Traitement Quantique de l'Information et Calcul Quantique

Draft version 2022 - 2023

Nicolas Macris

Contents

Part I

Introduction à la Théorie Quantique

1 La Dualité Onde-Particule

La mécanique quantique fut établie à travers un long processus expérimental et conceptuel au début du 20ème siècle ($\sim$ 1900-1930). En 1900, l'édifice de la physique classique semblait complet avec d'une part les lois de la mécanique Newtonienne décrivant le mouvement des particules matérielles, et d'autre part la théorie de Maxwell décrivant tous les phénomènes électromagnétiques. La distinction entre particule et onde était nette. Par exemple, on considérait que l'électron est une particule possédant une position $\vec{x} = (x, y, z) \in \mathbb{R}^3$ bien définie, une vitesse et quantité de mouvement (ou impulsion) bien définie $\vec{p} = m\vec{v}$ ($\vec{v} = \frac{\mathrm{d}\vec{x}}{\mathrm{d}t}$). Étant donné des conditions initiales $\vec{x}(0)$ et $\vec{v}(0)$ on peut décrire grâce à l'équation différentielle de Newton la trajectoire $\vec{x}(t)$ (et aussi $\vec{v}(t)$). La précision attribuée à cette description n'est qu'une "affaire de technologie" et on peut espérer repousser toute incertitude grâce à des instruments de plus en plus précis. D'autre part, suite aux travaux de Maxwell et aux expériences de Hertz, il était établi que la lumière visible est une onde électromagnétique, de même nature que les ondes radio; la seule différence étant l'ordre de grandeur de la longueur d'onde et la fréquence. Les ondes électromagnétiques sont des vibrations ondulatoires des champs électriques et magnétiques décrites par les équations (aux dérivées partielles) de Maxwell. Celles-ci sont parfaitement déterministes et l'évolution du champ électromagnétique est connu en tout temps, si on sait fixer les conditions initiales.

Les expériences qui furent à l'origine d'une véritable révolution conceptuelle, et menèrent à un changement complet de paradigme, concernent l'interaction de la matière et du rayonnement.

Tout d'abord en 1900, la distribution spectrale des fréquences dans un "corps noir" mena Planck à considérer que la lumière est absorbée et émise par les parois matérielles en quantités discrètes. La théorie électromagnétique des ondes prédisait un échange continu d'énergie et était incapable de reproduire le bon spectre de fréquence du corps noir[1].

Ensuite, "l'effet photoélectrique" (Section **1.2**) fut clairement mis en évidence par Lénard en 1902 (suite aux travaux antérieurs de Hertz $\approx$ 1885) et conduisit Einstein à postuler que la lumière est en fait constituée de corpuscules, aujourd'hui appelés "photons".

L'observation de raies spectrales (d'émission ou d'absorption) discrètes pour divers éléments chimiques indiquait aussi que l'échange d'énergie entre atomes et radiation est

[1] Un "corps noir" est une cavité matérielle (comme un four) dont les parois sont en équilibre thermique avec la radiation. Nous n'en dirons pas plus dans ce cours

de nature discrète. Bohr (1910) repris la nouvelle idée de photon et proposa le "modèle de Bohr de l'atome". Sa théorie permettait d'expliquer les raies spectrales connues et même d'en prédire de nouvelles qui furent observées beaucoup plus tard. (Section **1.6**)

Étant donné que la lumière semblait avoir une nature à la fois corpusculaire et ondulatoire, De Broglie postula que cela pourrait être le cas pour les électrons aussi; et en fait pour toute forme de matière. En 1924, il proposa une formule associant une longueur d'onde (appelée aujourd'hui longueur d'onde de De Broglie) aux électrons. Grâce à cette formule, les résultats de Bohr à propos des raies spectrales (de l'hydrogène) pouvaient être reproduit. Les idées de De Broglie furent confirmées expérimentalement par Davisson et Germer par des expériences de diffraction d'électrons sur des cristaux. Ces expériences (1927) confirmèrent de façon éclatante et surprenante que les électrons peuvent parfois se comporter comme des ondes.

Entre 1925-1930, la théorie quantique moderne, encore universellement utilisée aujourd'hui, fut développée par Schrödinger, Heisenberg, Born, Jordan, Dirac (et d'autres...). Schrödinger (1926) élabora (à partir des idées de De Broglie) l'équation régissant l'évolution temporelle de "l'onde de De Broglie" associée aux particules (Section **1.7**). Cela lui permit de calculer de façon fondamentale le spectre de l'atome d'hydrogène. Aujourd'hui, on sait que l'équation d'onde de Schrödinger est à la base de la chimie. Heisenberg (1925-26) développa une approche différente, plus directement basée sur les propriétés connues à l'époque de l'interaction entre atomes et rayonnement (par exemple les raies spectrales, etc.). Son approche était abstraite et conduisait à décrire la position et l'impulsion des électrons orbitant autour du noyau par des matrices. Sa mécanique était appelée "mécanique des matrices". Born, Jordan et Dirac montrèrent très rapidement que les approches de Schrödinger et Heisenberg étaient en fait des formalismes mathématiquement équivalents.

Aujourd'hui, les physiciens font à peine la distinction entre les deux approches, ondulatoire et algébrique. En fait, elles furent largement unifiées dans les travaux de Dirac et von Neumann ($\sim$ 1930-1932) qui présentèrent une formalisation et un cadre mathématiquement précis des lois de la mécanique quantique. Cette formalisation sera présentée dans le Chapitre **3**. Elle reste à ce jour essentiellement inchangée et donne un cadre mathématique clair pour la description des phénomènes. Ce Chapitre **1** ainsi que le Chapitre **2** donnent une première introduction à quelques concepts à partir de la phénoménologie expérimentale. Ainsi, la formalisation mathématique de la théorie quantique paraîtra moins arbitraire.

Le cheminement historique brièvement décrit ci-dessus est à la fois trop long et compliqué pour être décrit en détail ici. Nous allons donc introduire deux aspects essentiels de la phénoménologie quantique en sélectionnant deux expériences clés. Celles-ci sont : l'expérience des fentes de Young (Section **1.1**) et l'effet photoélectrique (Section **1.2**).

1.1 Les expériences de doubles fentes

L'expérience de Young - 1803

La nature ondulatoire de la lumière fut d'abord mise en avant par Hooke, Huygens et Euler. Néanmoins, Newton pensait que celle-ci était constituée de corpuscules et c'est cette conception qui domina jusqu'au 19ème siècle. Ce débat fut (provisoirement!) clos par Young en 1803 qui établit de manière expérimentale que la lumière est une onde et fut notamment capable de déterminer la longueur d'onde de la lumière visible (rouge, vert, bleu, etc).

Le schéma de l'expérience est esquissé sur la Figure **1.1**.

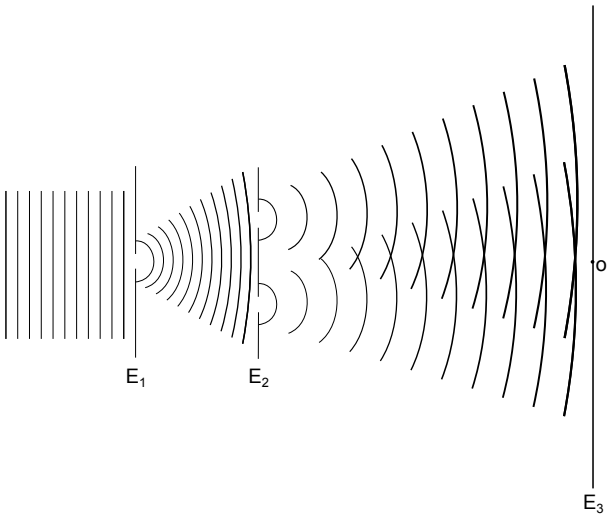


Figure 1.1 Expérience de doubles fentes.

Un faisceau monochromatique est envoyé sur un écran E_1 percé par une fente. Le faisceau est diffracté et arrive en E_2. Cette étape sert simplement à travailler avec une source cohérente ponctuelle constituée par la fente de E_1 . La lumière est ensuite diffractée par les deux fentes de E_2 qui se comportent comme deux sources ponctuelles. Si $\vec{r}_1$ et $\vec{r}_2$ sont les positions des deux fentes, les deux ondes sphériques sortantes des fentes ont la forme

$$\psi_1(\vec{r}) = A\frac{e^{i(k|\vec{r}-\vec{r_1}|-\omega t)}}{|\vec{r}-\vec{r_1}|} \tag{1.1}$$

et

$$\psi_2(\vec{r}) = A\frac{e^{i(k|\vec{r}-\vec{r_2}|-\omega t)}}{|\vec{r}-\vec{r_2}|} \tag{1.2}$$

Ici $\psi_1(\vec{r})$ et $\psi_2(\vec{r})$ sont les amplitudes des deux ondes au point $\vec{r}$ de l'espace. Le nombre d'onde k est relié à la longueur d'onde par $k = \frac{2\pi}{\lambda}$ et ω est relié à la fréquence de l'onde $\omega = 2\pi\nu$. Pour la lumière, on a $\lambda\nu = \frac{\omega}{k} = c$ où $c \approx 2.997 \times 10^8$ m/s est la vitesse de la lumière (dans le vide). La figure montre des cercles qui représentent les maxima des deux amplitudes (l'espace entre les cercles correspond aux minima). Ces deux ondes "se superposent". Le principe de superposition de la théorie des ondes stipule que l'amplitude totale au point $\vec{r}$ est donnée par

$$\psi(\vec{r}) = \psi_1(\vec{r}) + \psi_2(\vec{r}). \tag{1.3}$$

En d'autres termes, les maxima se renforcent quand les cercles s'intensifient. Cela produit les "figures d'interférence" avec lesquelles nous sommes tous familiers à condition d'être un peu observateur. En effet, des figures d'interférence peuvent être observées en jetant deux cailloux sur la surface d'un lac plat! Toujours est-il qu'une observation plus précise est obtenue, comme le fît Young, en récoltant l'intensité lumineuse sur l'écran E_3. L'intensité récoltée est donnée par $|\psi(\vec{r})|^2$ où $\vec{r} \in E_3$. Il peut être montré que

$$|\psi(\vec{r})|^2 \simeq \frac{4A^2}{D^2}\cos^2(\frac{\pi d}{\lambda}\frac{\rho}{D}), \qquad D >> d \tag{1.4}$$

où D est la distance entre E_2 et E_3, d la distance entre les deux fentes ($d = |\vec{r_1} - \vec{r_2}|$) et ρ la variable radiale mesurée à partir du centre de l'écran (point O). Les franges d'interférence sont circulaires. La Figure **1.2** représente une coupe radiale de l'intensité en fonction de ρ.

Les maxima de l'intensité se trouvent sur les cercles concentriques de rayons $\rho_n = n\lambda\frac{D}{d}$. La distance entre deux maxima est $\rho_{n+1} - \rho_n = \lambda\frac{D}{d}$. En mesurant cette distance, Young était capable de déterminer λ. Pour la lumière visible λ est de l'ordre de 600 à 400 nm (1 nm = 10^{-9} m).

Que se passerait-il si le faisceau lumineux était constitué d'un ensemble de projectiles ("des grains de lumière") ou de corpuscules? La prédiction "naïve classique" donnerait un élargissement du faisceau au passage des fentes à cause des collisions avec le bord des fentes. On s'attendrait à trouver plus de particules au centre de chaque faisceau diffracté et moins au bord (Figure **1.3**). L'intensité récoltée en E_3 ne présenterait pas de franges d'interférences. Vous pouvez essayer de faire l'expérience avec des balles de tennis ou de ping-pong. Vous pouvez lancer ces balles une par une, ou toutes ensemble à travers les fentes; cela ne change pas grand-chose aux résultats de la Figure **1.3**.

L'intensité récoltée sur E_3 (nombre de balles en fonction de la position) à la forme $N_1(\vec{r}) + N_2(\vec{r})$ (Figure **1.4**).

Mais la nature est très surprenante en fait! Dans le paragraphe suivant, nous discutons des expériences modernes de Young faites avec des électrons et même des molécules.

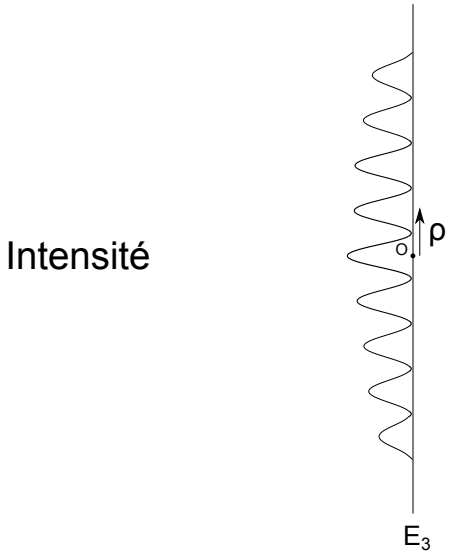


Figure 1.2 Intensité en E_3 de la figure d'interférence.

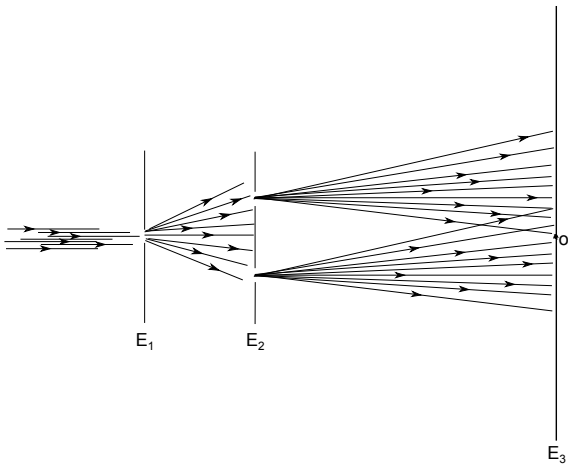


Figure 1.3 Expérience avec des projectiles.

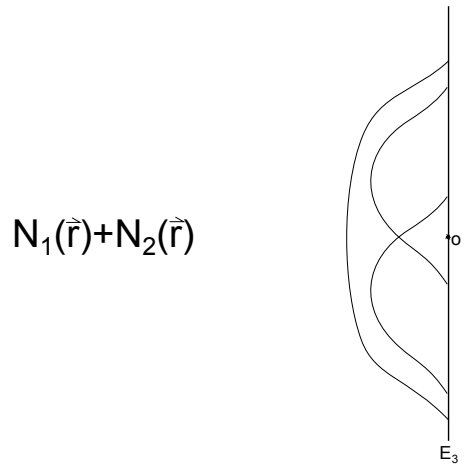


Figure 1.4 Intensité en E_3 avec des projectiles.

Expériences modernes de doubles fentes - post 1960

En 1909, G. Taylor répéta l'expérience de Young avec une source de lumière très faible, correspondant à la lumière d'une bougie placée à environ 1 km de l'écran! La figure d'interférence est toujours observée. à l'époque, on ne pouvait pas conclure grand-chose de cette expérience, mis à part le fait que la nature ondulatoire de la lumière est valable même pour des intensités extrêmement faibles. Nous allons voir qu'en fait l'expérience de Taylor préfigure des expériences modernes absolument remarquables faites avec des particules. En effet, comme nous le verrons dans le paragraphe sur l'effet photoélectrique, on sait aujourd'hui que les sources d'intensité très faibles sont des sources de photons (corpuscules de lumière).

En 1961, C. Jönsson parvint à réaliser l'expérience des doubles fentes avec des électrons. Les franges d'interférence sont observées et cela suggère que les électrons se comportent comme une onde. En fait, une chose très surprenante est aussi observée : si nous envoyons les électrons un par un à travers les fentes on observe des points d'absorption aléatoires sur l'écran E_3. En attendant un certain temps, on observe que l'ensemble de ces points forme une figure d'interférence. Ainsi, il semble que l'électron est ponctuel au moment où son absorption est observée sur l'écran. Mais la distribution statistique des points d'absorptions satisfait à la figure d'interférence! Les électrons se comportent donc aussi comme des ondes.

Cette expérience a été réalisée aussi avec des neutrons et en 1999 avec des molécules de Carbone 60. Ces molécules sont "grosses", elles contiennent 60 atomes de carbone arrangés de manière sphérique sur les sommets de 12 pentagones et 20 hexagones (elles sont donc des minis ballons de foot). Le diamètre d'une telle molécule est d'environ 0.7 nm (alors que l'on ne sait pas associer une dimension à l'électron; on dit qu'il est "ponctuel"). Encore plus récemment, des expériences ont été réalisées avec d'autres

types de molécules artificielles contenant entre 400 et 1000 atomes. À chaque fois, sous certaines conditions expérimentales, les franges d'interférence sont observées! Ces grosses molécules se comportent donc comme des ondes. La distance entre les maxima des franges d'interférence permet d'associer une longueur d'onde λ à ces molécules (en acceptant la formule $\rho_{n+1} - \rho_n = \lambda \frac{D}{d}$ dérivée précédemment). Étonnamment, on trouve une longueur d'onde beaucoup plus petite (quelques centaines de fois plus petite) que la taille des molécules elles-mêmes. La description ondulatoire du passage des molécules à travers ces fentes a-t-elle encore un sens? La physique quantique moderne stipule que oui! (Il est permis d'être sceptique; mais en même temps il faut savoir que la physique quantique passe tous les tests expérimentaux par ailleurs).

Ces expériences confirment de façon éclatante que les particules matérielles possèdent sous certaines conditions un comportement ondulatoire. Mais pourquoi est-ce que cela n'est pas le cas avec une balle de tennis ou de ping-pong ou un ballon de foot; pourquoi n'observe-t-on pas de franges d'interférence? Où se situe la limite entre l'électron, le C60, les molécules à 400-1000 atomes et les ballons de foot ou les corps macroscopiques? Cette question est profonde et mal comprise. On ne sait pas très bien répondre à la question de savoir où se situe la limite entre comportement quantique dual "ondulatoire-corpusculaire" et le comportement classique non dual (ondulatoire ou corpusculaire). Les expériences modernes des doubles fentes avec les grosses molécules permettent d'étudier cette question, et c'est là que réside tout l'intérêt de ces expériences.

Il a été mis en évidence que lorsque les molécules de C60 interagissent trop fortement avec leur environnement (par exemple en échangeant de la radiation avec leur environnement) les franges d'interférence disparaissent. En d'autres termes, la caractéristique ondulatoire est préservée à condition que les molécules de C60 soient suffisamment bien isolées de leur environnement. Lorsque ceci n'est pas le cas, les physiciens parlent de "décohérence". La décohérence est le processus de perte de cohérence des ondes quantiques à travers l'interaction d'un système avec son environnement. Une très grosse molécule, ou un ballon de football, est constamment en interaction avec l'environnement et possède donc un comportement classique.

Comme nous le verrons à la fin de ce cours, construire un ordinateur quantique serait un peu comme réaliser une expérience de Young avec des ballons de football tellement bien isolés de leur environnement que des franges d'interférence seraient observées! En effet, la plupart des physiciens pensent que les comportements quantiques s'appliquent à toute forme de matière, à toute échelle, tant que le système est suffisamment bien isolé pour que la décohérence n'opère pas.

1.2 L'effet photoélectrique

L'effet photoélectrique concerne l'interaction de la lumière avec la matière. Quand la fréquence de la radiation devient assez grande (typiquement à partir des ultraviolets, mais aussi dans le visible et dans les infrarouges) on observe que cette interaction est de nature discrète. Il existe plusieurs effets qui mettent cela en évidence (L'effet Compton,

la création de paires, etc), mais historiquement l'effet photoélectrique fut le premier à être découvert.

En 1885, Hertz étudie l'éclair produit par des ondes radio sur une bobine. Il observe que curieusement la longueur de l'éclair est plus courte quand l'installation est placée dans une chambre noire ; et plus longue quand l'expérience est réalisée à la lumière du jour. À l'époque, ces résultats étaient peu clairs et Hertz abandonna cette expérience. On sait aujourd'hui que les rayons ultraviolets contenus dans la lumière du jour contribuaient à intensifier l'éclair en arrachant des électrons aux atomes environnants. Ce n'est qu'en 1902 que cet effet fut observé clairement par Lénard qui illuminait des gaz avec de la lumière ultraviolette. Au-dessus d'une fréquence critique (dans l'ultraviolet) des électrons sont arrachés aux atomes de gaz, un courant est observé dans un circuit couplé au système. De plus, Lénard parvint à montrer que l'énergie cinétique des électrons arrachés augmentait avec la fréquence de la radiation. Cette énergie cinétique semblait indépendante de l'intensité de la radiation (le courant électrique, ou en d'autres termes le nombre d'électrons arrachés par unité de seconde, est lui proportionnel à l'intensité de la radiation).

Ces résultats furent interprétés par Einstein dans un de ses fameux articles en 1905. Celui-ci postula que la radiation est constituée de corpuscules - aujourd'hui appelés photons. Il associa à ces photons une énergie et une quantité de mouvement

$$E = h\nu \qquad \text{et} \qquad p = \frac{h}{\lambda}, \tag{1.5}$$

où λ et ν sont la longueur d'onde et la fréquence de radiation. Ici $h \simeq 6.63 \times 10^{-24}$ J·s est la constante de Planck. En fait, Planck avait déjà introduit une idée connexe, ainsi que la constante h, dans son étude du corps noir. Celui-ci supposait que les échanges d'énergie entre la matière et la radiation électromagnétique sont discrets et multiples de $h\nu$. Néanmoins, Planck ne concevait pas que la radiation électromagnétique puisse être constituée de particules; les photons. Ainsi, c'est Einstein qui a introduit le premier la dualité onde-corpuscule pour la radiation électromagnétique.

Selon Einstein, l'énergie cinétique des électrons arrachés au métal vaut (Figure **1.5**)

$$\frac{1}{2}mv^2 = \begin{cases} h\nu - W_0 & h\nu > W_0, \\ 0 & h\nu < W_0. \end{cases} \tag{1.6}$$

W_0 est l'énergie minimale qu'il faut pour arracher un électron au gaz. Si on pose $h\nu = W_0$ on trouve la fréquence critique $\nu_0 = \frac{W_0}{h}$ en deçà de laquelle il n'y a pas d'effet photoélectrique. L'expression ci-dessus repose sur 3 hypothèses; (i) la conservation de l'énergie; (ii) le fait que l'énergie d'un photon vaut $h\nu$ et (iii) W_0 est indépendant de la fréquence et de l'intensité de la radiation. La conservation de l'énergie est une loi universellement valable en physique et n'a jamais été remise en question jusqu'ici. En revanche (ii) et (iii) ne sont pas évidents a priori. L'hypothèse (iii) est valable dans un régime approprié et n'a pas le statut de loi fondamentale. D'ailleurs il serait assez difficile de déterminer W_0 par un calcul fondamental. Quant à (ii), Énergie du photon = $h\nu$ est une loi fondamentale de la physique.

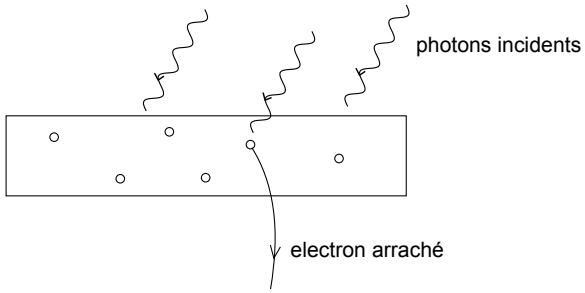


Figure 1.5 Électron arraché avec des photons.

Bien que simple, en 1905 cette formule était révolutionnaire. En effet, elle stipule que la radiation est constituée de quanta élémentaires (les photons) et que leur énergie dépend (linéairement) de la fréquence uniquement; et non pas de l'intensité lumineuse. La théorie de Maxwell, quant à elle, prédit que l'énergie est proportionnelle à l'intensité. Il est possible de réconcilier ces théories en réalisant qu'une onde électromagnétique classique doit être associée à plusieurs photons dont l'énergie est $Nh\nu$. On montre alors que le nombre de photons N est relié à l'intensité de l'onde.

La relation linéaire en fréquence qu'Einstein (suite aux travaux de Planck) postula n'était pas évidente à partir des résultats expérimentaux de Lénard. Ce n'est qu'en 1934 que R.A. Millikan réussit à faire des expériences bien contrôlées qui vérifièrent cette relation linéaire.

Millikan réussit aussi à déterminer la constante de Planck expérimentalement à partir de la pente de la courbe (une droite) de la Figure **1.6**. (La valeur numérique de $h = 6.63\times10^{-24}$ J $\cdot$ s. Notons que 1 J $= 6.2\times10^{18}$ eV et que $13,6$ eV est l'énergie nécessaire pour arracher un électron à l'atome d'hydrogène.)

L'expérience de Young et l'effet photoélectrique mettent en évidence des aspects complémentaires du comportement de la lumière. Ensemble, ils établissent que la lumière possède un comportement dual. Au début du 20ème siècle, ceci était tellement révolutionnaire que la théorie d'Einstein mis du temps à être acceptée, même après les expériences de Millikan.

Revenons un instant à l'expérience de Taylor de 1909. Celui-ci observait une figure d'interférence pour une lumière d'intensité si faible qu'il devait attendre environ six mois avant de voir les franges d'interférence se constituer. La raison est que les photons

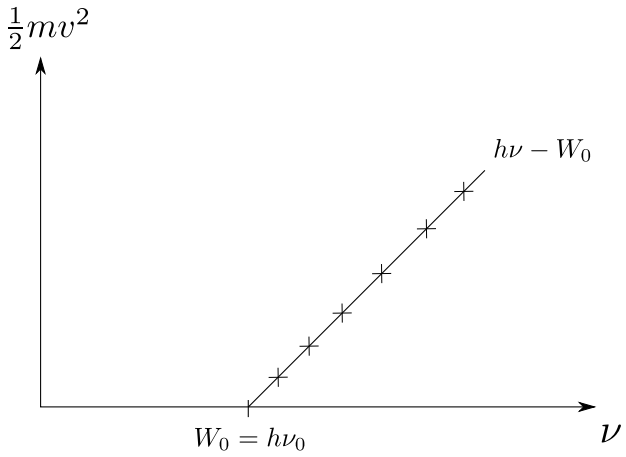


Figure 1.6 Relation linéaire de l'expérience de Millikan

arrivent par petits nombres sur les deux fentes puis l'écran. De nos jours, il est possible de contrôler assez précisément des sources de photons essentiellement uniques.

1.3 La "fonction d'onde": une révolution conceptuelle

L'effet photoélectrique et l'expérience de Young montrent que la lumière (le champ électromagnétique) possède un comportement dual. En 1924, De Broglie conjectura que ceci pouvait aussi être le cas pour des électrons et en fait pour toute matière. Les expériences d'interférence modernes (de double fente) montrent de façon spectaculaire que cela est bien le cas.

De Broglie associa à la particule une longueur d'onde λ et une fréquence ν données par la relation

$$\lambda = \frac{h}{p}, \qquad \text{et} \qquad \nu = \frac{E}{h}. \tag{1.7}$$

Ce sont essentiellement les mêmes relations que pour un photon. Mais ici, pour une particule non relativiste de masse m, on a $p = mv$ et $E = \frac{1}{2}mv^2 = \frac{p^2}{2m}$ (où v est la vitesse de la particule).

Plus généralement: *on peut associer à la particule une onde d'amplitude* $\psi(\vec{r}, t)$. Nous discutons d'abord deux exemples simples, puis l'interprétation de cette "fonction d'onde".

Si la particule n'est soumise à aucune force et que son mouvement est dans la direc-

tion z, il est naturel de lui associer l'onde plane (par analogie aux ondes électromagnétiques).

$$\psi(\vec{r},t) = Ae^{2\pi i(\frac{z}{\lambda}-\nu t)} = Ae^{\frac{i}{\hbar}(pz-Et)} \tag{1.8}$$

où $\hbar = \dfrac{h}{2\pi}$.

Pour une particule dans l'expérience des doubles fentes, l'onde sera sphérique:

$$\psi_{1,2}(\vec{r},t) = A\frac{e^{\frac{i}{\hbar}(p|\vec{r}-\vec{r_{1,2}}|-Et)}}{|\vec{r}-\vec{r_{1,2}}|} \tag{1.9}$$

La "fonction d'onde" ou "amplitude" totale est donnée par le principe de superposition

$$\psi(\vec{r},t) = \psi_1(\vec{r},t) + \psi_2(\vec{r},t) \tag{1.10}$$

La différence cruciale avec la théorie classique des ondes est qu'ici la "fonction d'onde" décrit une seule particule unique.

Quelle est l'interprétation de $\psi(\vec{r},t)$? Max Born suggéra peu après l'introduction de la fonction d'onde (la fonction d'onde prend des valeurs complexes et ci-dessous on prend le module au carré du nombre complexe) que

$$|\psi(\vec{r},t)|^2 \tag{1.11}$$

représente *la densité de probabilité* de trouver la particule en $\vec{r}$ au temps t. En d'autres termes, la quantité

$$\int_V \mathrm{d}^3\vec{r}\,|\psi(\vec{r},t)|^2 \tag{1.12}$$

est la probabilité de trouver la particule dans une région $V \subset \mathbb{R}^3$. Pour que cette interprétation soit consistante il faut bien sûr imposer la condition de normalisation

$$\int \mathrm{d}^3\vec{r}\,|\psi(\vec{r},t)|^2 = 1 \tag{1.13}$$

En général cette condition peut toujours être satisfaite en ajustant la constante de normalisation A.

Cette interprétation est parfaitement consistante avec les expériences de doubles fentes. Lorsque les particules sont envoyées une par une à travers les fentes, on observe des points d'absorption bien localisés sur l'écran. Ces points sont aléatoires. Quelle est leur distribution statistique? L'ensemble des points d'absorption forme des franges d'interférence d'intensité proportionnelle à $|\psi(\vec{r},t)|^2$. On en déduit que la distribution de probabilité des points d'absorption est donnée par la règle de Born $|\psi(\vec{r},t)|^2$.

1.4 Première notion "d'état quantique"

En un mot, l'"état quantique" est une abstraction ou bien une généralisation de la notion de "fonction d'onde". Nous discutons cette généralisation dans ce paragraphe. La fonction d'onde est un état associé aux degrés de liberté continus (telle que la position). Au Chapitre **2** nous allons introduire des états quantiques associés à des degrés de liberté discrets. Ce sont en fait ces degrés de liberté discrets qui nous intéressent vraiment en information quantique. Nous verrons aussi que les états associés aux degrés de liberté sont les "quantum bits".

Le concept de fonction d'onde et la règle de Born étaient vraiment révolutionnaires. On abandonne la notion de trajectoire pour les particules. Les concepts de position ou de vitesse bien déterminés font encore sens uniquement si on ne les mesure pas simultanément. Par exemple, lorsque l'on observe le point d'absorption sur l'écran, celui-ci est aléatoire : la position est observée, mais la direction de propagation au moment de l'absorption a perdu son sens.

En mécanique quantique, l'état de la particule est décrit par une fonction d'onde. On peut penser à cette fonction de façon plus abstraite comme à un vecteur de l'espace des fonctions. Ce vecteur ou état est noté

$$|\psi\rangle \quad \longleftrightarrow \quad \psi(\vec{r}). \tag{1.14}$$

Cette notation est la notation traditionnelle de la MQ introduite par Dirac; on pourrait aussi noter $\vec{\psi}$, mais $|\psi\rangle$ est conventionnel. Le mérite de cette notation est peut-être de nous rappeler que les vecteurs d'états ne vivent pas dans l'espace physique familier à trois dimensions. Le symbole $|\psi\rangle$ s'appelle aussi un "ket". Le vecteur transposé et complexe conjugué $\vec{\psi}^{T,*}$ s'appelle un "bra" et est noté

$$\langle\psi| \quad \longleftrightarrow \quad \psi^*(\vec{r}). \tag{1.15}$$

Le produit scalaire entre un vecteur et son propre transposé est la norme au carré de ce vecteur. Ici on a

$$\langle\psi|\psi\rangle = \int \mathrm{d}^3\vec{r}\ \psi^*(\vec{r})\,\psi(\vec{r}) = \int \mathrm{d}^3\vec{r}\ |\psi(\vec{r})|^2. \tag{1.16}$$

L'interprétation de Born impose $\int \mathrm{d}^3\vec{r}\ |\psi(\vec{r})|^2 = 1$ (probabilité totale égale à 1). Ainsi, un vecteur d'état doit satisfaire à la condition, dite de normalisation

$$\langle\psi|\psi\rangle = 1.$$

Ces considérations suggèrent que plus généralement le produit scalaire entre deux vecteurs $|\psi\rangle$ et $|\phi\rangle$ jouera un rôle important en MQ. Pour des vecteurs d'états correspondants à des fonctions d'onde, le produit scalaire naturel (naturel à cause de la règle de Born! En effet, du point de vue purement mathématique, on aurait pu définir d'autres produits scalaires!) est

$$\langle\phi|\psi\rangle = \int \mathrm{d}^3\vec{r}\ \phi^*(\vec{r})\,\psi(\vec{r}). \tag{1.17}$$

Ce produit scalaire satisfait aux règles usuelles de linéarité, symétrie et positivité.

Lorsqu'une particule est observée localisée en $\vec{r_1}$ son vecteur d'état est noté $|\vec{r_1}\rangle$. On peut, de façon un peu cavalière, penser à ce vecteur d'état comme étant celui qui correspond à la "fonction de Dirac" (ou distribution)

$$|\vec{r_1}\rangle \quad \longleftrightarrow \quad \delta(\vec{r} - \vec{r_1}) \tag{1.18}$$

Notons maintenant qu'en vertu du produit scalaire introduit ci-dessus

$$\langle \vec{r_1}|\psi\rangle = \int \mathrm{d}^3\vec{r}\, \delta(\vec{r} - \vec{r_1})\, \psi(\vec{r}) = \psi(\vec{r_1}). \tag{1.19}$$

Ainsi, nous avons la connexion suivante entre la notation des "bras" et "kets" et celle de la fonction d'onde

$$\langle \vec{r}|\psi\rangle = \psi(\vec{r}). \tag{1.20}$$

En fait la dérivation ci-dessus est un peu cavalière (essayez de dire pourquoi) et il faut prendre cette dernière relation comme la définition du bra $\langle \vec{r}|$. De même, on définit le ket $|\vec{r}\rangle$ via

$$\langle \psi|\vec{r}\rangle = \psi^*(\vec{r}). \tag{1.21}$$

Nous allons maintenant donner une formulation un peu plus générale de la règle de Born. Rappelons-nous que la densité de probabilité d'observer la particule en $\vec{r}$ quand son état est $|\psi\rangle$ est donnée par

$$|\psi(\vec{r})|^2 = |\langle \vec{r}|\psi\rangle|^2 \tag{1.22}$$

En mécanique quantique, le "postulat de la mesure" généralise cette règle. Nous donnerons au Chapitre **3** une formulation complètement précise de ce postulat. Jusqu'à là, nous allons nous contenter de la formulation suivante:

"La probabilité d'observer l'état final $|\phi\rangle$ juste après une mesure, lorsque l'état initial du système est $|\psi\rangle$ juste avant la mesure, est donnée par $|\langle\phi|\psi\rangle|^2$."

Lorsque nous étudierons la formalisation de la mécanique quantique, nous verrons que le bon cadre formel est celui des espaces de Hilbert. Nous verrons entre autres que:

(i). Les états quantiques sont des vecteurs. Il est possible d'additionner deux vecteurs, ce qui correspond à la superposition de deux fonctions d'onde.
(ii). Les produits scalaires donnent les probabilités observées lors des mesures.

Ces règles seront également précisées au Chapitre **3**.

