurs axes (Ox) , $(O'x')$ et $(O''x'')$ coïncident.

Dans le référentiel \mathcal{R} , A2 est en $x = -L_{A2}$ et G se trouve en $x = L_G$. On choisit l'origine des temps telle qu'à $t = 0$ les joueurs A1 et D se croisent en $x = 0$. C'est à cet instant précis, dans \mathcal{R} , que le joueur A2 frappe le ballon pour l'envoyer au joueur A1.

On définit donc les deux événements suivants : E_2 : « A2 frappe la balle », et E_1 : « A1 entre en position de hors-jeu (au croisement avec D) ».

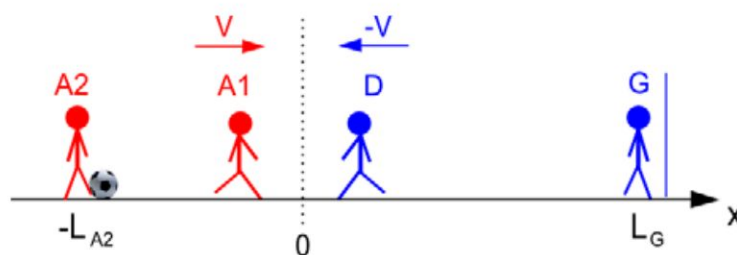


Figure 2 – Situation avant la passe quand $t < 0$, vu de \mathcal{R} .

- a) Donner les transformations de Lorentz qui permettent de passer des coordonnées (x, ct) d'un événement dans le référentiel \mathcal{R} à celles (x', ct') de ce même événement dans le référentiel \mathcal{R}' ; ainsi que celles qui permettent de passer des coordonnées (x, ct) dans \mathcal{R} aux coordonnées (x'', ct'') dans \mathcal{R}'' . Préciser également l'expression du facteur γ associé.
- b) 1. Quel est l'instant de croisement t'_c de A1 et D dans \mathcal{R}' ?
2. De même, à quel instant t''_c se croisent ces joueurs dans \mathcal{R}'' ?
- c) À quel point d'espace-temps (ct_p, x_p) de \mathcal{R} se trouve A2 quand il effectue la passe ?
- d) À quel point d'espace-temps (ct'_p, x'_p) de \mathcal{R}' se trouve A2 quand il effectue la passe ?
- e) Donner l'expression, en fonction de V , de la vitesse relative V_1 de A1 par rapport à D (égale à la vitesse de \mathcal{R}' par rapport à \mathcal{R}'').
- f) À quel point d'espace-temps (ct''_p, x''_p) de \mathcal{R}'' se trouve A2 quand il effectue la passe ?
- g) En utilisant la question 5), exprimer x''_p en fonction de x'_p et ct'_p .
- h) En comparant les temps t'_c et t'_p , dire si A1 se trouve en situation de hors-jeu dans \mathcal{R}' .
- i) Même question dans \mathcal{R}' en comparant les temps t''_c et t''_p .
- j) Conclure sur cette règle du hors-jeu dans le cas relativiste.
- k) 1. Qu'est-ce qu'un intervalle espace-temps entre deux événements ? Que pouvez-vous dire sur la valeur qu'il prend dans différents référentiels galiléens ?
2. Donner l'expression de l'intervalle d'espace-temps Δs^2 entre E_2 et E_1 . De quel *genre* est-il ?
3. En supposant que vous ne connaissiez que le genre de l'intervalle trouvé en 2., auriez-vous pu en déduire la présence ou non d'un lien de causalité entre les deux événements ? Justifier votre réponse.
- l) 1. Donner l'ordre chronologique dans lequel le gardien G voit les deux événements E_2 et E_1 (c'est-à-dire l'ordre dans lequel l'œil du gardien perçoit les rayons lumineux émis lors de ces événements). Justifier votre réponse.
2. Le gardien peut-il, à partir de ce qu'il perçoit, conclure à un lien de causalité entre les deux événements ? Justifier qualitativement votre réponse.
- m) 1. On suppose, dans cet exercice, que le ballon est également un objet relativiste. Sa vitesse selon x est $V_x = 0.8c$ et selon y , $V_y \ll c$. Vu depuis le référentiel \mathcal{R} du terrain, quelle est la forme du ballon ? En sachant que le rayon propre de ce ballon est $R_0 = 11$ cm calculer sa dimension R_x dans \mathcal{R} .

2. Dans certains dessins animés mettant en scène une partie de football, pour montrer sa grande vitesse le ballon a une forme ovale, aplatie selon l'axe Oy ($R_x > R_y$), qu'en pensez-vous ?

5 N'est pas gagnant celui qu'on croit !

Deux vaisseaux spatiaux décident de faire une course. Ils partent dans des directions opposées, à la même vitesse $v_0 = +0,8c$ par rapport au référentiel \mathcal{R} associé à la ligne de départ. On choisit la direction de la course comme axe (Ox) de ce référentiel, et la ligne de départ comme origine des coordonnées, où se tient, immobile, le juge de course. La ligne d'arrivée du vaisseau 1 se situe en $x_1 = 0,8$ années-lumière, celle du vaisseau 2 en $x_2 = -0,8$ années-lumière.

- Quel est le résultat de la course selon le juge de course ? Justifier.
- Soient E_1 et E_2 les événements associés à la fin de la course : E_1 correspondant au vaisseau 1 coupant sa ligne d'arrivée, et E_2 correspondant au vaisseau 2 coupant sa ligne d'arrivée. Déterminer Δs^2 , le carré de l'intervalle d'espace-temps entre ces deux événements. Donner l'expression littérale dans \mathcal{R} , puis faire l'application numérique.
- De quel type d'intervalle s'agit-il ? Existe-t-il un référentiel où ces deux événements seraient distants de moins de 1,6 années-lumière ? Justifier.
- On cherche maintenant à comprendre quel est le résultat de la course mais vu depuis chacun des vaisseaux.
 - Soit \mathcal{R}' le référentiel propre associé au vaisseau 1 (les horloges des deux référentiels \mathcal{R} et \mathcal{R}' sont ajustées de telle sorte que les origines O et O' des référentiels coïncident à l'instant $t = t' = 0$ du début de la course). Déterminer les coordonnées temporelles t'_1 et t'_2 des événements E_1 et E_2 dans \mathcal{R}' .
 - Choisir, en expliquant clairement la raison de votre choix, une des solutions ci-dessous :
 - Chaque vaisseau pense avoir gagné.
 - Chaque vaisseau pense avoir perdu.
 - Le vaisseau 1 pense avoir gagné et le vaisseau 2 pense avoir perdu.
 - Le vaisseau 2 pense avoir gagné et le vaisseau 1 pense avoir perdu.
 - Aucune de ces réponses.

6 Une autre course de vaisseaux

Deux vaisseaux, de vitesses respectives v_1 et v_2 non négligeables devant celle de la lumière, décident de faire la course pour arriver jusqu'à la Terre. Le vaisseau 2 est

deux fois plus rapide que le vaisseau 1 ($v_2 = 2v_1$), mais, il doit parcourir une distance L_2 deux fois plus importante que L_1 , celle que doit parcourir le vaisseau 1, soit $L_2 = 2L_1$. Les distances L_1 et L_2 sont mesurées dans un référentiel fixe par rapport à la Terre (référentiel R).

- a) Vu de la Terre (référentiel R), exprimer les temps t_1 et t_2 mis par les vaisseaux 1 et 2 pour arriver sur Terre en fonction de L_1 et v_1 . Quel est le résultat de la course selon un arbitre immobile sur Terre ?

Au courant de ce résultat prévisible, le pilote du vaisseau 1 a proposé d'embarquer l'arbitre de façon à ce que l'horloge du vaisseau 1 décide du vainqueur de la course. Avant d'accepter, le pilote du deuxième vaisseau souhaite vérifier que cela ne tourne pas à son désavantage.

- b) Soit R' le référentiel lié au vaisseau 1. Donner dans R' les distances L'_1 et L'_2 à parcourir par chacun des vaisseaux en fonction de c , L_1 et de v_1 .
- c) À quelles vitesses se déplacent la Terre et le vaisseau 1 dans R' ? En déduire le temps t'_1 mis par le vaisseau 1 pour arriver sur Terre dans R' .
- d) Exprimer en fonction de c , v_1 et v_2 la vitesse du vaisseau 2 dans R' .
- e) Déterminer t'_2 , le temps mis par le vaisseau 2 pour arriver sur Terre dans R' .
- f) Démontrer alors que :

$$\frac{t'_2}{t'_1} = \frac{2}{1 + \frac{1}{1 - \frac{v_1^2}{c^2}}}$$

- g) Quelles valeurs extrêmes peut prendre v_1 si on veut que les données du problème restent correctes ? Donner alors un encadrement de $\frac{t'_2}{t'_1}$.
- h) S'il veut gagner, le pilote du vaisseau 2 doit-il accepter que les temps d'arrivée soient déterminés dans le vaisseau 1 (R') ?
- i) On considère les deux événements *arrive du vaisseau 1* et *arrive du vaisseau 2*. Donner en justifiant le signe de l'intervalle spatio-temporel Δs^2 entre ces deux événements. De quel type d'intervalle s'agit-il ?

7 Croisement de fusées

Une fusée part de la Terre à l'instant $t = 0$ avec une vitesse constante $v_1 = \beta_1 c$, $\beta_1 = 4/5$, vers l'étoile la plus proche du Soleil, Proxima du Centaure, située à quatre années lumière de nous. La distance entre la Terre et cette étoile sera supposée constante par la suite.

Les référentiels R et R' sont, respectivement, liés à la Terre et à la fusée. Les axes de coordonnées sont parallèles et les origines sont confondues à l'instant $t = 0$.

L'axe (Ox) sera pris dans le sens du mouvement de la fusée. Les résultats des applications numériques s'exprimeront en années et années lumière.

- a) Calculer la durée du voyage pour un observateur Terrestre.
- b) En déduire la durée du voyage pour un observateur à bord de la fusée.

Simultanément (dans R) au départ de la première fusée, une deuxième fusée part de Proxima du Centaure vers la Terre, en mouvement rectiligne uniforme, à une vitesse $v_2 = \beta_2 c$, $\beta_2 = 1/2$. Appelons A et B les événements correspondant aux deux départs.

$$x_A = 0, t_A = 0, x_B = d, t_B = 0, d = 4 \text{ années lumière.}$$

- c) Exprimer les coordonnées x'_A, t'_A, x'_B, t'_B de ces deux événements dans R' en fonction de c , d et d'un des deux paramètres β_1 ou β_2 . Effectuer les applications numériques.
- d) Est-ce que les deux départs de fusées sont simultanés dans R' ? Justifier votre réponse.
- e) Trouver, dans R , la position x^* et le temps t^* correspondant au croisement des deux fusées.

Faire l'application numérique. Ce calcul est-il relativiste ou non-relativiste ? Justifier votre réponse.

8 Train et tunnel relativistes – ou – le paradoxe de la bombe

On considère un train « très » rapide traversant un tunnel rectiligne dont la longueur propre L_0 est identique à la longueur propre du train. La vitesse (constante et uniforme) V du train par rapport au tunnel est supposée proche de celle de la lumière c .

Conformément à la théorie de la relativité restreinte, la longueur L du train apparaîtra donc plus courte si elle est mesurée dans le référentiel \mathcal{R} du tunnel. Réciproquement, c'est le tunnel qui apparaîtra plus court si sa longueur est mesurée dans le référentiel \mathcal{R}' du train (phénomène de contraction des longueurs).

On appelle A et B les extrémités avant et arrière du train, et G et H les points d'entrée et de sortie du tunnel. Avant que le train ne pénètre dans le tunnel, ces quatre points se lisent donc $B - A - G - H$ de la gauche vers la droite (faire un schéma !).

On choisit B , l'arrière du train, comme origine de coordonnées spatiales du référentiel \mathcal{R}' , et G , l'entrée du tunnel, comme origine de coordonnées spatiales du référentiel \mathcal{R} . Les coordonnées spatiales des extrémités A et B du train dans le référentiel \mathcal{R}' sont donc :

$$x'_B = 0 ; x'_A = L_0$$

L'origine temporelle commune des deux référentiels correspond à l'instant où B et G coïncident. Ainsi, l'événement :

$$E_1 = \text{« l'arrière } B \text{ du train coïncide avec l'entrée } G \text{ du tunnel »}$$

a pour coordonnées $(x_1 = 0, t_1 = 0)$ dans \mathcal{R} et $(x'_1 = 0, t'_1 = 0)$ dans \mathcal{R}' .

On cherche, dans un premier temps, à déterminer les coordonnées de l'événement :

$$E_2 = \text{« l'avant } A \text{ du train coïncide avec la sortie } H \text{ du tunnel »}$$

- a)
 1. Donner l'expression de L en fonction de L_0 et γ , puis celle du facteur de Lorentz γ en fonction de V et c .
 2. Rappeler les formules de Lorentz exprimant les coordonnées (x, t) d'un événement dans \mathcal{R} en fonction des coordonnées (x', t') de ce même événement dans \mathcal{R}' .
- b)
 1. Quelle est la trajectoire $x_B(t)$ de l'extrémité B du train dans \mathcal{R} ?
 2. Exprimer les coordonnées spatio-temporelles de l'extrémité A du train dans \mathcal{R} (en utilisant les transformations de Lorentz).
 3. Éliminer la variable « t' » des deux équations obtenues à la question b.2. pour déterminer la trajectoire $x_A(t)$ de l'extrémité A du train dans \mathcal{R} .
 4. Donner l'expression de $x_A(t) - x_B(t)$ en fonction de L_0 et γ , puis comparer à la question a.1.
- c)
 1. Déterminer les coordonnées (x_2, t_2) de l'événement E_2 dans \mathcal{R} .
 2. Quelle durée Δt sépare les deux événements E_1 et E_2 dans \mathcal{R} ?
 3. Dans le référentiel \mathcal{R} , lequel des deux événements a lieu avant l'autre ?
- d)
 1. Déterminer les coordonnées (x'_2, t'_2) de l'événement E_2 dans \mathcal{R}' .
 2. Quelle durée sépare les deux événements E_1 et E_2 dans \mathcal{R}' ? Comparer au résultat de la question c.2.
 3. Dans le référentiel \mathcal{R}' , lequel des deux événements a lieu avant l'autre ? Y-a-t-il un paradoxe au vu de la question c.3 ?
- e)
 1. Donner l'expression de l'intervalle d'espace-temps Δs^2 entre les deux événements E_1 et E_2 .
 2. De quel signe est-il ? Comment appelle-t-on ce type d'intervalle d'espace-temps ?
 3. Est-ce que les deux événements E_1 et E_2 peuvent être reliés par un rayon lumineux dans un référentiel ? Justifier.

- f) Un signal lumineux (isotrope) est émis à l'arrière du train lorsqu'advient l'événement E_1 . Dans \mathcal{R} , ce signal atteint l'avant A du train à l'instant t_3 .
- g) 1. De quel signe est la quantité $t_3 - t_2$?
 2. Calculer $t_3 - t_2$.
 3. On présente parfois cet exercice sous la forme :

« Une bombe, située à l'avant du train, est réglée pour exploser à l'instant précis où l'avant du train sort du tunnel. Cette bombe peut néanmoins être désamorcée par un signal lumineux émis à l'arrière du train à l'instant précis où l'arrière du train entre dans le tunnel. »

À votre avis, la bombe explosera-t-elle ? Justifier.

9 Miroir, mon beau miroir...

Deux princesses (relativistes !) A et B courent toutes les deux, côte à côte, selon la même direction et dans le même sens, à la vitesse $V = 0,999998c$ par rapport au sol Terrestre. La princesse A tient à bout de bras un miroir leur faisant face.

- a) Que voit la princesse B dans le miroir ? Justifier clairement votre/vos choix !
1. Rien.
 2. Elle-même, mais courant au ralenti.
 3. Elle-même, mais avec des couleurs différentes (décalées vers le bleu par exemple).
 4. Elle-même, mais contractée dans la direction du mouvement.
 5. Aucune de ces réponses.
- b) Écrire V/c sous la forme $1 - \alpha$, et en déduire une expression approchée du facteur de Lorentz γ en fonction de α . Faire ensuite l'application numérique.
- c) Estimer un ordre de grandeur de « l'épaisseur » des deux princesses vue depuis le référentiel lié au sol Terrestre. (« épaisseur », ici, signifie « largeur, ou extension, selon la direction du mouvement »)

10 Flash rouge ou flash bleu ?

Un véhicule muni d'une horloge quitte l'origine $x = 0$ du référentiel \mathcal{R} à l'instant $t = 0$ avec une vitesse $v = +\frac{\sqrt{3}}{2}c$ dirigée le long de l'axe (Ox) . Quand l'horloge du véhicule indique le temps T_0 , le véhicule envoie un flash lumineux rouge, de longueur d'onde $0,8 \mu\text{m}$, vers l'origine de \mathcal{R} . On appellera \mathcal{R}' le référentiel lié au véhicule.

- a) Donner la valeur numérique du facteur de Lorentz γ associé au mouvement relatif des référentiels \mathcal{R} et \mathcal{R}' .