1.5 Principe d'incertitude de Heisenberg

Une caractéristique fondamentale de la mécanique quantique est l'impossibilité de déterminer avec une précision infinie la position et l'impulsion d'une particule. Une expression mathématique de ce fait est donnée par le principe d'incertitude que nous discutons ici.

Nous avons vu que pour une particule dans l'état $|\psi\rangle$, la probabilité de trouver la particule en x est $|\psi(x)|^2$ (On considère le cas à une dimension spatiale $x \in \mathbb{R}$). L'incertitude ou l'écart type obtenu lors de mesures répétées est alors

$$\sigma_x = \left\{ \int x^2 |\psi(x)|^2 \, \mathrm{d}x - \left(\int x \, |\psi(x)|^2 \, \mathrm{d}x \right)^2 \right\}^{\frac{1}{2}} \tag{1.23}$$

Les physiciens écrivent plutôt δx au lieu de σ_x.

Supposons maintenant que l'on mesure l'impulsion p de la particule. Quel est le ket $|p\rangle$ correspondant à une particule d'impulsion p ? En fait la fonction d'onde associée à ce ket $|p\rangle$ est l'onde plane $e^{\frac{i}{\hbar}px}$. Notez que la densité de probabilité associée est $|e^{\frac{i}{\hbar}px}|^2 = 1$, donc cette fonction d'onde est complètement délocalisée dans l'espace. Aussi, puisque par définition $\langle \vec{r} | \psi \rangle = \psi(\vec{r})$ nous avons

$$\langle x | p \rangle = e^{\frac{i}{\hbar}px}. \tag{1.24}$$

Si l'état observé est $|\psi\rangle$ la probabilité de trouver une impulsion mesurée p est donnée par la règle de Born généralisée

$$|\langle p | \psi \rangle|^2 = \int \mathrm{d}x \, e^{-\frac{i}{\hbar}px} \psi(x) = |\hat{\psi}(p)|^2 \tag{1.25}$$

où $\hat{\psi}$ est essentiellement la transformée de Fourier de ψ. Ainsi $|\hat{\psi}(p)|^2$ est une densité de probabilité pour l'impulsion de la particule. L'incertitude peut être définie comme

$$\sigma_p = \left\{ \int p^2 |\hat{\psi}(p)|^2 \, \mathrm{d}p - \left(\int p \, |\hat{\psi}(p)|^2 \, \mathrm{d}p \right)^2 \right\}^{\frac{1}{2}} \tag{1.26}$$

À nouveau, les physiciens notent δp au lieu de σ_p.

Étant donné une fonction de carré sommable, c.-à-d. $\int \mathrm{d}x \, |\psi(x)|^2$ fini[2], un théorème de mathématique affirme que

$$\sigma_x \, \sigma_p \geq \frac{\hbar}{2}. \tag{1.27}$$

Cette inégalité souvent écrite $\delta x \, \delta p \geqslant \frac{\hbar}{2}$ est le principe d'incertitude d'Heisenberg. L'interprétation physique est qu'il est impossible de mesurer avec précision infinie x et p en même temps avec un même appareil de mesure (par exemple un "microscope"). Si on gagne en précision pour x on perd en précision pour p et vice-versa. Ce n'est pas un problème de limitation dû à l'instrument de mesure, mais une limitation intrinsèque imposée par les lois quantiques.

[2] Notez que le théorème de Parseval affirme $\int \mathrm{d}x \, |\psi(x)|^2 = \int \mathrm{d}p \, |\hat{\psi}(p)|^2$

1.6 Le problème des raies spectrales

Lorsque la lumière visible passe à travers un prisme, on observe un spectre continu de couleurs. Plus généralement, le spectre des ondes électromagnétiques est continu.

En 1885, Balmer mis en évidence que l'hydrogène et les atomes en général, possèdent un spectre d'émission discret. Pour l'hydrogène, Balmer détectait 4 raies dans les longueurs d'onde de la lumière visible $\lambda = 656.3$ nm; 481.1 nm; 434.1 nm et 410 nm. Elles satisfont à la formule empirique

$$\frac{1}{\lambda} = R_H\,(\frac{1}{4} - \frac{1}{m^2}), \qquad m = 3, 4, 5, 6 \tag{1.28}$$

avec $R_h = 10973731,57\ \mathrm{m}^{-1}$.

Le problème qui se posait était d'expliquer cette formule qui donne des longueurs d'onde discrètes décrites par le nombre entier m. Nous allons voir qu'en fait cette formule est un cas particulier d'une formule générale proposée par Bohr. La formule de Bohr prédisait toute une série d'autres raies spectrales qui furent observées, bien plus tard pour certaines d'entre elles.

Les célèbres expériences de Rutherford (diffusion de particules α sur des feuilles d'or) avaient démontré en 1903 que les atomes sont constitués d'un noyau chargé positivement (charge Ze) et de Z électrons (charge e) "orbitant" autour du noyau. Néanmoins, la stabilité de ce "système planétaire" ne pouvait pas être expliqué par les lois de la physique classique. En effet, une charge orbitant autour d'un centre doit rayonner et perdre de l'énergie si bien qu'au bout d'un temps (très long) elle tombe sur le noyau.

Les idées quantiques naissantes permirent, sinon d'expliquer la stabilité des atomes, de justifier la formule de Balmer et de la généraliser. Cela fut d'abord établi par Bohr. Celui-ci postula que (i) l'électron peut orbiter autour de certaines "trajectoires permises" et (ii) les lois de la mécanique classique s'appliquent à ces trajectoires permises. Ensuite il donna une "règle de quantification" pour toutes ces "trajectoires permises". Cette règle sera discutée aux exercices; ici nous suivons une approche due à De Broglie basée sur le concept de fonction d'onde.

Sur une trajectoire permise supposée circulaire, on a :

$$m\,\frac{v^2}{R} = k\,\frac{Ze^2}{R^2} \tag{1.29}$$

et

$$E = \frac{1}{2}mv^2 - k\,\frac{Ze^2}{R} \tag{1.30}$$

où v est la vitesse, m la masse, k la constante de Coulomb, E l'énergie mécanique et R le rayon de la trajectoire. À partir de ces formules, il est facile de montrer que l'énergie associée à la trajectoire de rayon R est :

$$E = -\frac{k}{2}\frac{Ze^2}{R}. \tag{1.31}$$

Maintenant, il s'agit de trouver les rayons des trajectoires permises. Selon De Broglie l'onde associée à l'électron le long de sa trajectoire doit être stationnaire (c'est-à-dire que les nuds doivent être immobiles). Pour une trajectoire circulaire, la condition de stationnarité est :

$$n\lambda = 2\pi R, \qquad n \geq 1 \tag{1.32}$$

où λ est la longueur d'onde et n un entier[3]. Pour λ on pose suivant De Broglie:

$$\lambda = \frac{h}{p} = \frac{h}{mv} \tag{1.33}$$

Puisque $v = \left(\frac{k}{m}\frac{Ze^2}{R}\right)^2$ on obtient $\lambda = \frac{hR^{\frac{1}{2}}}{(kmZe^2)^{\frac{1}{2}}}$ et donc

$$R = \frac{\hbar^2}{kmZe^2}n^2, \qquad n \geq 1 \tag{1.34}$$

Pour les rayons permis. Cela donne

$$E_n = -\frac{k^2Z^2e^4m}{2\hbar^2}\frac{1}{n^2} = -\mathrm{R_y}\frac{Z^2}{n^2}; \qquad n \geq 1. \tag{1.35}$$

Ici $\mathrm{R_y} = \frac{k^2e^4m}{2\hbar^2} \simeq 13.6$ eV est la constante de Rydberg. Pour l'hydrogène, on a en particulier $Z = 1$.

Les énergies E_n des trajectoires permises s'appellent les "niveaux d'énergies" et forment le "spectre" de l'atome d'hydrogène (Pour $Z = 1$). Comme nous l'avons dit, la notion de trajectoire n'a pas vraiment de sens. Mais remarquablement cette formule est exacte pour l'hydrogène. Ceci fut établi par Schrödinger et est discuté brièvement à la Section **1.7**.

Revenons aux raies spectrales. Toujours selon Bohr, un électron peut transiter d'une orbite "numéro m" à une orbite "numéro n" en émettant un photon d'énergie $E_m - E_n$. La fréquence du photon sera donnée par la relation d'Einstein

$$h\nu_{m\to n} = E_m - E_n \tag{1.36}$$

Puisque $E_n = -\frac{\mathrm{R_y}}{n^2}$ (on prend $Z = 1$ pour l'hydrogène) on trouve

[3] Comme pour une corde de violon de longueur $2\pi R$: $n = 1$ donne la note fondamentale et $n = 2, 3, 4, ...$ donnent les harmoniques

$$h\nu_{m\to n} = -\left(\frac{\mathrm{Ry}}{m^2} - \frac{\mathrm{Ry}}{n^2}\right) \tag{1.37}$$

avec $\lambda\nu = c$ pour le photon, on trouve

$$\frac{1}{\lambda_{m\to n}} = \frac{\mathrm{Ry}}{hc}\left(\frac{1}{n^2} - \frac{1}{m^2}\right) \tag{1.38}$$

la constante $\dfrac{\mathrm{Ry}}{hc} = R_H$ (la constante de Balmer). Les raies de Balmer correspondent à la série des transitions $m \to 2$ ($n = 2$). La série $m \to 1$ s'appelle série de Lyman (et fut observée aussi tard qu'en 1914; ultraviolets). La série $m \to 3$ est la série de Paschen et correspond à l'infrarouge. Le nombre de ces raies spectrales est infini et le point important est qu'elles sont discrètes.

1.7 L'équation de Schrödinger

Comme nous l'avons vu l'approche dite "semi-classique" utilisée à la Section **1.6** n'est pas très satisfaisante conceptuellement, car elle fait encore appel à la notion classique de trajectoire.

Schrödinger dériva en 1926 sa fameuse équation qui décrit la dynamique de la fonction d'onde $\psi(\vec{r}, t)$. En résolvant son équation, il était capable de trouver la condition de stationnarité et les niveaux d'énergie. Remarquablement la formule $E_n = -Ry\dfrac{Z^2}{n^2}$ reste inchangée, mais il y a un gros bonus! Pour le même "nombre quantique" n caractérisant le niveau d'énergie, il y a plusieurs solutions possibles à l'équation et donc plusieurs fonctions d'onde ou états possibles. Cela signifie qu'il y a plusieurs distributions de probabilités possibles pour l'électron autour du noyau. On dit que les niveaux d'énergie sont dégénérés. Nous n'allons pas discuter ceci en détail ici, mais c'est ce qui permet d'expliquer plusieurs propriétés du tableau périodique. Ainsi l'équation de Schrödinger est à la base de la chimie.

Nous donnons ici une dérivation très heuristique de l'équation de Schrödinger. Pour une particule d'impulsion p et d'énergie E nous associons la fonction d'onde

$$\psi(z, t) = A\ \exp\{\frac{i}{\hbar}(pz - Et)\}. \tag{1.39}$$

Si la particule est libre $E = \dfrac{p^2}{2m}$. Supposons que la particule soit soumise à un potentiel $V(z)$. Alors la mécanique classique nous dit que $E = \dfrac{p^2}{2m} + V(z)$ et il est tentant de remplacer E par cette expression dans l'onde plane. Ceci n'est pas vraiment permis, mais on peut considérer que c'est une bonne approximation si $V(z)$ varie très lentement sur une échelle plus grande que $\lambda = \dfrac{h}{p}$:

$$\psi(z,t) \simeq A\ \exp\{\frac{i}{\hbar}(pz - \frac{p^2}{2m} + V(z))\}. \tag{1.40}$$

Il est facile de vérifier que cette fonction satisfait

$$i\hbar\frac{\partial\psi(z,t)}{\partial t} = -\frac{\hbar^2}{2m}\frac{d^2}{dz^2}\psi(z,t) + V(z)\psi(z,t) \tag{1.41}$$

Cette équation est en fait exacte (pour des particules non-relativistes). C'est l'équation dérivée par Schrödinger en 1926. À trois dimensions, elle se généralise aisément,

$$i\hbar\frac{\partial\psi(\vec{r},t)}{\partial t} = -\frac{\hbar^2}{2m}\Delta\psi(\vec{r},t) + V(\vec{r})\psi(\vec{r},t) \tag{1.42}$$

Avec $\Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$ l'opérateur Laplacien. Dans le cas de l'atome, l'électron orbite autour du noyau de charge Ze et on prend $V(\vec{r}) = -k\frac{Ze^2}{r}$.

Pour calculer les niveaux d'énergie, on suppose qu'il existe des solutions stationnaires. Ces solutions sont de la forme

$$\psi(\vec{r},t) = \phi(\vec{r})\exp\{-\frac{i}{\hbar}Et\} \tag{1.43}$$

Les nuds de $\psi(\vec{r},t)$ sont ceux de $\phi(\vec{r})$ qui ne dépend pas du temps. De plus la densité de probabilité de trouver la particule en $\vec{r}$ est $|\psi(\vec{r},t)|^2 = |\phi(\vec{r})|^2$ qui ne dépend pas du temps.

Pour un atome, si on suppose que l'électron est lié au noyau, on cherchera des solutions telles que $\phi(\vec{r}) \to 0$ pour $|\vec{r}| \to +\infty$. On se représente souvent $|\phi(\vec{r})|^2$ comme une sorte de "nuage électronique" autour du noyau.

En remplaçant l'ansatz ci-dessus dans l'équation de Schrödinger, on trouve

$$\left(-\frac{\hbar^2}{2m}\Delta + V(\vec{r})\right)\phi(\vec{r}) = E\,\phi(\vec{r}). \tag{1.44}$$

Cette équation est une équation aux valeurs propres pour l'opérateur linéaire

$$H = -\frac{\hbar^2}{2m}\Delta + V(\vec{r}). \tag{1.45}$$

La fonction d'onde stationnaire $\phi(\vec{r})$ est le "vecteur propre" et E la "valeur propre" associée. L'opérateur linéaire peut être vu comme une matrice infinie (celle-ci est infini, car l'espace des fonctions $\phi(\vec{r})$ est un espace vectoriel de dimension infinie). Cette matrice infinie ou opérateur est l'Hamiltonien quantique du système. Les valeurs propres

donnent les niveaux d'énergie. Pour l'hydrogène $V(\vec{r}) = -k\frac{e^2}{r}$ le calcul de Schrödinger donne $E_n = -\frac{R_y}{n^2}$. Remarquablement, c'est le même résultat que la théorie de Bohr. Cette coïncidence est limitée au cas spécial de l'atome d'hydrogène et pour d'autres systèmes, il faut recourir à l'équation fondamentale de Schrödinger.

1.8 Le principe de correspondance

Reprenons le problème de la particule dans un champ de forces décrit par un potentiel $V(\vec{r})$. Classiquement l'énergie est calculée comme $E = \frac{\vec{p}^2}{2m} + V(\vec{r})$. L'équation de Schrödinger nous enseigne que les niveaux d'énergie discrets sont donnés par les valeurs propres de $H = -\frac{\hbar^2}{2m}\Delta + V(\vec{r})$. Nous voyons que l'on peut obtenir l'Hamiltonien quantique en remplaçant $V(\vec{r})$ par $V(\vec{r})$ et $\frac{\vec{p}^2}{m}$ par $-\frac{\hbar^2}{2m}\Delta$. Ceci est une expression du "principe de correspondance".

Le principe de correspondance est un ensemble de règles qui permet de déduire la forme des lois quantiques à partir des lois classiques. Grâce à ce principe, on peut remplacer n'importe quelle observable classique $A(\vec{r}, \vec{p})$ (c.-à-d. une fonction de la position et de l'impulsion) par un opérateur ou une matrice (infinie) notée $\hat{A}$. La règle de base est donnée par la correspondance

$$\begin{cases} \vec{r} & \longrightarrow & \vec{r} = (x, y, z) \\ \vec{p} & \longrightarrow & i\hbar\vec{\nabla} = \left(i\hbar\frac{\partial}{\partial x}, i\hbar\frac{\partial}{\partial y}, i\hbar\frac{\partial}{\partial z}\right). \end{cases} \tag{1.46}$$

Par exemple, $\frac{\vec{p}^2}{2m} = \frac{p_x^2}{2m} + \frac{p_y^2}{2m} + \frac{p_z^2}{2m}$ devient $-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} - \frac{\hbar^2}{2m}\frac{\partial^2}{\partial y^2} - \frac{\hbar^2}{2m}\frac{\partial^2}{\partial z^2} = -\frac{\hbar^2}{2m}\Delta$. L'application de la règle ci-dessus donne bien

$$\frac{p^2}{2m} + V(\vec{r}) \to -\frac{\hbar^2}{2m}\Delta + V(\vec{r}). \tag{1.47}$$

Ce principe souffre d'une ambiguïté qui est la suivante. Classiquement $xp_x - p_x x = 0$ car on multiplie des nombres. Néanmoins, on peut montrer que

$$-x\left(i\hbar\frac{\partial}{\partial x}\right) + \left(i\hbar\frac{\partial}{\partial x}\right)x = i\hbar \tag{1.48}$$

En effet:

$$-x\left(i\hbar\frac{\partial}{\partial x}\phi(x)\right)+i\hbar\frac{\partial}{\partial x}\left(x\phi(x)\right)=-xi\hbar\phi'(x)+i\hbar x\phi'(x)+i\hbar\phi(x) \tag{1.49}$$

$$=i\hbar\phi(x). \tag{1.50}$$

Ainsi x et $p_x \equiv -i\hbar\frac{\partial}{\partial x}$ ne commutent pas en mécanique quantique (il faut y penser comme à des opérateurs ou des matrices). La quantité

$$xp_x - p_x x \equiv [x, p_x] \tag{1.51}$$

s'appelle le commutateur de x et p_x. La relation

$$[x, p_x] = i\hbar, \qquad \text{(idem pour } y, p_y \text{ et } z, p_z) \tag{1.52}$$

s'appelle la relation de commutation canonique. En fait cette relation est intimement liée au principe d'incertitude $\delta x \delta p \geq \frac{\hbar}{2}$.

Le principe de correspondance ne précise pas quel est l'ordre correct des produits entre x et p pour des observables $A(x, p)$ qui contiennent des termes mixtes. Le bon choix est guidé par des considérations physiques spécifiques au problème donné.

2 Degrés de Liberté Discrets: Polarisation et Spin

2.1 Polarisation des ondes électromagnétiques

Les équations de Maxwell dans le vide possèdent des solutions qui sont des ondes planes. Pour une onde plane se propageant dans la direction z les champs électriques $\vec{E}$ et magnétiques $\vec{B}$ sont perpendiculaires à la direction de propagation

$$\vec{E} = \mathrm{Re}\left\{\vec{E_0}\, e^{i(kz-\omega t)}\right\} \qquad \text{et} \qquad \vec{B} = \frac{1}{c}\hat{z} \times \vec{E} \tag{2.1}$$

avec $k = \dfrac{2\pi}{\lambda}$, $\omega = 2\pi\nu$ et $\lambda\nu = c$ la vitesse de la lumière. Il suffit de considérer le vecteur $\vec{E}$ car $\vec{B}$ est automatiquement $\perp$ à $\vec{E}$.

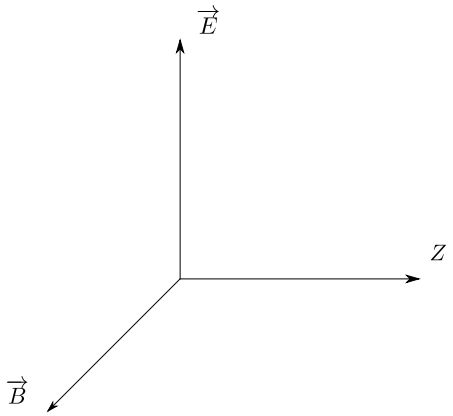


Figure 2.1 Directions des champs électrique et magnétique.

En général $\vec{E_0} = E_0\left((\cos\theta)e^{i\delta_x}; (\sin\theta)e^{i\delta_y}; 0\right)$. L'orientation de $\vec{E}$ dans le plan $\perp$ à z s'appelle la polarisation de l'onde. On peut toujours poser $\delta_x = 0$ et garder le paramètre

δ_y ouvert (cela revient à changer l'origine du temps). Il est possible de montrer que pour z fixé, le vecteur champ électrique $\vec{E} = (E_x, E_y, 0)$ trace une ellipse en fonction du temps dans le plan $xy \perp z$. Ici nous allons considérer uniquement quelques cas particuliers importants.

Exemple 2.1: Polarisation linéaire

Soit $\delta_x = 0$ et $\delta_y = 0$. On a pour $z = 0$

$$\vec{E} = \mathrm{Re}\left\{ E_0 \begin{pmatrix} \cos\theta \\ \sin\theta \\ 0 \end{pmatrix} e^{-i\omega t} \right\} = E_0 \begin{pmatrix} \cos\theta \\ \sin\theta \\ 0 \end{pmatrix} \cdot \cos(\omega t) \tag{2.2}$$

Puisque $\dfrac{E_y}{E_x} = \tan\theta$, le champ électrique oscille le long de la direction θ ou bien $-\theta$. C'est ce que l'on appelle la polarisation linéaire le long de θ. (Figure **2.2**)

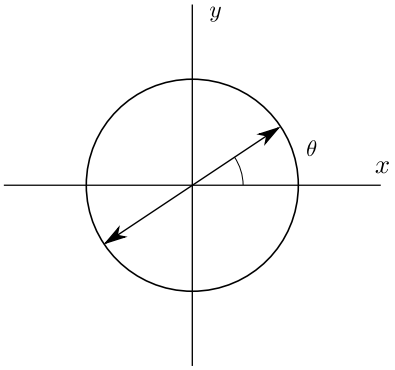


Figure 2.2 Polarisation linéaire le long de θ.

Exemple 2.2: Polarisation circulaire

Considérons le cas $\theta = \frac{\pi}{4}$ et $\delta_x = 0$ avec $\delta_y = \frac{\pi}{2}$ ou bien $\delta_y = -\frac{\pi}{2}$. Le champ électrique devient :

$$\vec{E} = \mathrm{Re}\left\{ E_0 \begin{pmatrix} \cos\frac{\pi}{4} \\ \sin\frac{\pi}{4} e^{\pm i\frac{\pi}{2}} \\ 0 \end{pmatrix} e^{-i\omega t} \right\} = E_0 \frac{1}{\sqrt{2}} \begin{pmatrix} \cos(\omega t) \\ \pm \sin(\omega t) \\ 0 \end{pmatrix} \tag{2.3}$$

Cette fois $E_y^2 + E_x^2 = E_0^2$, donc le champ électrique effectue un mouvement circulaire droite ($\delta_y = +\frac{\pi}{2}$) ou gauche ($\delta_y = -\frac{\pi}{2}$). (Figure **2.3**).

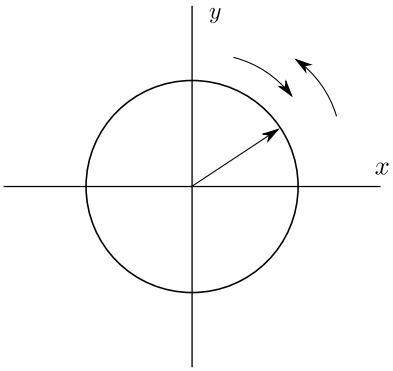


Figure 2.3 Polarisation circulaire du champ électrique.

Le cas général (θ, δ_y) quelconque (on peut toujours prendre $\delta_x = 0$) est celui de la polarisation elliptique : le champ électrique parcourt une ellipse dans la direction droite ou gauche. Les cas linéaires et circulaires sont des formes dégénérées de l'ellipse. On peut mettre en évidence la polarisation des ondes électromagnétiques grâce à des filtres. Par exemple, un "**polarisateur linéaire**" (Figure **2.4**) permet de sélectionner la composante du champ électrique dans une direction donnée disons θ.

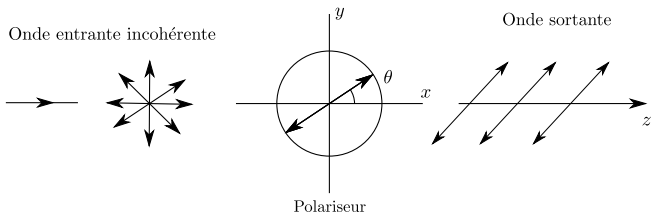


Figure 2.4 Polariseur linéaire.

Le champ électrique sortant est simplement la composante du champ entrant le long de θ. Donc le champ de l'onde sortante est ici (en $z = 0$ mettons)

$$\vec{E} = E_0 \begin{pmatrix} \cos\theta \\ \sin\theta \\ 0 \end{pmatrix} \tag{2.4}$$

On peut placer un "deuxième filtre en série" le long de la direction α. Celui-ci s'appelle un "analyseur" (Figure **2.5**) car il sert à analyser la polarisation de l'onde.

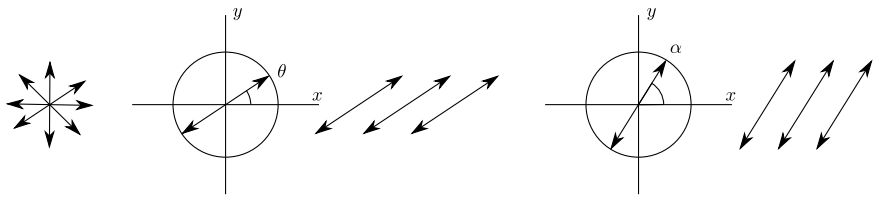


Figure 2.5 Analyseur.

L'onde transmise par l'analyseur possède un champ électrique dans la direction α. Celui-ci est simplement la composante du champ entrant dans la direction α.

$$\vec{E} = E_0 \cos(\alpha - \theta) \begin{pmatrix} \cos\alpha \\ \sin\alpha \\ 0 \end{pmatrix} \tag{2.5}$$

Ici l'amplitude est obtenue en faisant le produit scalaire

$$(\cos\alpha, \sin\alpha) \cdot \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix} = \cos\alpha\cos\theta + \sin\alpha\sin\theta = \cos(\alpha - \theta) \tag{2.6}$$

L'intensité de l'onde avant l'analyseur (et après le premier polariseur) est $\sim E_0^2$ alors

que celle après l'analyseur est $\sim E_0^2(\cos(\alpha-\theta))^2$. Le rapport des intensités transmises et incidentes est donc

$$\cos^2(\alpha-\theta) \tag{2.7}$$

C'est la loi de **Malus**.

2.2 Polarisation du photon

Le photon possède un "degré de liberté interne" qui ressemble à la polarisation du champ électrique. C'est ce qui s'appelle la "polarisation du photon".

Au Chapitre **1** nous avons introduit le concept de fonction d'onde. En particulier l'onde plane associée à une particule libre se propageant dans la direction z est $\psi(z,t) = A\, e^{i(kz-\omega t)}$. Pour des photons, cette fonction d'onde est à valeurs vectorielles tout comme le champ électrique :

$$A\, e^{i(kz-\omega t)} \begin{pmatrix} \cos\theta \\ (\sin\theta)e^{i\phi} \\ 0 \end{pmatrix} \tag{2.8}$$

Ici nous avons posé $\delta_x = 0$ et $\delta_y = \phi$ comme il est usuellement fait pour des photons.

Le choix $\phi = 0, \pi$ correspond à un photon avec la polarisation linéaire dans la direction θ ou $\theta_\perp$ et le choix $\theta = \dfrac{\pi}{4}$ et $\phi = \dfrac{\pi}{2}$ ou $-\dfrac{\pi}{2}$ correspond à un photon avec polarisation circulaire droite ou gauche respectivement.

En notation de Dirac, l'état général d'un photon libre est :

$$e^{-i\omega t}|k\rangle \otimes (\cos\theta|x\rangle + (\sin\theta)e^{i\phi}|y\rangle). \tag{2.9}$$

La correspondance avec la notation usuelle est

$$|k\rangle \leftrightarrow e^{ikz}; \qquad |x\rangle \leftrightarrow \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \text{ et } \qquad |y\rangle \leftrightarrow \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \tag{2.10}$$

En fait on posera $|x\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ et $|y\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ et l'on fera abstraction de la composante z qui est toujours nulle.

Dans ce chapitre nous nous intéressons uniquement au "degré de liberté de polarisation" du photon:

$$|\theta,\phi\rangle = \cos\theta|x\rangle + (\sin\theta)e^{i\phi}|y\rangle = \begin{pmatrix} \cos\theta \\ (\sin\theta)e^{i\phi} \end{pmatrix} \tag{2.11}$$

et laissons tomber le "degré de liberté orbital" $|k\rangle$. Nous laissons aussi tomber la dépendance temporelle $e^{-i\omega t}$. Ces vecteurs forment l'espace vectoriel $\mathbb{C}^2$ et satisfont $\langle\theta,\phi|\theta,\phi\rangle = 1$. Pour vérifier cela, on utilise

$$\langle\theta,\phi| = \cos\theta\langle x| + e^{-i\phi}\sin\theta\langle y| \tag{2.12}$$

et

$$\langle x|x\rangle = \langle y|y\rangle = 1, \qquad \langle x|y\rangle = \langle y|x\rangle = 0. \tag{2.13}$$

On peut aussi le vérifier en composante grâce à :

$$\langle x| = (1,0) \qquad , \qquad \langle y| = (0,1) \tag{2.14}$$

et

$$\langle\theta,\phi| = (\cos\theta, e^{-i\phi}\sin\theta). \tag{2.15}$$

Les états de polarisation linéaire:

$$\begin{cases} |\theta\rangle = \cos\theta|x\rangle + \sin\theta|y\rangle = \begin{pmatrix}\cos\theta\\ \sin\theta\end{pmatrix} \\ |\theta_\perp\rangle = \cos\theta_\perp|x\rangle + \sin\theta_\perp|y\rangle = \sin\theta|x\rangle - \cos\theta|y\rangle = \begin{pmatrix}\sin\theta\\ -\cos\theta\end{pmatrix} \end{cases} \tag{2.16}$$

forment une base orthonormée pour $\mathbb{C}^2$. Ceci est aussi le cas pour les deux états de polarisation circulaire:

$$\begin{cases} |R\rangle = \frac{1}{\sqrt{2}}(|x\rangle + i|y\rangle) = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\ i\end{pmatrix} \\ |L\rangle = \frac{1}{\sqrt{2}}(|x\rangle - i|y\rangle) = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\ -i\end{pmatrix} \end{cases} \tag{2.17}$$

2.3 Expériences sur la polarisation des photons

Nous allons maintenant considérer une source de photons uniques préparés dans l'état $|\theta\rangle$ de polarisation linéaire. Cela peut par exemple être réalisé grâce à une source de très basse intensité devant laquelle on place un filtre polarisant placé selon l'angle θ.

Photodétection après un analyseur

Les photons uniques sont envoyés sur un analyseur α puis enregistrés dans un photodétecteur D. (Figure **2.6**)

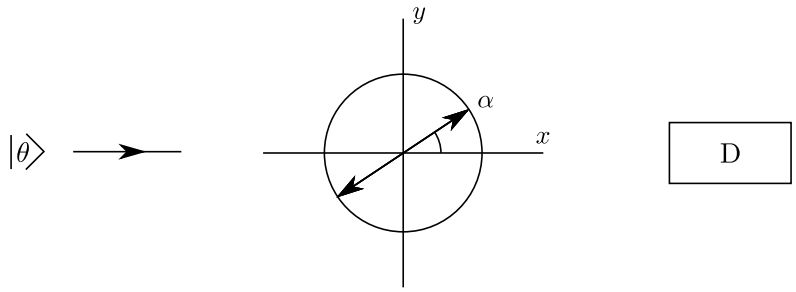


Figure 2.6 Analyseur et photodétecteur.

L'observation expérimentale est la suivante: le photodétecteur enregistre 1 ou 0 photons. C'est-à-dire que le photon traverse l'analyseur α ou bien est absorbé et ne parvient pas à D. Si l'expérience est répétée plusieurs fois, on observe une séquence aléatoire

$$100101110010101 \tag{2.18}$$

et il n'est pas possible de prévoir si le photon traverse ou non l'analyseur. La deuxième observation expérimentale est la fréquence empirique des 1: la probabilité empirique de voir un photon dans le photodétecteur D est

$$\cos^2(\theta - \alpha). \tag{2.19}$$

Ces observations peuvent être expliquées très facilement grâce à la règle de Born introduite au chapitre 1. L'état du photon avant l'analyseur est $|\theta\rangle$. Le système analyseur + détecteur joue ici le rôle d'appareil de mesure. Si le détecteur enregistre un photon, c'est que celui-ci est observé dans l'état $|\alpha\rangle$ (il a traversé l'analyseur). La probabilité de transition est :

$$\mathbb{P}\left[|\theta\rangle \to |\alpha\rangle\right] = |\langle\alpha|\theta\rangle|^2 \tag{2.20}$$

en vertu de la règle de Born, il est facile de voir que

$$\begin{aligned}\langle\alpha|\theta\rangle &= (\cos\alpha\langle x| + \sin\alpha\langle y|)(\cos\theta|x\rangle + \sin\theta|y\rangle) \\ &= \cos\alpha\cdot\cos\theta + \sin\alpha\cdot\sin\theta \\ &= \cos(\alpha-\theta).\end{aligned} \tag{2.21}$$

Décomposition par une lame biréfringente

Une lame biréfringente décompose la lumière en deux parties. L'une possède une polarisation verticale et l'autre une polarisation horizontale. (Figure **2.7**)

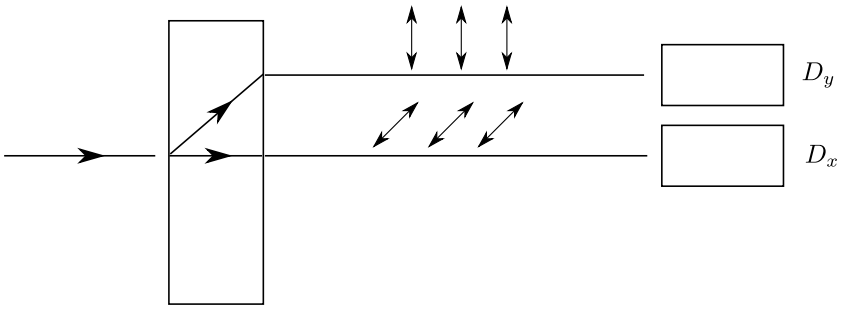


Figure 2.7 Lame biréfringente

Pour une onde avec champ électrique

$$\vec{E} = E_0 \begin{pmatrix}\cos\theta\\ \sin\theta\\ 0\end{pmatrix} \mathrm{Re}\left\{e^{i(kz-\omega t)}\right\} \tag{2.22}$$

On obtient deux ondes après la lame biréfringente :

$$E_y = E_0 \begin{pmatrix}0\\ \sin\theta\\ 0\end{pmatrix} \mathrm{Re}\left\{e^{i(kz-\omega t)}\right\} \tag{2.23}$$

et

$$E_x = E_0 \begin{pmatrix}\cos\theta\\ 0\\ 0\end{pmatrix} \mathrm{Re}\left\{e^{i(kz-\omega t)}\right\} \tag{2.24}$$

L'intensité mesurée dans les deux détecteurs (divisée par l'intensité incidente) est pour $D_y \sim \sin^2\theta$ et pour $D_x \sim \cos^2\theta$. La somme des intensités est égale à l'intensité totale incidente.

Que se passe-t-il si on envoie des photons uniques? On observe que soit D_x ou D_y

enregistre un photon; ces évènements sont exclusifs. Si la séquence des enregistrements pour D_x est:

$$1010011010101, \tag{2.25}$$

pour D_y elle est:

$$0101100101010. \tag{2.26}$$

Ces séquences sont complémentaires, mais aléatoires. on peut seulement connaître la statistique des 1 et 0. La probabilité empirique d'observer 1 dans D_x est $\cos^2\theta$ et elle est $\sin^2\theta$ pour D_y.