- b) Rappeler les formules de passage des coordonnées spatio-temporelles d'un événement dans \mathcal{R} en fonction de celles du même événement dans \mathcal{R}' . Donner également les formules de passage inverse.
- c) Soit l'événement : « *émission du flash lumineux rouge par le véhicule* », déterminer ses coordonnées (ct'_1, x'_1) dans \mathcal{R}' , puis ses coordonnées (ct_1, x_1) dans \mathcal{R} .
- d) Quelle est la trajectoire, dans \mathcal{R} , du flash lumineux rouge ?
- e) Soit l'événement : « *réception du flash lumineux rouge par l'origine de \mathcal{R}* », déterminer ses coordonnées (ct_2, x_2) dans \mathcal{R} , puis ses coordonnées (ct'_2, x'_2) dans \mathcal{R}' .
- f) Quand l'horloge du référentiel \mathcal{R} indique le temps T_0 , un flash lumineux bleu, de longueur d'onde $0,4 \mu\text{m}$, est émis par l'origine O de \mathcal{R} en direction de l'origine O' de \mathcal{R}' . Déterminer, dans \mathcal{R} , l'instant t_3 de croisement des flashes lumineux rouge et bleu.
- g) Est-ce que le flash bleu détecté par le véhicule est dans le domaine visible (c'est-à-dire entre $0,4 \mu\text{m}$ et $0,8 \mu\text{m}$) ? Justifier.

11 Plus vite que la lumière ?

On considère trois référentiels liés à trois observateurs se déplaçant le long de l'axe (Ox) : un observateur O , un observateur S allant vers la droite de O en mouvement rectiligne uniforme à la vitesse $v = 0,8c$ par rapport à O , et un observateur S' allant vers la gauche de O en mouvement rectiligne uniforme à la même vitesse v par rapport à O . Ces trois référentiels coïncident à l'instant $t = 0$.

- a) L'observateur O dit : « La distance entre S et S' étant $d = 1,6ct$ (en fonction du temps propre t de l'observateur O), elle augmente dans le temps à une vitesse supérieure à c , ce qui contredit les principes de la relativité restreinte. » Est-ce que cette conclusion est correcte ? Expliquer pourquoi.
- b) Rappeler la formule relativiste d'addition des vitesses en précisant la signification des différents termes, et donner l'expression de u , la vitesse de S' par rapport à S , en fonction de v . Faire l'application numérique.

12 Absorption-émission de photons, effet Mössbauer

Les noyaux atomiques peuvent absorber, émettre ou diffuser des photons.

Absorption. Considérons donc un noyau de masse au repos m_0 qui absorbe un photon incident d'énergie Q . L'état final du système est alors constitué d'un noyau excité de masse au repos $m' \neq m_0$ et de vitesse v (appelée vitesse de recul).

- a) Faire un schéma pour représenter l'état du système avant et après l'absorption, et écrire les relations de conservation de l'énergie et de l'impulsion relativistes.

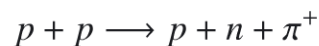
- b) Exprimer ensuite m' et v en fonction de m_0 et Q .
- c) En fait, les lois de la mécanique quantique prédisent qu'il n'existe qu'une liste discrète des valeurs possibles de m' . Si m' n'appartient pas à cette liste « quantique », que devient le photon après son interaction avec le noyau ?

Émission. Considérons maintenant le processus inverse : l'émission d'un photon d'énergie Q par un noyau excité de masse au repos m_0 . Après l'émission, la masse au repos du noyau s'écrira m'_0 , son énergie E' et son impulsion \vec{p}' .

- d) Suivre la même démarche que dans le cas de l'absorption : schéma et relations de conservation.
- e) Que représente physiquement la quantité $\Delta E = m_0 c^2 - m'_0 c^2$?
- f) Utiliser la relation fondamentale $E'^2 = \vec{p}'^2 c^2 + m'^2_0 c^4$ pour exprimer Q en fonction de m_0 et ΔE .
- g) La question précédente a mis en lumière le fait que l'énergie d'un photon est diminuée par effet de recul. Même si cet effet est faible (comme vous pouvez le vérifier facilement), il ne peut pas être négligé dans les calculs de mécanique quantique. Cependant, lorsque le noyau concerné fait partie d'un assemblage régulier de noyaux comme dans un cristal, on peut observer des photons émis (ou absorbés) avec l'énergie ΔE , c'est-à-dire sans effet de recul (c'est l'effet Mössbauer). Comment expliquer simplement cet effet ?

13 Création de Pion par collisions proton-proton

Dans les accélérateurs-collisionneurs de particules, on peut convertir de l'énergie cinétique en masse et produire ainsi de nouvelles particules. Considérons par exemple une collision proton-proton qui, dans certaines conditions, peut créer une particule appelée Pion, responsable de la force nucléaire forte :



Plaçons-nous d'abord dans le référentiel du centre de masse, le proton « de gauche » allant vers la droite et le proton « de droite » allant vers la gauche, tous deux à la vitesse u . Après la collision, les particules sont supposées au repos. Les masses impliquées sont $m_\pi/m_p \approx 0,149$ et $m_n \approx m_p \approx 938 \text{ MeV}/c^2$.

- a) En utilisant la relation de conservation de l'énergie, exprimer et calculer le rapport u/c nécessaire à la création d'un Pion.
- b) Calculer également l'énergie cinétique relativiste des deux protons avant la collision.

Considérons maintenant une collision proton-proton où un des protons est immobile. Pour éviter de faire de nouveaux calculs, reprenons le cas précédent

mais plaçons-nous dans un référentiel allant vers la gauche à la vitesse u tel que le proton de droite soit immobile.

- c) Exprimer la vitesse v du proton de gauche en fonction de u et c .
- d) Exprimer et calculer l'énergie cinétique relativiste de ce proton mobile.
- e) Comparer cette énergie à celle obtenue à la question b) (lorsque les deux protons étaient mobiles). Pourquoi est-elle plus importante ? Conclusion.

14 Énergie minimale de création d'un boson de Higgs

On ne considèrera, dans cet exercice, qu'une seule direction spatiale, et l'on notera « c » la vitesse de la lumière.

a) Généralités

- 1. Rappeler l'expression relativiste qui lie l'énergie E et la quantité de mouvement p d'une particule de masse m .
- 2. Exprimer également E en fonction du facteur de Lorentz γ , de m et de c .
- 3. Rappeler l'expression de γ en fonction de c et de la vitesse v de la particule, puis exprimer le rapport v/c en fonction de γ .
- 4. En déduire l'expression de v/c au premier ordre en $1/\gamma^2$ dans le cas ultra-relativiste ($\gamma \gg 1$).

b) On considère, dans le référentiel du laboratoire, la collision frontale entre deux électrons de même vitesse absolue v_0 .

- 1. Déterminer l'énergie minimale E_0 , de *chaque* électron, nécessaire à la création d'un boson de Higgs ($m_{Higgs} = 125 \text{ GeV}/c^2$) lors de l'annihilation des deux électrons suite à leur collision.
- 2. Déterminer la valeur numérique du facteur de Lorentz γ_0 associé à v_0 (rappel : $m_{\text{électron}} = 0,5 \text{ MeV}/c^2$).
- 3. En déduire la valeur numérique de $1 - v_0/c$.

c) On suppose maintenant que l'un des deux électrons précédents est au repos dans le référentiel du laboratoire, et que l'autre vient le percuter à la vitesse v_1 .

- 1. Déterminer l'énergie minimale E_1 de l'électron mobile, nécessaire à la création d'un boson de Higgs.
- 2. Déterminer la valeur numérique du facteur de Lorentz associé γ_1 .
- 3. En déduire la valeur numérique de $1 - v_1/c$.
- 4. Comparer ce résultat avec celui de la question b.3.

d) On se place dans le cas ultra-relativiste ($\gamma \gg 1$). Donner l'expression approchée de E en fonction de p et c , puis exprimer γ en fonction du rapport $\frac{pc}{mc^2}$.

15 Empoisonnement d'un agent secret au Polonium 210

En novembre 2006, à Londres, un ancien agent des services secrets est mort empoisonné par ingestion de Polonium ^{210}Po . On se propose, dans ce problème, d'étudier les réactions nucléaires mises en jeu et d'évaluer la dose de radiations absorbées qui a conduit à la mort de cet homme.

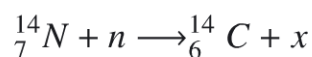
- a) Le ^{210}Po peut être obtenu à partir du Bismuth $^{209}_{83}\text{Bi}$ après absorption (provoquée) d'un neutron (= réaction 1), puis émission d'un électron (= réaction 2). Écrire et préciser le nom de ces deux réactions, et donner le nombre de protons d'un noyau de ^{210}Po .
- b) Calculer le défaut de masse (en u) de chacun des deux isotopes du Bismuth apparaissant dans les réactions précédentes.
On donne : $m_p = 1,007276u$, $m_n = 1,008665u$, $m(^{209}\text{Bi}) = 208,934860u$, $m(^{210}\text{Bi}) = 209,938584u$.
- c) Sachant qu'un noyau atomique est d'autant plus stable que son énergie de liaison « par nucléon » est élevée, lequel de ces deux isotopes est le plus stable ?
- d) Un noyau de ^{210}Po n'est pas stable et se désintègre spontanément en émettant une particule alpha pour donner un noyau de plomb.
 1. Écrire la réaction de désintégration correspondante et calculer l'énergie cinétique maximale de la particule alpha émise (rappel : $1u * c^2 = 931,5\text{MeV}$).
 2. Est-ce un « rayonnement ionisant » très énergétique ? (comparé, par exemple, aux quelques eV typiques d'un photon du domaine visible). On donne : $m(^{210}\text{Po}) = 209,936800u$, $m(^4\text{He}) = 4,001506u$, $m(^{206}_{82}\text{Pb}) = 205,929489u$.
- e) La quantité de ^{210}Po ingérée par l'agent secret a été estimée à $10\text{ }\mu\text{g}$ par les autorités britanniques. Calculer le nombre N_0 d'atomes de ^{210}Po absorbés (rappel : $1u = 1,66054 \cdot 10^{-27}\text{ kg}$).
- f) Préciser la relation donnant l'activité A_0 de ces N_0 atomes de Polonium en fonction de N_0 et de la période radioactive du Polonium ^{210}Po .
 1. Calculer numériquement l'activité A_0 sachant que la période radioactive du ^{210}Po est de 138,38 jours.
 2. Combien de particules alpha – dues aux $10\text{ }\mu\text{g}$ de ^{210}Po radioactif – sont alors émises par seconde ?
 3. Expliquer pourquoi l'enquête policière a été facilitée par la valeur importante de la période radioactive du Polonium 210.

g) On rappelle que la « dose » radioactive est définie comme étant l'énergie déposée (par les rayonnements ionisants) dans une unité de masse de matière (unité : le Gray, $1 \text{ Gy} = 1 \text{ J/kg}$).

1. Quelle était la dose journalière (en Gray) reçue par l'agent secret ($m = 75 \text{ kg}$) au début de son empoisonnement ? (on supposera l'activité constante)
2. Pour obtenir la « dose équivalente » correspondante (en Sievert Sv), il faut multiplier le résultat précédent par 20 (facteur 20 pour particules alpha). Sachant que la dose (équivalente) létale est de 10 Sv , conclure sur la cause du décès de l'agent secret.

16 Datation au carbone 14

Un isotope du carbone, le ^{14}C , est formé dans les plus hautes couches de l'atmosphère par l'impact d'un neutron n sur un noyau d'azote, suivant la réaction :



On donne les valeurs suivantes : masses au repos pour le proton p , le neutron n , et pour les noyaux de ^{14}N et de ^{14}C : $m(n) = 1,009u$; $m(p) = 1,007u$; $m(^{14}\text{N}) = 14,003u$; $m(^{14}\text{C}) = 14,007u$, avec $1u = 1,6 \cdot 10^{-27} \text{ kg}$. Charge élémentaire $e = 1,6 \cdot 10^{-19} \text{ C}$

- a) Déterminer les numéros Z et A de la particule x . De quelle particule s'agit-il ?
- b) Calculer les défauts de masse Δm de ^{14}N et de ^{14}C .
- c) Exprimer alors l'énergie de liaison par nucléon que possèdent les noyaux ^{14}N et ^{14}C en fonction de c . Lequel des deux noyaux est le plus stable ?
- d) Exprimer l'énergie de la réaction ΔE en fonction de $m(^{14}\text{N})$, $m(^{14}\text{C})$, $m(n)$, $m(p)$ et c . La réaction est-elle endothermique ou exothermique ? Faire l'application numérique en eV.
- e) On suppose que l'azote est initialement au repos. Déterminer l'énergie cinétique minimale que doit avoir le neutron pour que la réaction puisse avoir lieu.
- f) Dans ce cas, en déduire l'expression et l'ordre de grandeur de $\frac{v}{c}$ sa vitesse par rapport à celle de la lumière.

17 Effet photoélectrique sur les satellites

Un satellite en orbite peut se charger par effet photoélectrique quand la lumière du Soleil éjecte des électrons de sa surface. Les ingénieurs doivent donc concevoir les satellites de façon à ce que cet effet soit réduit au minimum. Supposez qu'un satellite est couvert de platine, métal ayant un travail d'extraction très élevé ($5,32 \text{ eV}$). Trouvez la plus grande longueur d'onde de lumière solaire incidente pouvant éjecter un électron du platine.

18 Force électromagnétique exercée sur un miroir

- Rappeler l'expression de l'énergie et de la quantité de mouvement d'un des photons associés à une onde électromagnétique plane de champ électrique « $E_0 \exp^{i(\omega t - kz)}$ ».
- Ces photons sont incidents sur un miroir parfaitement réfléchissant se trouvant dans le plan xOy (c'est-à-dire orthogonal à la direction (Oz)). Déterminer la direction et l'amplitude de la quantité de mouvement acquise par le miroir lorsqu'un photon se réfléchit sur sa surface.
- Exprimer la force exercée sur le miroir par un faisceau lumineux dont le flux de photons est de N photons par seconde.
- En déduire la force exercée sur le miroir par un faisceau lumineux de puissance $P = 3 \text{ mW}$.

19 Pression de radiation solaire sur la Terre

L'intensité I du rayonnement solaire au niveau de la Terre est $I = 1,35 \text{ kW.m}^{-2}$.

Données supplémentaires : distance Terre-Soleil $D = 1,5 \cdot 10^{11} \text{ m}$, rayon Terrestre $R = 6,4 \cdot 10^6 \text{ m}$, masse du Soleil $m_S = 2 \cdot 10^{30} \text{ kg}$, masse de la Terre $m_T = 6 \cdot 10^{24} \text{ kg}$, constante de gravitation $G = 6,67 \cdot 10^{-11} \text{ N.m}^2.\text{kg}^{-2}$, vitesse de la lumière dans le vide $c = 3 \cdot 10^8 \text{ m.s}^{-1}$.

- Exprimer la puissance totale P émise par le Soleil en fonction de I et D . Faire l'application numérique.
- Soit ϕ_ν le flux (en s^{-1}) de photons solaires de fréquence ν reçus par la Terre. Exprimer la puissance P_ν associée à ces photons tombant sur la Terre (en fonction de ϕ_ν , h et ν). Faire l'application numérique.
- Exprimer la norme de la quantité de mouvement portée par un photon de fréquence ν en fonction de h , ν et c . Faire l'application numérique.
- Exprimer la norme de la force (dite « de pression de radiation ») F_ν qu'exercent les photons solaires de fréquence ν sur la Terre (en fonction de ϕ_ν , h , ν et c). Faire l'application numérique.
- Exprimer la norme de la force totale F qu'exerce l'ensemble des photons solaires sur la Terre (en fonction de I , R et c). Faire l'application numérique.
- Comparer cette force à la force d'attraction gravitationnelle entre le Soleil et la Terre.

20 Un, deux, trois, quatre... polariseurs

Des photons polarisés verticalement traversent un polariseur dont l'axe passant fait un angle orienté de $+60^\circ$ avec l'axe horizontal. On notera $|H\rangle$ et $|V\rangle$ les états de polarisation horizontal et vertical.

- a) Faire un rapide schéma (à 2D) représentant la polarisation du photon, l'axe horizontal et l'axe passant du polariseur.
- b) Quelle est la probabilité qu'un photon soit transmis par le polariseur ? (faire l'application numérique)
- c) Exprimer l'état de polarisation $|\psi_1\rangle$ d'un photon transmis en fonction de $|H\rangle$ et $|V\rangle$.

On place maintenant un deuxième polariseur (après le premier) sur le trajet des photons. L'axe passant de ce polariseur fait un angle orienté de -30° par rapport à l'axe passant du premier polariseur.

- d) Quelle est la probabilité qu'un photon soit transmis à travers les deux polariseurs ? Faire l'application numérique.

On ajoute un troisième polariseur (*après* les deux autres).

- e) Déterminer l'angle orienté θ de son axe passant (par rapport à l'axe horizontal) pour qu'aucun photon ne soit transmis à travers l'ensemble des trois polariseurs.
- f) On suppose maintenant que l'angle θ de ce troisième polariseur varie dans le temps de façon uniforme : $\theta(t) = \omega t$, où la pulsation ω est constante. Exprimer la probabilité $P_6(t)$ qu'un photon soit transmis à travers l'ensemble des trois polariseurs.

- g) On ajoute un quatrième polariseur (*après* les trois autres) dont l'axe passant fait un angle orienté $\varphi(t) = \omega t + \pi/2$. Exprimer la probabilité $P_7(t)$ qu'un photon soit transmis à travers l'ensemble des quatre polariseurs.

- h) Déterminer les instants t_n pour lesquels la probabilité $P_7(t)$ est maximale (en fonction de ω et d'un nombre entier n).

- i) Reprendre les cinq questions b, d, e, f et g dans le cas où la lumière incidente est composée de photons non-polarisés.

- j) Donner un exemple physique de lumière non-polarisée.

21 Photons polarisés, lame demi-onde

Dans cet exercice, on examine l'effet combiné de polariseurs et d'une lame demi-onde sur des photons qu'on suppose polarisés rectilignement. Les vecteurs d'état $|H\rangle$ et $|V\rangle$ décrivent respectivement un photon polarisé horizontalement et verticalement.