L'interprétation quantique de ce résultat est la suivante. Avant la lame biréfringente, l'état du photon est $|\theta\rangle$. Après la lame biréfringente, l'état orbital est différent (il existe "deux chemins possibles") mais l'état de polarisation est toujours $|\theta\rangle = \cos\theta|x\rangle + \sin\theta|y\rangle$. La probabilité d'enregistrer un photon dans D_x est la probabilité d'observer une polarisation $|x\rangle$:

$$\mathbb{P}\left[|\theta\rangle \to |x\rangle\right] = |\langle x|\theta\rangle|^2 = \cos^2\theta \tag{2.27}$$

La probabilité d'enregistrer un photon dans D_y est la probabilité d'observer une polarisation $|y\rangle$:

$$\mathbb{P}\left[|\theta\rangle \to |y\rangle\right] = |\langle y|\theta\rangle|^2 = \sin^2\theta \tag{2.28}$$

Décomposition - Recombinaison

Cette fois-ci, au lieu d'observer les photons juste après la lame biréfringente, on recombine l'onde (ou les photons) grâce à une lame symétrique. Ensuite on analyse les photons avec le système analyseur α + détecteur. (Figure **2.8**)

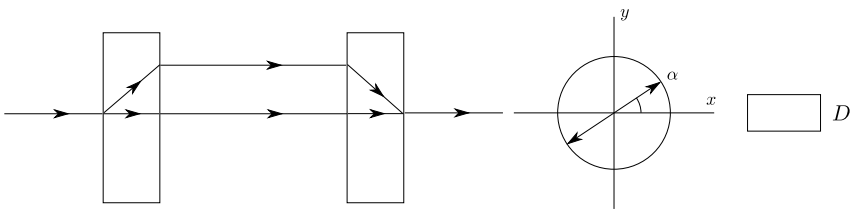


Figure 2.8 Décomposition et recombinaison.

Si l'expérience est faite avec une onde électromagnétique, celle-ci est d'abord décomposée entre les lames, puis recomposée. Après la deuxième lame, le champ électrique est donné par la superposition

$$E_x + E_y = \mathrm{Re}\left\{e^{i(kz-\omega t)}\right\} E_0 \begin{pmatrix} \cos\theta \\ \sin\theta \\ 0 \end{pmatrix} \tag{2.29}$$

Après l'analyseur, l'intensité enregistrée dans le détecteur sera donc $\cos^2(\theta - \alpha)$.

Avec des photons uniques, on enregistre ou non un photon dans le détecteur. À nouveau, lorsque l'expérience est répétée on obtient une suite aléatoire de 1 et 0. La fréquence empirique des 1 est $\cos^2(\theta - \alpha)$.

Ce résultat est "évident" si l'on accepte l'interprétation quantique. En effet, après la seconde lame biréfringente, l'état du photon est $|\theta\rangle$ (à nouveau!). La probabilité que celui-ci soit détecté par l'appareil $\alpha + D$ est donc

$$\mathbb{P}\left[|\theta\rangle \to |\alpha\rangle\right] = |\langle\alpha|\theta\rangle|^2 = \cos^2(\alpha - \theta). \tag{2.30}$$

Il est instructif de faire un calcul "purement classique" en supposant que les photons se comportent comme des particules classiques ayant des trajectoires et des polarisations uniques bien définies. Nous allons voir que le résultat n'est pas en accord avec l'expérience.

Si la "particule" est classique elle suit le chemin supérieur avec probabilité $\sin^2\theta$ et le chemin inférieur avec probabilité $\cos^2\theta$. Quand elle suit le chemin supérieur, sa polarisation est "y" et la probabilité de détection après l'analyseur doit être $\cos^2(\frac{\pi}{2} - \alpha) = \sin^2\alpha$. Quand elle suit le chemin inférieur, sa polarisation est "x" et la probabilité de détection après l'analyseur doit être $\cos^2(0 - \alpha) = \cos^2\alpha$. Ainsi:

$$\begin{aligned} \mathbb{P}\left[\text{détection}\right] &= \mathbb{P}\left[\text{détection} \mid \text{chemin sup}\right] \cdot \mathbb{P}\left[\text{chemin sup}\right] \\ &\quad + \mathbb{P}\left[\text{détection} \mid \text{chemin inf}\right] \cdot \mathbb{P}\left[\text{chemin inf}\right] \\ &= \sin^2\theta \sin^2\alpha + \cos^2\theta \cos^2\alpha \\ &\neq \cos^2(\theta - \alpha). \end{aligned} \tag{2.31}$$

La différence entre le résultat d'une interprétation classique et le (vrai) résultat quantique est égale à $2\sin\theta\sin\alpha\cos\theta\cos\alpha$. En effet:

$$\begin{aligned} \cos^2(\theta - \alpha) &= (\cos\theta\cos\alpha + \sin\theta\sin\alpha)^2 \\ &= \sin^2\theta\sin^2\alpha + \cos^2\theta\cos^2\alpha + 2\cos\theta\sin\theta\cos\alpha\sin\alpha. \end{aligned} \tag{2.32}$$

La situation est en fait très similaire à l'expérience des fentes de Young. Le terme qui est absent dans le calcul classique est un terme d'interférence entre les "deux chemins possibles": avant d'être observés dans le photodétecteur, les photons ont un comportement ondulatoire et on ne peut pas leur associer des trajectoires et des états de polarisations

"x" et "y" bien définis. Leur état est $|\theta\rangle = \cos\theta|x\rangle + \sin\theta|y\rangle$ tant qu'ils ne sont pas détectés.

2.4 Observables associées à la polarisation

Reprenons l'expérience de photodétection avec le système Analyseur + Détecteur. (Figure **2.9**)

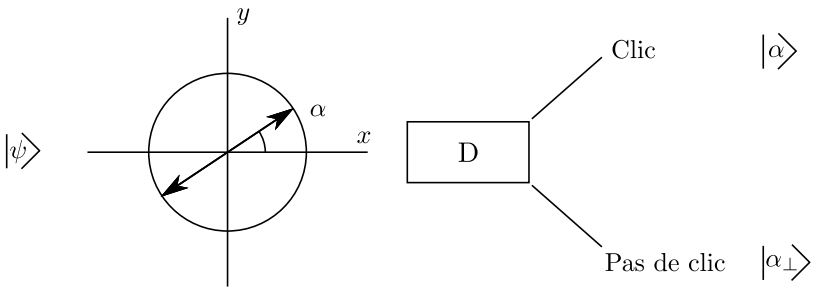


Figure 2.9 Expérience de photodétection

Nous avons vu que si le photon entrant est dans l'état $|\theta\rangle$, la $\mathbb{P}$ [clic] $= \cos^2(\theta-\alpha) = |\langle\alpha|\theta\rangle|^2$ et la $\mathbb{P}$ [pas de clic] $= \sin^2(\theta-\alpha) = |\langle\alpha_\perp|\theta\rangle|^2$. Le clic détecte une transition $|\theta\rangle \to |\alpha\rangle$ et l'absence de clic détecte une transition $|\theta\rangle \to |\alpha_\perp\rangle$. Nous pouvons enregistrer les résultats de l'expérience dans une variable (aléatoire) $p_\alpha = +1$ (clic) et $p_\alpha = -1$ (pas de clic). La valeur moyenne de cette variable aléatoire est

$$\mathbb{E}[p_\alpha] = (+1)|\langle\alpha|\theta\rangle|^2 + (-1)|\langle\alpha_\perp|\theta\rangle|^2 \tag{2.33}$$

.

Plus généralement, si l'état entrant dans le système Analyseur + Détecteur est $|\psi\rangle$ un état de $\mathbb{C}^2 = \left\{\begin{pmatrix}a\\b\end{pmatrix};\ a \text{ et } b \in \mathbb{C} \text{ et } |a|^2 + |b|^2 = 1\right\}$ on a:

$$\mathbb{E}[p_\alpha] = (+1)|\langle\alpha|\psi\rangle|^2 + (-1)|\langle\alpha_\perp|\psi\rangle|^2 \tag{2.34}$$

Cette expression peut se mettre sous la forme

$$\mathbb{E}[p_\alpha] = (+1)\langle\psi|\alpha\rangle\langle\alpha|\psi\rangle + (-1)\langle\psi|\alpha_\perp\rangle\langle\alpha_\perp|\psi\rangle \tag{2.35}$$

(Ici on utilise $\overline{\langle\alpha|\psi\rangle} = \langle\psi|\alpha\rangle$ qui est une propriété du produit scalaire). En d'autres termes

$$\begin{aligned}\mathbb{E}[p_\alpha] &= \langle\psi|\ (|\alpha\rangle\langle\alpha| - |\alpha_\perp\rangle\langle\alpha_\perp|)\ |\psi\rangle \\ &\equiv \langle\psi|P_\alpha|\psi\rangle\end{aligned} \tag{2.36}$$

où on a **<u>défini</u>** "l'observable polarisation"

$$P_\alpha = (+1)|\alpha\rangle\langle\alpha| + (-1)|\alpha_\perp\rangle\langle\alpha_\perp|. \tag{2.37}$$

Avant de discuter la signification physique de P_α, nous discutons sa signification mathématique. Ici $|\alpha\rangle\langle\alpha|$ est un ket fois un bra, c'est-à-dire un vecteur fois son transposé: ceci est un projecteur sur le vecteur $|\alpha\rangle$. De même $|\alpha_\perp\rangle\langle\alpha_\perp|$ est un projecteur sur $|\alpha_\perp\rangle$. Ces projecteurs ne sont rien d'autre que des matrices:

$$\begin{aligned}|\alpha\rangle\langle\alpha| &= \begin{pmatrix}\cos\alpha \\ \sin\alpha\end{pmatrix}\begin{pmatrix}\cos\alpha & \sin\alpha\end{pmatrix} \\ &= \begin{pmatrix}\cos^2\alpha & \cos\alpha\sin\alpha \\ \sin\alpha\cos\alpha & \sin^2\alpha\end{pmatrix}\end{aligned} \tag{2.38}$$

$$\begin{aligned}|\alpha_\perp\rangle\langle\alpha_\perp| &= \begin{pmatrix}-\sin\alpha \\ \cos\alpha\end{pmatrix}\begin{pmatrix}-\sin\alpha & \cos\alpha\end{pmatrix} \\ &= \begin{pmatrix}\sin^2\alpha & -\cos\alpha\sin\alpha \\ -\sin\alpha\cos\alpha & \cos^2\alpha\end{pmatrix}\end{aligned} \tag{2.39}$$

et donc

$$P_\alpha = \begin{pmatrix}\cos(2\alpha) & \sin(2\alpha) \\ \sin(2\alpha) & -\cos(2\alpha)\end{pmatrix} \tag{2.40}$$

On peut vérifier que cette matrice possède les valeurs propres ± 1 associées aux vecteurs propres $|\alpha\rangle = \begin{pmatrix}\cos\alpha \\ \sin\alpha\end{pmatrix}$ et $|\alpha_\perp\rangle = \begin{pmatrix}-\sin\alpha \\ \cos\alpha\end{pmatrix}$. C'est-à-dire:

$$P_\alpha|\alpha\rangle = (+1)|\alpha\rangle \tag{2.41}$$

$$P_\alpha|\alpha_\perp\rangle = (-1)|\alpha_\perp\rangle \tag{2.42}$$

En fait, il est très instructif de faire cette vérification en notation de Dirac plutôt qu'en travaillant avec le tableau matriciel.

$$\begin{aligned} P_\alpha|\alpha\rangle &= (|\alpha\rangle\langle\alpha| - |\alpha_\perp\rangle\langle\alpha_\perp|)\,|\alpha\rangle \\ &= |\alpha\rangle \underbrace{\langle\alpha|\alpha\rangle}_{1} - |\alpha_\perp\rangle \underbrace{\langle\alpha_\perp|\alpha\rangle}_{0} \\ &= |\alpha\rangle \end{aligned} \tag{2.43}$$

$$\begin{aligned} P_\alpha|\alpha_\perp\rangle &= (|\alpha\rangle\langle\alpha| - |\alpha_\perp\rangle\langle\alpha_\perp|)\,|\alpha_\perp\rangle \\ &= |\alpha\rangle \underbrace{\langle\alpha|\alpha_\perp\rangle}_{0} - |\alpha_\perp\rangle \underbrace{\langle\alpha_\perp|\alpha_\perp\rangle}_{1} \\ &= -|\alpha_\perp\rangle \end{aligned} \tag{2.44}$$

Quelle est l'interprétation physique de la matrice ou "observable" P_α? Cette matrice caractérise la quantité mesurée, ici "la polarisation du photon dans les directions (α, $\alpha_\perp$)". L'appareil qui sert à mesurer cette quantité est l'Analyseur + Détecteur. Le résultat de la mesure est donné par les valeurs propres et vecteurs propres de la matrice. Ici, il y a deux résultats possibles (+1, $|\alpha\rangle$) et (-1, $|\alpha_\perp\rangle$). La probabilité d'obtenir (+1, $|\alpha\rangle$) est $|\langle\alpha|\psi\rangle|^2$ ou bien $\langle\psi|\alpha\rangle\,\langle\alpha|\psi\rangle$. La probabilité d'obtenir (-1, $|\alpha_\perp\rangle$) est $|\langle\alpha_\perp|\psi\rangle|^2$ ou bien $\langle\psi|\alpha_\perp\rangle\,\langle\alpha_\perp|\psi\rangle$. La valeur moyenne de "l'observable" est $\langle\psi|P_\alpha|\psi\rangle$.

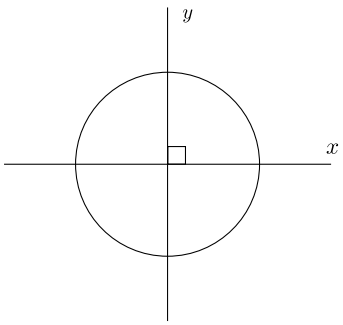
Si l'analyseur fait un angle β (avec x) et non pas α, l'appareil de mesure est différent. L'observable mesurée est alors aussi différente, notamment $P_\beta = |\beta\rangle\langle\beta| - |\beta_\perp\rangle\langle\beta_\perp|$.

Il existe trois observables qui jouent un rôle privilégié et que nous rencontrerons plus tard souvent.

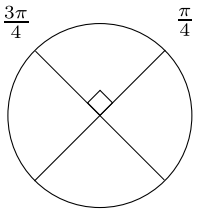
Si $\alpha = 0$, l'analyseur mesure la polarisation dans les directions x et y et

$$P_{\alpha=0} = |x\rangle\langle x| - |y\rangle\langle y| = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{2.45}$$

On symbolise souvent cet analyseur ou cet "appareil de mesure" par



Si $\alpha = \dfrac{\pi}{4}$, l'analyseur mesure la polarisation dans les directions $\dfrac{\pi}{4}$ et $\dfrac{3\pi}{4}$ et cet appareil de mesure est souvent symbolisé par



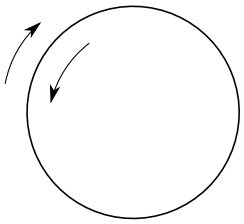
L'observable correspondant est:

$$P_{\alpha=\frac{\pi}{4}} = |\frac{\pi}{4}\rangle\langle\frac{\pi}{4}| - |\frac{3\pi}{4}\rangle\langle\frac{3\pi}{4}| = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \tag{2.46}$$

Si on utilise un analyseur qui mesure la polarisation circulaire, l'observable correspondante est:

$$P_{\text{circ}} = |R\rangle\langle R| - |L\rangle\langle L| = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}. \tag{2.47}$$

Le symbole pour cet analyseur est



Il est instructif de vérifier cette identité en utilisant

$$|R\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ i \end{pmatrix} \qquad \langle R| = \frac{1}{\sqrt{2}}\begin{pmatrix} 1, & -i \end{pmatrix} \tag{2.48}$$

$$|L\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -i \end{pmatrix} \qquad \langle L| = \frac{1}{\sqrt{2}}\begin{pmatrix} 1, & i \end{pmatrix} \tag{2.49}$$

Nous verrons dans les axiomes de la MQ qu'une quantité mesurable est toujours représentée par une matrice hermitienne. Les résultats d'une mesure de cette quantité sont les valeurs propres et vecteurs propres de cette matrice. Si λ_i et $|v_i\rangle$ sont valeur propre et vecteur propre de la matrice, la règle de Born stipule que

$$\mathbb{P}\,[\text{observer } \lambda_i \text{ et } |v_i\rangle] = |\langle v_i|\psi\rangle|^2 \tag{2.50}$$

quand $|\psi\rangle$ est l'état initial avant la mesure.

2.5 Moments magnétiques classiques

Dans les quelques paragraphes qui suivent, nous discutons un autre type de degré de liberté classique: le spin 1/2. Tout d'abord nous devons faire quelques rappels sur la notion de moment magnétique classique.

Considérons une boucle de courant plongée dans un champ magnétique uniforme. (Figure **2.10**)

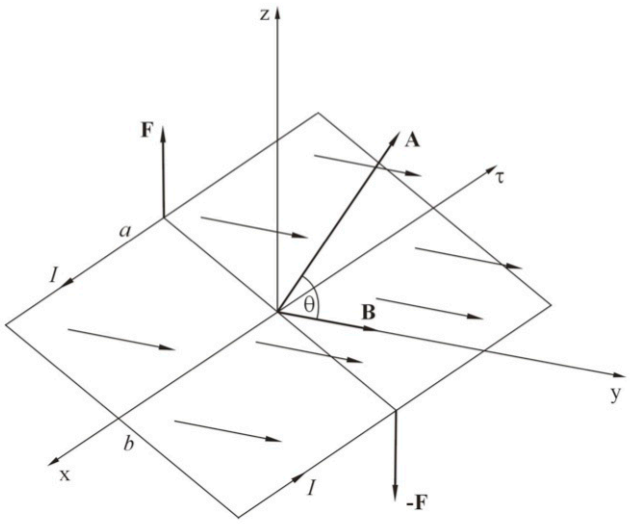


Figure 2.10 Boucle rectangulaire porteuse d'un courant électrique et plongée dans un champ magnétique uniforme

Si cette boucle est traversée par un courant, la force de Laplace qui s'exerce sur les sections du fil aura tendance à ramener la boucle dans la position d'équilibre. Cette position d'équilibre correspond à la boucle de courant $\perp$ à $\vec{B}$ de façon à ce que les forces de Laplace s'équilibrent. L'origine microscopique de la force de Laplace est en fait la force de Lorentz. Pour une particule de charge q, de vitesse $\vec{v}$, dans un champ magnétique $\vec{B}$, la force de Lorentz qui s'exerce sur cette particule est (produit vectoriel ici): $\vec{F} = q\vec{v} \times \vec{B}$. Si δq est la quantité de charge traversant une section du fil pendant un temps δt le terme la force s'exerçant sur une longueur $\delta\vec{l}$ du fil est $\delta\vec{F} = \delta q \frac{\delta\vec{l}}{\delta t} \times \vec{B} = I\delta\vec{l} \times \vec{B}$. Cette dernière expression est celle de la force de Laplace. On peut calculer le

travail total de la force de Laplace s'exerçant sur la boucle de courant et en déduire que celle-ci possède une énergie potentielle dans le champ magnétique. La boucle est à l'équilibre lorsque cette énergie est minimale (cas ou la boucle est $\perp$ au champ). Un calcul (que nous omettons ici) montre que cette énergie potentielle est donnée par

$$E = -\vec{M} \cdot \vec{B} \tag{2.51}$$

où $\vec{M}$ est le *moment magnétique* associé à la boucle de courant $\vec{M} = I\vec{S}$, avec $\vec{S}$ le vecteur unité perpendiculaire à la surface de longueur égale à la surface. On peut comprendre intuitivement cette formule en remarquant que (force de Laplace)×(distance) possède les mêmes unités. Ce calcul montre aussi que l'on peut associer un moment magnétique à une charge q en rotation autour de la boucle de courant

$$\vec{M} = \frac{q}{2}\vec{r} \times \vec{v} = \frac{q}{2m}\vec{L} \tag{2.52}$$

où $\vec{L} = \vec{r} \times \vec{p}$ est le moment cinétique. L'énergie est minimale lorsque le moment cinétique $\vec{L}$ ou bien le moment magnétique $\vec{M}$ pointent dans la direction du champ $\vec{B}$; et est maximale lorsque $\vec{L}$ ou $\vec{M}$ pointent dans la direction opposée au champ.

Il existe aussi d'autres types de "moments magnétiques" dans la nature qui ne sont pas associés au mouvement de charges, mais sont "intrinsèques aux particules". Par exemple, l'électron, le proton, le neutron (et par conséquent les noyaux atomiques) possèdent des moments magnétiques intrinsèques. Tout se passe comme si ces particules étaient des petites toupies en rotation sur elles-mêmes, ce qui produit des boucles de courant. Néanmoins, cette image classique est trop naïve et n'est finalement pas très utile pour comprendre le formalisme nécessaire à la description des moments magnétiques intrinsèques. Nous allons voir que les formules $E = -\vec{M} \cdot \vec{B}$ et $\vec{M} \sim \frac{q}{2m}\vec{L}$ sont toujours valables sauf que $\vec{L}$ et $\vec{M}$ sont des vecteurs dont les composantes sont des matrices! Dans ce contexte, le vecteur (à composantes matricielles) $\vec{L}$ s'appelle le *spin* (et on utilise plutôt la notation $\vec{S}$ à la place de $\vec{L}$).

2.6 L'expérience de Stern-Gerlach

L'expérience célèbre de Stern et Gerlach mis en évidence le "moment magnétique intrinsèque" de l'électron. À ce moment magnétique intrinsèque est associé un "moment cinétique intrinsèque" que l'on appelle le spin.

L'expérience consiste à préparer un faisceau d'atomes d'Argent qui sortent d'un four et à faire passer ce faisceau à travers un champ magnétique possédant un gradient dans la direction z. (Figure **2.11**)

Lorsque les particules passent à travers l'aimant, le faisceau est séparé en deux et on observe deux taches séparées sur l'écran.

Ce résultat expérimental est étonnant à plusieurs titres. Tout d'abord l'atome d'Argent est neutre si bien que la force de Lorentz ne devrait pas affecter la trajectoire du faisceau. On peut très bien imaginer que, bien que neutres, les atomes d'Argent possèdent un moment magnétique $\vec{M}$ non nul. Alors la force qui s'exerce entre eux vaut $\vec{\nabla}(\vec{M} \cdot B) = \vec{M} \cdot \vec{\nabla} B$

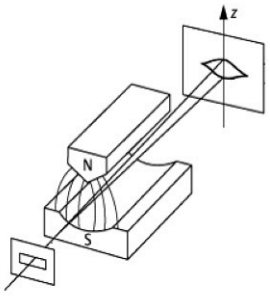


Figure 2.11 Expérience de Stern et Gerlach.

et on conçoit que le faisceau soit dévié. Mais on s'attendrait à ce qu'à la sortie du four $\vec{M}$ soit "incohérent" et pointe dans des directions aléatoires. Puisque $\vec{M} \cdot \vec{\nabla} B = M_z (\vec{\nabla} B)_z$ et M_z est continu (prend des valeurs aléatoires) on s'attendrait à observer une tache plus ou moins uniforme étalée sur l'écran. Mais l'observation consiste en deux petites taches séparées (le long de l'axe z). *Cela indique qu'en fait M_z prend deux valeurs possibles*.

Cette *quantification du moment magnétique* ne peut pas être expliquée par la physique classique. Les électrons de l'atome d'Argent (et tous les électrons dans la nature) possèdent un moment magnétique intrinsèque qui n'a rien à voir avec leur mouvement orbital. Ce moment magnétique intrinsèque prend deux valeurs possibles. Pour les atomes d'Argent, le nombre total d'électrons est impair et il se trouve que les moments magnétiques intrinsèques des électrons se compensent deux à deux sauf pour l'électron de la couche atomique externe de l'atome d'Argent. Cet électron sur la couche atomique externe confère à l'atome un moment magnétique quantifié prenant deux valeurs possibles.

Dans le paragraphe suivant nous discutons le moment magnétique intrinsèque et le spin de l'électron.

2.7 Spin $\frac{1}{2}$ et moments magnétiques quantiques

Les particules élémentaires possèdent un moment cinétique intrinsèque appelé "spin" et un moment magnétique intrinsèque associé. Le spin est une sorte d'analogue du moment cinétique $\vec{L}$. Néanmoins, il serait trop naïf de considérer la particule (ici l'électron) comme une boule minuscule tournant sur elle-même. Rappelons-nous que nous avons déjà abandonné la notion de trajectoire bien définie.

Le "vecteur" associé au spin est noté $\vec{S}$. De façon analogue à $\vec{L} = (L_x, L_y, L_z)$ il possède trois composantes (S_x, S_y, S_z). L'unité de $\vec{L}$ est "quantité de mouvement $\times$ position" = "J·s" ≡ unité de $\hbar$. Pour cette raison, on posera

$$\vec{S} = \frac{\hbar}{2}\vec{\sigma} \tag{2.53}$$

où $\vec{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$ sont des composantes sans dimension. Comme nous allons le voir ces composantes sont chacune des matrices: en effet le spin (tout comme $\vec{L} = \vec{r} \times \vec{p} \rightarrow \vec{r} \times \frac{\hbar}{i}\vec{\nabla}$) est une observable donc une matrice en MQ.

Le moment magnétique associé au spin est

$$\vec{M} = \gamma \vec{S} \tag{2.54}$$

tout comme $\vec{M} = \frac{q}{2m}\vec{L}$ pour une particule de charge q (et masse m). Ici γ est une constante qui dépend aussi de q et m; pour une particule chargée telle que l'électron $\gamma = g\frac{q}{2m}$ avec $g \approx 2.002...$; pour le proton $g \approx 5$.

L'énergie associée à l'interaction entre le moment magnétique et le champ magnétique est comme dans le cas classique donnée par

$$-\vec{M} \cdot \vec{B} = -\gamma \frac{\hbar}{2}\vec{\sigma} \cdot \vec{B} \tag{2.55}$$

Comme pour toutes les observables en MQ, nous allons voir que cette quantité est une matrice (qui s'appelle l'Hamiltonien du spin dans le champ $\vec{B}$).

En bloquant un des deux faisceaux dans l'appareil de Stern-Gerlach, on peut fabriquer un **filtre** qui est l'analogue des filtres polariseurs et/ou analyseurs. Ce filtre sélectionne un des deux états possibles pour les particules de "spin 1/2". On peut alors procéder à des expériences similaires à celles faites avec les photons.

Cela mène alors à la conclusion suivante. Pour des particules telles que l'électron, les matrices $\sigma_x, \sigma_y, \sigma_z$ sont des matrices 2×2 similaires aux observables de polarisation linéaires et circulaires:

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}; \qquad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}; \qquad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{2.56}$$

Lors de l'expérience de Stern-Gerlach, on mesure en fait la composante z du spin. Le résultat de la mesure donne deux valeurs possibles correspondant aux valeurs propres de $M_z = \gamma\frac{\hbar}{2}\sigma_z$. Ces deux valeurs propres sont égales à ± 1 multipliées par la constante $\gamma\frac{\hbar}{2}$. Les vecteurs propres correspondants sont

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix} \equiv |\uparrow\rangle \qquad \text{et} \qquad \begin{pmatrix} 0 \\ 1 \end{pmatrix} \equiv |\downarrow\rangle \tag{2.57}$$

et ce sont les deux états possibles obtenus lors de la mesure de σ_z. On peut vérifier qu'en notation de Dirac:

$$\sigma_z = (+1)|\uparrow\rangle\langle\uparrow| + (-1)|\downarrow\rangle\langle\downarrow| \tag{2.58}$$

$$= \begin{pmatrix}1 & 0\end{pmatrix}\begin{pmatrix}1\\0\end{pmatrix} - \begin{pmatrix}0 & 1\end{pmatrix}\begin{pmatrix}0\\1\end{pmatrix} \tag{2.59}$$

Si on tourne l'appareil le long de l'axe x, on mesure l'observable σ_x ou le moment magnétique $M_x = \gamma\frac{\hbar}{2}\sigma_x$. Les valeurs propres sont à nouveau ± 1 et les états propres correspondants sont

$$\frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\end{pmatrix} = \frac{1}{\sqrt{2}}(|\uparrow\rangle + |\downarrow\rangle) \equiv |+\rangle \tag{2.60}$$

$$\frac{1}{\sqrt{2}}\begin{pmatrix}1\\-1\end{pmatrix} = \frac{1}{\sqrt{2}}(|\uparrow\rangle - |\downarrow\rangle) \equiv |-\rangle \tag{2.61}$$

En notation de Dirac

$$\sigma_x = (+1)|+\rangle\langle+| + (-1)|-\rangle\langle-| \tag{2.62}$$

De même on a pour σ_y les valeurs propres ± 1 avec les états propres

$$\frac{1}{\sqrt{2}}\begin{pmatrix}1\\i\end{pmatrix} = \frac{1}{\sqrt{2}}(|\uparrow\rangle + i|\downarrow\rangle) \equiv |\circlearrowleft\rangle \tag{2.63}$$

$$\frac{1}{\sqrt{2}}\begin{pmatrix}1\\-i\end{pmatrix} = \frac{1}{\sqrt{2}}(|\uparrow\rangle - i|\downarrow\rangle) \equiv |\circlearrowright\rangle \tag{2.64}$$

et

$$\sigma_y = (+1)|\circlearrowleft\rangle\langle\circlearrowleft| + (-1)|\circlearrowright\rangle\langle\circlearrowright| \tag{2.65}$$

2.8 L'espace de Hilbert du spin $\frac{1}{2}$

Nous avons vu que les photons possèdent un degré de liberté de polarisation. Les états quantiques possibles de la polarisation du photon sont des vecteurs de l'espace de Hilbert $\mathbb{C}^2 = \left\{\begin{pmatrix}a\\b\end{pmatrix};\ a \text{ et } b \in \mathbb{C} \text{ et } |a|^2 + |b|^2 = 1\right\}$.

Un état général peut s'écrire en notation de Dirac $a|x\rangle + b|y\rangle$ où $|x\rangle = \begin{pmatrix}1\\0\end{pmatrix}$ et $|y\rangle = \begin{pmatrix}0\\1\end{pmatrix}$. De plus, pour les photons, une paramétrisation naturelle qui est analogue à la paramétrisation de la polarisation du champ électrique consiste à prendre $a = \cos\theta$ et $b = (\sin\theta)e^{i\phi}$. Ainsi les états de polarisation du photon sont en général

$$|\theta, \phi\rangle = \cos\theta|x\rangle + e^{i\phi}\sin\theta|y\rangle \tag{2.66}$$

avec $0 \leq \theta \leq \pi$ et $0 \leq \phi \leq 2\pi$.

Les particules de "spin 1/2" possèdent un degré de liberté interne analogue à la polarisation pour un photon. Des exemples de particules possédant un spin 1/2 sont l'électron, le proton, le neutron, etc. Les noyaux atomiques possèdent un spin total qui est la somme des spins des protons et neutrons. Souvent ceux-ci se compensent entre eux et si le nombre de protons et neutrons est impair, le spin résultant du noyau classique est à nouveau de "type 1/2". Comme nous l'avons vu, ce degré de liberté peut prendre essentiellement deux valeurs lors d'une mesure (par exemple avec un appareillage de Stern-Gerlach).

Ainsi l'espace des vecteurs d'état du spin 1/2 est à nouveau l'espace à deux dimensions $\mathbb{C}^2 = \left\{\begin{pmatrix} a \\ b \end{pmatrix};\ \ a \text{ et } b \in \mathbb{C} \text{ et } |a|^2 + |b|^2 = 1\right\}$. Cette fois, on préfère noter $\begin{pmatrix} 1 \\ 0 \end{pmatrix} = |\uparrow\rangle$ et $\begin{pmatrix} 0 \\ 1 \end{pmatrix} = |\downarrow\rangle$.

La paramétrisation naturelle est presque la même que pour les photons. Un état général de spin est

$$|\theta, \phi\rangle = \cos\left(\frac{\theta}{2}\right)|\uparrow\rangle + e^{i\phi}\sin\left(\frac{\theta}{2}\right)|\downarrow\rangle. \tag{2.67}$$

avec $0 \leq \theta \leq \pi$ et $0 \leq \phi \leq 2\pi$. La présence de $\frac{\theta}{2}$ au lieu de θ signifie entre autres que

$$|\theta = 0, \phi = 0\rangle = |\uparrow\rangle \quad \text{et} \quad |\theta = \pi, \phi = 0\rangle = |\downarrow\rangle. \tag{2.68}$$

Cette paramétrisation est naturelle, car si on renverse le champ magnétique dans l'appareil de Stern-Gerlach, on échange les deux taches sur l'écran. Renverser le champ magnétique revient à faire $\theta : 0 \to \pi$ et échanger les deux taches correspond à $|\uparrow\rangle \to |\downarrow\rangle$.

Notez que pour les photons, tourner un polariseur d'un angle π ne change pas la direction de polarisation. De même pour les photons $|\theta = 0, \phi = 0\rangle = |x\rangle$ et $|\theta = \pi, \phi = 0\rangle = -|x\rangle$ qui est équivalent à $|x\rangle$ (à une phase $e^{i\pi}$ près).

2.9 Notion de Bit Quantique

Nous avons vu que "l'espace de Hilbert" à deux dimensions

$$\mathbb{C}^2 = \left\{\begin{pmatrix} a \\ b \end{pmatrix};\ \ a \text{ et } b \in \mathbb{C}\right\} \tag{2.69}$$

muni du produit scalaire

$$\begin{pmatrix} \overline{c} & \overline{d} \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix} = \overline{c}\,a + \overline{d}\,b \tag{2.70}$$

intervient dans la description de deux systèmes physiques: la polarisation du photon et le spin 1/2 (de l'électron ou de certains noyaux atomiques).

Les degrés de libertés décrits par cet espace de Hilbert s'appelle aussi des systèmes à deux niveaux. La nature nous offre toute une variété de systèmes à deux niveaux décrits par l'espace des états $\mathbb{C}^2$. Parfois $\mathbb{C}^2$ est une description exacte du système comme c'est le cas pour la polarisation du photon et le spin de l'électron (ou du proton, neutron, noyaux atomiques). Parfois, c'est une description approximative qui consiste à retenir la partie importante des degrés de liberté plus compliqués. C'est le cas par exemple avec la molécule de Benzène C_6H_6. Dans des conditions normales (température ambiante) la molécule de Benzène existe dans l'état $\frac{1}{\sqrt{2}}(|1\rangle + |2\rangle)$ de superposition des deux états de base (Figure **2.12**). C'est l'état stable d'énergie la plus basse. Quand la molécule absorbe de la lumière ultraviolette (photons) elle peut passer dans l'état excité d'énergie plus élevée $\frac{1}{\sqrt{2}}(|1\rangle - |2\rangle)$. Les états de base de la figure ne sont pas stable: on peut se faire l'image d'une molécule qui résonne ou oscille entre ces deux états.

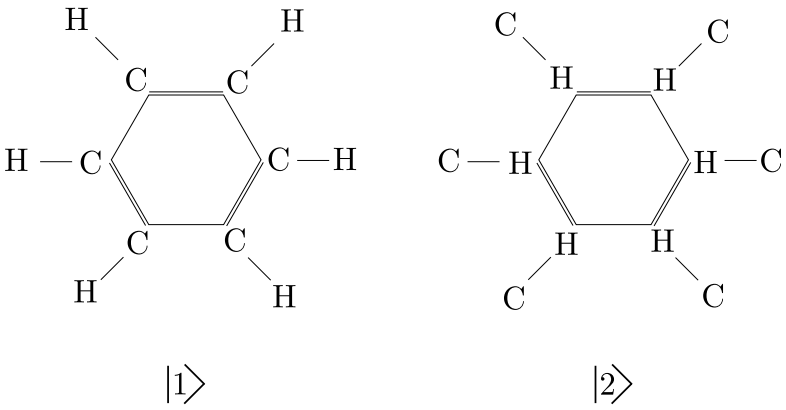


Figure 2.12 Molécule de Benzène. Les barres représentent les liaisons chimiques impliquant une paire d'électrons. Les doubles barres sont des doubles liaisons impliquant deux paires. L'état stable du Benzène est $\frac{1}{\sqrt{2}}(|1\rangle + |2\rangle)$.

On pourrait donner beaucoup d'autres exemples de systèmes à deux niveaux dans la nature, dans le domaine de la chimie, de la physique moléculaire ou atomique, de la physique nucléaire et des particules élémentaires.

Le bit quantique est simplement l'abstraction de la notion de système à deux niveaux du contexte physique détaillé. Un bit quantique est un état du type

$$|\psi\rangle = a|0\rangle + b|1\rangle \tag{2.71}$$

où $|0\rangle = \begin{pmatrix}1\\0\end{pmatrix}$, $|1\rangle = \begin{pmatrix}0\\1\end{pmatrix}$ et $a, b \in \mathbb{C}$ avec $|a|^2 + |b|^2 = 1$.

On adopte souvent (par convention) la convention relative au spin 1/2, c.-à-d. $a =$

$\cos\frac{\theta}{2}$ et $b = e^{i\phi}\sin\frac{\theta}{2}$ (mais pour la notion abstraite du bit quantique cela n'est pas obligatoire).

Le bit quantique est un degré de liberté qui possède une nature duale. Il est discret dans la mesure ou l'espace $\mathbb{C}^2$ possède la dimension 2 et les résultats de mesure sont binaires. Il est continu dans la mesure ou α et β sont des nombres complexes continus. Nous reviendrons sur ces considérations.

2.10 La sphère de Bloch

L'espace de Hilbert du bit quantique $\mathbb{C}^2$, t.q. $|\psi\rangle = a|0\rangle + b|1\rangle$, $|a|^2 + |b|^2 = 1$ est abstrait. La sphère de Bloch (Figure **2.13**) est une représentation géométrique très utile. Celle-ci est basée sur la paramétrisation

$$|\psi\rangle = |\theta,\phi\rangle = (\cos\frac{\theta}{2})|0\rangle + e^{i\phi}(\sin\frac{\theta}{2})|1\rangle \tag{2.72}$$

ou

$$|\psi\rangle = |\theta,\phi\rangle = (\cos\frac{\theta}{2})|\uparrow\rangle + e^{i\phi}(\sin\frac{\theta}{2})|\downarrow\rangle \tag{2.73}$$

On représente $|\theta,\phi\rangle$ par un vecteur unité sur une sphère où θ est l'angle par rapport à z et ϕ est l'angle par rapport à x dans le plan (x,y). Ici θ et ϕ ne sont rien d'autres que les coordonnées sphériques.

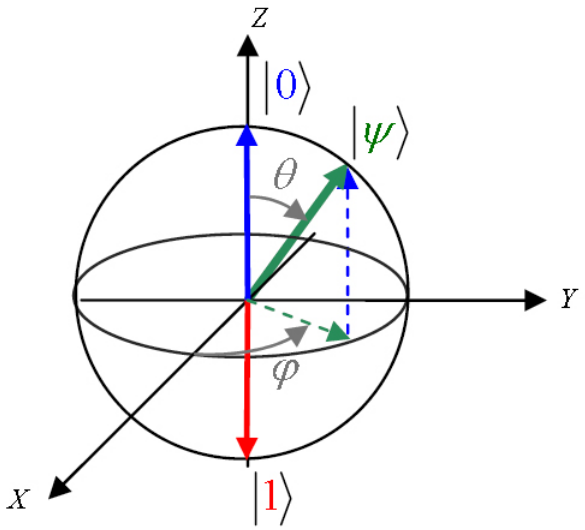


Figure 2.13 La sphère de Bloch

Les états des trois bases orthonormées ci-dessous

$$\{|0\rangle \equiv |\uparrow\rangle \;\; ; \;\; |1\rangle \equiv |\downarrow\rangle\} \tag{2.74}$$

$$\{\frac{1}{\sqrt{2}}(|\uparrow\rangle + |\downarrow\rangle) \equiv |+\rangle \;\; ; \;\; \frac{1}{\sqrt{2}}(|\uparrow\rangle - |\downarrow\rangle) \equiv |-\rangle\} \tag{2.75}$$

$$\{\frac{1}{\sqrt{2}}(|\uparrow\rangle + i|\downarrow\rangle) \equiv |\circlearrowleft\rangle \; ; \;\; \frac{1}{\sqrt{2}}(|\uparrow\rangle - i|\downarrow\rangle) \equiv |\circlearrowright\rangle\} \tag{2.76}$$

sont représentés sur la sphère de Bloch sur la Figure **2.14**.

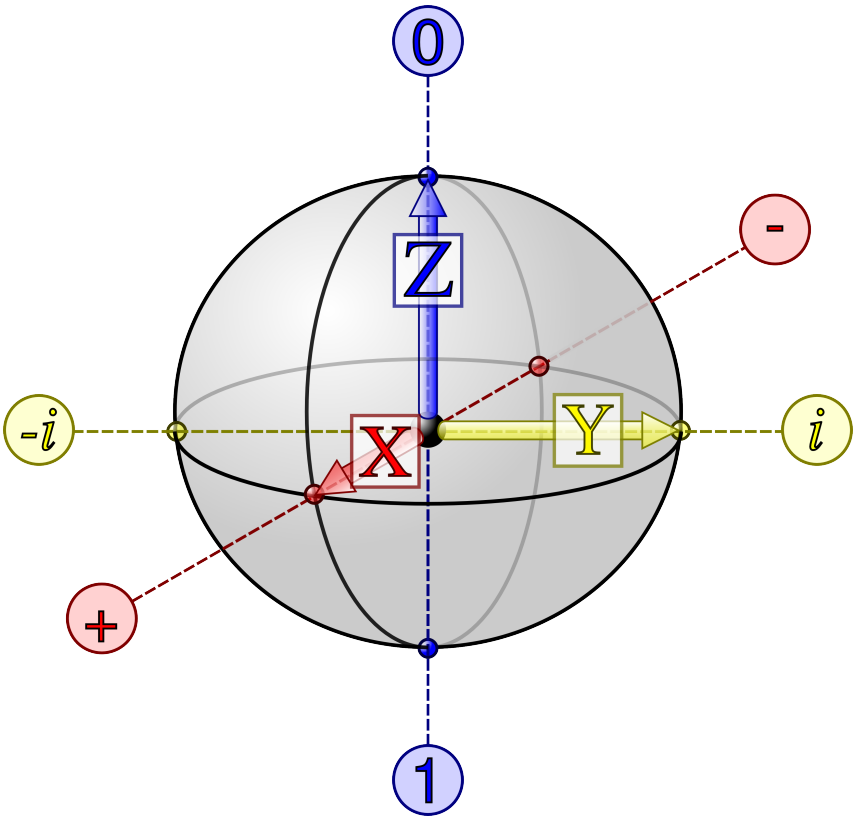


Figure 2.14 Représentation des trois états de bases orthonormées sur la sphère de Bloch où $|i\rangle = |\circlearrowleft\rangle$ et $|-i\rangle = |\circlearrowright\rangle$.

Ces trois bases s'appellent, pour des raisons évidentes les bases Z, X et Y en information quantique. Elles correspondent aux bases des états propres des trois matrices de Pauli du spin σ_z, σ_x et σ_y. Ces matrices sont aussi appelées souvent Z, X et Y.

3 Principes de la Mécanique Quantique

La physique moderne est fondée sur la Mécanique Quantique. Cette théorie a été élaborée suite à plusieurs expériences [p.ex: raies spectrales, corps noir, effet photoélectrique, effet Compton, diffraction des électrons sur les cristaux, physique atomique, expérience de Stern Gerlach, etc] et travaux des pères fondateurs [p.ex: Planck sur le corps noir 1900, Einstein sur le photon 1905, Bohr sur l'atome 1913, De Broglie sur la fonction d'onde 1924, Schrödinger sur l'évolution de la fonction d'onde 1926, Born sur l'interprétation de la fonction d'onde 1926, Heisenberg sur la formulation algébrique 1925, Dirac sur la mécanique quantique relativiste 1930, etc]. Un bref aperçu de quelques-uns de ces sujets a été donné au Chapitre **1**.

En 1930, les grands principes physiques de la mécanique quantique étaient essentiellement connus. Leur formulation mathématique précise et un cadre cohérent fut donné par Dirac et von Neumann. Leurs livres "Principles of Quantum Mechanics" (Dirac, 1930) et "Mathematische Grundlagen der Quantenmechanik" (von Neumann, 1932) jouèrent un rôle fondamental. Aujourd'hui même, les grands principes sont inchangés, et, combinés avec les principes de la relativité, décrivent avec succès une gamme impressionnante de phénomènes sur les échelles de distances, d'énergies et de températures associées à la physique de la matière condensée, atomique et moléculaire, nucléaire, sub-nucléaire. Malheureusement, on ne sait toujours pas combiner de façon cohérente la théorie classique de la gravitation avec la mécanique quantique, et, dans ce domaine, il n'existe aussi pas d'expériences pouvant guider les physiciens (car la force de gravité est en fait très faible).

Le cadre général de la MQ est l'espace de Hilbert. L'espace de Hilbert est essentiellement un espace vectoriel sur le corps des nombres complexes muni d'un produit scalaire. Nous allons donc commencer par donner quelques rappels d'algèbre linéaire sur ces espaces. En même temps, ceci est l'occasion d'introduire la notation de Dirac des "bras" et "kets" de façon un peu plus formelle. Ensuite nous formulons 5 postulats qui ensemble forment les grands principes de la MQ.

3.1 Algèbre linéaire en notation de Dirac

Un espace de Hilbert $\mathcal{H}$ est un espace vectoriel sur le corps $\mathbb{C}$, muni d'un produit scalaire. Pour un espace de dimension fini, cette définition est suffisante. Pour des espaces de dimension infinie il faut préciser des conditions qui permettent de prendre des

limites; mais dans le cadre de ce cours nous resterons en dimension finie comme cela est le plus souvent le cas en information quantique.

Les vecteurs sont notés $|\psi\rangle$ (prononcé "ket psi"). L'hermitien conjugué (transposé et complexe conjugué) est noté $\langle\psi|$ (prononcé "bra psi"). Le produit scalaire est noté $\langle\phi|\psi\rangle$. C'est le produit scalaire entre les vecteurs $|\phi\rangle$ et $|\psi\rangle$. On appelle aussi $\langle\phi|\psi\rangle$ un "braket". Le produit scalaire satisfait à:

1. *Positivité*: $\langle\phi|\phi\rangle \geq 0$ avec égalité si et seulement si $|\phi\rangle = 0$.
2. *Linéarité*: $\langle\phi|(\alpha|\psi_1\rangle + \beta|\psi_2\rangle) = \alpha\langle\phi|\psi_1\rangle + \beta\langle\phi|\psi_2\rangle$, $\alpha, \beta \in \mathbb{C}$
3. *Symétrie*: $\langle\phi|\psi\rangle = \overline{\langle\psi|\phi\rangle}$ où la barre dénote la conjugaison complexe.

Exemple 3.1: Bit quantique ou système à deux niveaux

$\mathcal{H} = \mathbb{C}^2 = \left\{\begin{pmatrix}\alpha\\ \beta\end{pmatrix} \text{ with } \alpha, \beta \in \mathbb{C}\right\}$. Le produit scalaire est $(\overline{\gamma}, \overline{\delta})\begin{pmatrix}\alpha\\ \beta\end{pmatrix} = \overline{\gamma}\alpha + \overline{\delta}\beta$. En notation de Dirac

$$\begin{pmatrix}\alpha\\ \beta\end{pmatrix} = \alpha|0\rangle + \beta|1\rangle$$

où $|0\rangle = \begin{pmatrix}1\\ 0\end{pmatrix}$, $|1\rangle = \begin{pmatrix}0\\ 1\end{pmatrix}$. De plus

$$(\bar{\gamma}, \bar{\delta}) = \bar{\gamma}\langle 0| + \bar{\delta}\langle 1|$$

et

$$(\bar{\gamma}\langle 0| + \bar{\delta}\langle 1|)(\alpha|0\rangle + \beta|1\rangle) = \overline{\gamma}\alpha\langle 0|0\rangle + \overline{\gamma}\beta\langle 0|1\rangle + \overline{\delta}\alpha\langle 1|0\rangle + \overline{\delta}\beta\langle 1|1\rangle = \overline{\gamma}\alpha + \overline{\delta}\beta$$

Exemple 3.2: Particule dans l'espace à trois dimensions

$\mathcal{H} = L^2(\mathbb{R}^3) = \{f : \mathbb{R}^3 \to \mathbb{C},\ \int \mathrm{d}^3\vec{x}\,|f(\vec{x})|^2 < \infty\}$. Le produit scalaire $\langle f|g\rangle = \int \mathrm{d}^3\vec{x}\,\overline{f(\vec{x})}g(\vec{x})$ et la norme induite $\|f\|_2 = \langle f|f\rangle^{1/2} = \int \mathrm{d}^3\vec{x}\,|f(\vec{x})|^2$. Cet espace joue un rôle fondamental en MQ mais nous n'en parlerons quasiment plus dans ce cours car nous nous occuperons uniquement de degrés de liberté discrets.

Produit tensoriel. Nous aurons besoin de la notion de *produit tensoriel*. Soit $\mathcal{H}_1$ et $\mathcal{H}_2$ deux espaces de Hilbert avec deux bases finies. Soit $|i\rangle_1$, $i = 1, ..., n_1$ la première base de $\dim\mathcal{H}_1 = n_1$ et $|j\rangle_2$, $j = 1, ..., n_2$ celle de $\dim\mathcal{H}_2 = n_2$. Nous pouvons former "l'espace produit"

$$\mathcal{H}_1 \otimes \mathcal{H}_2$$

qui est simplement le nouvel espace de Hilbert engendré par la base des vecteurs

$$|i\rangle_1 \otimes |j\rangle_2$$

(aussi notés $|i, j\rangle$ or $|i\rangle_1|j\rangle_2$). Il y a $n_1 \cdot n_2$ vecteurs dans cette base, donc

$$\dim \mathcal{H}_1 \otimes \mathcal{H}_2 = n_1 n_2$$

Un vecteur général de l'espace produit est

$$|\psi\rangle = \sum_{i=1}^{n_1}\sum_{j=1}^{n_2} c_{ij}|i, j\rangle = \sum_{i=1}^{n_1}\sum_{j=1}^{n_2} c_{ij}|i\rangle_1 \otimes |j\rangle_2$$

Le produit scalaire dans l'espace produit est par définition:

$$\langle i', j'|i, j\rangle = (\langle i'|_1 \otimes \langle j'|_2)(|i\rangle_1 \otimes |j\rangle_2) = \langle i'|i\rangle_1 \langle j'|j\rangle_2$$

.

Exemple 3.3: Produit tensoriel de bits quantiques

Pour un bit quantique, l'espace de Hilbert est $\mathbb{C}^2$. Nous verrons que l'espace de Hilbert de deux bits quantiques est $\mathbb{C}^2 \otimes \mathbb{C}^2$. Les vecteurs de base de $\mathbb{C}^2 \otimes \mathbb{C}^2$ sont $\{|0\rangle \otimes |0\rangle, |0\rangle \otimes |1\rangle, |1\rangle \otimes |0\rangle, |1\rangle \otimes |1\rangle\}$ ou bien $\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$. Un état général est

$$|\psi\rangle = \alpha_{00}|00\rangle + \alpha_{01}|01\rangle + \alpha_{10}|10\rangle + \alpha_{11}|11\rangle$$

On a $\dim \mathbb{C}^2 \otimes \mathbb{C}^2 = 4$ et bien sûr $\mathbb{C}^2 \otimes \mathbb{C}^2$ est isomorphe à $\mathbb{C}^4$.
Voici quelques exemples de produits scalaires:

- $\langle 00|00\rangle = \langle 0|0\rangle \langle 0|0\rangle = 1$
- $\langle 01|01\rangle = \langle 0|0\rangle \langle 1|1\rangle = 1$
- $\langle 01|11\rangle = \langle 0|1\rangle \langle 1|1\rangle = 0$

À partir de là on peut calculer le produit scalaire de $|\psi\rangle$ et $|\phi\rangle = \beta_{00}|00\rangle + \beta_{01}|01\rangle + \beta_{10}|10\rangle + \beta_{11}|11\rangle$. On trouve comme attendu $\langle\phi|\psi\rangle = \overline{\beta}_{00}\alpha_{00} + \overline{\beta}_{01}\alpha_{01} + \overline{\beta}_{10}\alpha_{10} + \overline{\beta}_{11}\alpha_{11}$. Il peut être utile de travailler dans une base canonique de $\mathbb{C}^4$

$$\begin{pmatrix}1\\0\\0\\0\end{pmatrix} = |00\rangle\ ,\ \begin{pmatrix}0\\1\\0\\0\end{pmatrix} = |01\rangle\ ,\ \begin{pmatrix}0\\0\\1\\0\end{pmatrix} = |10\rangle\ ,\ \begin{pmatrix}0\\0\\0\\1\end{pmatrix} = |11\rangle$$

Une fois cette correspondance (conventionnelle) fixée, on peut inférer les règles du produit tensoriel en composantes:

$$\begin{pmatrix}1\\0\end{pmatrix} \otimes \begin{pmatrix}1\\0\end{pmatrix} = \begin{pmatrix}1\\0\\0\\0\end{pmatrix}, \qquad \begin{pmatrix}1\\0\end{pmatrix} \otimes \begin{pmatrix}0\\1\end{pmatrix} = \begin{pmatrix}0\\1\\0\\0\end{pmatrix}$$

$$\begin{pmatrix}0\\1\end{pmatrix}\otimes\begin{pmatrix}1\\0\end{pmatrix}=\begin{pmatrix}0\\0\\1\\0\end{pmatrix},\qquad \begin{pmatrix}0\\1\end{pmatrix}\otimes\begin{pmatrix}0\\1\end{pmatrix}=\begin{pmatrix}0\\0\\0\\1\end{pmatrix}$$

Ces règles se généralisent à $\mathbb{C}^2\otimes\mathbb{C}^2\otimes\mathbb{C}^2$ etc.

Inégalités de Cauchy-Schwarz. Comme d'habitude:

$$|\langle\phi|\psi\rangle| \leq \langle\phi|\phi\rangle^{1/2}\,\langle\psi|\psi\rangle^{1/2}$$

Relation de Fermeture. Soit $|i\rangle$, $i = 1,...,n$ une base orthonormée d'un espace de Hilbert à n-dimensions. Un vecteur peut être développé

$$|\phi\rangle = \sum_{i=1}^{n} c_i|i\rangle, \qquad c_i = \langle i|\phi\rangle$$

Les composantes c_i sont obtenues en projetant $|\phi\rangle$ sur les vecteurs de base. Le développement devient

$$|\phi\rangle = \sum_{i=1}^{n}|i\rangle\,\langle i|\phi\rangle$$

Notez que $|i\rangle\langle i|$ la (matrice du) projecteur sur $|i\rangle$. On peut penser à $\sum_{i=1}^{n}|i\rangle\langle i|$ comme à la matrice identité agissant sur $|\phi\rangle$. On obtient donc la *relation de fermeture*

$$\sum_{i=1}^{n}|i\rangle\langle i| = \mathbb{1}$$

Observables. En MQ les observables ("quantités mesurables") sont représentées par des matrices hermitiennes agissant sur $\mathcal{H}$. Rappelons quelques propriétés importantes.

L'application $A : \mathcal{H} \to \mathcal{H}$, $|\psi\rangle \to A|\psi\rangle$ est linéaire si

$$A(\alpha|\phi_1\rangle + \beta|\phi_2\rangle) = \alpha(A|\phi_1\rangle) + \beta(A|\phi_2\rangle)$$

une application linéaire peut être représentée par une matrice aussi notée A.

Les éléments de matrice de A dans la base orthonormée $\{|i\rangle, i = 1,...,n\}$ de $\mathcal{H}$ sont notés $\langle i|A|j\rangle$ ou A_{ij}. Étant donné A, *l'adjoint* de A est noté $A^\dagger$ et défini par

$$\langle\phi|A^\dagger|\psi\rangle = \overline{\langle\psi|A|\phi\rangle}$$

Donc l'adjoint (ou l'hermitien conjugué) est l'application linéaire avec la matrice transposée et complexe conjuguée. On a pour les éléments de matrice

$$\langle i|A^\dagger|j\rangle = \overline{\langle j|A|i\rangle}, \qquad (A^\dagger)_{ij} = \overline{A_{ij}}$$

On dit que A est hermitienne si $A = A^\dagger$. On peut vérifier que $(A+B)^\dagger = A^\dagger + B^\dagger$ and $(AB)^\dagger = B^\dagger A^\dagger$.

On définit aussi le *commutateur*

$$[A,B] = AB - BA$$

et l'*anticommutateur*

$$\{A, B\} = AB + BA$$

Projecteurs en notation de Dirac. L'opération linéaire

$$|i\rangle\langle i| = P_i$$

est un projecteur sur le vecteur de base $|i\rangle$. Si P_i est un projecteur, on a $P_i^\dagger = P_i$ et $P_i^2 = P_i$. Voici comment on peut le vérifier en notation de Dirac:

$$P_i^\dagger = (|i\rangle\langle i|)^\dagger = (\langle i|)^\dagger(|i\rangle)^\dagger = |i\rangle\langle i| = P_i$$

$$P_i^2 = (|i\rangle\langle i|)(|i\rangle\langle i|) = |i\rangle \underbrace{\langle i|i\rangle}_{=1}\langle i| = |i\rangle\langle i| = P_i$$

Puisque $|i\rangle$ and $|j\rangle$ sont orthogonaux pour $i \neq j$ on a $P_iP_j = P_jP_i = 0$. En effet,

$$P_iP_j = (|i\rangle\langle i|)(|j\rangle\langle j|) = |i\rangle \underbrace{\langle i|j\rangle}_{=0}\langle j|) = 0$$

$$P_jP_i = (|j\rangle\langle j|)(|i\rangle\langle i|) = |j\rangle \underbrace{\langle j|i\rangle}_{=0}\langle i|) = 0$$

Si $|\phi\rangle$ est n'importe quel vecteur sur l'espace de Hilbert, alors $P_\phi = |\phi\rangle\langle\phi|$ est le projecteur sur $|\phi\rangle$.

Décomposition Spectrale. Les matrices hermitiennes sur l'espace de Hilbert ont une *décomposition spectrale*,

$$A = \sum_n a_n P_n$$

où $a_n \in \mathbb{R}$ sont les valeurs propres et P_n les projecteurs propres de A. Dans le cas non-dégénéré, on a

$$P_n = |\phi_n\rangle\langle\phi_n|$$

où $|\phi_n\rangle$ est le vecteur propre associé à la valeur propre a_n:

$$A|\phi_n\rangle = a_n|\phi_n\rangle$$

Les vecteurs propres et projecteurs associés à des valeurs propres différentes sont orthogonaux. De plus, ils satisfont à la relation de fermeture

$$\mathbb{1} = \sum_n P_n = \sum_n |\phi_n\rangle\langle\phi_n|$$

Nous écrirons souvent la décomposition spectrale sous la forme

$$A = \sum_n a_n|\phi_n\rangle\langle\phi_n|$$

3.2 Principes de la mécanique quantique

Dans ce paragraphe, nous expliquons les 5 grands principes de la MQ:

- les systèmes isolés sont décrits par *des vecteurs (kets) états d'un espace de Hilbert*,
- le vecteur d'état *évolue dans le temps de façon unitaire*,
- les *observables* sont décrites par des *matrices hermitiennes*,
- l'opération de mesure est un processus distinct de l'évolution temporelle: c'est une *projection aléatoire*,
- on peut *composer* des systèmes: leur espace de Hilbert est un espace produit (tensoriel).

Principe 1: Vecteurs d'états. L'état d'un système - isolé du reste de l'univers - est *complètement* spécifié par un vecteur de l'espace de Hilbert. Le vecteur $|\psi\rangle \in \mathcal{H}$ doit être normalisé $\langle\psi|\psi\rangle = 1$.

Exemple 3.4: Quelques vecteurs d'états

- La polarisation du photon (Section **2.2**) est décrite par $\mathcal{H} = \mathbb{C}^2$. Les vecteurs d'état de $\mathbb{C}^2$ sont $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$, $|\alpha|^2 + |\beta|^2 = 1$. Un état de polarisation linéaire $|\theta\rangle = \cos\theta|0\rangle + \sin\theta|1\rangle$, un état de polarisation circulaire $|\tilde{\theta}\rangle = \cos\theta|0\rangle + i\sin\theta|1\rangle$, et un état général
$$\cos\theta|0\rangle + e^{i\delta}\sin\theta|1\rangle$$
- Le spin $\frac{1}{2}$ (de l'électron par exemple: Section **2.7**) est décrit par le même espace de Hilbert. La paramétrisation naturelle est
$$\cos\frac{\theta}{2}|0\rangle + e^{i\delta}\sin\frac{\theta}{2}|1\rangle$$
La sphère de Bloch (Section **2.10**) est une représentation géométrique naturelle de l'espace de Hilbert de ces états.
- Pour une particule dans $\mathbb{R}^3$ on a $\mathcal{H} = L^2(\mathbb{R}^3)$. Les vecteurs d'états sont les fonctions d'ondes normalisées $\int \mathrm{d}^3\vec{x}\,|\psi(\vec{x})|^2 = 1$.

Principe 2: Évolution temporelle. Un système isolé évolue (au cours du temps) de façon unitaire. Cela signifie que si $|\psi\rangle$ est l'état au temps 0, l'état au temps t est de la forme $U_t|\psi\rangle$ où U_t est une matrice unitaire de $\mathcal{H} \to \mathcal{H}$. Ici unitaire signifie que $U_t^\dagger U_t = U_t U_t^\dagger = \mathbb{1}$ ou bien de façon équivalente $U_t^{-1} = U_t^\dagger$.

L'évolution unitaire forme un groupe (ou plutôt la représentation du groupe des translations temporelles) au sens suivant:

$$U_{t=0} = \mathbb{1}, \qquad U_{t_1}U_{t_2} = U_{t_1+t_2}$$

La MQ nous indique comment calculer U_t pour un système donné: il faut résoudre

l'équation de Schrödinger (Section **1.7**). En information quantique, nous ne nous intéressons pas (en général) à cette équation. On suppose (de façon optimiste) qu'un ingénieur ou un physicien saura construire un appareil (appelé circuit quantique) qui réalise l'opération unitaire U_t voulue. L'opération voulue sera spécifiée par l'algorithme quantique. Nous reviendrons sur ce point plus tard dans le cours.

Exemple 3.5: Porte logique de Hadamard

Un miroir semi-transparent décompose un rayon incident en rayon réfléchi et rayon transmis. Soit $\mathcal{H} = \mathbb{C}^2$ l'espace de Hilbert avec la base $|T\rangle$, $|R\rangle$. Le miroir semi-transparent agit de façon unitaire

$$|T\rangle \rightarrow \boxed{\text{H}} \rightarrow H|T\rangle = \frac{1}{\sqrt{2}}(|T\rangle + |R\rangle)$$

$$|R\rangle \rightarrow \boxed{\text{H}} \rightarrow H|R\rangle = \frac{1}{\sqrt{2}}(|T\rangle - |R\rangle)$$

La matrice unitaire H s'appelle matrice de Hadamard ou "porte logique de Hadamard"

$$H = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}.$$

Principe 3: Quantités observables. En mécanique quantique, une quantité observable (énergie, moment magnétique, moment cinétique, position, impulsion, vitesse, etc) est représentée par une matrice hermitienne.

Il n'est pas forcément évident de savoir comment choisir la matrice (ou opérateur) correspondante. Il existe un "principe de correspondance" (Section **1.8**) qui est une sorte de règle pratique pour construire l'opérateur à partir de la quantité classique. En fait, cette règle est parfois ambiguë, car les matrices sont des objets qui ne commutent pas. D'autre part, il existe des observables (comme le spin) qui n'ont pas d'analogue classique.

Exemple 3.6: Quelques observables

- Position x, impulsion $p = \frac{\hbar}{i}\frac{\partial}{\partial x}$, énergie ou Hamiltonien $\frac{p^2}{2m} + V(x)$. dans ce cours, nous n'aurons pas besoin de ces observables.
- Polarisation du photon. On envoie un photon à travers une lame biréfringente (Section **2.3**). Si D_y clique on enregistre -1 alors que si D_x clique on enregistre $+1$. Les observations sont décrites par l'observable

$$\mathcal{P} = (+1)|x\rangle\langle x| + (-1)|y\rangle\langle y|$$

cette observable est la matrice hermitienne $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ (exprimée dans la base $\{|x\rangle, |y\rangle\}$).
- Toute observable (matrice hermitienne) de $\mathcal{H} = \mathbb{C}^2$ peut être représentée par une matrice 2×2

$$A = \begin{pmatrix} \alpha & \beta \\ \overline{\beta} & \gamma \end{pmatrix}$$

ou en notation de Dirac

$$A = \alpha|0\rangle\langle 0| + \beta|0\rangle\langle 1| + \overline{\beta}|1\rangle\langle 0| + \gamma|1\rangle\langle 1|$$

Toutes ces matrices peuvent être représentées par une combinaison linéaire des matrices

$$I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, Y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

Les matrices hermitiennes X, Y, Z sont appelées les matrices de Pauli.

Les matrices de Pauli servent entre autres à décrire le spin $\frac{1}{2}$: il s'agit d'un vecteur à 3 composantes $\Sigma = (X, Y, Z)$. Dans la littérature physique $\Sigma = (\sigma_x, \sigma_y, \sigma_z)$. Les propriétés importantes de ces matrices sont

$$X^2 = Y^2 = Z^2 = I,\ XY = -YX,\ XZ = -ZX,\ YZ = -ZY$$

et

$$[X, Y] = 2iZ,\ [Y, Z] = 2iX,\ [Z, X] = 2iY$$

Principe 4: Postulat de la mesure. Soit un système préparé dans l'état $|\psi\rangle$. On veut mesurer une observable du système grâce à un appareil. L'appareil est modélisé par un ensemble de projecteurs orthogonaux $\{P_n\}$ satisfaisant $\sum_n P_n = \mathbb{1}$. Une mesure *projette*[1] l'état ψ du système qui devient juste après la mesure

$$|\phi_n\rangle = \frac{P_n|\psi\rangle}{||\, P_n|\psi\rangle||} = \frac{P_n|\psi\rangle}{\langle\psi|P_n|\psi\rangle^{1/2}}$$

Pour une mesure unique, *il n'y a pas moyen de prédire* l'état résultant $|\phi_n\rangle$: celui-ci est

[1] les physiciens disent que l'état ou la fonction d'onde est "réduit(e)"

aléatoire. Si l'expérience de mesure est répétée plusieurs fois, la probabilité (interprétation fréquentiste de la probabilité) d'observer n est

$$\mathbb{P}\ [\text{resultat } n] = |\langle \phi_n|\psi\rangle|^2 = \langle \psi|P_n|\psi\rangle$$

Remarque 1. Puisque $\sum_j P_j = \mathbb{1}$ et $|\psi\rangle$ sont normalisés on a

$$\sum_j \mathbb{P}\ [\text{resultat } j] = 1$$

Remarque 2. Avec $P_j = |j\rangle\langle j|$ la probabilité de l'état résultant j est

$$\mathbb{P}\ [\text{resultat } j] = \langle \psi|P_j|\psi\rangle = |\langle j|\psi\rangle|^2$$

et l'état juste après la mesure est $|j\rangle$.

Conséquence importante concernant la mesure des observables. Ce point est fondamental, car ce sont les observables que l'on mesure dans une expérience. L'appareil de mesure modélisé par un ensemble de projecteurs $\{P_n\}$ permet de mesurer toutes les observables de la forme $A = \sum_j a_j P_j$. Une mesure donne $|\psi\rangle \to |\phi_n\rangle$ pour un certain n. Puisque $A|\phi_n\rangle = a_n|\phi_n\rangle$, la valeur de A donnée par la mesure est précisément a_n.

La valeur moyenne des mesures de A si l'état du système est $|\psi\rangle$

$$\sum_j a_j \langle \psi|P_j|\psi\rangle = \langle \psi|A|\psi\rangle$$

et la variance

$$\sum_j a_j^2 \langle \psi|P_j|\psi\rangle - (\sum_j a_j \langle \psi|P_j|\psi\rangle)^2 = \langle \psi|A^2|\psi\rangle - \langle \psi|A|\psi\rangle^2$$

En pratique, on utilise le membre de droite de ces formules pour calculer les valeurs moyennes et variances.

Après une mesure, le vecteur d'état est réduit à $|\psi\rangle \to |\phi_n\rangle$, pour un certain n, et la valeur moyenne dans le nouvel état (i.e $|\phi_n\rangle$) devient a_n, et la variance devient nulle.

De plus, on ne peut mesurer avec le même appareil que des observables ayant les mêmes projecteurs propres. En particulier, la mesure simultanée (avec un même appareil) n'a de sens que si les observables ont les mêmes projecteurs et vecteurs propres. Elles peuvent avoir des valeurs propres différentes mais doivent commuter.

Exemple 3.7: Mesure de la polarisation du photon

Pour mesurer l'observable

$$\mathcal{P} = |x\rangle\langle x| - |y\rangle\langle y|$$

on utilise l'appareil constitué d'un analyseur orienté le long de x et un détecteur. Cet appareil est la réalisation physique de la base de mesure $\{|x\rangle, |y\rangle\}$. Si le photon traverse l'analyseur, l'état juste après la mesure est $|x\rangle$, et si le photon est absorbé, l'état juste après la mesure est $|y\rangle$. Les probabilités associées sont

$$\mathbb{P}\left[\text{resultat } +1\right] = |\langle x|\psi\rangle|^2, \quad \mathbb{P}\left[\text{resultat } -1\right] = |\langle y|\psi\rangle|^2$$

Si la préparation initiale de la polarisation des photons est $|\psi\rangle = \cos\theta|x\rangle + \sin\theta|y\rangle$ ces probabilités sont simplement $\cos^2\theta$ et $\sin^2\theta$. Supposons que l'analyseur soit tourné d'un angle γ. cela signifie que l'on mesure l'observable

$$\mathcal{P} = |\gamma\rangle\langle\gamma| - |\gamma_\perp\rangle\langle\gamma_\perp|$$

Les probabilités associées à cette mesure sont

$$\mathbb{P}\left[\text{resultat } +1\right] = |\langle\gamma|\psi\rangle|^2 = \cos^2(\theta-\gamma)$$

$$\mathbb{P}\left[\text{resultat } -1\right] = |\langle\gamma_\perp|\psi\rangle|^2 = \sin^2(\theta-\gamma)$$

Finalement, notons que dans le premier cas l'observable mesurée est la matrice

$$\mathcal{P} = Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

et dans le second cas

$$\mathcal{P} = \begin{pmatrix} \cos 2\gamma & \sin 2\gamma \\ \sin 2\gamma & -\cos 2\gamma \end{pmatrix} = (\cos 2\gamma)Z + (\sin 2\gamma)X$$

.