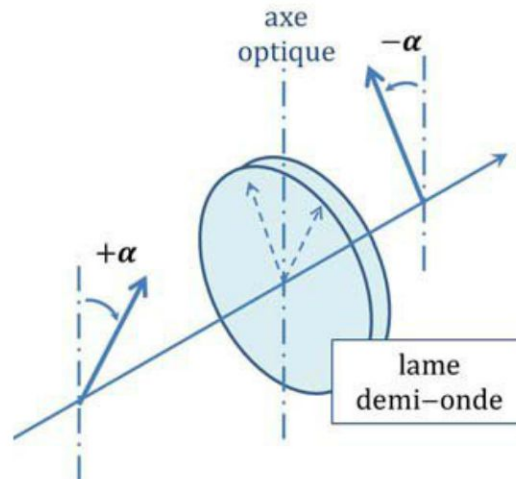


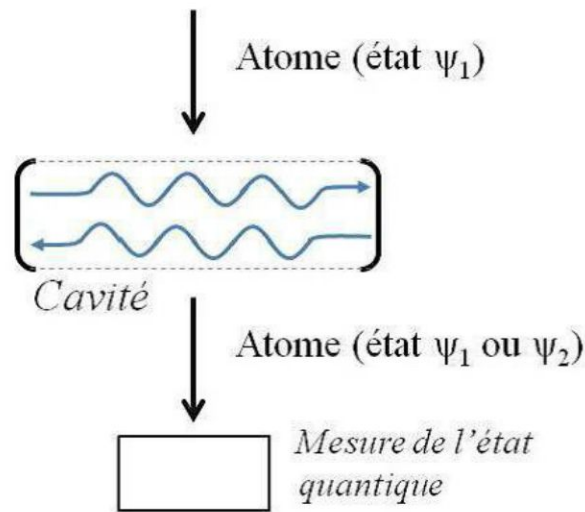
Figure 3 – Une « lame demi-onde » (LDO) est un élément optique taillé dans un cristal biréfringent, c'est-à-dire possédant un axe privilégié qu'on appelle « axe optique ». Cet élément modifie la polarisation d'une onde lumineuse sans en altérer son intensité (contrairement à un polariseur par exemple). Ainsi, lorsqu'une onde lumineuse polarisée rectilignement traverse une lame demi-onde, il n'y a pas d'absorption et la direction de polarisation de l'onde sortante est symétrique de la direction de polarisation incidente par rapport à l'axe optique de la lame demi-onde. Dit autrement, si la polarisation de l'onde incidente fait un angle $+\alpha$ avec l'axe optique de la lame demi-onde, alors la polarisation de l'onde sortante fait un angle $-\alpha$ avec cet axe optique.

- a) La polarisation des photons incidents fait un angle θ_0 avec l'horizontale. Donner l'expression du vecteur d'état $|0\rangle$ décrivant la polarisation de ces photons dans la base $\{|H\rangle, |V\rangle\}$.
- b) On met, sur le trajet de ces photons, un premier polariseur P_1 faisant un angle θ_1 avec l'horizontale.
 1. Donner le vecteur d'état $|1\rangle$ décrivant un photon transmis dans la base $\{|H\rangle, |V\rangle\}$.
 2. Donner la probabilité de passage d'un photon $|0\rangle$ à travers P_1 en fonction de θ_0 et θ_1 .
- c) On enlève ce polariseur P_1 et on le remplace par une lame demi-onde « LDO » dont l'axe optique fait un angle θ_2 avec l'horizontale.
 1. Faire un schéma pour représenter les polarisations entrante et sortante ainsi que la lame demi-onde LDO.

2. Donner le vecteur d'état $|2\rangle$ décrivant la polarisation d'un photon transmis par LDO dans la base $\{|H\rangle, |V\rangle\}$.
- d) On laisse la lame demi-onde en place, et on ajoute le polariseur P_1 (faisant toujours un angle θ_1 avec l'horizontale) après la lame demi-onde.
1. Donner le vecteur d'état $|3\rangle$ décrivant la polarisation d'un photon transmis par LDO et P_1 dans la base $\{|H\rangle, |V\rangle\}$.
 2. Donner la probabilité totale de passage d'un photon à travers ces deux éléments optiques (LDO et P_1), d'abord sous la forme d'un produit scalaire d'états $|i\rangle$, puis en fonction de θ_0 , θ_1 et θ_2 .
 3. Déterminer les angles θ_2 pour lesquels cette probabilité est maximale.
- e) On ajoute maintenant un deuxième polariseur P_2 , mais positionné « avant » la lame demi-onde LDO. Les deux polariseurs sont en configuration « croisée », c'est-à-dire que la direction passante du nouveau polariseur P_2 est orthogonale à celle du premier polariseur P_1 . On fait alors tourner, à vitesse angulaire ω constante, la direction de l'axe optique de la lame demi-onde intercalée entre les deux polariseurs : $\theta_2 = \omega t$. La probabilité de passage d'un photon à travers l'ensemble de ces trois éléments optiques (P_2 , LDO et P_1) varie donc périodiquement dans le temps.
1. Donner le vecteur d'état $|4\rangle$ décrivant la polarisation d'un photon transmis par P_2 seul dans la base $\{|H\rangle, |V\rangle\}$.
 2. Donner le vecteur d'état $|5\rangle$ décrivant la polarisation d'un photon transmis par P_2 et LDO seuls dans la base $\{|H\rangle, |V\rangle\}$.
 3. Donner le vecteur d'état $|6\rangle$ décrivant la polarisation d'un photon transmis par l'ensemble des trois éléments P_2 , LDO et P_1 dans la base $\{|H\rangle, |V\rangle\}$.
 4. Donner la probabilité totale de passage d'un photon à travers ces trois éléments, d'abord sous la forme d'un produit scalaire d'états $|i\rangle$, puis en fonction des angles θ_i du problème.
 5. Déterminer la période d'oscillation de cette probabilité en fonction de ω .

22 Photon et atome dans une cavité

Des expériences récentes de physique quantique, qui ont valu le prix Nobel à S. Haroche, ont permis de piéger un photon entre deux miroirs (une cavité), et de tester sa présence en faisant passer des atomes uniques entre les deux miroirs. Le but de ce problème est de donner les ordres de grandeur de cette expérience, et d'analyser quelques-uns des états quantiques que l'on peut produire avec cette technique.



Photon piégé dans une cavité. Le dispositif central de l'expérience est une cavité : deux miroirs se faisant face, situés à 2,7 cm l'un de l'autre.

- Si l'on injecte de la lumière dans la cavité, elle se retrouve piégée en faisant des allers-retours entre les deux miroirs. Qu'est ce qui limite le temps pendant lequel la lumière reste piégée ?
- Pour que le photon reste piégé dans la cavité, il faut que l'onde associée soit résonnante avec la cavité, donc que la longueur entre les deux miroirs soit un multiple entier de sa longueur d'onde. Dans le cas présent, la longueur de la cavité est égale à trois fois la longueur d'onde du photon. Quelle est la fréquence du photon ? À quel domaine du spectre électromagnétique cela correspond-il ?
- Dans cette expérience, la lumière reste piégée en moyenne pendant 0,18 s. Quelle distance parcourt la lumière pendant ce temps ? Combien effectue-t-elle de rebonds sur chacun des miroirs avant de s'échapper ?

Un atome pour détecter un photon. On considère un atome dont l'électron possède deux états quantiques possibles, $|g\rangle$ et $|e\rangle$, d'énergie respective E_g et E_e (avec $E_g < E_e$). On fait passer l'atome à travers la cavité (voir le dessin plus haut).

- L'état initial de l'atome est $|\psi_1\rangle = \frac{1}{\sqrt{2}}(|g\rangle + |e\rangle)$. Si on mesure l'énergie de l'atome, quelle est la probabilité de mesurer E_g ? Et celle de mesurer E_e ?
- Après son passage dans la cavité, l'état de l'atome devient $|\psi_2\rangle = \frac{1}{\sqrt{2}}(|g\rangle - |e\rangle)$ si un photon était présent dans la cavité, ou reste $|\psi_1\rangle$ sinon. En mesurant l'énergie de l'atome qui sort de la cavité, peut-on savoir si un photon était présent ou non ?
- Expliquer comment, en principe, il est possible de construire un appareil de mesure capable de discriminer entre $|\psi_1\rangle$ et $|\psi_2\rangle$.

- g) L'ensemble du dispositif (passage de l'atome dans la cavité + mesure de l'état des atomes sortant de la cavité) est un appareil de mesure de la lumière dans la cavité. Quelle quantité physique mesure-t-il ? Quelles sont ses limitations ?

Mesures et états possibles. On prépare maintenant la cavité avec un champ faible, de telle sorte que l'état du champ dans la cavité est $|\phi\rangle = \frac{1}{\sqrt{3}}|0\rangle + \sqrt{\frac{2}{3}}|1\rangle$, où $|0\rangle$ représente l'état à zéro photon et $|1\rangle$ l'état à un photon.