Principe d'incertitude Prenons un système dans l'état $|\psi\rangle$ et considérons deux observables A et B. Elles ont chacune une représentation spectrale

$$A = \sum_j a_j P_j, \quad B = \sum_j b_j Q_j$$

Comme expliqué précédemment, dans l'état $|\psi\rangle$, chaque observable possède la valeur moyenne $\langle\psi|A|\psi\rangle$, $\langle\psi|B|\psi\rangle$ et l'écart type $\Delta A = \sqrt{\langle\psi|A^2|\psi\rangle - \langle\psi|A|\psi\rangle^2}$, $\Delta B = \sqrt{\langle\psi|B^2|\psi\rangle - \langle\psi|B|\psi\rangle^2}$.

La relation ou inégalité d'incertitude de Heisenberg stipule

$$\Delta A \cdot \Delta B \geq \frac{1}{2}\langle\psi|[A,B]|\psi\rangle$$

L'interprétation de cette inégalité est la suivante. Si $[A,B] \neq 0$ il n'est pas possible de mesurer A et B simultanément avec précision infinie. Si $\Delta A = 0$ alors $\Delta B = \infty$.

L'exemple le plus frappant est $A = x$ (position) and $B = p = \frac{\hbar}{i}\frac{\partial}{\partial x}$ (impulsion ou quantité de mouvement). Dans ce cas $\Delta x \Delta p \geq \frac{h}{4\pi}$ et on ne peut pas mesurer avec une précision infinie la position et l'impulsion de la particule: ce n'est pas une limitation technologique, mais une limitation imposée par les lois de la nature.

Si $[A, B] = 0$ il existe une base commune de l'espace de Hilbert dans laquelle A et B sont toutes deux diagonales. En mesurant dans cette base, le postulat de la mesure indique que les observables peuvent être mesurées avec précision infinie. Il n'y a pas de contradiction avec le principe d'incertitude, car le membre de droite de l'inégalité de Heisenberg s'annule quand $[A, B] = 0$.

Principe 5: systèmes quantiques composés. Prenons deux systèmes $\mathcal{A}$ et $\mathcal{B}$ avec espaces de Hilbert $\mathcal{H}_\mathcal{A}$ and $\mathcal{H}_\mathcal{B}$. L'espace de Hilbert du système composé $\mathcal{AB}$ est donné par le produit tensoriel

$$\mathcal{H}_\mathcal{A} \otimes \mathcal{H}_\mathcal{B}$$

Les états de $\mathcal{AB}$ sont les vecteurs $|\psi\rangle \in \mathcal{H}_\mathcal{A} \otimes \mathcal{H}_\mathcal{B}$. Les postulats précédents s'appliquent aux systèmes composés.

Ce postulat est en fait non trivial et nous étudierons quelques conséquences. En particulier, Einstein, Podolsky et Rosen ainsi que Schrödinger furent les premiers à analyser la signification de ce postulat. Ces études menèrent aux inégalités de Bell, à la téléportation et au dense coding qui jouent aujourd'hui un rôle important en information quantique.

Exemple 3.8: Système composé

N bits quantiques possèdent l'espace de Hilbert

$$\underbrace{\mathbb{C}^2 \otimes \mathbb{C}^2 \otimes \mathbb{C}^2 \otimes ... \otimes \mathbb{C}^2}_{N \text{ copies}}$$

Si $\{|0\rangle, |1\rangle\}$ est une base pour $\mathbb{C}^2$, une base du système composé est donnée par

$$|b_1\rangle \otimes |b_2\rangle ... \otimes |b_N\rangle = |b_1 \dots b_N\rangle$$

où $b_i = \{0, 1\}$. Il y a 2^N états de base en correspondance avec 2^N suites de longueur N de bits classiques. Un état de N bits quantiques est une superposition des états de base:

$$|\psi\rangle = \sum_{b_1 \dots b_N} c_{b_1 \dots b_N} |b_1 \dots b_N\rangle$$

ou les coefficients $c_{b_1 \dots b_N}$ satisfont

$$\sum_{b_1 \dots b_N} |c_{b_1 \dots b_N}|^2 = 1$$

3.3 États produits et états intriqués

Les états d'un système composé $\mathcal{AB}$ appartiennent à $\mathcal{H}_\mathcal{A} \otimes \mathcal{H}_\mathcal{B}$. Un état est de type produit s'il peut être représenté comme

$$|\psi\rangle = |\phi\rangle_\mathcal{A} \otimes |\gamma\rangle_\mathcal{B}$$

On dit aussi que $|\psi\rangle$ est *séparable*.
Un état *intriqué* $|\psi\rangle \in \mathcal{H}_\mathcal{A} \otimes \mathcal{H}_\mathcal{B}$ est un état pour lequel il est impossible de trouver $|\phi\rangle_\mathcal{A} \in \mathcal{H}_\mathcal{A}$ et $|\gamma\rangle_\mathcal{B} \in \mathcal{H}_\mathcal{B}$ telle que ψ soit de la forme produit.

Les états intriqués engendrent des corrélations très spéciales entre les parties $\mathcal{A}$ et $\mathcal{B}$. Nous verrons que ces corrélations n'ont aucune contrepartie classique (et jouent un rôle important dans la téléportation par exemple).

Exemple 3.9: Quelques états produits

Considérons le cas de deux bits quantiques avec $\mathcal{A} \otimes \mathcal{B} = \mathbb{C}^2 \otimes \mathbb{C}^2$.

- Quelques états produits simples:

$$|0\rangle_\mathcal{A} \otimes |0\rangle_\mathcal{B} = |00\rangle$$
$$|0\rangle_\mathcal{A} \otimes |1\rangle_\mathcal{B} = |01\rangle$$
$$|1\rangle_\mathcal{A} \otimes |0\rangle_\mathcal{B} = |10\rangle$$
$$|1\rangle_\mathcal{A} \otimes |1\rangle_\mathcal{B} = |11\rangle$$

- Quelques états produits moins évidents:

$$\frac{1}{\sqrt{2}}(|0\rangle_\mathcal{A} + |1\rangle_\mathcal{B}) \otimes |0\rangle_\mathcal{B} = \frac{1}{2}(|00\rangle + |10\rangle)$$
$$\frac{1}{\sqrt{2}}(|0\rangle_\mathcal{A} + |1\rangle_\mathcal{B}) \otimes \frac{1}{\sqrt{2}}(|0\rangle_\mathcal{B} - |1\rangle_\mathcal{B}) = \frac{1}{2}(|00\rangle - |01\rangle + |10\rangle - |11\rangle)$$

- Il existe des états intriqués qui ne peuvent pas se mettre sous forme produit:

$$\frac{1}{\sqrt{2}}(|0\rangle_\mathcal{A} \otimes |0\rangle_\mathcal{B} + |1\rangle_\mathcal{A} \otimes |1\rangle_\mathcal{B}) = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$$
$$\frac{1}{\sqrt{2}}(|0\rangle_\mathcal{A} \otimes |0\rangle_\mathcal{B} - |1\rangle_\mathcal{A} \otimes |1\rangle_\mathcal{B}) = \frac{1}{\sqrt{2}}(|00\rangle - |11\rangle)$$
$$\frac{1}{\sqrt{2}}(|1\rangle_\mathcal{A} \otimes |0\rangle_\mathcal{B} + |0\rangle_\mathcal{A} \otimes |1\rangle_\mathcal{B}) = \frac{1}{\sqrt{2}}(|10\rangle + |01\rangle)$$
$$\frac{1}{\sqrt{2}}(|0\rangle_\mathcal{A} \otimes |0\rangle_\mathcal{B} - |1\rangle_\mathcal{A} \otimes |0\rangle_\mathcal{B}) = \frac{1}{\sqrt{2}}(|01\rangle + |10\rangle)$$

Ces quatre états jouent un rôle particulier et s'appellent états de Bell.

Production d'états intriqués. Soit un système composé avec état initial $|\phi\rangle_{\mathcal{A}} \otimes |\chi\rangle_{\mathcal{B}}$. Ce pourrait être par exemple deux électrons dans l'état de spin $|\uparrow\rangle \otimes |\downarrow\rangle$. Si on les laisse évoluer séparément sans interaction, l'opérateur d'évolution unitaire est de la forme $U_{\mathcal{A}} \otimes U_{\mathcal{B}}$ et

$$U_{\mathcal{A}} \otimes U_{\mathcal{B}}(|\uparrow\rangle \otimes |\downarrow\rangle) = U_{\mathcal{A}}|\uparrow\rangle \otimes U_{\mathcal{B}}|\downarrow\rangle$$

si bien que l'état reste dans un état produit.

Pour produire des états intriqués $\mathcal{A}$ and $\mathcal{B}$ doivent interagir pendant l'évolution temporelle, pour que $U_{\mathcal{AB}} \neq U_{\mathcal{A}} \otimes U_{\mathcal{B}}$. Toutes les interactions physiques connues sont locales dans l'espace et le temps: deux systèmes dans un état intriqué ont nécessairement été en contact dans le passé.

3.4 Impossibilité de "cloner" un état quantique

Les bits classiques peuvent être copiés. Par exemple, un texte peut être dupliqué ou copié avec une machine à photocopier "universelle": la même machine peut copier tous les textes.

Soit un ensemble d'états quantiques $|\psi\rangle \in \mathcal{H}$ et supposons que nous voulions construire une "machine universelle" qui "copie" tout $|\psi\rangle \in \mathcal{H}$. cette machine quantique devrait être décrite par un opérateur ou matrice unitaire U (ceci est vrai pour tout processus physique sauf pour celui de la mesure). L'espace de Hilbert est $\mathcal{H}_{\mathcal{A}} \otimes \mathcal{H}_{\mathcal{B}}$ où $\mathcal{A}$ est l'espace des états que l'on désire copier et $\mathcal{B}$ celui des copies. On commence par l'état initial

$$|\psi\rangle \otimes |\text{blank}\rangle$$

La machine produit la sortie:

$$|\psi\rangle \otimes |\text{blank}\rangle \rightarrow \boxed{\text{U}} \rightarrow |\psi\rangle \otimes |\psi\rangle$$

En termes mathématiques la question est: peut-on trouver un opérateur unitaire tel que pour un ensemble raisonnablement large d'états ψ

$$U(|\psi\rangle \otimes |\text{blank}\rangle) = |\psi\rangle \otimes |\psi\rangle$$

La réponse est non. Ce fait s'appelle parfois le "no cloning theorem". Par contre, il est possible de cloner/copier un ensemble d'états orthogonaux avec un U approprié (qui dépend de l'ensemble spécifique en question).

Preuve du théorème de non-clonage. Supposons qu'il existe U tel que $U^\dagger U = UU^\dagger = 1$ avec

$$U(|\phi_1\rangle \otimes |\text{blank}\rangle) = |\phi_1\rangle \otimes |\phi_1\rangle$$

$$U(|\phi_2\rangle \otimes |\text{blank}\rangle) = |\phi_2\rangle \otimes |\phi_2\rangle$$

En prenant l'hermitien conjugué de la deuxième équation

$$(\langle\phi_2| \otimes \langle\text{blank}|)U^\dagger = \langle\phi_2| \otimes \langle\phi_2|$$

En prenant le produit scalaire avec la première équation

$$\langle\phi_2| \otimes \langle \text{blank}|U^\dagger U|\phi_1\rangle \otimes |\text{blank}\rangle = (\langle\phi_2| \otimes \langle\phi_2|)(|\phi_1\rangle \otimes |\phi_1\rangle)$$

ce qui implique

$$\langle\phi_2|\phi_1\rangle \langle \text{blank}|\text{blank}\rangle = \langle\phi_2|\phi_1\rangle^2$$

donc

$$\langle\phi_2|\phi_1\rangle = 0 \text{ or } \langle\phi_2|\phi_1\rangle = 1$$

Nous concluons qu'il n'est pas possible de copier $|\phi_1\rangle$ and $|\phi_2\rangle$ qui ne sont sont identiques ou bien pas orthogonaux, avec le même U. En fait il est possible de copier une base orthogonale ou des états orthogonaux.

Les états non-orthogonaux ne peuvent pas être parfaitement distingués.

Il existe plusieurs variantes et raffinements du no-cloning theorem. Ici nous discutons une de ces variantes. Donnons-nous deux états $|\psi\rangle$ et $|\phi\rangle$ et essayons de construire une machine (unitaire) qui permet de les distinguer. Mathématiquement on cherche une matrice unitaire U telle que

$$U|\psi\rangle \otimes |a\rangle = |\psi\rangle \otimes |v\rangle$$

$$U|\phi\rangle \otimes |a\rangle = |\phi\rangle \otimes |v'\rangle$$

où les sorties $|v\rangle$ et $|v'\rangle$ sont différents. Le produit scalaire entre ces deux équations donne

$$\langle\phi| \otimes \langle a|U^\dagger U|\psi\rangle \otimes |a\rangle = (\langle\phi| \otimes \langle v'|)(|\psi\rangle) \otimes |v\rangle)$$

ceci implique

$$\langle\phi|\psi\rangle \langle a|a\rangle = \langle\phi|\psi\rangle \langle v'|v\rangle$$

Si $|\phi\rangle$ n'est pas orthogonal à $|\psi\rangle$ nous avons $\langle\phi|\psi\rangle \neq 0$ donc

$$\langle v'|v\rangle = \langle a|a\rangle = 1$$

Ainsi $|v\rangle = |v'\rangle$ et il n'y a pas d'information dans $|v\rangle$ et $|v'\rangle$ qui permet de distinguer $|\psi\rangle$ et $|\phi\rangle$.

4 La matrice Densité

Nous avons formulé les règles de la mécanique quantique pour des systèmes isolés. En particulier, l'état d'un système isolé est décrit par un vecteur de l'espace de Hilbert $|\psi\rangle \in \mathcal{H}$, avec $\||\psi\|| = 1$. Souvent, cette description n'est en fait pas suffisante. Tout d'abord, un système n'est jamais complétement isolé et cette hypothèse est, tout au plus, une idéalisation. Quand l'influence de l'environnement sur le système à son importance, il faut étendre la description grâce au formalisme de la *matrice densité*, qui remplace le vecteur d'état. Une autre situation qui requiert cette extension est la description d'un *ensemble statistique*. Considérons par exemple un gaz de photons dans une boîte contenant 30% des photons dans l'état de polarisation $|H\rangle$, 20% des photons dans l'état $|V\rangle$ et 50% des photons dans l'état $\frac{1}{\sqrt{2}}\Big(|H\rangle + |V\rangle\Big)$. Imaginons que l'on ouvre un trou dans la boîte qui laisse passer les photons aléatoirement un par un. Ce système constitue une source émettant des photons dans un état aléatoire parmi $\left\{|H\rangle, |V\rangle, \frac{1}{\sqrt{2}}\Big(|H\rangle + |V\rangle\Big)\right\}$ avec probabilités respectives $\left\{\frac{3}{10}, \frac{1}{5}, \frac{1}{2}\right\}$. C'est ce qu'on appelle un ensemble statistique et nous verrons que la matrice densité est l'outil approprié pour ce type de situation.

4.1 États mixtes et matrice densité

Soit $\mathcal{H}$ un espace de Hilbert. Les vecteurs $|\psi\rangle \in \mathcal{H}$, $\||\psi\|| = 1$ ou, de façon équivalente, les projecteurs $|\psi\rangle\langle\psi|$ sont appelé des états *purs*. Si A est une observable, c'est-à-dire une matrice hermitienne $A : \mathcal{H} \longrightarrow \mathcal{H}$, l'état pur défini une fonctionnelle linéaire semi-définie positive de l'ensemble des observables $\mathcal{B}(\mathcal{H})$[1] dans $\mathbb{C}$

$$\mathrm{Av}_\psi : \begin{array}{ccc} \mathcal{B}(\mathcal{H}) & \longrightarrow & \mathbb{C} \\ A & \mapsto & \mathrm{Av}_\psi(A) = \langle\psi|A|\psi\rangle \end{array}$$

qui donne la valeur moyenne (average) de A lors de mesures répétées de systèmes dans l'état $|\psi\rangle$. Nous notons que

$$\mathrm{Av}_\psi(A) = \langle\psi|A|\psi\rangle = \mathrm{Tr}\,(A|\psi\rangle\langle\psi|) \tag{4.1}$$

[1] $\mathcal{B}(\mathcal{H})$ est l'ensemble des transformations linéaires $\mathcal{L}(\mathcal{H}) \equiv \mathcal{H} \longrightarrow \mathcal{H}$ qui sont bornées. Dans le cadre de l'information quantique, l'espace de Hilbert est typiquement de dimension finie, les observables sont donc des matrices de taille finie et sont automatiquement bornées. Cependant, en général, les observables peuvent être des matrices de "dimension infinie" ou des opérateurs non-bornés. L'analyse de genre d'observables fait intervenir des outils d'analyse fonctionnelle.

Cette fonctionnelle satisfait à trois propriétés importantes.

1. $\text{Av}_\psi(\lambda A + \mu B) = \lambda \text{Av}_\psi(A) + \mu \text{Av}_\psi(B) \qquad \forall \lambda, \mu \in \mathbb{C},\ \forall A, B \in \mathcal{B}(\mathcal{H})$
2. $\text{Av}_\psi(\mathbb{1}) = 1$
3. Si $A \geq 0$[2], alors $\text{Av}_\psi(A) \in \mathbb{R}_+$

Nous généralisons la notion d'état comme suit:

Définition 1: état quantique

Étant donné un espace de Hilbert $\mathcal{H}$ et une observable $A \in \mathcal{B}(\mathcal{H})$, un "état quantique" est en général défini par une fonctionnelle linéaire positive:

$$\text{Av} : \begin{array}{ccc} \mathcal{B}(\mathcal{H}) & \longrightarrow & \mathbb{C} \\ A & \mapsto & \text{Av}(A) \end{array}$$

c'est-à-dire $\text{Av}(\lambda A + \mu B) = \lambda \text{Av}(A) + \mu \text{Av}(B)$, $\text{Av}(\mathbb{1}) = 1$ et $\text{Av}(A) \in \mathbb{R}_+$ pour $A \geq 0$.

Cette définition semble abstraite, mais un théorème général d'algèbre linéaire affirme qu'une telle fonctionnelle linéaire positive peut toujours être représentée par une matrice ρ telle que

$$\text{Av}(A) = \text{Tr}(A\rho)$$

avec $\rho = \rho^\dagger, \rho \geq 0$ et $\text{Tr}(\rho) = 1$. La matrice ρ s'appelle *la matrice densité*. On peut facilement voir que l'ensemble des matrices densité forme un ensemble convexe.

Exemple 4.1: État pur

Si un système est décrit par un vecteur d'état $|\psi\rangle$, on a $\rho = |\psi\rangle\langle\psi|$. Remarquez que dans ce cas, ρ est une matrice de projection de rang 1 (projecteur sur $|\psi\rangle$). On a bien $\rho = \rho^\dagger$, les valeurs propres sont 1 et 0, donc $\rho \geq 0$ et la condition $\langle\psi|\psi\rangle = \||\psi\|^2 = 1$ est équivalent à $\text{Tr}(\rho) = 1$. Dans ce cas, on dit que ρ est un état pur.

[2] Si A est semi-définie positive. Une matrice A est dite semi-définie positive si elle satisfait $\vec{x}^\dagger A \vec{x} \geq 0 \ \ \forall \vec{x}$. De manière équivalente, A est semi-définie positive si $A^\dagger = A$ et toutes ses valeurs propres $\lambda_i \geq 0$.

Exemple 4.2: Gaz de photons

Dans le cas du gaz de photons décrit dans l'introduction, si on fait une mesure d'observable pour chaque photon qui sort de la boîte (p.ex. on pourrait mesurer la polarisation), on trouve en moyenne

$$\begin{aligned}&\frac{3}{10}\langle H|A|H\rangle + \frac{1}{5}\langle V|A|V\rangle + \frac{1}{2}\left(\frac{\langle H| + \langle V|}{\sqrt{2}}\right)A\left(\frac{|H\rangle + |V\rangle}{\sqrt{2}}\right)\\ &= \mathrm{Tr}\left(A\left\{\frac{3}{10}|H\rangle\langle H| + \frac{1}{5}|V\rangle\langle V| + \frac{1}{2}\left(\frac{|H\rangle + |V\rangle}{\sqrt{2}}\right)\left(\frac{\langle H| + \langle V|}{\sqrt{2}}\right)\right\}\right)\end{aligned}$$

En d'autres termes, cette source de photons est décrite par la matrice densité

$$\rho = \frac{3}{10}|H\rangle\langle H| + \frac{1}{5}|V\rangle\langle V| + \frac{1}{2}\left(\frac{|H\rangle + |V\rangle}{\sqrt{2}}\right)\left(\frac{\langle H| + \langle V|}{\sqrt{2}}\right)$$

Définition 2: état quantique

Un état mixte est une matrice densité de la forme

$$\rho = \sum_{i=1}^{k} p_i|\varphi_i\rangle\langle\varphi_i|$$

où $\left\{|\varphi_1\rangle, \ldots, |\varphi_k\rangle\right\}$ est un ensemble de k états normalisés non nécessairement orthogonaux de l'espace de Hilbert et $\{p_1, \ldots, p_k\}$ un ensemble de probabilités $p_i \geq 0$ et $\sum_{i=1}^{k} p_i = 1$.

L'état est dit pur quand $k = 1$ et les états purs sont les points extrémaux de l'ensemble convexe des matrices densités. Il est dit mixte pour $k \geq 2$ (avec $|\varphi_1\rangle \neq |\varphi_2\rangle$).

On peut montrer que les définitions 1 et 2 d'un état quantique sont équivalentes. Il est très facile de voir sur la deuxième définition que $\rho = \rho^\dagger$ car chacun des k termes est auto-adjoint. De même, $\rho \geq 0$ car chacun des k termes est proportionnel à un projecteur multiplié par $p_i \geq 0$. Enfin, $\mathrm{Tr}(\rho) = \sum_{i=1}^{k} p_i = 1$ car $\mathrm{Tr}(|\varphi_i\rangle\langle\varphi_i|) = \langle\varphi_i|\varphi_i\rangle = 1$. Ainsi, la deuxième définition implique la première. Réciproquement, puisque toute fonctionnelle linéaire positive peut être représentée par une matrice densité hermitienne, on peut utiliser la décomposition spectrale

$$\rho = \sum_{i=1}^{d} \lambda_i|\chi_i\rangle\langle\chi_i|$$

où $\rho|\chi_i\rangle = \lambda_i|\chi_i\rangle$ et $\left\{|\chi_1\rangle, \ldots, |\chi_d\rangle\right\}$ forme une base orthonormée et $\lambda_i \in \mathbb{R}$. Puisque $\rho \geq 0$ et $\mathrm{Tr}(\rho) = 1$, on doit aussi avoir $\lambda_i \geq 0$ et $\sum_{i=1}^{d} \lambda_i = 1$, c'est-à-dire qu'on peut interpréter

$\{\lambda_1, \ldots, \lambda_d\}$ comme un ensemble de probabilités.

Remarque 1. Nous remarquons néanmoins que cette décomposition spectrale est utilisée ici pour démontrer que la définition 1 implique la définition 2 mais ne correspond pas forcément à la préparation physique d'un ensemble statistique. Plus généralement, la donnée d'une matrice densité ρ avec $\rho = \rho^\dagger$, $\rho \geq 0$ et $\mathrm{Tr}(\rho) = 1$ ne contient pas forcément l'information sur la préparation physique sous-jacente du système physique.

Remarque 2. En général, dans la donnée de $\rho = \sum_{i=1}^{k} p_i |\varphi_i\rangle\langle\varphi_i|$, les états $|\varphi_1\rangle, \ldots, |\varphi_k\rangle$ ne forment pas forcément une base orthonormée. De plus, cette formule peut correspondre à la préparation physique d'une source statistique, ou pas. En particulier, cette décomposition n'est pas unique (par exemple, ρ possède aussi une décomposition spectrale).

Remarque 3. Nous verrons plus loin que la matrice densité permet également de décrire des systèmes qui ne sont pas isolés de leur environnement. Cette situation est physiquement très différente d'une source statistique.

Mesure sur une matrice densité.

Il reste à expliquer comment fonctionne le postulat de la mesure avec le formalisme de la matrice densité. Il est évidement nécessaire que les résultats obtenus en effectuant la mesure d'un système décrit par un état pur soient cohérents, indépendamment de la description mathématique utilisée (vecteurs d'états ket's ou matrice densité).

Soit $A \in \mathcal{B}(\mathcal{H})$ une observable dont la décomposition spectrale est $A = \sum_i \alpha_i |\varphi_i\rangle\langle\varphi_i| = \sum_i \alpha_i P_i$, où $\left\{|\varphi_i\rangle\right\}$ est une base orthonormée de $\mathcal{H}$ et soit un état pur représenté par le ket $|\psi\rangle$ ou, de manière équivalente, par la matrice densité $\rho = |\psi\rangle\langle\psi|$. Grâce au Chapitre **3**, nous savons que la valeur moyenne de A lors de mesures répétées sur un état décrit par un ket $|\psi\rangle$ est donnée par

$$\langle\psi|A|\psi\rangle$$

Par ailleurs, la matrice ρ a été définie de manière que

$$\mathrm{Av}(A) = \mathrm{Tr}(A\rho) \tag{4.2}$$

L'équation (4.1) prouve que ces deux quantités sont bien égales.

Si l'on ne s'intéresse qu'à une unique mesure de A sur l'état $|\psi\rangle$, nous savons que le résultat obtenu est l'une des valeurs propres α_i de A. La probabilité d'un tel événement est donnée par

$$\mathbb{P}\left[\alpha_i\right] = |\langle\varphi_i|\psi\rangle|^2$$

Or, $|\langle\varphi_i|\psi\rangle|^2 = \langle\psi|\varphi_i\rangle\langle\varphi_i|\psi\rangle = \langle\psi|P_i|\psi\rangle$ n'est rien d'autre que la valeur moyenne du

projecteur P_i sur l'état $|\psi\rangle$. Il est donc naturel de définir

$$\mathbb{P}\left[\alpha_i\right] = \mathrm{Tr}(P_i\rho) \tag{4.3}$$

dans le formalisme des matrices densités.

Enfin, la théorie quantique affirme que l'état $|\psi\rangle$ est projeté sur l'état propre $|\varphi_i\rangle$ associé à α_i lorsque cette valeur propre est mesurée. Avec la notation des vecteurs d'états, on trouve que l'état pur après la mesure est donné par[3]

$$|\psi'\rangle = \frac{P_i|\psi\rangle}{\sqrt{\mathbb{P}\left[\alpha_i\right]}}$$

La matrice densité associée à ce ket est

$$\rho' = |\psi'\rangle\langle\psi'| = \frac{P_i|\psi\rangle\langle\psi|P_i^\dagger}{\sqrt{\mathbb{P}\left[\alpha_i\right]^2}} = \frac{P_i\rho P_i}{\mathbb{P}\left[\alpha_i\right]} \tag{4.4}$$

Les équations (4.2), (4.3) et (4.4) permettent de faire les calculs en utilisant le formalisme de la matrice densité. Elles sont, par ailleurs, également valide pour les calculs sur des états mixtes.

4.2 Matrice densité pour un qubit

L'ensemble des états d'un seul bit quantique peut être décrit par des matrices densités 2×2 que nous caractériserons entièrement dans ce paragraphe. Les matrices de Pauli

$$\mathbb{1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \qquad X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad Y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \qquad Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

forment une base pour l'espace vectoriel des matrices 2×2 hermitiennes complexes. Ainsi,

$$\rho = \frac{a_0}{2}\mathbb{1} + \frac{a_1}{2}X + \frac{a_2}{2}Y + \frac{a_3}{2}Z$$

De plus, étant donné que ρ est une matrice hermitienne, les coefficients a_0, a_1, a_2 et $a_3 \in \mathbb{R}$. Comme on doit avoir $\mathrm{Tr}(\rho) = 1$, on déduit que $a_0 = 1$. On peut donc écrire

$$\rho = \frac{1}{2}\left(\mathbb{1} + \vec{a}\cdot\vec{\Sigma}\right) = \frac{1}{2}\begin{pmatrix} 1 + a_3 & a_1 - ia_2 \\ a_1 + ia_2 & 1 - a_3 \end{pmatrix}$$

avec $\vec{a} = (a_1, a_2, a_3)$ et $\vec{\Sigma} = (X, Y, Z)$ le vecteur des 3 matrices de Pauli. Pour avoir $\rho \geq 0$, il faut nécessairement que $\det(\rho) \geq 0$. Puisque $\mathrm{Tr}(\rho) = 1 \geq 0$, la condition sur le déterminant est même suffisante. Ainsi, il faut que

$$\det(\rho) = 1 - \|\vec{a}\|^2 \geq 0$$

[3] Une manière d'interpréter cette équation est la suivante; $P_i|\psi\rangle$ projette $|\psi\rangle$ sur le sous-espace engendré par $|\varphi_i\rangle$ et le terme $\sqrt{\mathbb{P}\left[\alpha_i\right]}$ sert à normaliser $|\psi'\rangle$.

c.-à-d. $\|\vec{a}\|^2 \leq 1$. En conclusion, l'ensemble des matrices densité pour un qubit est $\left\{\rho = \frac{1}{2}\left(\mathbb{1} + \vec{a}\cdot\vec{\Sigma}\right) \mid \|\vec{a}\|^2 \leq 1\right\}$.

La boule $\|\vec{a}\|^2 \leq 1$ est un ensemble convexe dont les points extrémaux sont sur la sphère $\|\vec{a}\|^2 = 1$. Ceux-ci correspondent à des états purs. En effet,

$$\begin{aligned}
\rho^2 &= \frac{1}{4}\left(\mathbb{1} + \vec{a}\cdot\vec{\Sigma}\right)^2 \\
&= \frac{1}{4}\left(\mathbb{1} + a_1^2X^2 + a_2^2Y^2 + a_3^2Z^2\right) + \frac{1}{4}2\vec{a}\cdot\vec{\Sigma} \\
&+ \frac{1}{4}a_1a_2\underbrace{(XY+YX)}_{0} + \frac{1}{4}a_1a_3\underbrace{(XZ+ZX)}_{0} + \frac{1}{4}a_2a_3\underbrace{(YZ+ZY)}_{0} \\
&= \frac{\mathbb{1}}{4}\underbrace{(1+\|\vec{a}\|^2)}_{2} + \frac{1}{2}\vec{a}\cdot\vec{\Sigma} \\
&= \frac{1}{2}\left(\mathbb{1} + \vec{a}\cdot\vec{\Sigma}\right) \\
&= \rho
\end{aligned}$$

donc ρ est un projecteur si et seulement si $\|\vec{a}\|^2 = 1$.

La boule $\|\vec{a}\|^2 \leq 1$ s'appelle la boule de Bloch et est une représentation géométrique de l'ensemble des matrices densité qui généralise la sphère de Bloch, elle-même n'étant rien d'autre que la surface de la boule. Les états purs sont paramétrés par deux angles uniquement, vu qu'ils se trouvent à la surface de la boule, tandis que les états mixtes requièrent en plus la longueur du vecteur dans la boule. Notons que le vecteur $\vec{a} = (0,0,0)$ correspond à l'état mixte trivial $\begin{pmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{pmatrix}$ correspondant à un ensemble statistique totalement aléatoire. Les vecteurs $\vec{a} = (0,0,1)$ et $\vec{a} = (0,0,-1)$ correspondent aux états purs orthogonaux $\{|\uparrow\rangle, |\downarrow\rangle\}$ et les vecteurs $\vec{a} = (1,0,0)$ et $\vec{a} = (-1,0,0)$ aux états purs orthogonaux $\{|+\rangle, |-\rangle\}$.

4.3 Matrice densité réduite

Une situation physiquement commune et très importante est celle d'un système S en interaction avec son environnement $\mathcal{E}$. Lorsque ρ n'est pas suffisamment bien isolé de $\mathcal{E}$, les interactions ont tendance à intriquer l'état (pur) de $S \cup \mathcal{E}$. On ne peut alors plus décrire le système S lui-même par un simple état pur et il faut faire appel à sa matrice densité, parfois appelée *matrice densité réduite*. Pour introduire cette notion, nous devons d'abord expliquer ce qu'est une *trace partielle*.

Notion de trace partielle

Soit l'espace de Hilbert $\mathcal{H}_S \otimes \mathcal{H}_\mathcal{E}$. Cet espace possède une base formée par le produit tensoriel des deux bases orthonormées de $\mathcal{H}_S$ et $\mathcal{H}_\mathcal{E}$, notées $\{|a_1\rangle, \ldots, |a_N\rangle\}$ et $\{|b_1\rangle, \ldots, |b_M\rangle\}$. Tout état $|\psi\rangle \in \mathcal{H}_S \otimes \mathcal{H}_\mathcal{E}$ peut donc être développé comme

$$|\psi\rangle = \sum_{i=1}^{N} \sum_{j=1}^{M} c_{ij} |a_i\rangle \otimes |b_j\rangle$$

et pour tout observable $A \in \mathcal{B}(\mathcal{H}_S \otimes \mathcal{H}_\mathcal{E})$, on a

$$A = \sum_{i,j,k,l} A_{ij;kl} \left(|a_i\rangle \otimes |b_j\rangle\right)\left(\langle a_k| \otimes \langle b_l|\right) = \sum_{i,j,k,l} A_{ij;kl} \left(|a_i\rangle\langle a_k|\right) \otimes \left(|b_j\rangle\langle b_l|\right)$$

L'état $|\psi\rangle$ possède NM composantes (un vecteur colonne c_{ij}) et la matrice A est un tableau de taille $NM \times NM$ avec les éléments $A_{ij;kl}$.

L'opération de *trace partielle* sur "l'environnement" $\mathcal{E}$ est définie par

$$\begin{aligned} \mathrm{Tr}_\mathcal{E}(A) &\equiv \sum_{i,j,k,l} A_{ij;kl} \, |a_i\rangle\langle a_k| \, \mathrm{Tr}_\mathcal{E}\left(|b_j\rangle\langle b_l|\right) \\ &= \sum_{i,j,k,l} A_{ij;kl} \, |a_i\rangle\langle a_k| \underbrace{\langle b_l|b_j\rangle}_{\delta_{jl}} \\ &= \sum_{i,k} \Big(\sum_{j} A_{ij;kj} \Big) |a_i\rangle\langle a_k| \end{aligned} \tag{4.5}$$

On voit que $\mathrm{Tr}_\mathcal{E}(A)$ est une matrice $N \times N$ d'éléments $\sum_j A_{ij;kj}$. Cette matrice agit sur $\mathcal{H}_S$ (i.e. $A \in \mathcal{B}(\mathcal{H}_S)$).

L'opération de trace partielle sur "le système" S est définie par

$$\begin{aligned} \mathrm{Tr}_S(A) &\equiv \sum_{i,j,k,l} A_{ij;kl} \; \mathrm{Tr}_S\left(|a_i\rangle\langle a_k|\right) |b_j\rangle\langle b_l| \\ &= \sum_{i,j,k,l} A_{ij;kl} \; \underbrace{\langle a_i|a_k\rangle}_{\delta_{ik}} |b_j\rangle\langle b_l| \\ &= \sum_{j,l} \Big(\sum_{i} A_{ij;il} \Big) |b_j\rangle\langle b_l| \end{aligned} \tag{4.6}$$

Ainsi, $\mathrm{Tr}_S(A)$ est une matrice $M \times M$ d'éléments $\sum_i A_{ij;il}$ qui agit sur $\mathcal{H}_\mathcal{E}$.

Exemple 4.3: Trace partielle

Soient $\mathcal{H}_S = \mathbb{C}^2$ et $\mathcal{H}_\mathcal{E} = \mathbb{C}^2$ deux espaces de Hilbert d'un qubit. Le système total est décrit par $\mathcal{H}_S \otimes \mathcal{H}_\mathcal{E} = (\mathbb{C}^2)^{\otimes 2}$ qui possède la base computationnelle

$$\Big\{|0\rangle_S \otimes |0\rangle_\mathcal{E}, \quad |0\rangle_S \otimes |1\rangle_\mathcal{E}, \quad |1\rangle_S \otimes |0\rangle_\mathcal{E}, \quad |1\rangle_S \otimes |1\rangle_\mathcal{E}\Big\}$$

Considérons l'état pur intriqué

$$|\psi\rangle = \frac{1}{\sqrt{2}}\Big(|0\rangle_S \otimes |0\rangle_\mathcal{E} + |1\rangle_S \otimes |1\rangle_\mathcal{E}\Big)$$

La matrice densité associée est

$$\rho = |\psi\rangle\langle\psi| = \frac{1}{\sqrt{2}}\Big\{|0\rangle\langle 0|_S \otimes |0\rangle\langle 0|_\mathcal{E} + |0\rangle\langle 1|_S \otimes |0\rangle\langle 1|_\mathcal{E} \\ +|1\rangle\langle 0|_S \otimes |1\rangle\langle 0|_\mathcal{E} + |1\rangle\langle 1|_S \otimes |1\rangle\langle 1|_\mathcal{E}\Big\}$$

En appliquant les formules pour trouver les deux traces partielles à ρ, on obtient

$$\mathrm{Tr}_S(\rho) = \frac{1}{2}|0\rangle\langle 0|_S + \frac{1}{2}|1\rangle\langle 1|_S$$
$$\mathrm{Tr}_\mathcal{E}(\rho) = \frac{1}{2}|0\rangle\langle 0|_\mathcal{E} + \frac{1}{2}|1\rangle\langle 1|_\mathcal{E}$$

Sous forme de matrices, si la base est

$$|0\rangle \otimes |0\rangle \equiv \begin{pmatrix}1\\0\\0\\0\end{pmatrix}, \quad |0\rangle \otimes |1\rangle \equiv \begin{pmatrix}0\\1\\0\\0\end{pmatrix}, \quad |1\rangle \otimes |0\rangle \equiv \begin{pmatrix}0\\0\\1\\0\end{pmatrix}, \quad |1\rangle \otimes |1\rangle \equiv \begin{pmatrix}0\\0\\0\\1\end{pmatrix}$$

on a

$$\rho = \frac{1}{2}\begin{pmatrix}1&0&0&1\\0&0&0&0\\0&0&0&0\\1&0&0&1\end{pmatrix} = \begin{pmatrix}A_{0000}&A_{0001}&A_{0010}&A_{0011}\\A_{0100}&A_{0101}&A_{0110}&A_{0111}\\A_{1000}&A_{1001}&A_{1010}&A_{1011}\\A_{1100}&A_{1101}&A_{1110}&A_{1111}\end{pmatrix}$$

et les traces partielles sont obtenues en faisant les opérations

$$\mathrm{Tr}_S(\rho) = \begin{pmatrix}A_{0000}+A_{1010}&A_{0001}+A_{1011}\\A_{0100}+A_{1110}&A_{0101}+A_{1111}\end{pmatrix} = \frac{1}{2}\begin{pmatrix}1&0\\0&1\end{pmatrix}$$
$$\mathrm{Tr}_\mathcal{E}(\rho) = \begin{pmatrix}A_{0000}+A_{0101}&A_{0010}+A_{0111}\\A_{1000}+A_{1101}&A_{1010}+A_{1111}\end{pmatrix} = \frac{1}{2}\begin{pmatrix}1&0\\0&1\end{pmatrix}$$

On voit bien avec cet exemple que la notation de Dirac est plus pratique !

Matrice densité réduite et son interprétation physique

L'exemple précédent capture la notion de matrice densité réduite. Si $\mathcal{H}_{\text{tot}} = \mathcal{H}_S \otimes \mathcal{H}_\mathcal{E}$ et si l'état de $\mathcal{H}_{\text{tot}}$ est défini par un état pur $|\psi\rangle$, la matrice densité réduite de S est donnée par la trace partielle sur l'environnement:

$$\rho_S = \text{Tr}_\mathcal{E}(|\psi\rangle\langle\psi|)$$

De façon similaire, la matrice densité réduite de l'environnement est donnée par la trace partielle sur le système:

$$\rho_\mathcal{E} = \text{Tr}_S(|\psi\rangle\langle\psi|)$$

Quelle est l'interprétation physique de ρ_S ou $(\rho_\mathcal{E})$? Supposons que l'on fait une expérience pour mesurer une observable de S, c'est-à-dire $A = A_S \otimes \mathbb{1}_\mathcal{E}$. D'après les principes de la mécanique quantique, la valeur moyenne de l'observable lors de la mesure répétée de $|\psi\rangle$ est

$$\begin{aligned} \text{Av}_\psi(A) &= \langle\psi|A_S \otimes \mathbb{1}_\mathcal{E}|\psi\rangle = \text{Tr}_{\mathcal{H}_S\otimes\mathcal{H}_\mathcal{E}}\Big(A_S \otimes \mathbb{1}_\mathcal{E}|\psi\rangle\langle\psi|\Big) \\ &= \text{Tr}_{\mathcal{H}_S}\left(A_S\left\{\text{Tr}_{\mathcal{H}_\mathcal{E}}\Big(|\psi\rangle\langle\psi|\Big)\right\}\right) = \text{Tr}_{\mathcal{H}_S}\Big(A_S\rho_S\Big) \end{aligned} \tag{4.7}$$

Ainsi, $\text{Av}_\psi(A) = \text{Tr}_{\mathcal{H}_S}\Big(A_S\rho_S\Big)$, ce qui signifie que le résultat expérimental et entièrement défini par la matrice densité ρ_S.

Allons un peu plus loin. Supposons que $A_S = \sum_n \alpha_n P_n$ où $P_n = |n\rangle\langle n|$ sont les projecteurs propres de A_S et α_n les valeurs propres. La décomposition spectrale de $A = A_S \otimes \mathbb{1}_\mathcal{E}$ est

$$A = A_S \otimes \mathbb{1}_\mathcal{E} = \sum_n \alpha_n\Big(P_n \otimes \mathbb{1}_\mathcal{E}\Big)$$

où $P_n \otimes \mathbb{1}_\mathcal{E}$ sont des projecteurs propres de A. D'après le postulat de la mesure, si $|\psi\rangle$ est l'état avant la mesure, alors l'état après la mesure est $\propto P_n \otimes \mathbb{1}_\mathcal{E}|\psi\rangle$ avec probabilité $\mathbb{P}[n] = \langle\psi|P_n \otimes \mathbb{1}_\mathcal{E}|n\rangle$. Cette probabilité peut s'écrire

$$\begin{aligned} \mathbb{P}[n] = \langle\psi|P_n \otimes \mathbb{1}_\mathcal{E}|n\rangle &= \text{Tr}_{\mathcal{H}_S\otimes\mathcal{H}_\mathcal{E}}\Big(P_n \otimes \mathbb{1}_\mathcal{E}|\psi\rangle\langle\psi|\Big) \\ &= \text{Tr}_{\mathcal{H}_S}\Big(P_n\Big\{\text{Tr}_{\mathcal{H}_\mathcal{E}}(|\psi\rangle\langle\psi|\Big\}\Big) \\ &= \text{Tr}_{\mathcal{H}_S}\Big(P_n\,\rho_S\Big) \end{aligned}$$

La matrice densité totale dans $\mathcal{H}_S \otimes \mathcal{H}_\mathcal{E}$ après la mesure est proportionnelle à

$$\rho_{\text{après}} \propto P_n \otimes \mathbb{1}_\mathcal{E}(|\psi\rangle\langle\psi|)P_n \otimes \mathbb{1}_\mathcal{E}$$

et la matrice densité réduite est proportionnelle à

$$\begin{aligned} \rho_{S,\text{après}} \propto \text{Tr}_\mathcal{E}\Big(\rho_{\text{après}}\Big) &= \text{Tr}_\mathcal{E}\Big(P_n \otimes \mathbb{1}_\mathcal{E}(|\psi\rangle\langle\psi|)P_n \otimes \mathbb{1}_\mathcal{E}\Big) \\ &= P_n\Big(\text{Tr}_\mathcal{E}(|\psi\rangle\langle\psi|)\Big)P_n \\ &= P_n\,\rho_S\,P_n \end{aligned}$$

Il faut encore normaliser cette expression, si bien que

$$\rho_{S,\text{après}} = \frac{P_n \, \rho_S \, P_n}{\text{Tr}\Big(P_n \, \rho_S \, P_n\Big)}$$

Exemple 4.4: Matrice densité réduite d'un état pur

Considérons un état pur. Si c'est un état produit (séparable), alors la trace partielle résulte en un état qui est encore pur. Si au contraire l'état est intriqué, alors la trace partielle résulte en un état mixte. Ici, nous ne prouvons pas cette affirmation formellement, mais nous l'illustrons avec un exemple explicite (un état pur maximalement intriqué).

Reprenons l'espace de Hilbert de deux qubits $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2 = \mathbb{C}^2 \otimes \mathbb{C}^2$. Pour un état produit $|\varphi_1\rangle \otimes |\varphi_2\rangle$, on a

$$\rho = \Big(|\varphi_1\rangle \otimes |\varphi_2\rangle\Big)\Big(\langle\varphi_1| \otimes \langle\varphi_2|\Big) = |\varphi_1\rangle\langle\varphi_1| \otimes |\varphi_2\rangle\langle\varphi_2|$$

et les matrices densités réduites obtenues par trace partielle sont

$$\rho_1 = |\varphi_1\rangle\langle\varphi_1| \quad \text{et} \quad \rho_2 = |\varphi_2\rangle\langle\varphi_2|$$

On voit que chaque qubit est bien décrit par un état pur (ce qui était déjà bien clair avec l'état produit).

Prenons maintenant un état intriqué, par exemple $|\psi\rangle = \frac{1}{\sqrt{2}}\Big(|0\rangle \otimes |0\rangle + |1\rangle \otimes |1\rangle\Big)$. Les matrices densités réduites obtenues par trace partielle sont (comme discuté auparavant)

$$\rho_1 = \frac{1}{2}\Big(|0\rangle\langle 0| + |1\rangle\langle 1|\Big) = \frac{1}{2}\mathbb{1} \quad \text{et} \quad \rho_2 = \frac{1}{2}\Big(|0\rangle\langle 0| + |1\rangle\langle 1|\Big) = \frac{1}{2}\mathbb{1}$$

qui, les deux, sont des états mixtes correspondant au vecteur $\vec{a} = (0, 0, 0)$ de la boule de Bloch. Nous notons que cet état peut correspondre à n'importe quel état mixte fabriqué avec deux vecteurs quelconques orthonormés $\{|e_1\rangle, |e_2\rangle\}$ pris avec probabilités $\{\frac{1}{2}, \frac{1}{2}\}$. En effet, la matrice $\frac{1}{2}\mathbb{1}$ est diagonale dans n'importe quelle base orthonormée. En d'autres termes, les états ρ_1 et ρ_2 sont totalement aléatoires et individuellement, ils ne contiennent aucune information !

4.4 Décomposition de Schmidt et Purification

Nous présentons dans ce paragraphe deux outils mathématiques qui s'avèrent très utiles en information quantique.

Décomposition de Schmidt

Si un système est "bipartite", c'est-à-dire avec un espace de Hilbert de la forme $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$ et que l'état du système est pur $|\psi\rangle$, alors les matrices densités réduites $\rho_A = \mathrm{Tr}_B(|\psi\rangle\langle\psi|)$ et $\rho_B = \mathrm{Tr}_A(|\psi\rangle\langle\psi|)$ satisfont à des propriétés remarquables.

Théorème

Soit $|\psi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$ un vecteur d'état pur. Alors, ρ_A et ρ_B possèdent les mêmes valeurs propres non-nulles avec les mêmes multiplicités. En conséquence, les matrices densités réduites ont une décomposition spectrale

$$\rho_A = \sum_i \lambda_i |\varphi_i\rangle_A\langle\varphi_i|_A \quad \text{et} \quad \rho_B = \sum_i \lambda_i |\chi_i\rangle_B\langle\chi_i|_B$$

avec le même nombre de termes $\lambda_i > 0$ et $\sum_i \lambda_i = 1$. Des valeurs propres nulles peuvent exister avec des multiplicités différentes pour ρ_A et ρ_B. Ici, $|\varphi_i\rangle_A$ et $|\chi_i\rangle_B$ sont des vecteurs orthonormés de $\mathcal{H}_A$ et $\mathcal{H}_B$ respectivement avec $\rho_A|\varphi_i\rangle_A = \lambda_i|\varphi_i\rangle_A$ et $\rho_B|\chi_i\rangle_A = \lambda_i|\chi_i\rangle_B$.

De plus, on a

$$|\psi\rangle = \sum_i \sqrt{\lambda_i}\, |\varphi_i\rangle_A \otimes |\chi_i\rangle_B$$

Cette dernière représentation s'appelle la décomposition de Schmidt.

Remarque. Si les valeurs propres non-nulles $\lambda_i > 0$ sont non-dégénérées, alors les vecteurs propres associés $|\varphi_i\rangle_A$ et $|\chi_i\rangle_B$ sont uniques à une phase près. Sinon, ils ne sont pas uniques car on peut effectuer des rotations dans le sous-espace propre associé à λ_i. De plus, ces vecteurs propres ne forment pas une base complète de $\mathcal{H}_A$ et $\mathcal{H}_B$ à moins de leur adjoindre les vecteurs propres associés aux valeurs propres nulles de ρ_A et ρ_B.

Preuve du théorème

Soient $\{|n\rangle_A\}$ et $\{|n'\rangle_B\}$ deux bases orthonormales de $\mathcal{H}_A$ et $\mathcal{H}_B$. On peut développer l'état pur comme

$$|\psi\rangle = \sum_{n,\, n'} a_{nn'} |n\rangle_A \otimes |n'\rangle_B$$

Pour chaque n, on pose $|\tilde{n}\rangle_B = \sum_{n'} a_{nn'} |n'\rangle_B$ si bien que

$$|\psi\rangle = \sum_n |n\rangle_A \otimes |\tilde{n}\rangle_B$$

Notons que $\left\{|\tilde{n}\rangle_B\right\}$ ne forme pas nécessairement une base orthonormale. Pour la matrice densité réduite de A, on trouve

$$\rho_A = \mathrm{Tr}_B(|\psi\rangle\langle\psi|) = \sum_{n_1,\, n_2} \langle\tilde{n}_2|\tilde{n}_1\rangle_B \ |n_1\rangle_A\langle n_2|_A$$

En particulier, pour la base $\left\{|n\rangle_A\right\}$, on peut choisir la base $\left\{|\varphi_i\rangle_A\right\}$ des vecteurs propres de ρ_A tels que $\rho_A|\varphi_i\rangle_A = \lambda_i|\varphi_i\rangle_A$, si bien que

$$\rho_A = \sum_{i_1,i_2} \langle\tilde{\varphi}_{i_2}|\tilde{\varphi}_{i_1}\rangle_B \, |\varphi_{i_1}\rangle_A\langle\varphi_{i_2}|_A = \sum_{i_1} \lambda_{i_1} |\varphi_{i_1}\rangle_A\langle\varphi_{i_2}|_A$$

où la deuxième égalité n'est autre que la décomposition spectrale. Donc, pour toute valeur propre $\lambda_{i_1} \neq 0$, on doit avoir

$$\langle\tilde{\varphi}_{i_2}|\tilde{\varphi}_{i_1}\rangle_B = \lambda_{i_1}\delta_{i_1 i_2}$$

et donc les états $|\tilde{\varphi}_i\rangle_B$ sont orthogonaux. On peut les normaliser en définissant $|\chi_i\rangle_B = \lambda_i^{-\frac{1}{2}}|\tilde{\varphi}_i\rangle_B$. En conséquence

$$|\psi\rangle = \sum_i |\varphi_i\rangle_A \otimes |\tilde{\varphi}_i\rangle_B = \sum_i \lambda_i^{-\frac{1}{2}} |\varphi_i\rangle_A \otimes |\chi_i\rangle_B$$

qui est la décomposition de Schmidt. En prenant les traces partielles, on obtient aussi

$$\rho_A = \mathrm{Tr}_B(|\psi\rangle\langle\psi|) = \sum_i \lambda_i |\varphi_i\rangle_A\langle\varphi_i|_A$$

$$\rho_B = \mathrm{Tr}_A(|\psi\rangle\langle\psi|) = \sum_i \lambda_i |\chi_i\rangle_B\langle\chi_i|_B$$

Définition. On appelle *nombre de Schmidt* de l'état pur $|\psi\rangle$ le nombre de coefficients non-nuls λ_i dans la décomposition de Schmidt.

Pour un état produit $|\psi\rangle = |\varphi_1\rangle \otimes |\varphi_2\rangle$, le nombre de Schmidt est égal à 1. Si le nombre de Schmidt est supérieur ou égal à 2, l'état est intriqué. Ce nombre est notre première mesure quantitative de la quantité d'intrication dans un système bipartite.

On note que le nombre de Schmidt est un invariant sous les transformations unitaires locales de la forme $U = U_A \otimes U_B$. En effet

$$U|\psi\rangle = \sum_i \sqrt{\lambda_i} U_A|\varphi_i\rangle_A \otimes U_B|\chi_i\rangle_B$$

et $\left\{U_A|\varphi_i\rangle_A\right\}$, $\left\{U_B|\varphi_i\rangle_B\right\}$ sont des systèmes orthonormés. Ainsi, le nombre de Schmidt peut changer seulement si les parties A et B interagissent (auquel cas, $U \neq U_A \otimes U_B$).

Purification d'une matrice densité

Nous avons vu que pour un système total $S \cup \mathcal{E}$, si $|\psi\rangle \in \mathcal{H}_S \otimes \mathcal{H}_\mathcal{E}$, la description réduite du système S est donnée par une matrice densité

$$\rho_S = \mathrm{Tr}_{H_\mathcal{E}}(|\psi\rangle\langle\psi|)$$

Réciproquement, étant donné un système S décrit par une matrice densité ρ_S (ce système pourrait être une source de photons dans des états aléatoires comme décrit au début du chapitre), il est possible de construire un système plus grand $\mathcal{H}_S \otimes \mathcal{H}_\mathcal{E}$ et un état pur $|\psi\rangle$ tel que $\rho_S = \mathrm{Tr}_{H_\mathcal{E}}(|\psi\rangle\langle\psi|$. Cette construction n'est pas unique et ne correspond pas forcément à la préparation physique de S.

Donnons une construction explicite (qui s'avère être une technique mathématique utile) en utilisant la décomposition de Schmidt. Tout d'abord, étant donné ρ_S, on considère sa décomposition spectrale

$$\rho_S = \sum_i \lambda_i |\varphi_i\rangle_S \langle\varphi_i|_S$$

et on prend pour l'environnement une copie de $\mathcal{H}_S$ si bien que $\mathcal{H}_{\mathrm{tot}} = \mathcal{H}_S \otimes \mathcal{H}_\mathcal{E}$ avec $\mathcal{H}_\mathcal{E}$ isomorphe à $\mathcal{H}_S$. Chaque vecteur $|\varphi_i\rangle_S$ possède sa copie $|\varphi_i\rangle_\mathcal{E}$ et on pose

$$|\psi\rangle = \sum_i \sqrt{\lambda_i}|\varphi_i\rangle_S \otimes |\varphi_i\rangle_\mathcal{E}$$

4.5 Matrice densité pour deux qubits

Nous avons vu comment exprimer la matrice densité pour un qubit à la Section **4.2**. Une représentation similaire peut être obtenue pour deux qubits. En effet, puisque $\left\{\mathbb{1}, X, Y, Z\right\}$ forme une base de l'espace vectoriel des matrices hermitiennes 2×2, le produit tensoriel de deux copies de cette base forme une base pour l'espace vectoriel réel des matrices hermitiennes 4×4. En d'autres termes, toute matrice densité pour deux qubits peut s'exprimer comme

$$\rho = \frac{1}{4}\left(\mathbb{1} \otimes \mathbb{1} + \vec{a} \cdot \vec{\Sigma} \otimes \mathbb{1} + \mathbb{1} \otimes \vec{b} \cdot \vec{\Sigma} + \sum_{n,m=1}^{3} t_{nm}\, \sigma_n \otimes \sigma_m\right) \tag{4.8}$$

avec $\vec{\Sigma} = (\sigma_1, \sigma_2, \sigma_3) = (X, Y, Z)$ les matrices de Pauli, $\vec{a}, \vec{b} \in \mathbb{R}^3$ et (t_{nm}) une matrice 3×3 réelle.

Les conditions $\rho = \rho^\dagger$ et $\mathrm{Tr}\,\rho = 1$ sont bien satisfaites, car les matrices de Pauli sont hermitiennes et de trace nulle. Ainsi, ρ est paramétrisée par 15 paramètres réels. Il faut encore adjoindre la condition $\rho \geq 0$ (la matrice densité doit être semi-définie positive), qui est moins aisée à exprimer analytiquement. Toujours est-il que l'ensemble des matrices densité à 2 qubits forme un sous-ensemble borné et convexe de $\mathbb{R}^{15}$. Cet ensemble est bien plus riche que la simple boule de Bloch dans $\mathbb{R}^3$ pour le cas à 1 qubit.

On note par ailleurs que

$$\mathrm{Tr}_B\,\rho = \frac{1}{2}\left(\mathbb{1} + \vec{a}\cdot\vec{\Sigma}\right) \quad \text{et} \quad \mathrm{Tr}_A\,\rho = \frac{1}{2}\left(\mathbb{1} + \vec{b}\cdot\vec{\Sigma}\right)$$

Le cas particulier $\vec{a} = \vec{b} = \vec{0}$ correspond à des états réduits complétement aléatoires. Néanmoins, nous verrons que ce sous-ensemble d'états (à deux qubits) possède une représentation géométrique riche est intéressante.

Remarque. Les éléments de ρ sont complexes car Y est imaginaire. De plus, les matrices de Pauli sont hermitiennes et il suit de la contrainte $\rho = \rho^\dagger$ que $\vec{a}, \vec{b}, t_{nm}$ sont réels. Cette représentation se généralise aisément pour n qubits:

$$\rho = \frac{1}{2^n} \sum_{\substack{(i_1,i_2,\ldots,i_n)\\ \in(0,1,2,3)^n}} t_{i_1 i_2 \ldots i_n}\, \sigma_{i_1} \otimes \sigma_{i_2} \otimes \ldots \otimes \sigma_{i_n} \tag{4.9}$$

où $(\sigma_0, \sigma_1, \sigma_2, \sigma_3) = (\mathbb{1}, X, Y, Z)$ et $t_{i_1 i_2 \ldots i_n} \in \mathbb{R}$. La contrainte $\mathrm{Tr}\,\rho = 1$ impose $t_{00\ldots0} = 1$ et le nombre total de paramètres est $4^n - 1$. La condition $\rho \geq 0$ restreint l'espace des matrices densité à un sous-ensemble convexe borné de $\mathbb{R}^{4^n-1}$.

États séparables et intriqués

La distinction entre états produits et intriqués au niveau des matrices densité est plus subtile que pour les états purs. À première vue, on pourrait définir un état produit comme une matrice densité $\rho = \rho_A \otimes \rho_B$ mais cela n'est pas naturel, si on pense à un état de mélange, qui représente une "source". En effet, une source peut émettre des états "produits" $\rho_A^i \otimes \rho_B^i$ avec probabilité p_i. Cette source est donc décrite par l'état donné par la combinaison convexe

$$\rho = \sum_i p_i\, \rho_A^i \otimes \rho_B^i \tag{4.10}$$

avec $0 \leq p_i \leq 1$ et $\sum_i p_i = 1$. Or, les états de la forme (4.10) ne peuvent pas être mis sous la forme $\rho_A \otimes \rho_B$, bien que chacun des termes i soit un état produit. Cette grosse différence par rapport à des états purs fait que la définition "état produit" utilisée pour ceux-ci n'est pas suffisante. Les états de la forme (4.10) sont donc appelés *séparables* (on fait ici une distinction avec le vocabulaire utilisé pour les états purs).

Définition: États séparables et intriqués

Un état est dit séparable s'il est de la forme (4.10). Si un état ne peut pas se mettre sous la forme (4.10), il est dit intriqué.

Cette définition s'applique à tout système bipartite et en particulier aux matrices densités de deux qubits, auquel cas ρ_A^i et ρ_B^i sont de la forme $\rho_A^i = \frac{1}{2}\left(\mathbb{1} + \vec{a}^i \cdot \vec{\Sigma}\right)$ et $\rho_B^i = \frac{1}{2}\left(\mathbb{1} + \vec{b}^i \cdot \vec{\Sigma}\right)$. Dans ce cas, en comparant (4.8) et (4.10), on trouve $\vec{a} = \sum_i p_i \vec{a}^i$, $\vec{b} = \sum_i p_i \vec{b}^i$ et $t_{nm} = \sum_i p_i a_n^i b_m^i$. On note qu'il est possible d'avoir $\vec{a} = \vec{b} = \vec{0}$ et $t_{nm} \neq 0$.

En général, étant donné une matrice ρ, il n'est pas évident de décider s'il s'agit d'un état séparable ou intriqué. Dans le cas particulier de deux qubits $\mathcal{H}_A \otimes \mathcal{H}_B = \mathbb{C}^2 \otimes \mathbb{C}^2$, un critère dû à Peres donne une condition nécessaire et suffisante aisément vérifiable. Cette condition s'appelle *Positive Partial Transpose criterion (PPT)*. Soit ρ^{T_B} la matrice obtenue par transposition partielle du système B. Si l'état est séparable, on a

$$\rho^{T_B} = \sum_i p_i \, \rho_A^i \otimes {\rho_B^i}^T$$

où ${\rho_B^i}^T$ est la matrice ρ_B^i transposée. Puisque $\rho_B^i \geq 0$, on a aussi ${\rho_B^i}^T \geq 0$. Ainsi, si ρ est séparable, alors ρ^{T_B} est aussi une matrice densité et ses valeurs propres doivent être positives ou nulles. En d'autres termes, si ρ^{T_B} possède une ou plusieurs valeurs propres négatives, alors ρ ne peut pas être séparable et est forcément intriquée.

Critère de Peres-Horodecki

Soit ρ une matrice densité d'un système bipartite général $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$. Si la matrice transposée partielle ρ^{T_B} possède des valeurs propres négatives, alors ρ est intriqué. En particulier, ρ ne peut pas se mettre sous la forme (4.10).

Ce critère est toujours nécessaire mais il n'est pas suffisant en général. Néanmoins, remarquablement, on sait qu'il est bien suffisant pour $\mathcal{H}_A \otimes \mathcal{H}_B = \mathbb{C}^2 \otimes \mathbb{C}^2$ et $\mathbb{C}^2 \otimes \mathbb{C}^3$. En particulier, pour l'ensemble des matrices densité à deux qubits, le critère de Peres-Horodecki est nécessaire et suffisant.

États Bell-diagonaux et leur géométrie

Le sous-ensemble des matrices densité (4.8) avec $\vec{a} = \vec{b} = \vec{0}$ possède une caractérisation géométrique aisée. Notons que cet ensemble correspond à des matrices densité réduites à 1 qubit complètement mixtes et contient donc certainement les états de Bell.

Soit ρ une matrice densité appartenant à cet ensemble. On peut toujours construire

une transformation unitaire $U_A \otimes U_B$ telle que $(U_A \otimes U_B)\rho(U_A^\dagger \otimes U_B^\dagger)$ soit de la forme[4]

$$\tilde{\rho} = \frac{1}{4}\left(\mathbb{1} \otimes \mathbb{1} + \sum_{i=1}^{3} t_i\, \sigma_i \otimes \sigma_i\right)$$

avec $(t_1, t_2, t_3) \in \mathbb{R}^3$. Ces matrices densité forment donc un ensemble convexe de $\mathbb{R}^3$. Pour le déterminer, il faut examiner la condition $\rho \geq 0$.

De plus, on vérifie explicitement que les vecteurs propres et valeurs propres sont donnés par les états de Bell: $\tilde{\rho}|B_{ij}\rangle = p_{ij}|B_{ij}\rangle$:

$$\begin{aligned}
|B_{00}\rangle &= \tfrac{|00\rangle+|11\rangle}{\sqrt{2}} \quad \text{avec} \quad p_{00} = \tfrac{1+t_1-t_2+t_3}{4} \\
|B_{01}\rangle &= \tfrac{|01\rangle+|10\rangle}{\sqrt{2}} \quad \text{avec} \quad p_{01} = \tfrac{1+t_1+t_2-t_3}{4} \\
|B_{10}\rangle &= \tfrac{|00\rangle-|11\rangle}{\sqrt{2}} \quad \text{avec} \quad p_{10} = \tfrac{1-t_1+t_2+t_3}{4} \\
|B_{11}\rangle &= \tfrac{|01\rangle-|10\rangle}{\sqrt{2}} \quad \text{avec} \quad p_{11} = \tfrac{1-t_1-t_2-t_3}{4}
\end{aligned}$$

On peut donc exprimer $\tilde{\rho}$ dans la base propre des états de Bell et on obtient la représentation

$$\tilde{\rho} = \sum_{i,j=0}^{1} p_{ij}|B_{ij}\rangle\langle B_{ij}|$$

Dans cette base (orthonormée), la matrice densité est diagonale avec valeurs propres p_{ij}. On note que ρ est diagonale avec les mêmes valeurs propres dans la base $(U_A \otimes U_B)|B_{ij}\rangle$ qui est aussi une base orthonormée d'états intriqués formés à partir de $U_A|0\rangle$, $U_A|1\rangle$, $U_B|0\rangle$ et $U_B|1\rangle$. Finalement, on vérifie que $\sum_{i,j=0}^{1} p_{ij} = 1$ et pour obtenir le bon domaine convexe des (t_1, t_2, t_3) permis, on exprime la condition $0 \leq p_{ij} \leq 1$ $i, j = 0, 1$. À partir des expressions ci-dessus, on trouve les conditions

$$\begin{aligned}
t_1 - t_2 + t_3 &\geq -1 \\
t_1 + t_2 - t_3 &\geq -1 \\
t_1 - t_2 - t_3 &\leq 1 \\
t_1 + t_2 + t_3 &\leq 1
\end{aligned}$$

Ces quatre inégalités définissent une région bornée par quatre plans représentés à la Figure **4.1**.

Il s'agit d'un tétraèdre (qui est bien convexe). Un point contenu dans le tétraèdre correspond à un état appelé BDS (pour *Bell-diagonal state*). Les sommets du tétraèdre ne peuvent pas être écrits comme des combinaisons convexes non triviales et sont donc les états purs $|B_{ij}\rangle$, c'est-à-dire les quatre états de Bell. Tous les autres points correspondent à des états mixtes. Pour les arêtes, nous avons des mélanges de deux états de Bell, pour les quatre faces, des mélanges de trois états de Bell et pour l'intérieur du tétraèdre, des

[4] Le but de cette transformation est d'appliquer une rotation à ρ de façon à n'avoir plus que 3 paramètres libres (t_1, t_2, t_3) au lieu des 9 paramètres (linéairement dépendants) contenus dans la matrice (t_{nm}). Ceci permet de pouvoir représenter géométriquement ρ en trois dimensions. Notez que trouver la bonne transformation $U_A \otimes U_B$ n'est pas forcément facile.

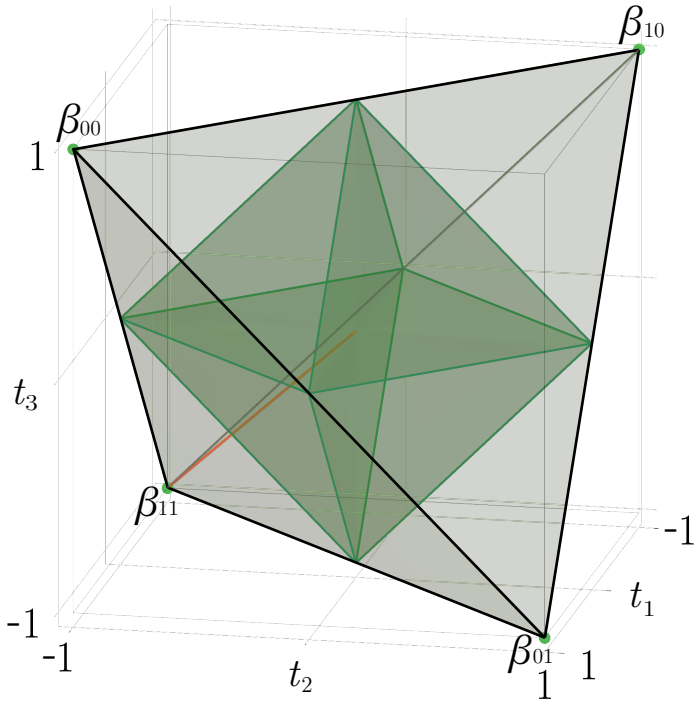


Figure 4.1 Représentation géométrique des états BDS par un tétraèdre. L'octaèdre intérieur correspond au sous-ensemble des BDS séparables. La ligne rouge correspond aux états de Werner.

mélanges des quatre états de Bell.

Ce tétraèdre contient des états séparables et intriqués. Intuitivement, il est clair que les états séparables devraient être proches de l'origine $(t_1, t_2, t_3) = (0, 0, 0)$ et les états intriqués proches des quatre sommets. Grâce au critère de Peres, il est facile de déterminer ces régions exactement. En effet, la matrice transposée partielle d'un BDS est

$$\begin{aligned}\rho^{T_B} &= \frac{1}{4}\left(\mathbb{1}\otimes\mathbb{1}^T + t_1\,\sigma_1\otimes\sigma_1^T + t_2\,\sigma_2\otimes\sigma_2^T + t_3\,\sigma_3\otimes\sigma_3^T\right)\\ &= \frac{1}{4}\left(\mathbb{1}\otimes\mathbb{1} + t_1\,\sigma_1\otimes\sigma_1 - t_2\,\sigma_2\otimes\sigma_2 + t_3\,\sigma_3\otimes\sigma_3\right)\end{aligned}$$

Si (t_1, t_2, t_3) est dans le tétraèdre, alors ρ^{T_B} forme un ensemble de matrices correspondant à un tétraèdre réfléchi autour du plan (t_1, t_3). Les points du tétraèdre réfléchi qui correspondent à une bonne matrice densité (avec v.p. positives) sont tels que $(t_1, -t_2, t_3)$ reste dans le tétraèdre de départ. Ainsi, les états séparables doivent correspondre à l'intersection du tétraèdre de départ avec le tétraèdre réfléchi. Cette intersection correspond à un octaèdre contenant l'origine (une région bornée par huit faces) également

montrée à la Figure **4.1**. Il s'agit de la région

$$\text{SEP} = \left\{(t_1, t_2, t_3) \mid |t_1| + |t_2| + |t_3| \leq 1\right\}$$

Le complément de cette région correspond aux états intriqués et est donnée par quatre pyramides disjointes contenant chacune un état de Bell (sommet du tétraèdre de départ).

Pour conclure, nous mentionnons encore que la ligne paramétrée par $w = -t_1 = -t_2 = -t_3$ des états dits *de Werner* joue un rôle spécial dans l'étude des BDS en raison de la simplicité de ces états. On peut facilement voir que ces états sont une simple combinaison convexe de l'état totalement mixte et $|B_{11}\rangle$, c.-à-d.

$$\tilde{\rho} = \frac{1-w}{4}\mathbb{1} \otimes \mathbb{1} + w\,|B_{11}\rangle\langle B_{11}|$$

Il est facile de calculer $\tilde{\rho}^{T_B}$ et de voir par le critère de Peres que pour $0 \leq w \leq \frac{1}{3}$, les états sont séparables et pour $\frac{1}{3} \leq w \leq 1$, ils sont intriqués, Le point $w = \frac{1}{3}$ se trouve sur une face de l'octaèdre.