- h) On utilise le dispositif décrit précédemment pour analyser l'état du champ dans la cavité. Quelle est la probabilité de mesurer l'atome dans l'état $|\psi_1\rangle$? Et dans l'état $|\psi_2\rangle$?
- i) On suppose qu'on a mesuré $|\psi_2\rangle$. On fait passer un deuxième atome dans la cavité. Quelle est alors la probabilité de mesurer ce deuxième atome dans l'état $|\psi_1\rangle$? Et dans l'état $|\psi_2\rangle$? Justifier votre réponse.
- j) Dans l'expérience, on envoie à intervalles réguliers des atomes à travers la cavité, à un rythme de 900 atomes par seconde. Dans une réalisation particulière, on mesure les 450 premiers atomes dans l'état $|\psi_2\rangle$, puis tous les autres dans l'état $|\psi_1\rangle$. En déduire le nombre de photons présents dans la cavité en fonction du temps, et comparer avec le temps donné en question 3. Commenter.

23 Cryptographie quantique et espionnage

Le but de cet exercice est de calculer en détail l'effet d'une attaque d'un espion sur un protocole de cryptographie quantique.

On rappelle que lors d'un protocole de cryptographie quantique, « Alice » envoie des photons à « Bob ». La polarisation du photon est utilisée pour coder les informations 0 ou 1. Alice peut alors choisir de coder sur une base $\{H, V\}$, où H est 0 et V est 1, ou de coder sur une base $\{D, \bar{D}\}$, où D est 0 et \bar{D} est 1. On rappelle qu'en notation de Dirac $|D\rangle = \frac{1}{\sqrt{2}}(|V\rangle + |H\rangle)$ et $|\bar{D}\rangle = \frac{1}{\sqrt{2}}(|V\rangle - |H\rangle)$.

- a) On suppose qu'Alice envoie le bit 1 codé dans la base $\{H, V\}$. Par quel ket ce photon peut-il être représenté ? Bob fait sa mesure dans la même base, avec quelle probabilité va-t-il trouver 1 ? Et 0 ?
- b) Si Bob avait fait sa mesure dans la base $\{D, \bar{D}\}$, avec quelle probabilité aurait-il trouvé 1 (on demande de faire explicitement le calcul) ? Et 0 ?

Comme, après le protocole, Alice et Bob ne conservent que les événements pour lesquels ils ont utilisé la même base, on considère, pour faire le calcul, qu'Alice et Bob utilisent tous les deux systématiquement la même base $\{H, V\}$.

Une espionne, « Eve », décide de mettre en œuvre une attaque. Pour cela, elle intercepte le photon émis par Alice et le mesure. Après sa mesure, Eve renvoie à Bob un photon correspondant à ce qu'elle a mesuré, en espérant ainsi ne pas

être détectée par Bob ou Alice. Comme Alice, Eve a quatre possibilités pour émettre un photon : $|H\rangle$, $|V\rangle$, $|D\rangle$ ou $|\overline{D}\rangle$.

- c) On suppose d'abord qu'Eve effectue sa mesure dans la base $\{H, V\}$. Avec quelle probabilité mesurera-t-elle correctement le bit envoyé par Alice ? Justifier.
- d) Après sa mesure, Eve renvoie un photon correspondant à celui qu'elle a mesuré. Si Alice avait envoyé 1 (i.e. $|V\rangle$), avec quelle probabilité Eve va renvoyer $|H\rangle$ ou $|V\rangle$? Avec quelle probabilité Bob mesurera-t-il correctement le bit envoyé par Alice ? Eve pourra-t-elle être détectée dans ce cas ?
- e) On suppose maintenant qu'Eve effectue sa mesure dans la base $\{D, \overline{D}\}$, et qu'Alice envoie toujours le bit 1. Avec quelle probabilité mesurera-t-elle 0 ou 1 ?
- f) Si Eve mesure 0, quel état quantique de photon va-t-elle envoyer à Bob ? Avec quelle probabilité Bob mesurera-t-il le même bit que celui envoyé par Alice ? Qu'en serait-il si Eve avait mesuré 1 ?
- g) On cherche maintenant à faire un bilan complet. Sachant qu'Eve a une chance sur deux d'utiliser une base identique à celle d'Alice et Bob lors de sa mesure, avec quelle probabilité Bob mesurera correctement le bit envoyé par Alice ?
- h) À votre avis, comment doivent agir Alice et Bob pour détecter la présence d'Eve ? Comment prendre en compte le fait que même en absence d'Eve, il y a des erreurs expérimentales dues à l'imperfection des appareils ?

24 Effet de la gravité sur la longueur d'onde d'un jet de neutrons

Soit un jet de neutrons ($m_N = 1,7 \cdot 10^{-27}$ kg) vertical dirigé vers le haut suivant l'axe des z croissants. On donne sa longueur d'onde de de Broglie en $z = 0$: $\lambda_0 = 10$ nm.

- a) Calculer la vitesse en $z = 0$ du jet.
- b) Donner l'expression de la variation relative de la longueur d'onde de de Broglie en fonction de z . On mettra le résultat sous la forme $\frac{\lambda(z) - \lambda_0}{\lambda_0}$.
- c) Faire l'application numérique pour $z = 1$ m.

25 Notion de Qubit, porte cNOT

Soient deux qubits, a et b , pouvant prendre les valeurs 0 ou 1 lors de leur mesure. Leurs états respectifs s'expriment de la façon suivante :

$$|a\rangle = \frac{|0\rangle + |1\rangle}{\sqrt{2}}$$

$$|b\rangle = \frac{|0\rangle - |1\rangle}{\sqrt{2}}$$

dans la base orthonormée $\{|0\rangle, |1\rangle\}$ de l'espace des états du problème.

- a) Question préliminaire : dans le cas où les vecteurs d'état $|0\rangle$ et $|1\rangle$ représentent les états de polarisation horizontal et vertical de photons polarisés, quelles sont les directions de polarisation des états $|a\rangle$ et $|b\rangle$?
- b) Exprimer l'état des deux qubits $|ab\rangle = |a\rangle|b\rangle$ en fonction des états $|00\rangle$, $|01\rangle$, $|10\rangle$, $|11\rangle$ (où le premier chiffre représente l'état du qubit a et le second chiffre celui du qubit b), en utilisant la linéarité de l'espace de Hilbert des états de deux qubits.
- c) Dans une mesure du système dans l'état $|ab\rangle$, quelle est la probabilité :
 1. que le qubit a ait la valeur 0 et le qubit b la valeur 1 ?
 2. que le qubit a ait la valeur 0 ?
 3. que les deux qubits aient des valeurs différentes ?
- d) La réalisation d'un calcul informatique nécessite de définir des « opérations logiques ». Un exemple important d'opération logique est la porte cNOT (controlled-NOT) qui agit sur un ensemble de deux qubits, le premier servant de « qubit de contrôle » pour changer, de façon conditionnelle, le second qubit, appelé « qubit cible ». L'opération est simple : si le qubit de contrôle (« a » dans notre cas) est dans l'état $|0\rangle$, le qubit cible (« b » dans notre cas) n'est pas modifié ; par contre, si le qubit de contrôle est dans l'état $|1\rangle$, le qubit cible change d'état. Par exemple :

$$\text{cNOT } |10\rangle = |11\rangle$$

1. Exprimer l'état **cNOT** $|ab\rangle$ en fonction des états $|00\rangle$, $|01\rangle$, $|10\rangle$, $|11\rangle$.
2. Dans une mesure du système dans l'état **cNOT** $|ab\rangle$, quelle est la probabilité que les deux qubits aient des valeurs différentes ?

26 Particule dans un piège harmonique

On considère un électron, de masse m et de quantité de mouvement p , piégé dans un potentiel de type oscillateur harmonique. En mécanique classique, on peut montrer que l'énergie potentielle de l'électron en fonction de sa position x est donnée par $E_p(x) = \frac{1}{2}m\omega^2x^2$, avec ω la pulsation propre de l'oscillateur.

On donne $\hbar = 1.10^{-34} \text{ J.s.rad}^{-1}$, $\omega = 6,4.10^8 \text{ rad.s}^{-1}$ et $e = 1,6.10^{-19} \text{ C}$

On cherche dans un premier temps à déterminer quelle est l'énergie minimale, en mécanique quantique, que peut prendre l'électron dans ce potentiel.

- a) Écrire l'énergie totale $E(x, p)$ de l'électron en fonction de ω , m , x et p .
- b) Écrire la relation d'incertitude de Heisenberg Position-Quantité de mouvement à une dimension.