Part II

Information et Calcul Quantiques

5 Cryptographie Quantique

Une des premières applications de la MQ à la théorie de l'information quantique est le protocole inventé par Bennett et Brassard en 1984 pour la distribution d'une clé secrète entre deux acteurs distants (Alice et Bob). Depuis, d'autres protocoles ont vu le jour et une nouvelle discipline, "la cryptographie quantique", a émergée. À proprement parler, comme nous le verrons, il ne s'agit pas vraiment de cryptographie, mais plutôt de méthodes de génération de clé secrète commune.

L'idée générale du protocole BB84 est la suivante. Alice envoie une suite de bits classiques - la clé secrète - à Bob en utilisant des qubits intermédiaires[1]. Toute tentative, de la part d'un troisième acteur (Eve) d'extraction d'information à propos de la clé nécessite d'observer les qubits. selon les postulats de la MQ cette observation perturbe l'état des qubits. Nous verrons qu'Alice et Bob sont capables de détecter cette perturbation, et donc la présence d'Eve. Dans un tel cas de figure, la communication est arrêtée.

Le sujet est bien plus compliqué que le traitement exposé dans ce chapitre. En réalité, le canal de communication (la fibre optique) est bruité et il n'est pas évident de distinguer les perturbations d'Eve de celles associées au bruit. D'autre part, les opérations d'Alice et Bob ne sont pas parfaites, au niveau de la préparation des états ainsi qu'au niveau de leurs mesures. La preuve mathématique de la sécurité du protocole de BB84 repose sur des hypothèses qui peuvent en pratique être violées. Néanmoins, si l'on accepte certaines hypothèses, on peut démontrer la sécurité du protocole. Une telle preuve est hautement non triviale et dépasse largement le cadre de ce cours. Nous discuterons néanmoins deux attaques simplifiées de la part d'Eve, ce qui sera suffisant pour comprendre pourquoi les principes de la MQ assurent la sécurité de la clé.

La cryptographie quantique n'est pas seulement une idée théorique, c'est également un sujet véritablement expérimental. La génération de clé secrète commune a été réalisée dans les laboratoires (d'abord chez IBM en 1989, sur une distance de 32 cm!) et plus tard à l'extérieur des laboratoires sur des distances de quelques dizaines à des centaines de kilomètres (Genève, Los Alamos ...). Aujourd'hui, il existe des sociétés proposant des systèmes commerciaux[2]. Des implémentations récentes permettent la génération de clés secrètes communes sur des distances de 100 km (resp. 250 km) à un taux de de 6000 (resp. 15) bits par seconde. Celles-ci exigent une connaissance approfondie de l'optique et ne seront pas discutées ici. Récemment, ces systèmes ont été violés en exploitant les

[1] Ici nous pouvons penser au qubit associé à la polarisation du photon; bien qu'en pratique le protocole est implémenté avec des degrés de libertés associés à la phase des photons.

[2] IdQuantique

limites physiques des photo-détecteurs du côté de Bob. En illuminant un photodétecteur de façon appropriée, celui-ci fonctionne alors en mode classique et les avantages liés à la MQ sont perdus.

5.1 La génération des clés selon BB84

Le protocole comporte quatre phases essentielles: la procédure d'encodage d'Alice, la procédure de décodage de Bob, une communication publique entre les deux parties, et enfin la génération de la clé secrète commune. La Figure **5.1** illustre le set-up général.

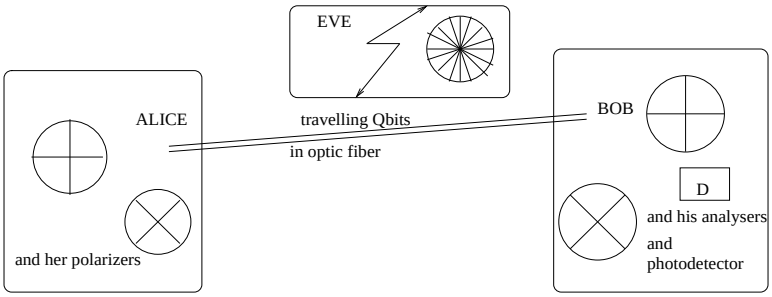


Figure 5.1 Alice et Bob génèrent une clé secrète sur le canal d'une fibre optique

Procédure de codage d'Alice. Elle génère une suite binaire aléatoire $x_1, ..., x_n$, $x_i \in \{0, 1\}$ qu'elle garde secrète. La clé commune sera un sous-ensemble de ces bits. Elle génère également une deuxième suite binaire aléatoire $e_1, ..., e_N$, $e_i \in \{0, 1\}$ qu'elle garde secrète *pour l'instant*. Alice *encode* les bits classiques x_i en qubits comme suit:

- Pour $e_i = 0$ elle génère un qubit dans l'état $|x_i\rangle$. Concrètement, elle prépare les photons avec un polariseur dans la base Z (Figure **18.2**).

$$\Big\{|0\rangle, |1\rangle\Big\}$$

  Pour $x_i = 0$ (resp. $x_i = 1$) le polariseur est orienté horizontalement (resp. verticalement). Ainsi les photons sont préparés dans l'état de polarisation $|0\rangle$ (resp. $|1\rangle$). Un seul photon est ensuite sélectionné dans le faisceau sortant (ce qui bien-sûr est une idéalisation).

Figure 5.2 Orientations des polariseurs pour la préparation des photons dans la base Z.

- Pour $e_i = 1$, elle génère un qubit dans l'état $H|x_i\rangle$[3]. Concrètement, cela peut se faire en envoyant des photons à travers un polariseur dans la base X (Figure **18.3**)

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

ce qui les prépare dans un état de polarisation $\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$ (resp. $\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$.

Figure 5.3 Orientations du polariseur pour la préparation des photons dans la base X.

En résumé, Alice envoie une chaîne de qubits $|A_{e_i,x_i}\rangle = H^{e_i}|x_i\rangle$, $i = 1, ..., N$ par un canal (dans la pratique, une fibre optique).

Procédure de décodage de Bob . Bob génère une suite binaire aléatoire $d_1, ..., d_n$, $d_i \in \{0, 1\}$ qu'il garde secrète *pour l'instant* . Il décode les qubits reçus d'Alice comme suit:

- Si $d_i = 0$ il effectue une mesure des qubits reçus $|A_{e_i,x_i}\rangle$ dans la base Z

$$\Big\{|0\rangle, |1\rangle\Big\}.$$

L'état photon après la mesure

$$|y_i\rangle \in \Big\{|0\rangle, |1\rangle\Big\}.$$

est enregistré dans des bits classiques y_i. Pour ce faire, concrètement, il utilise l' appareil de mesure analyseur-détecteur (Figure **5.4**) décrit dans le Chapitre **2**: l'analyseur est placé horizontalement; si le détecteur clique, cela signifie que l'état des photons est projeté sur $|0\rangle$; et si le détecteur ne clique pas, cela signifie que l'état photon est projeté sur $|1\rangle$. Nous soulignons que, selon le postulat de la mesure, ces résultats sont *vraiment aléatoire* . C'est uniquement Bob qui les connaît.

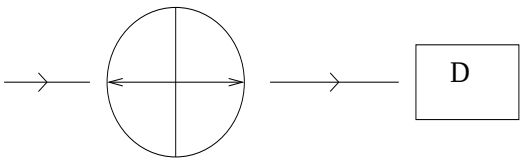


Figure 5.4 Dispositif analyseur-détecteur pour la mesure de la polarisation dans la base Z.

[3] Ici H est la matrice de Hadamard $\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$.

- Si $d_i = 1$, il effectue une mesure des qubits reçus $|A_{e_i,x_i}\rangle$ dans la base X.

$$\left\{\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle, \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle\right\}.$$

L'état du photon après la mesure appartient à

$$H|y_i\rangle \in \left\{\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle), \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)\right\}$$

Pour un état $H|y_i\rangle$, Bob enregistre le bit classique y_i .

Pour ce faire, concrètement, il utilise l'appareil analyseur-détecteur décrit dans le Chapitre **2**: l'analyseur est tourné vers la droite (Figure **5.5**) à 45 degrés; si le détecteur clique, cela signifie l'état des photons est projeté sur l'état $H|0\rangle$; tandis que si le détecteur ne clique pas cela signifie que l'état photon est projeté sur $H|1\rangle$. Nous soulignons à nouveau que, selon le postulat de la mesure, ces résultats sont *vraiment aléatoire*. Seul Bob les connaît.

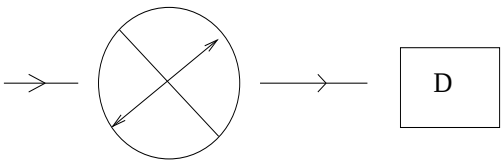


Figure 5.5 Dispositif analyseur-détecteur pour la mesure de la polarisation dans la base X.

En résumé, Bob a décodé les qubits envoyés par Alice, en une suite binaire classique $y_1, ..., y_n$. cette suite est le résultat des mesures de Bob et ne peut être prédite.

Communication Publique. Alice possède à sa disposition deux suites binaires: $e_1, ..., e_N$ utilisée pour encoder; et $x_1, ..., x_N$ qui est mappée sur les qubits. Bob aussi possède deux suites binaires: $d_1, ..., d_N$ pour choisir une base de mesure et $y_1, ..., y_N$ qui sont les résultats des mesures.

Alice et Bob communiquent $e_1, ..., e_N$ et $d_1, ..., d_N$ sur un canal public classique, et gardent les deux suites $x_1, ..., x_N$ et $y_1, ..., y_N$ secrètes. *Il importe que la communication publique ne commence qu'après la phase de mesures de Bob*. Alice et Bob peuvent déduire les informations suivantes (en fait quiconque entendant la communication publique peut déduire ces informations):

- Si $d_i = e_i$, c.-à-d. s'ils ont utilisé la même base, alors certainement $y_i = x_i$ (on peut s'en convaincre avec quelques exemples; en fait si Bob et Alice ont utilisé la même base, c'est comme s'ils vivaient dans un monde classique).
- Si $d_i \neq e_i$, c.-à-d. s'ils n'ont pas utilisé la même base, de véritables effets quantiques entrent en jeu quand Bob fait la mesure. Selon le postulat de la mesure $y_i \neq x_i$ avec probabilité $\frac{1}{2}$ et $y_i = x_i$ avec probabilité $\frac{1}{2}$. Prouvons-le. Bob reçoit le qubit

$$|A_{e_i,x_i}\rangle = H^{e_i}|x_i\rangle$$

et mesure dans la base

$$\{H^{d_i}|0\rangle, H^{d_i}|1\rangle\}.$$

Le résultat sera un des deux vecteurs de base:

$$H^{d_i}|0\rangle, \qquad \text{avec proba } |\langle 0|H^{d_i}H^{e_i}|x_i\rangle|^2$$

ou bien

$$H^{d_i}|1\rangle, \qquad \text{avec proba } |\langle 1|H^{d_i}H^{e_i}|x_i\rangle|^2.$$

Le lecteur peut vérifier que si $e_i \neq d_i$ les deux probabilités correspondantes sont égales à $\frac{1}{2}$ (et si $e_i = d_i$ elles valent 0 ou 1).

Génération de la clé commune. Bob et Alice effacent tous les bits x_i et y_i correspondant à i tel que $e_i \neq d_i$. Ils gardent les bits x_i et y_i restants indexés par i tels que $e_i = d_i$. Ils sont assurés que ces deux suites de bits sont identiques $x_i = y_i$, donc cela peut potentiellement constituer la clé secrète commune. La longueur de cette sous-suite est proche de $\frac{N}{2}$, car $\mathbb{P}[e_i \neq d_i] = \frac{1}{2}$. Enfin, Alice et Bob effectuent un test de sécurité: selon la mécanique quantique, on doit avoir[4]

$$\mathbb{P}[x_i = y_i | e_i = d_i] = 1$$

Alice et Bob testent cette condition en échangeant une petite fraction (disons $\epsilon\frac{N}{2}$) de la sous-suite commune sur le canal public. Si le test réussi, ils gardent le reste de sous suite commune: ils ont réussi à générer une clé secrète commune de longueur $(1-\epsilon)\frac{N}{2}$.

5.2 Attaques de la part d'Eve - discussion simplifiée

Nous supposons qu'Alice a une source de photon unique parfait, que la préparation des qubits est parfaite, qu'il n'y a pas de bruit et que les appareils de mesure de Bob sont parfaits. Dans ces conditions, lorsque Eve est absente, le critère de sécurité absolue $\mathbb{P}[x_i = y_i | e_i = d_i] = 1$ est satisfait.

En outre, nous supposons qu'Eve peut seulement attaquer en effectuant des opérations sur un qubit à la fois, sur les photons capturés le long de la fibre optique et qu'elle n'a pas accès aux laboratoires d'Alice et Bob. Nous supposons néanmoins qu'Eve a une connaissance parfaite des orientations X et Z des polariseurs et analyseurs d'Alice et Bob (mais pas des choix aléatoires successifs).

Nous considérons deux attaques possibles: "l'attaque basée sur une mesure" et "l'attaque basée sur des opérations unitaires". Les deux attaques se composent de deux étapes. Premièrement, Eve capture un photon, fait une mesure ou applique une opération unitaire, puis transmet le photon à Bob, ou alors lui transmet un autre photon (Figure 5.6). Nous allons voir que les postulats de base de QM impliquent que le critère de sécurité est violé. Lorsque Alice et Bob constatent cette violation, ils découvrent la présence d'Eve et interrompent la communication.

[4] Sans bruit et sans la présence d'Eve.

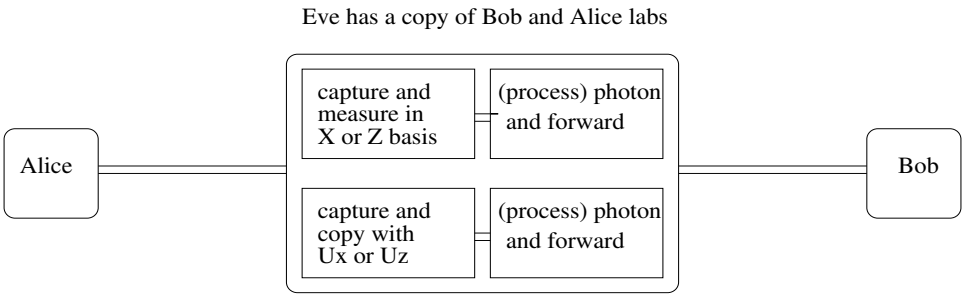


Figure 5.6 Laboratoire d'interception de photons d'Eve sur le chemin de la fibre optique

Attaque de type mesure. Supposons qu'Eve capture un photon dans la fibre optique. Le photon capturé est dans l'un des états

$$|A_{e_i,x_i}\rangle \in \{|0\rangle, |1\rangle, H|0\rangle, H|1\rangle\}$$

Eve effectue une mesure. Si elle utilise la base Z le résultat appartient à $\{|0\rangle, |1\rangle\}$ et elle enregistre un bit $y_i^E \in \{0, 1\}$. Si elle utilise la base X son résultat est dans $\{H|0\rangle, H|1\rangle\}$ et elle enregistre le bit correspondant $y_i^E \in \{0, 1\}$. Une fois qu'elle a terminé la mesure, elle envoie le photon à Bob dans le nouvel état laissé par la mesure[5]. Deux possibilités peuvent se présenter:

- Eve a utilisé la même base qu'Alice: alors $y_i^E = x_i$ et le photon est reçu par Bob dans l'état correct.
- Eve a utilisé une base différente qu'Alice: alors $y_i^E = x_i$ seulement la moitié du temps, et elle envoie à Bob le photon dans un état "correct" seulement la moitié du temps.

Voyons ce qu'Alice et Bob trouvent quand ils effectuent le test de sécurité. On note EA pour l'évènement "Eve utilise la même base que Alice".

$$\begin{aligned}\mathbb{P}\left[x_i = y_i | e_i = d_i\right] &= \mathbb{P}\left[x_i = y_i | e_i = d_i, EA\right]\mathbb{P}\left[EA\right] \\ &\quad + \mathbb{P}\left[x_i = y_i | e_i = d_i, \text{not } EA\right]\mathbb{P}\left[\text{not } EA\right] \\ &= 1 \cdot \mathbb{P}\left[EA\right] + \frac{1}{2} \cdot (1 - \mathbb{P}\left[EA\right]) \\ &= \frac{1}{2}(1 + \mathbb{P}\left[EA\right])\end{aligned}$$

où nous avons utilisé

$$\mathbb{P}\left[x_i = y_i | e_i = d_i, EA\right] = 1, \qquad \mathbb{P}\left[x_i = y_i | e_i = d_i, \text{not EA}\right] = \frac{1}{2} \tag{5.1}$$

En supposant qu'Eve n'a aucune information sur les choix de base d'Alice, nous prenons

[5] Elle pourrait aussi envoyer un autre photon dans cet état ou un autre état, mais cela ne peut pas améliorer sa performance.

$\mathbb{P}[EA] = \frac{1}{2}$. Alors

$$\mathbb{P}[x_i = y_i | e_i = d_i] = \frac{3}{4}.$$

Alice et Bob savent qu'un quart des bits ne sont pas corrects même quand ils ont utilisés la même base: ils concluent qu'un espion est à l'uvre!

Attaque unitaire. Avec l'attaque précédente, lorsque Eve fait une mesure, elle n'a aucune information sur la base qu'Alice a choisie. Une solution possible serait de copier les qubits $|A_{e_i,x_i}\rangle$ et laisser le photon dans l'état original arriver à Bob. La copie pourra être utilisée par Eve (du moins, le croit-elle) après la phase de communication publique. Quand Alice et Bob entrent dans la phase de communication publique, Eve connaît les choix de base de Bob. Donc pour i tel que $e_i = d_i$ elle obtient les mêmes résultats que Bob $y_i^E = y_i = x_i$. Elle partage donc le secret d'Alice et Bob.

Toutefois, il y a une erreur dans le raisonnement ci-dessus. Le *théorème de non-clonage* (Section **3.4**) (qui est une conséquence du postulat de l'évolution unitaire) garanti qu'il n'existe pas de machine (unitaire) telle que

$$U(|A_{e_i,x_i}\rangle \otimes |\text{blank}\rangle) = |A_{e_i,x_i}\rangle \otimes |A_{e_i,x_i}\rangle$$

Le point important est que $|A_{e_i,x_i}\rangle$ appartient à

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

qui est un ensemble d'états non-orthogonaux! (A et B ont bien travaillé leur MQ).

Eve pourrait essayer d'utiliser deux machines de copie: une pour la copie des deux états de la base Z et une autre pour la copie des deux états de la base X. Mais cette fois, elle n'a aucun moyen de savoir laquelle des deux utiliser quand un photon est capturé. Elle utilisera la mauvaise machine la moitié du temps. Une analyse similaire à la précédente montre qu'à nouveau

$$\mathbb{P}[x_i = y_i | e_i = d_i] = \frac{3}{4}.$$

Le critère de sécurité est violé: le quart des bits communs sont corrompus.

5.3 Le protocole de Bennet 1992

Dans BB84, Alice prépare les photons dans 4 états possibles. En fait, l'analyse précédente montre que le point important est que ces états ne sont pas deux à deux tous orthogonaux. Bennet inventa en 1992 un protocole qui utilise uniquement deux états non-orthogonaux: $|0\rangle$ et $H|0\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$. Ce protocole ainsi qu'une analyse détaillée de sa sécurité est l'objet d'un exercice.

6 Intrication Quantique

Dans ce chapitre, nous étudions la nature des corrélations présentes dans les états "intriqués". Ce sont de véritables corrélations de type quantique intrinsèquement présentes dans des systèmes quantiques à plusieurs qubits. Ces corrélations n'ont pas d'équivalent classique, en d'autres termes, elles ne peuvent pas être décrites par des distributions de probabilité classiques. On utilise le terme intrication pour désigner ce type spécial de corrélations dont la nature est purement quantique.

Tout d'abord, nous allons discuter le prototype des états intriqués pour deux qubits: les états de Bell (Section **6.1**). Nous abordons ensuite le sujet des inégalités de Bell[1] (Section **6.2**). Ces inégalités, qui furent d'abord proposées par John Bell ($\sim$ 1964), donnent un critère d'intrication qui peut être vérifié expérimentalement. Ces expériences, qui furent réalisées en premier par Aspect et Grangier, puis dans divers autres contextes, confirment les prédications théoriques de la MQ.

L'intrication est une ressource importante dans le traitement quantique de l'information. Elle joue un rôle important dans plusieurs protocoles importants pour la communication quantique. Ici nous décrirons deux applications, sous leur forme la plus simple possible: la *téléportation* (Section **6.3**) et le *codage superdense* (Section **6.4**). L'intrication a aussi été utilisée dans des protocoles de cryptographie quantique (Ekert 1991). Toutes ces applications ont été réalisées expérimentalement. L'intrication joue aussi un rôle important dans le calcul quantique que nous aborderons plus tard dans le cours.

6.1 États de Bell

Le prototype des états intriqués est constitué des états de Bell. Ceux-ci forment une base orthonormée de $\mathbb{C}^2 \otimes \mathbb{C}^2$ qui est un espace à 4 dimensions. Les 4 états de Bell sont

$$|B_{00}\rangle = \frac{1}{\sqrt{2}}\Big(|00\rangle + |11\rangle\Big) = U|00\rangle$$

$$|B_{01}\rangle = \frac{1}{\sqrt{2}}\Big(|01\rangle + |10\rangle\Big) = U|01\rangle$$

$$|B_{10}\rangle = \frac{1}{\sqrt{2}}\Big(|00\rangle - |11\rangle\Big) = U|10\rangle$$

[1] Nous allons en fait discuter une version plus transparente due à Clauser, Horne, Shimony, Holt (CHSH).

$$|B_{11}\rangle = \frac{1}{\sqrt{2}}\Big(|01\rangle - |10\rangle\Big) = U|11\rangle$$

Ici U est une matrice unitaire 4×4 qui peut facilement être construite. Dans ces états, les deux qubits sont en quelque sorte corrélés: en effet, dans l'état $|B_{00}\rangle$, les deux degrés de libertés de polarisation des deux photons sont parallèles, ou bien les deux états de spin sont parallèles. En fait, il serait faux de penser que la direction (des deux spins p.ex) est up-up ou down-down. En effet, le lecteur peut vérifier que

$$|B_{00}\rangle = \frac{1}{\sqrt{2}}\Big(|00\rangle + |11\rangle\Big) = \frac{1}{\sqrt{2}}\Big(|\gamma\gamma\rangle + |\gamma_\perp\gamma_\perp\rangle\Big) \tag{6.1}$$

où $|\gamma\rangle = \cos\gamma|0\rangle + \sin\gamma|1\rangle$. Ainsi les degrés de libertés ne pointent pas dans des directions précises. Néanmoins, les directions des deux degrés de liberté sont corrélées.

Une paire de photons ou une paire de moments magnétiques (spins) peut être préparée dans un état de Bell. Pour cela, il faut amener les deux degrés de liberté suffisamment près l'un de l'autre dans l'espace et les faire interagir. Si l'interaction est appropriée, une corrélation est induite. La paire de particules peut ensuite être spatialement séparée. Si les particules ne sont pas affectées par leur environnement, la corrélation des degrés de liberté de polarisation ou de spin est préservée.

Comme nous allons le voir dans ce chapitre, les corrélations dont nous parlons n'ont pas d'analogue classique. On utilise le terme "intrication" pour désigner ce type spécial de corrélations. Les paires dans les états de Bell sont aussi appelées "paires EPR" car Einstein, Podolski et Rosen (ainsi que Schrödinger) furent parmi les premiers à attirer l'attention sur certaines propriétés en apparence paradoxales de ces états (pour les degrés de liberté de position et d'impulsion en fait).

Pour nous familiariser avec la subtilité de ces états, nous commençons par discuter des scénarios de mesures possibles. Nous supposons qu'une source produit des paires de photons EPR, qu'Alice a capturé un photon dans son laboratoire, et que Bob a capturé l'autre photon dans son laboratoire (Figure **6.1**). Quelle que soit la distance entre les deux laboratoires, les photons (leur polarisation) restent intriqués dans l'état $|B_{00}\rangle$. Regardons le résultat de plusieurs mesures simples qu'Alice et Bob pourrait faire, chacun dans leur propre laboratoire. *Dans ce paragraphe, nous supposons qu'ils ne peuvent pas communiquer entre eux les résultats de ces mesures* . Nous examinons trois situations précises où: Alice mesure avant / Bob mesure après; Bob mesure avant / Alice mesure après; Alice et Bob mesurent simultanément.

- *Alice mesure avant et Bob après*. L'appareil de mesure d'Alice est formé par les projecteurs $\{|\alpha\rangle\langle\alpha|\otimes\mathbb{1}, |\alpha_\perp\rangle\langle\alpha_\perp|\otimes\mathbb{1}\}$. Selon le postulat de la mesure, l'état de Bell est projeté sur l'un des états

$$|\alpha\rangle\langle\alpha|\otimes\mathbb{1}|B_{00}\rangle = \frac{1}{\sqrt{2}}|\alpha\alpha\rangle \to |\alpha\rangle\otimes|\alpha\rangle \qquad \text{avec proba } \frac{1}{2}$$

$$|\alpha_\perp\rangle\langle\alpha_\perp|\otimes\mathbb{1}|B_{00}\rangle = \frac{1}{\sqrt{2}}|\alpha_\perp\alpha_\perp\rangle \to |\alpha_\perp\rangle\otimes|\alpha_\perp\rangle \qquad \text{avec proba } \frac{1}{2}$$

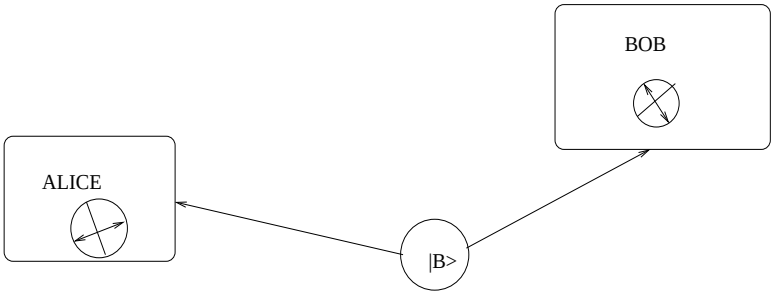


Figure 6.1 Alice et Bob partagent une paire intriquée.

Par conséquent, Alice observe *son photon* dans l'état $|\alpha\rangle$ ou $|\alpha_\perp\rangle$ avec probabilité $1/2$ (Bob, de son côté, ne sait rien, et ne sait même pas qu'Alice a effectué une mesure!). Pour effectuer sa mesure, Bob choisi une base $\{|\beta\rangle, |\beta_\perp\rangle\}$. Si son photon est dans l'état $|\alpha\rangle$ avant la mesure, celui-ci est projeté sur $|\beta\rangle$ avec probabilité $\cos^2(\alpha-\beta)$ ou $|\beta_\perp\rangle$ avec probabilité $\sin^2(\alpha-\beta)$. De même, si son photon est dans l'état $|\alpha_\perp\rangle$, il obtient le même résultat avec $\cos^2$ et $\sin^2$ interchangés. Le fait que Bob ne connaisse pas l'état initial de son photon, ou qu'il ne sache même pas si Alice a déjà mesuré, ne devrait pas vous déranger : le point est qu'il fait une expérience spécifique (mesure dans la base $|\beta\rangle, |\beta_\perp\rangle$) et trouve un résultat net. Le résultat net dans le laboratoire de Bob est: le photon est dans l'état $|\beta\rangle$ avec probabilité $\frac{1}{2}$ ou $|\beta_\perp\rangle$ avec probabilité $\frac{1}{2}$.

- *Bob mesure d'abord et Alice après.* La même discussion montre que, si Bob effectue des mesures en premier (dans la base $|\beta\rangle, |\beta_\perp\rangle$), pendant qu'Alice dort, puis Alice mesure après (dans la base $|\alpha\rangle, |\alpha_\perp\rangle$), le résultat net de chacune des parties est le même que précédemment.
- *Bob et Alice mesurent simultanément.* Vous pensez peut-être (?) que les résultats sont différents si les deux parties effectuent des mesures simultanées. Essayons. Supposons qu'Alice et Bob effectuent des mesures simultanées dans la base

$$\{|\alpha\beta\rangle; |\alpha\beta_\perp\rangle; |\alpha_\perp\beta\rangle; |\alpha_\perp\beta_\perp\rangle\}.$$

L'état de Bell

$$|B_{00}\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle) = \frac{1}{\sqrt{2}}(|\gamma\gamma\rangle + |\gamma_\perp\gamma_\perp\rangle)$$

va être projeté sur l'un des quatre états de base. Donc, Alice sera en possession d'un photon dans l'état $|\alpha\rangle$ ou $|\alpha_\perp\rangle$ et Bob en possession d'un photon dans l'état $|\beta\rangle$ ou $|\beta_\perp\rangle$. La situation est exactement la même que précédemment! Il est très instructif de calculer les probabilités des états projetés respectifs. Alice constate que la probabilité de ses résultats $|\alpha\rangle$ (resp. $|\alpha_\perp\rangle$) valent

$$\frac{1}{2}\cos^2(\alpha-\beta) + \frac{1}{2}\sin^2(\alpha-\beta) = \frac{1}{2}$$

La même chose vaut pour Bob. Par conséquent, les conclusions qu'Alice et Bob

déduisent de leurs mesures simultanées sont les mêmes que dans les cas non-simultanés ci-dessus. En fait, l'ordre des mesures importe peu.

Résumons la situation. Lorsque Alice et/ou Bob effectuent des mesures locales successives ou simultanées sur leurs photons, quel que soit leur choix de base, ils trouvent le photon dans l'un des deux états de la base choisie avec une probabilité $\frac{1}{2}$. En d'autres termes, l'entropie de la distribution de probabilité de leurs résultats locaux est maximale (elle est égale à $\ln 2 = 1$ bit). Alice et Bob déduisent que leur photon est dans un état "maximalement désordonné". Ceci est remarquable. En fait, s'ils ne savent pas que la source a produit une paire intriquée, ou si personne ne leur dit que les deux photons sont intriqués, et qu'ils n'ont aucun moyen de communiquer leurs mesures, ils n'ont aucun moyen de détecter l'intrication. Comme nous le verrons, la situation est encore plus subtile. Alice et Bob peuvent affirmer que leurs photons sont intriqués s'ils sont autorisés à communiquer. Par communiquer, nous entendons la transmission d'un message classique. Il semble donc que nous n'ayons aucun moyen de savoir si nous sommes intriqués avec des extraterrestres lointains dans l'univers en faisant uniquement des expériences locales dans notre partie de l'univers. Pour le savoir, il faudrait communiquer avec eux.

6.2 Inégalités de Bell

Nous avons vu que s'il n'y a pas de communication entre Alice et Bob, ils ne peuvent pas déduire que les photons sont dans un état intriqué. Chacun de son côté peut seulement déduire des mesures locales que l'état de son photon est hautement désordonné. Dans cette section, nous allons voir qu'en communiquant les résultats des mesures entre eux, Alice et Bob peuvent détecter l'intrication.

Les idées décrites ici ont été initiées par John Bell (1964), et motivée par un célèbre papier d'Einstein-Podolsky-Rosen (1935). Ces derniers affirmaient que les états intriqués ne constituent pas une "description complète" du système des deux particules et pensaient qu'il devait exister une théorie de "type classique" qui donne la description complète du système. L'approche de Bell donne un critère expérimental pour décider si les corrélations d'une paire EPR *peut être décrite* ou *ne peut pas être décrite* par *une théorie classique* . L'idée générale est la suivante: si une paire de photons est décrite par une théorie classique, alors certaines fonctions de corrélation appropriées des mesures d'Alice et Bob doivent satisfaire à des contraintes très particulières. Ces contraintes sont violées si la paire satisfait aux lois de la MQ. L'approche de Bell est capable de discriminer entre un vaste ensemble de théories classiques et la MQ. Les expériences fameuses d'Aspect-Grangier-Roger ont montré que la MQ gagne !

Le protocole expérimental. Une source S produit, à chaque instant n, une paire de photons. Un photon vole vers le laboratoire d'Alice et l'autre vers le laboratoire de Bob. Dans chaque laboratoire, nos deux protagonistes fonctionnent de façon indépendante: les deux laboratoires sont distants, ne communiquent pas, et ne se soucient pas ce que

l'autre fait. De plus, chaque laboratoire possède deux bases de mesures différentes (Figure **6.2**).

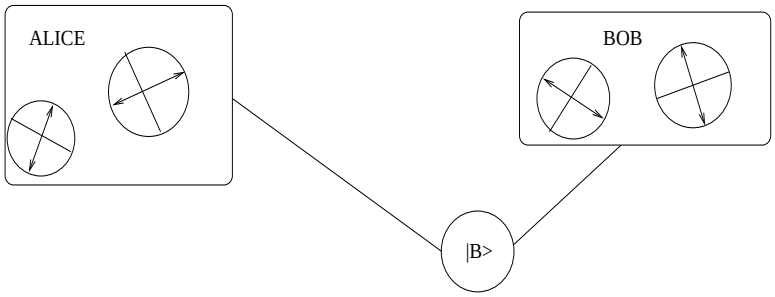


Figure 6.2 Protocole expérimental mis en place.

- À chaque instant n, Alice utilise au hasard les analyseurs (Bob ne connaît pas les choix d'Alice)

$$\{|\alpha\rangle, |\alpha_\perp\rangle\} \qquad \text{ou} \qquad \{|\alpha'\rangle, |\alpha'_\perp\rangle\}$$

  pour mesurer la polarisation de son photon. Quand elle enregistre un clic dans le détecteur, elle définit $a_n = +1$ ou $a'_n = +1$. Lorsque le détecteur ne clique pas, elle enregistre $a_n = -1$ ou $a'_n = -1$.
- À chaque instant n, Bob utilise au hasard les analyseurs (Alice ne connaît pas les choix de Bob)

$$\{|\beta\rangle, |\beta_\perp\rangle\} \qquad \text{ou} \qquad \{|\beta'\rangle, |\beta'_\perp\rangle\}$$

  pour mesurer la polarisation de son photon. Quand il enregistre un clic dans le détecteur, il enregistre $b_n = +1$ ou $b'_n = 1$. Lorsque le détecteur ne clique pas, il enregistre $b_n = -1$ ou $b'_n = -1$.
- Une fois les mesures terminées Alice et Bob passent à une phase de communication classique. Par exemple, ils se rencontrent ou se téléphonent (ou bien tweetent des messages) et discutent de leurs mesures. Ils classent les résultats selon les quatre configurations expérimentales. À chaque instant temps n, les arrangements possibles des analyseurs étaient

$$1 = (\alpha, \beta);\ 2 = (\alpha, \beta');\ 3 = (\alpha'\beta);\ 4 = (\alpha'\beta')$$

  Pour chaque arrangement, ils calculent les moyennes empiriques suivantes

$$\frac{1}{N_1}\sum_{n_1} a_{n_1} b_{n_1},\ \frac{1}{N_2}\sum_{n_2} a_{n_2} b'_{n_2},\ \frac{1}{N_3}\sum_{n_1} a'_{n_3} b_{n_3},\ \frac{1}{N_4}\sum_{n_4} a'_{n_4} b'_{n_4}$$

  Ensuite, ils calculent la fonction de corrélation suivante

$$X_{exp} = \frac{1}{N_1}\sum_{n_1} a_{n_1} b_{n_1} + \frac{1}{N_2}\sum_{n_2} a_{n_2} b'_{n_2} - \frac{1}{N_3}\sum_{n_1} a'_{n_3} b_{n_3} + \frac{1}{N_4}\sum_{n_4} a'_{n_4} b'_{n_4}$$

Pour calculer cette corrélation, Alice et Bob *doivent* communiquer pour échanger les variables a et b.