- c) Les valeurs moyennes $\langle x \rangle$ et $\langle p \rangle$ étant nulles, on peut écrire que $\Delta x = \sqrt{\langle x^2 \rangle}$ et $\Delta p = \sqrt{\langle p^2 \rangle}$. En déduire d'après la relation de la question b que :

$$\langle p^2 \rangle \geq \frac{\hbar^2}{4\langle x^2 \rangle}$$

- d) Écrire alors la valeur moyenne de l'énergie totale $\langle E(x, p) \rangle$ en fonction de $\langle x^2 \rangle$ et $\langle p^2 \rangle$. Montrer que cette énergie possède une valeur minimale, E_{\min} , que l'on exprimera.
- e) On peut montrer que, au minimum de l'énergie, on a $\langle x^2 \rangle = \frac{\hbar}{2m\omega}$. Déterminer l'expression du minimum de l'énergie totale moyenne de l'électron E_{\min} . Faire l'application numérique en J et en eV.
- En réalité, l'énergie totale accessible par l'électron est quantifiée dans le potentiel harmonique. Les énergies de l'électron sont données par $E_n = \hbar\omega(n + \frac{1}{2})$, avec n un entier naturel. L'état propre associé à chaque énergie E_n est représenté par le ket $|n\rangle$.
- f) On prépare l'électron dans une superposition d'états $|\phi\rangle = a|0\rangle + b|1\rangle$, avec a et b deux réels positifs. À quelle relation doivent obéir ces coefficients a et b ?
- g) Si l'on fait une mesure unique de l'énergie du système, peut-on mesurer $\frac{5}{2}\hbar\omega$?
- h) Un ensemble de mesures de l'énergie montrent que l'on a quatre fois plus de chances de mesurer E_0 que E_1 . En utilisant la réponse de la question f, déterminer les valeurs des coefficients a et b .
- i) Donner alors l'énergie moyenne E_{moy} du système en fonction de \hbar et de ω . Faire l'application numérique en eV.
- À chaque état $|n\rangle$ on peut associer une fonction d'onde $\psi_n(x)$ qui est reliée à la position de la particule dans l'état considéré. Pour l'état fondamental $|0\rangle$, la fonction d'onde associée à l'électron est :

$$\psi_0(x) = Ae^{-\frac{m\omega}{2\hbar}x^2}$$

où A est un paramètre indépendant de x .

- j) On se place dans l'état $|0\rangle$. Écrire la probabilité P de trouver la particule entre deux positions x_1 et x_2 .
- k) Selon vous, cette probabilité évolue-t-elle avec le temps ? Justifier.
- l) Pour $x_1 = -\infty$ et $x_2 = +\infty$, que doit obligatoirement valoir cette probabilité P ?
- m) En déduire une expression du coefficient A .

On donne :

$$\int_{-\infty}^{+\infty} e^{-\alpha x^2} dx = \sqrt{\frac{\pi}{\alpha}}$$

27 Évolution quantique et facteur de phase relatif

On considère une particule libre de masse m se propageant dans la direction Ox . On appelle E son énergie totale, égale dans ce cas à son énergie cinétique. On appelle $|\psi_E\rangle$ l'état quantique de la particule d'énergie E et $\psi_E(x) = \langle x|\psi_E\rangle$ la fonction d'onde associée.

- Rappeler l'expression d'une onde classique réelle plane $A(x,t)$ se propageant dans la direction Ox de vecteur d'onde k et de pulsation ω . Donner l'expression complexe de la même onde, telle que $A(x,t) = \Re(\underline{A}(x,t))$, \Re désignant la partie réelle (on donnera l'expression où le coefficient de t est imaginaire négatif).
- On fait l'hypothèse que la fonction d'onde $\psi(x,t)$ décrivant la particule libre est identique à l'expression complexe $\underline{A}(x,t)$ de l'onde associée trouvée dans la question précédente. Rappeler les deux relations exprimant la dualité onde-corpuscule reliant vecteur d'onde et pulsation de l'onde à quantité de mouvement p et énergie E de la particule. Donner l'expression reliant E et p pour une particule libre classique. En déduire la relation liant pulsation et quantité de mouvement pour l'onde associée.
- Montrer qu'on a pour la fonction d'onde $\psi_E(x,t)$ décrivant une particule d'énergie E : $\psi_E(x,t) = e^{-iEt/\hbar}\psi_E(x,0)$. En déduire que si à l'instant $t = 0$ la particule est décrite par le vecteur d'état $|\psi_E\rangle$, elle sera à tout instant ultérieur t décrite par le vecteur d'état $|\psi_E(t)\rangle = e^{-iEt/\hbar}|\psi_E\rangle$.
- On suppose que la particule se trouve à l'instant $t = 0$ décrite par la fonction d'onde :

$$\psi_0(x) = \frac{1}{\sqrt{2}}(\psi_{E_1}(x,0) + \psi_{E_2}(x,0))$$

avec $E_1 \neq E_2$. Donner l'expression de $\psi_0(x,t)$.

- Calculer la probabilité de présence de la particule en un point x à tout instant t . Montrer que celle-ci évolue périodiquement. Exprimer la période en fonction de E_1 , E_2 et \hbar .
- On fait à l'instant t une mesure de l'énergie de la particule décrite par la fonction d'onde $\psi_0(x,t)$. Quels sont les résultats possibles, et avec quelle probabilité ? Ces résultats dépendent-ils du temps ?

28 Boîte quantique, effet de recul et horloge atomique

On considère, dans ce problème, un système physique S de masse totale M dont on supposera qu'il peut se déplacer librement selon l'axe (Ox) d'un référentiel Galiléen \mathcal{R} . Ce système est susceptible d'absorber un photon de fréquence ν dont la direction de propagation est suivant (Ox). Après absorption, le photon a disparu.

- a) 1. Donner la définition de l'énergie interne en relativité. Rappeler les expressions de l'énergie cinétique, de l'énergie interne et de l'énergie totale relativistes. La vitesse de S par rapport à O dans le référentiel \mathcal{R} sera notée v .
2. Dans quelles circonstances la masse d'un système physique est-elle susceptible de changer ? Donner un exemple.
3. Quelle est la formule reliant énergie totale, quantité de mouvement et masse en relativité ?
4. Donner l'expression de la quantité de mouvement du photon en fonction de sa fréquence ν .
- b) 1. On suppose tout d'abord que S est une particule élémentaire, comme l'électron, sans structure interne. Que peut-on dire de sa masse ?
2. Avant l'absorption du photon, S est supposé au repos, et son énergie totale vaut E . Après absorption, p' est sa quantité de mouvement (selon Ox) et E' son énergie totale. Écrire les conservations de l'énergie et de la composante le long de (Ox) de la quantité de mouvement dans le processus d'absorption du photon.
3. Montrer que ces deux équations ne peuvent pas être simultanément satisfaites. Conclusion ?
- c) 1. S est maintenant un système ayant une structure interne, un atome ou une boîte quantique par exemple. Cette structure est susceptible de changer lors de l'absorption du photon. On appelle respectivement M et M' sa masse avant et après son interaction avec le photon. Écrire à nouveau les conservations de l'énergie et de la composante le long de (Ox) de la quantité de mouvement dans le processus d'absorption du photon par l'atome, en supposant toujours l'atome initialement au repos.
2. Montrer que ces deux équations peuvent maintenant être simultanément satisfaites, et donner l'équation permettant de déterminer M' à partir de M et ν .
3. On suppose que M' est proche de M , et on pose $M' = M + \delta M$. Trouver δM par un développement limité à l'ordre le plus bas. Quelle est la vitesse v (supposée petite devant c) acquise par S après absorption ?
4. Déterminer numériquement δM et v pour un photon jaune de longueur d'onde $\lambda = 0.6 \mu\text{m}$ et un système de masse $M = 10^{-25} \text{ kg}$. Les hypothèses utilisées sont-elles vérifiées ?
- d) 1. Le système mobile à structure interne S a maintenant des propriétés quantiques : c'est par exemple une boîte quantique, ou un atome. Pour simplifier, S est traité comme une boîte quantique à une dimension de masse M

enfermant une particule de masse $m \ll M$ dans un puits de potentiel infini à une dimension de largeur a . Quelles sont les longueurs d'onde possibles de l'onde de de Broglie associée à la particule enfermée dans la boîte quantique ? L'énergie de cette particule peut-elle prendre n'importe quelle valeur ? Rappeler brièvement le raisonnement qui permet de déterminer l'énergie des deux premiers états E_1 et E_2 de la boîte quantique et donner leur expression. On pose $E_2 - E_1 = h\nu_0$.

2. Déterminer numériquement à 1 % près ν_0 pour $m = 10^{-30}$ kg, $a = 10^{-9}$ m.
3. On revient au problème de l'absorption du photon par la boîte quantique susceptible de se déplacer dans son ensemble et initialement dans son état interne de plus basse énergie E_1 . À cause de la quantification de l'énergie étudiée dans la question d.1 la boîte après absorption du photon se trouve dans son état excité d'énergie E_2 , et sa masse après absorption ne peut prendre que la valeur $M' = M + h\nu_0/c^2$. Dédurre de la question c.2 que la fréquence du photon susceptible d'être absorbé par la boîte quantique libre de se déplacer est donnée par :

$$\nu = \nu_0 + \frac{h\nu_0^2}{2Mc^2}$$

Justifier le nom d'« effet de recul » donné à cet effet.

4. La précision relative actuelle des mesures de fréquence de la lumière est de 10^{-12} . L'effet de recul est-il mesurable avec la valeur numérique déjà trouvée pour ν_0 et $M = 10^{-25}$ kg ?
- e) 1. On appelle respectivement $|\psi_1\rangle$ et $|\psi_2\rangle$ les états quantiques de la boîte quantique d'énergies respectives E_1 et E_2 , et on dispose d'un dispositif de mesure de cette énergie sur chaque boîte quantique individuelle. On peut ajuster les propriétés de la lumière absorbée par la boîte pour que l'état de la boîte après absorption, à $t = 0$, soit :

$$|\psi\rangle = \cos\theta|\psi_1\rangle + \sin\theta|\psi_2\rangle$$

Déterminer la probabilité de trouver la boîte dans son niveau d'énergie E_2 . Dans quel état se trouvera le système après la mesure ? Pour quelles valeurs de θ peut-on prédire avec certitude le résultat d'une mesure d'énergie sur une boîte donnée ?

2. En déduire l'énergie moyenne \bar{E} du système lorsqu'on effectue un très grand nombre de mesures successives sur des boîtes quantiques identiques toutes préparées dans l'état $|\psi\rangle$. Déterminer aussi l'incertitude de cette mesure, en calculant sa variance $\Delta E^2 = (\bar{E}^2) - (\bar{E})^2$. Pour quelle valeur de θ l'incertitude est-elle maximale ? minimale ?

- f) Vous disposez d'une boîte quantique, d'une source de photons de fréquence contrôlable, et d'un dispositif de mesure de l'énergie de la boîte après interaction avec les photons. Indiquez comment construire à partir de ces éléments une source de lumière de fréquence parfaitement contrôlée (« horloge atomique »).

BIBLIOGRAPHIE

Pour l'ensemble du cours :

- BENSON H., *Physique 3 : Ondes, Optique et Physique Moderne*, ed. De Boeck Université (2009). Bonne introduction, livre très bien illustré.
- MANLY S., *Relativity and Quantum Physics for beginners*, For beginners LLC (en anglais) (2009). Présenté sous forme de bande dessinée, une excellente introduction au sujet.
- TIPLER P., *Physics for scientists and engineers*, V. 3 (en anglais) (2003). Couvre bien l'ensemble du cours.

Pour la partie relativité :

- EINSTEIN A., *Œuvres choisies - Relativités I*, Édition du Seuil (1989). La traduction en français de certains articles d'Einstein : lisible et passionnant.
- GOURGOULHON E., *Relativité restreinte, des particules à l'astrophysique*, coll. « Savoirs Actuels », EDP Science (2010). Tout y est ! De plus, un site web associé à cet ouvrage donne notamment de nombreux liens très intéressants : <http://relativite.obspm.fr/>
- HULIN M., HULIN N., MOUSSELIN L., *Relativité restreinte*, Dunod (1998). L'essentiel est exposé en peu de pages.
- SCHWINGER J., *L'héritage d'Einstein*, Pour La Science/Belin (1988). Écrit par l'un des plus grands physiciens du XX^e siècle, un ouvrage d'introduction bien illustré et à la portée de tous, tout en étant rigoureux.
- SIMON Y., *Relativité restreinte*, Vuibert (2004). Très complet mais plus technique.

Pour la partie quanta :

- BASDEVANT J.-L., DALIBARD J., *Mécanique quantique*, Éditions de l'École Polytechnique (2006).
- COHEN-TANNOUDJI C., DIU B., LALOE F., *Mécanique quantique*, Hermann (1997). Livre de niveau plus élevé, lire seulement le chapitre 1.
- LE BELLAC M., *Physique quantique*, EDP Sciences (2007). Livre de niveau plus élevé, lire seulement les chapitres 1 et 3.
- LE BELLAC M., *Le monde quantique*, coll. « Une introduction à... », EDP Sciences (2010). Une excellente introduction, qualitative, aux concepts de la physique quantique et à ses récents développements.
- SCARANI V., *Initiation à la Physique quantique*, Vuibert (2006).

Ouvrages de vulgarisation ou de culture générale en physique moderne :

BRÉZIN É., *Demain la physique*, Odile Jacob 2004). Une très bonne description, à la portée de tous, des avancées et des perspectives de la physique moderne, notamment sur la physique quantique.

PETIT J.-P., *Tout est relatif*, téléchargeable gratuitement sur Internet. Agréable à lire, permet de se faire une bonne idée du domaine.

PIETRYK G. (dir.), *Panorama de la Physique*, Pour La Science/Belin (2012). Une mine de renseignements sur la physique actuelle mise à la portée de la Licence.

INDEX

A

aberration
classique, 17
relativiste, 60
activité 108

B

biprisme de Fresnel 153, 156
boîte quantique 203, 206, 212, 230, 231

C

champ
électrique, 20, 141, 152, 159, 160, 269
magnétique, 20, 21, 65, 159, 172
collisions relativistes 87
composition des vitesses
classique, 12
relativiste, 55
contraction des longueurs 43, 48, 49
cryptographie 177, 179, 180, 274

D

de Broglie 191, 196, 200, 202, 249
décalage vers le rouge 119, 121
défaut de masse 97, 100
déterminisme 152, 239, 241, 252
diffraction
d'ondes de matière, 197, 198
d'ondes lumineuses, 142
de neutrons, 193
dilatation des durées 46, 48, 52, 121
Dirac 142, 167, 171, 249

E

effet Compton 92, 93
effet Doppler
classique, 16
relativiste, 60
effet photoélectrique 141, 143

Einstein 10, 31, 81, 114, 122, 141, 143
émission/absorption de photons 145
énergie cinétique
classique, 70
relativiste, 78
équation d'onde 15, 38, 250
équation de Dirac 250
équation de Klein-Gordon 250
équation de Schrödinger 183, 185, 186, 215, 218, 220, 222, 233, 238
équations de Maxwell 20, 141
espace de Hilbert 168, 174, 185, 202, 216, 235–237, 242
espace-temps 37, 41, 43, 85, 86, 112, 122, 123, 250
état après la mesure 177, 216, 240, 241
état quantique 166, 167, 171
événement 10, 43

F

fission nucléaire 104
fonction d'onde 15, 191, 200–203, 208, 210, 211, 215, 217, 218, 226, 232, 237
forces d'inertie 10, 113
Fresnel 14, 15, 201
fusion nucléaire 105

G

Galilée 8, 9, 13, 22
GPS 4, 34, 63

H

Hamiltonien 186, 232, 233, 238, 239, 242
Heisenberg 210
Hertz 141, 143
horloge 3, 9, 11, 32–34, 46, 48, 50, 51, 53, 63, 65, 121

I

inégalités de Heisenberg 126, 210–213, 231, 246, 252
interaction
 électromagnétique, 127
 faible, 129
 forte, 127
 gravitationnelle, 111, 112
interférences
 électroniques, 199
 d'ondes de matière, 196, 197
 lumineuses, 153
 moléculaires, 199
 neutroniques, 193
 photon par photon, 154
interféromètre
 à ondes de matière, 199
 de Mach-Zehnder, 158
 de Michelson, 24, 26, 125, 151, 157, 255
interprétation de Born 201, 216
intervalle d'espace-temps 41, 42
invariant relativiste 38, 41, 59, 72, 84, 89

J

jumeaux de Langevin 48

L

loi de conservation 71, 76, 125
loi de Malus 161, 165, 208
loi de Poisson 150, 151
longueur d'onde de de Broglie 192–194, 197, 249, 275
longueur propre 43, 256
Lorentz 35, 141

M

marche de potentiel 224
masse au repos 128
Maxwell 14, 141, 159
Michelson 24, 27, 31

N

nature probabiliste 106, 151, 152, 165, 168, 171, 202, 215, 238, 252
neutron 91
Newton xi, 8, 14, 111, 141

O

observable 227, 232, 238, 242, 243
onde de matière 191, 192, 199, 200, 217, 224
ondes gravitationnelles 123
oscillation de neutrinos 186

P

période radioactive 107
photon 56
Planck 141
polarisation
 description classique, 159
 description quantique, 164
polariseur 161
principe d'équivalence 114–117, 121
principe d'inertie 122
principe de relativité
 générale, 121
 galiléenne, 9
 restreinte, 31
principe de superposition 15, 167, 174, 179, 216, 236
principe fondamental de la dynamique
 classique, 9
 relativiste, 77
puits carré infini 220, 221, 231

Q

qubit 205, 275

R

réactions nucléaires 99
réduction de la fonction d'onde 216, 241
référentiel galiléen 7
radioactivité 95, 105
relativité générale 111, 121