Prédiction des "théories classiques". Nous supposons que les quantités qu'Alice et Bob mesurent correspondent à des observables A, A', B, B' qui prennent simultanément des valeurs précises a, a', b, b', indépendamment de la mesure. De façon analogue, une particule classique possède une certaine position et vitesse, toutes deux bien définies (ici analogue à *a et a′*) indépendamment de l'observation (ou mesure). C'est une hypothèse habituelle de la physique classique. En outre, nous supposons que les résultats d'Alice et Bob peuvent être modélisés par une distribution de probabilité jointe[2]

$$\mathbb{P}\left[a, a', b, b'\right]$$

La prédiction théorique correspondant à *chaque* moyenne empirique ci-dessus est

$$\mathbb{E}\left[ab\right], \mathbb{E}\left[ab'\right], \mathbb{E}\left[a'b\right], \mathbb{E}\left[a'b'\right]$$

La linéarité de l'espérance implique

$$\begin{aligned} X_{\text{theorique}} &= \mathbb{E}\left[ab\right] + \mathbb{E}\left[ab'\right] - \mathbb{E}\left[a'b\right] + \mathbb{E}\left[a'b'\right] \\ &= \mathbb{E}\left[ab + ab' - a'b + a'b'\right] \end{aligned}$$

Remarquez maintenant que

$$ab + ab' - a'b + a'b' = a(b + b') + a'(b' - b)$$

et que

$$a(b + b') + a'(b' - b) = \pm 2$$

En effet, si $b = b'$, alors seul le premier terme survit et vaut ± 2, tandis que si $b \neq b'$ seulement le second terme survit et vaut aussi ± 2. La moyenne est forcément comprise dans l'intervalle $[-2, +2]$, et ainsi nous avons pour la prédiction des théories classiques

$$-2 \leq X_{\text{theorique}} \leq 2$$

C'est l'une des inégalités "de type Bell" obtenue par Clauser-Horne-Shimony-Holt (CHSH).

Afin d'obtenir ce résultat, nous avons supposé l'existence d'une distribution conjointe $\mathbb{P}\left[a, a', b, b'\right]$ pour des valeurs des observables A, A', B et B' (c'est ce qui permet d'écrire $X_{\text{theorique}}$ comme la valeur moyenne d'une quantité comprise dans $[-2, +2]$). En fait, cette hypothèse n'est pas évidente a priori. Du point de vue expérimental, lorsque Alice et Bob se rencontrent, ils peuvent construire quatre histogrammes qui correspondent aux quatre arrangements possibles des analyseurs. Ces histogrammes correspondent à quatre distributions de probabilité:

$$\mathbb{P}_1\left[a, b\right],\ \mathbb{P}_2\left[a', b\right],\ \mathbb{P}_3\left[a, b'\right],\ \mathbb{P}_4\left[a', b'\right]$$

[2] Cette seconde hypothèse sera justifiée ci-dessous. Elle suit d'une hypothèse de localité des résultats de mesure. Elle englobe un vaste ensemble de théories classiques possibles, déterministes ou non.

Il n'est à priori pas évident que ces 4 distributions soient les marginales d'une distribution commune $\mathbb{P}[a, a', b, b']$. Nous allons voir que pour une théorie classique locale cela doit effectivement être le cas.

Admettons que les lois de la physique soient "locales". Nous entendons par là que lorsque Alice (resp. Bob) effectuent ses mesures, les résultats d'Alice (resp. Bob) ne dépendent que de son propre choix local des analyseurs. C'est dire qu'étant donné un état du système décrit par un ensemble de variables classiques λ (appelées parfois variables cachées), les résultats des expériences d'Alice et Bob doivent être indépendantes. Elles sont modélisées par des distributions dépendant uniquement de la configuration locale des analyseurs et de l'état du système:

$$p_{\mathcal{A}}(a|\alpha, \lambda),\ \ p_{\mathcal{A}}(a'|\alpha', \lambda),\ \ p_{\mathcal{B}}(b|\beta, \lambda),\ \ p_{\mathcal{B}}(b'|\beta', \lambda)$$

Alors pour les choix $\alpha, \alpha', \beta, \beta'$ fixes, les histogrammes d'Alice et Bob sont donnés par

$$\mathbb{P}_1[a, b] = \int \mathrm{d}\lambda\ h(\lambda) p_{\mathcal{A}}(a|\alpha, \lambda) p_{\mathcal{B}}(b|\beta, \lambda)$$

$$\mathbb{P}_2[a', b] = \int \mathrm{d}\lambda\ h(\lambda) p_{\mathcal{A}}(a|\alpha, \lambda) p_{\mathcal{B}}(b'|\beta'; \lambda)$$

$$\mathbb{P}_3[a, b'] = \int \mathrm{d}\lambda\ h(\lambda) p_{\mathcal{A}}(a'|\alpha'; \lambda) p_{\mathcal{B}}(b|\beta, \lambda)$$

$$\mathbb{P}_4[a', b'] = \int \mathrm{d}\lambda\ h(\lambda) p_{\mathcal{A}}(a|\alpha, \lambda) p_{\mathcal{B}}(b|\beta, \lambda)$$

Ce sont les marginales d'une distribution de probabilité jointe

$$\mathbb{P}_{\text{classique}}[a, a', b, b'] = \int \mathrm{d}\lambda\ h(\lambda) p_{\mathcal{A}}(a|\alpha, \lambda) p_{\mathcal{A}}(a'|\alpha'; \lambda) p_{\mathcal{B}}(b|\beta, \lambda) p_{\mathcal{B}}(b'|\beta'; \lambda)$$

Remarquez que ce formalisme englobe les théories déterministes aussi. En effet, dans les distributions ci-dessus, h, $p_{\mathcal{A}}$ et $p_{\mathcal{B}}$ pourraient être des distributions de Dirac.

Prédiction de la MQ pour un État de Bell. Tout d'abord nous remarquons que selon le formalisme quantique, les mesures d'Alice et Bob sont des mesures des 4 observables (matrices hermitiennes)

$$A = (+1)|\alpha\rangle\langle\alpha| + (-1)|\alpha_\perp\rangle\langle\alpha_\perp| \qquad A' = (+1)|\alpha'\rangle\langle\alpha'| + (-1)|\alpha'_\perp\rangle\langle\alpha'_\perp|$$

et

$$B = (+1)|\beta\rangle\langle\beta| + (-1)|\beta_\perp\rangle\langle\beta_\perp| \qquad B' = (+1)|\beta'\rangle\langle\beta'| + (-1)|\beta'_\perp\rangle\langle\beta'_\perp|$$

À chaque instant n, l'état de la paire de photons est décrit par certains ket $|\Psi\rangle \in \mathbb{C}^2 \otimes \mathbb{C}^2$. La prédiction de la mécanique quantique pour les quatre moyennes empiriques d'Alice et Bob est

$$\langle\Psi|A \otimes B|\Psi\rangle, \langle\Psi|A \otimes B'|\Psi\rangle, \langle\Psi|A' \otimes B|\Psi\rangle, \langle\Psi|A' \otimes B'|\Psi\rangle$$

Pour la fonction de corrélation

$$X_{MQ} = \langle\Psi|A \otimes B|\Psi\rangle + \langle\Psi|A \otimes B'|\Psi\rangle - \langle\Psi|A' \otimes B|\Psi\rangle + \langle\Psi|A' \otimes B'|\Psi\rangle$$

Maintenant, nous calculons cette quantité pour l'état de Bell

$$|\Psi\rangle = |B_{00}\rangle$$

La première moyenne est calculée en exprimant l'état de Bell comme $\frac{1}{\sqrt{2}}(|\alpha\alpha\rangle+|\alpha_\perp\alpha_\perp\rangle)$.

$$\begin{aligned}
\langle B_{00}|A\otimes B|B_{00}\rangle &= \frac{1}{2}\langle\alpha\alpha|A\otimes B|\alpha\alpha\rangle + \frac{1}{2}\langle\alpha_\perp\alpha_\perp|A\otimes B|\alpha_\perp\alpha_\perp\rangle \\
&+ \frac{1}{2}\langle\alpha\alpha|A\otimes B|\alpha_\perp\alpha_\perp\rangle + \frac{1}{2}\langle\alpha_\perp\alpha_\perp|A\otimes B|\alpha\alpha\rangle \\
&= \frac{1}{2}\langle\alpha|A|\alpha\rangle\langle\alpha|B|\alpha\rangle + \frac{1}{2}\langle\alpha_\perp|A|\alpha_\perp\rangle\langle\alpha_\perp|B|\alpha_\perp\rangle \\
&= \frac{1}{2}\cdot 1\cdot\left(|\langle\alpha|\beta\rangle|^2 - |\langle\alpha|\beta_\perp\rangle|^2\right) + \frac{1}{2}\cdot(-1)\cdot\left(|\langle\alpha_\perp|\beta\rangle|^2 - |\langle\alpha_\perp|\beta_\perp\rangle|^2\right) \\
&= \frac{1}{2}\left(\cos^2(\alpha-\beta) - \sin^2(\alpha-\beta)\right) - \frac{1}{2}\left(\sin^2(\alpha-\beta) - \cos^2(\alpha-\beta)\right) \\
&= \cos^2(\alpha-\beta) - \sin^2(\alpha-\beta) = \cos 2(\alpha-\beta)
\end{aligned}$$

Effectuant des calculs similaires pour les autres moyennes, nous trouvons

$$X_{MQ} = \cos 2(\alpha-\beta) + \cos 2(\alpha-\beta') - \cos 2(\alpha'-\beta) + \cos 2(\alpha'-\beta')$$

Cette quantité est maximisée pour le choix suivant des angles (et toutes les rotations globales de ce choix, Figure **6.3**):

$$\alpha = 0,\ \alpha' = -\frac{\pi}{4},\ \beta = \frac{\pi}{8},\ \beta' = -\frac{\pi}{8}$$

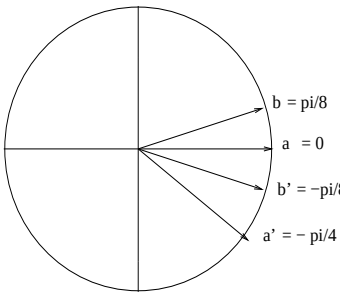


Figure 6.3 Choix optimal de l'orientation des analyseurs.

et cette quantité maximale est égale à

$$X_{MQ} = \cos\frac{\pi}{4} + \cos\frac{\pi}{4} - \cos\frac{3\pi}{4} + \cos\frac{\pi}{4} = 2\sqrt{2}$$

Nous constatons que l'inégalité CHSH est violée car $2\sqrt{2} > 2$! Pour les trois autres états Bell on trouve le même résultat.

En fait, la MQ prédit que les quatre histogrammes de Bob et Alice sont

$$\mathbb{P}_1[a,b] = \frac{1}{4}(1 + ab\cos 2(\alpha - \beta))$$
$$\mathbb{P}_2[a',b] = \frac{1}{4}(1 + ab'\cos 2(\alpha - \beta'))$$
$$\mathbb{P}_3[a,b'] = \frac{1}{4}(1 + a'b\cos 2(\alpha' - \beta))$$
$$\mathbb{P}_4[a',b'] = \frac{1}{4}(1 + a'b'\cos 2(\alpha' - \beta'))$$

Par exemple: $\mathbb{P}_2[+1,-1] = |\langle\alpha\beta'_\perp|B_{00}\rangle|^2 = \frac{1}{4}(1 - \cos 2(\alpha - \beta))$. Il y a des choix particuliers des angles $\alpha, \beta, \alpha'\beta'$ pour lesquels ces histogrammes ne sont pas les marginales d'une distribution commune $\mathbb{P}[a,b,a',b')]$. En effet, si c'était le cas nous aurions dû trouver $-2 \leq X \leq 2$. Ainsi les corrélations présentes dans les résultats de mesures ne peuvent pas être décrites par une distribution de probabilité classique. Elles sont décrite par l'état intriqué de Bell!

Remarque. On dit souvent que les états intriqués possèdent des "corrélation non-locales" car les deux photons peuvent être séparés d'une distance arbitraire, bien que l'inégalité de Bell soit violée (c.-à-d. $X = 2\sqrt{2} > 2$). Néanmoins, toutes les interactions connues sont locales au sens où les forces décroissent avec la distance et les mesures faites dans les laboratoires sont locales. Il faut donc faire attention lorsque l'on manipule les termes "local" et "non-local" en physique quantique. Les états intriqués possèdent des corrélations non-locales mais les interactions sont locales. De plus, on ne peut pas mettre en évidence l'intrication en faisant uniquement des opérations locales.

Expériences. Aspect-Grangier-Roger ont montré dans les années 1980 que l'expérience est en accord avec la MQ et non pas avec les théories classiques. Ces expériences nous forcent à abandonner la description classique. Une des difficultés expérimentales est de faire tourner les analyseurs d'Alice et Bob suffisamment rapidement pour que les événements de mesure soient séparés par un intervalle de temps plus court que le temps que mettrait la lumière pour parcourir la distance séparant Alice et Bob. Sinon, on pourrait argumenter qu'une certaine forme de communication classique ou interaction conspire pour établir les corrélations non-locales pendant la mesure.

6.3 La téléportation quantique

Supposons qu'Alice et Bob soient séparés dans l'espace et qu'Alice possède un qubit dans l'état

$$|\Phi\rangle = \alpha|0\rangle + \beta|1\rangle \qquad |\alpha|^2 + |\beta|^2 = 1$$

L'état (c.-à-d. α et β) n'est pas nécessairement connu pour Alice et n'est pas connu pour Bob. Ils partagent également une paire intriquée

$$|B_{00}\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$$

et ont aussi à leur disposition un canal de communication classique.

Nous allons expliquer que par l'envoi de seulement deux bits classiques d'information sur le canal classique, Alice peut *téléporter* l'état de son qubit vers Bob. Ici, *téléportation* signifie que $|\Phi\rangle$ est détruit dans le laboratoire d'Alice et est reconstruit dans le laboratoire de Bob. Notez que la destruction de $|\Phi\rangle$ dans le laboratoire d'Alice est nécessaire à cause du théorème de non-clonage (Section **3.4**). Après le processus de téléportation, Bob sait qu'il possède l' état $|\Phi\rangle$, mais ne connaît toujours pas l'état lui-même (c'est-à-dire qu'il ne connaît pas α et β). Nous soulignons que le processus de téléportation nécessite un transport physique d'information classique (stockée dans de la matière) dans la phase de communication classique entre Alice et Bob. Bien entendu, cette phase de communication classique ne peut pas se produire à une vitesse supérieure à celle de la lumière, de sorte que l'ensemble du processus de téléportation ne viole pas les principes de la relativité. Nous notons également que le support matériel de l'état $|\Phi\rangle$ (par exemple, la polarisation du photon, le spin de l'électron, les degrés de liberté atomiques ou moléculaires) n'est pas nécessairement le même dans les laboratoires d'Alice et de Bob.

On résume parfois la téléportation par la "loi" suivante

Téléporter 1 Qubit = Communiquer 2 Cbits + Partager 1 paire EPR

La téléportation peut être considérée comme une forme de communication entre Alice et Bob qui partagent un canal classique et un "canal constitué de paires intriquées" (canal EPR).

Le protocole.

- Une source produit une paire EPR de particules dans l'état de Bell $|B_{00}\rangle_{23}$. La particule, appelée "particule 2" est envoyée à Alice et la particule, appelée "particule 3" est envoyée à Bob. L'espace de Hilbert du système intriqué 23 est $\mathcal{H}_2 \otimes \mathcal{H}_3 = \mathbb{C}^2 \otimes \mathbb{C}^2$.
- Alice prépare une particule, appelée "particule 1", dans l'état $|\Phi\rangle_1 = \alpha|0\rangle + \beta|1\rangle$. L'espace de Hilbert de la particule 1 est $\mathcal{H}_1 = \mathbb{C}^2$.
- L'espace total de Hilbert du système composite 123 est $\mathcal{H}_1 \otimes \mathcal{H}_2 \otimes \mathcal{H}_3 = \mathbb{C}^2 \otimes \mathbb{C}^2 \otimes \mathbb{C}^2$. L'état complet du système est

  $$|\Psi\rangle = |\Phi\rangle_1 \otimes |B_{00}\rangle_{23}$$

  À ce stade, un bref calcul facilitera la discussion qui suit:

  $$|\Psi\rangle = \frac{\alpha}{\sqrt{2}}|000\rangle + \frac{\beta}{\sqrt{2}}|100\rangle + \frac{\alpha}{\sqrt{2}}|011\rangle + \frac{\beta}{\sqrt{2}}|111\rangle$$

- Alice fait une mesure locale dans son laboratoire, à savoir sur les particules 12. Elle utilise un appareil modélisé par la base de mesure dans $\mathcal{H}_1 \otimes \mathcal{H}_2$

  $$\{|B_{00}\rangle_{12}, |B_{01}\rangle_{12}, |B_{10}\rangle_{12}, |B_{11}\rangle_{12}\}$$

  Les projecteurs associés pour l'ensemble du système sont

  $$P_{00} = |B_{00}\rangle\langle B_{00}|\otimes\mathbb{1}_3, P_{01} = |B_{01}\rangle\langle B_{01}|\otimes\mathbb{1}_3, P_{10} = |B_{10}\rangle\langle B_{10}|\otimes\mathbb{1}_3, P_{11} = |B_{11}\rangle\langle B_{11}|\otimes\mathbb{1}_3$$

  Comme d'habitude, le résultat de la mesure est l'un des quatre états projetés possibles (à une normalisation près; vérifier ce calcul et aussi que la probabilité de chaque résultat est $\frac{1}{4}$)

  $$P_{00}|\Psi\rangle = \frac{1}{2}|B_{00}\rangle_{12} \otimes (\alpha|0\rangle_3 + \beta|1\rangle_3)$$

  $$P_{10}|\Psi\rangle = \frac{1}{2}|B_{10}\rangle_{12} \otimes (\alpha|0\rangle_3 - \beta|1\rangle_3)$$

  $$P_{01}|\Psi\rangle = \frac{1}{2}|B_{01}\rangle_{12} \otimes (\alpha|1\rangle_3 + \beta|0\rangle_3)$$

  $$P_{11}|\Psi\rangle = \frac{1}{2}|B_{11}\rangle_{12} \otimes (\alpha|1\rangle_3 - \beta|0\rangle_3)$$

- Pour chacun de ces résultats possibles, Bob possède l'un des quatre états

  $$\alpha|0\rangle_3 + \beta|1\rangle_3 = |\Phi\rangle$$
  $$\alpha|0\rangle_3 - \beta|1\rangle_3 = Z|\Phi\rangle$$
  $$\alpha|1\rangle_3 + \beta|0\rangle_3 = X|\Phi\rangle$$
  $$\alpha|1\rangle_3 - \beta|0\rangle_3 = -iY|\Phi\rangle$$

  mais il ne sait pas lequel, étant donné qu'il ne connait pas le résultat de la mesure d'Alice.
- Alice connait le résultat de sa mesure (dans son laboratoire). C'est l'un des quatre états de Bell. Ce résultat peut être codé par deux bits classiques, puis communique à Bob sur le canal classique,

  $$00, 01, 10, 11$$

  Dès que Bob reçoit le message d'Alice, il sait qu'elle a terminé ses opérations et possède deux bits d'information nécessaires pour décider quelle est l'opération *unitaire* qu'il doit effectuer sur son état afin de récupérer $|\Phi\rangle$,

  $$\mathbb{1}(\alpha|0\rangle_3 + \beta|1\rangle_3) = |\Phi\rangle$$
  $$Z(\alpha|0\rangle_3 - \beta|1\rangle_3) = |\Phi\rangle$$
  $$X(\alpha|1\rangle_3 + \beta|0\rangle_3) = |\Phi\rangle$$
  $$iY(\alpha|1\rangle_3 - \beta|0\rangle_3) = |\Phi\rangle$$

Notez que l'état $|\Phi\rangle$ n'existe plus dans le laboratoire d'Alice (elle l'a détruit en faisant sa mesure) et Bob l'a bien récupéré sur le qubit dans son laboratoire. L'état a donc bien été téléporté d'Alice à Bob.

6.4 Codage superdense

Supposons qu'Alice et Bob ont mis en place un canal quantique sur lequel ils peuvent envoyer des qubits (par exemple une fibre optique sur laquelle les photons voyagent). On suppose aussi qu'Alice et Bob partagent une paire EPR. Combien de bits d'information classique peuvent-ils communiquer en envoyant un seul qubit à travers le canal quantique?

La réponse: 2 bits d'information classique peuvent être transmis par Alice à Bob, en envoyant seulement 1 qubit tant qu'ils partagent une paire EPR. Le protocole qui réalise ceci s'appelle "codage superdense" (dense coding).

La "loi" du codage superdense peut être résumée ainsi:

Communiquer 2 Cbits = Envoyer 1 Qbit + Partager 1 paire EPR

Le protocole.

- Une paire EPR dans l'état $|B_{00}\rangle$ est préparée par une source. Chaque particule de la paire est envoyée à Alice et Bob.
- Alice veut communiquer deux bits d'information à Bob :
  - Pour envoyer 00 elle laisse sa particule intacte (ou applique la matrice unitaire $\mathbb{1}$) et envoie sa particule à Bob. Bob reçoit la particule et est maintenant en possession de l'état $|B_{00}\rangle$ tout entier
  $$|B_{00}\rangle$$
  - Pour envoyer 01 elle applique la matrice unitaire X à sa particule, puis envoie sa particule à Bob. Bob est maintenant en possession de la paire dans l'état
  $$X_1 \otimes \mathbb{1}_2|B_{00}\rangle = |B_{01}\rangle$$
  - Pour envoyer 10, elle applique la matrice unitaire Z à sa particule, puis envoie sa particule. Bob est maintenant en possession de la paire dans l'état
  $$Z_1 \otimes \mathbb{1}_2|B_{00}\rangle = |B_{10}\rangle$$
  - Pour envoyer 11, elle applique la matrice unitaire iY à sa particule, puis envoie physiquement sa particule. Bob est maintenant en possession de la paire dans l'état
  $$(iY)_1 \otimes \mathbb{1}_2|B_{00}\rangle = |B_{11}\rangle$$
- Bob a maintenant la paire EPR 12 dans un état $|B_{xy}\rangle$. Afin de déterminer les deux bits d'information classique qu'Alice a envoyés, il doit décider quel est l'état de Bell dont il dispose. Comme Bob sait qu'il possède l'un des quatre états de Bell dans son laboratoire, il peut faire une mesure locale dans la base de Bell, et accéder aux informations xy .

Expériences. La téléportation et le codage superdense ont été réalisés expérimentalement. Un résumé du sujet peut être trouvé dans "Les dossiers de la Recherche" no 18, février 2005 "L'étrange Pouvoir de l'intrication quantique", par N. Gisin.

7 Entropie quantique

Nous avons vu que la matrice densité permet de décrire un ensemble statistique d'états ainsi que l'état réduit d'une partie pour un système multipartite. Cette matrice constitue un analogue quantique d'une distribution de probabilités classiques et nous pouvons lui associer une entropie qui est l'analogue de l'entropie de Shannon. L'entropie quantique associée à la matrice densité est communément appelée *entropie de von Neumann*, qui introduit le concept dans le cadre de la mécanique statistique quantique en 1927 (en faisant une analogie avec l'entropie de Gibbs). De façon similaire à la théorie de l'information classique fondée par Shannon en 1948, l'entropie de von Neumann est à la base de plusieurs mesures de l'information, nécessaires à la description de la compression, transmission et accessibilité de l'information contenue dans un état. Dans ce chapitre, nous exposons uniquement quelques rudiments de cette théorie qui a pris son essor suite aux travaux d'Holevo dans les années 1970. Certains aspects de l'entropie quantique n'ont pas d'analogue classique. L'entropie de von Neumann donne notamment une mesure utile de la quantité d'intrication présente dans un état tripartite.

7.1 Entropie de Shannon

Soit X une variable aléatoire prenant des valeurs x dans un alphabet discret et fini avec probabilités $p_x = \mathbb{P}_{\mathrm{X}}[X = x]$. Par définition, l'entropie de Shannon associée à X (c.-à-d. à la loi de probabilité de X) est

$$H(X) = -\sum_x p_x \log p_x \tag{7.1}$$

Cette quantité est une mesure de *l'incertitude moyenne* à propos de X. Ici, nous ne spécifions pas la base utilisée pour le log. Habituellement, on considère le log en base 2 ou le log néperien selon le contexte.

L'entropie jointe pour deux variables aléatoires (X, Y) de loi $p_{x,y} = \mathbb{P}_{\mathrm{X,Y}}[X = x, Y = y]$ est

$$H(X, Y) = -\sum_{x,y} p_{x,y} \log p_{x,y} \tag{7.2}$$

et l'entropie conditionnelle est

$$H(X|Y) = -\sum_y p_y \sum_x p_{x|y} \log p_{x|y} \tag{7.3}$$

où $p_{x|y} = \frac{p_{x,y}}{p_y}$. En particulier, on peut écrire $H(X|Y) = \sum_y p_y H(X|Y = y)$ avec $H(X|Y = y) = -\sum_x p_{x|y} \log p_{x|y}$ l'entropie de X étant donné que $Y = y$ est observé. On voit facilement que

$$H(X|Y) = H(X, Y) - H(Y)$$

Cette formule est consistante avec l'interprétation suivante: lorsque Y est observé, l'incertitude jointe $H(X, Y)$ est réduite de la quantité $H(Y)$.

L'information mutuelle entre X et Y est le complément de l'incertitude restante pour X lorsque Y est observé:

$$I(X; Y) = H(X) - H(X|Y) \tag{7.4}$$

Cette quantité est symétrique sous l'échange de X et Y et on vérifie

$$\begin{aligned} I(X; Y) &= H(X) - H(X|Y) \\ &= H(X) + H(Y) - H(X, Y) \\ &= H(Y) - H(Y/X) \\ &= I(Y; X) \end{aligned}$$

On peut vérifier que $I(X; Y) = 0$ si et seulement si $p_{x,y} = p_x p_y$. L'information mutuelle est aussi une mesure de la corrélation entre X et Y.

Une quantité importante est la divergence de Kullback-Leibler, aussi appelée entropie relative entre deux distributions de probabilités $\mathbb{P}_X$ et $\mathbb{Q}_X$:

$$D(\mathbb{P}_X||\mathbb{Q}_X) = -\sum_x p_x \log q_x + \sum_x p_x \log p_x = \sum_x p_x \log\left(\frac{p_x}{q_x}\right)$$

Cette quantité est une mesure de "distance" entre deux distributions bien qu'elle ne soit pas symétrique: $D(\mathbb{P}_X||\mathbb{Q}_X) \neq D(\mathbb{Q}_X||\mathbb{P}_X)$. Elle est directement reliée à l'information mutuelle via

$$I(X; Y) = I(Y; X) = D(\mathbb{P}_{X,Y}||\mathbb{P}_X\mathbb{P}_Y)$$

Ainsi, l'information mutuelle est aussi une mesure de "distance" entre une distribution jointe et le produit de ses marginales.

7.2 Propriétés de base de l'entropie de Shannon

Dans ce paragraphe, nous donnons sans démonstration des inégalités essentielles satisfaites par les mesures classiques de l'information et spécifier lesquelles ne sont plus vraies dans le cas quantique. Certaines des démonstrations seront exposées directement

dans le cas quantique.

• Le maximum de $H(X)$ est atteint par la distribution uniforme. Pour un alphabet de cardinalité d, on a $0 \leq H(X) \leq \log(d)$ et la borne supérieure est atteinte pour $p_x = \frac{1}{d}$.

• H(X) est une fonctionnelle concave de $\mathbb{P}_X$. En d'autres termes, si $\mathbb{P}_Z = \sum_k a_k \mathbb{P}_{X_k}$ avec $0 \leq a_k \leq 1$ et $\sum_k a_k = 1$, on a

$$H(Z) \geq \sum_k a_k H(X_k)$$

• L'entropie est sous-additive. Si $\mathbb{P}_{X,Y}$ est une distribution jointe et $\mathbb{P}_X, \mathbb{P}_Y$ les marginales, on a:

$$H(X,Y) \leq H(X) + H(Y)$$

En particulier, cela implique aussi que $I(X;Y) \geq 0$. Il suit de la définition de l'entropie conditionnelle que

$$H(X|Y) \leq H(X) \qquad \text{et} \qquad H(Y|X) \leq H(Y)$$

• L'entropie conditionnelle est positive: $H(X|Y) \geq 0$.

De façon équivalente, $H(X,Y) \geq H(X)$ et $H(X,Y) \geq H(Y)$, c'est-à-dire que "l'entropie du tout est plus grande que l'entropie d'une partie". Cette affirmation, naturelle dans le cas classique, est fausse pour l'entropie quantique. En effet, nous verrons que pour un système bipartite intriqué, l'entropie du tout peut être nulle alors que l'entropie d'une partie est maximale (c'est le cas de paires EPR par exemple).

• Conditionnement et sous-additivité forte. Le conditionnement réduit l'entropie

$$H(X|Y,Z) \leq H(X|Y)$$

ou de façon équivalente

$$H(X,Y,Z) + H(Y) \leq H(X,Y) + H(Y,Z)$$

Cette dernière inégalité s'appelle "inégalité de sous-additivité forte". L'égalité est atteinte si et seulement si $X \to Y \to Z$ forme une chaîne de Markov, c.-à-d. ssi $p_{x,y,z} = p_{z|y} p_{y|x} p_x$. Nous notons que ceci est équivalent à $Z \to Y \to X$ forme une chaîne de Markov, c.-à-d. $p_{x,y,z} = p_{x|y} p_{y|z} p_z$. Ceci est aussi équivalent à l'indépendance de X et Z étant donné Y, c.-à-d. $p_{x,z|y} = p_{z|y} p_{x|y}$.

• Si $X \to Y \to Z$ forme une chaîne de Markov, on a l'inégalité de "data processing":

$$H(X|Z) \geq H(X|Y)$$

En effet, par sous-additivité forte, $H(X|Z) \geq H(X|Z,Y)$ et la condition de Markov entraîne $H(X|Z,Y) = H(X|Y)$ puisque X et Z sont indépendants lorsque l'on conditionne

sur Y.

• L'entropie relative est toujours positive: $D(\mathbb{P}_X \| \mathbb{Q}_X) \geq 0$. Ceci provient d'un argument de convexité qui reste valable dans le cas quantique.

• Les identités algébriques suivantes proviennent de la définition de l'entropie et de l'information mutuelle.

$$H(X_1, \ldots, X_n) = \sum_{i=1}^{n} H(X_i | X_{i+1}, \ldots, X_n)$$

$$I(X_1, \ldots, X_n; Y) = \sum_{i=1}^{n} I(X_i; Y | X_{i+1}, \ldots, X_n)$$

7.3 Entropie de von Neumann

Nous considérons maintenant un système quantique dans un état statistique ou un système qui n'est pas isolé de son environnement. Le système est décrit par une matrice densité $\rho : \mathcal{H} \longrightarrow \mathcal{H}$ telle que $\rho = \rho^\dagger$, $\rho \geq 0$ et $\mathrm{Tr}(\rho) = 1$. L'entropie de von Neumann est, par définition:

$$S(\rho) = -\mathrm{Tr}(\rho \log \rho) \tag{7.5}$$

Plus concrètement, si la décomposition spectrale de ρ est $\rho = \sum_i \lambda_i |\chi_i\rangle\langle\chi_i|$, avec $\rho|\chi_i\rangle = \lambda_i|\chi_i\rangle$, on a

$$S(\rho) = -\sum_i \lambda_i \log \lambda_i$$

qui peut s'interpréter comme l'entropie de Shannon pour la distribution de probabilités discrètes $\{\lambda_1, \ldots, \lambda_d\}$, où $d = \dim(\mathcal{H})$.

Remarque: Néanmoins, pour une préparation de l'état mixte $\{|\varphi_i\rangle, p_i\}$, où les $|\varphi_i\rangle$ ne forment en général pas une base orthonormale, l'entropie de von Neumann n'est pas égale à l'entropie de Shannon associé aux probabilités p_i. Si X est la variable aléatoire de loi $p_i = \mathbb{P}_X[X = x_i]$, on a $S(\rho) \leq H(X)$. En d'autres termes, l'état mixte contient moins d'incertitudes que X. Ceci provient du fait que des états non-orthogonaux ne peuvent pas être parfaitement distingués lors de mesures quantiques. Si les états $|\varphi_i\rangle$ sont orthonormés, alors $S(\rho) = H(X)$.

Pour un système bipartite d'espace de Hilbert $\mathcal{H}_A \otimes \mathcal{H}_B$ décrit par la matrice densité ρ_{AB}, l'entropie de von Neumann est $S(\rho_{AB}) = -\mathrm{Tr}_{\mathcal{H}_A \otimes \mathcal{H}_B}(\rho_{AB} \log \rho_{AB})$. On peut associer une entropie à chaque matrice densité réduite $\rho_A = \mathrm{Tr}_{\mathcal{H}_B} \rho_{AB}$ et $\rho_B = \mathrm{Tr}_{\mathcal{H}_A} \rho_{AB}$, c'est-à-dire

$$S(\rho_A) = -\mathrm{Tr}(\rho_A \log \rho_A) \qquad \text{et} \qquad S(\rho_B) = -\mathrm{Tr}(\rho_B \log \rho_B)$$

Nous verrons que ces quantités jouent un rôle important comme mesure quantitative de l'intrication.

Exemple: Entropie de systèmes simples. Pour un état pur $|\psi\rangle$, on a $\rho = |\psi\rangle\langle\psi|$ qui possède une valeur propre égale à 1 et les autres qui sont nulles. Donc, $S(|\psi\rangle\langle\psi|) = 0$. Un état pur ne contient aucune "incertitude".

Prenons maintenant un état pur intriqué et regardons les matrices densité réduites. Le prototype typique est l'état d'une paire EPR: $|\psi\rangle_{AB} = \frac{1}{\sqrt{2}}\Big(|0\rangle_A\otimes|0\rangle_B+|1\rangle_A\otimes|1\rangle_B\Big)$. Puisque $\rho_A = \rho_B = \frac{1}{2}\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}$, nous trouvons $S(\rho_{AB}) = 0$ et $S(\rho_A) = S(\rho_B) = \log 2$. Ainsi, cet état pur intriqué ne contient pas d'entropie mais les parties ont une entropie maximale. Ce phénomène est général et l'entropie des parties d'un système bipartite dans un état pur est souvent appelée *entropie d'intrication* (*entanglement entropy*). L'entropie d'intrication est proprement quantique et n'est pas associée à un ensemble statistique.

Exemple: Entropie d'un qubit. Pour un qubit, la matrice densité est de la forme $\rho = \frac{1}{2}\Big(\mathbb{1} + \vec{a}\cdot\vec{\Sigma}\Big)$ où $\|\vec{a}\| \leq 1$ est un vecteur de la boule de Bloch. Il est clair que pour obtenir les valeurs propres, on peut faire une transformation unitaire qui effectue une rotation des axes tel que $\vec{a} \to (0, 0, \|\vec{a}\|)$. Les deux valeurs propres sont donc $\lambda_1 = \frac{1}{2}(1 + \|\vec{a}\|)$ et $\lambda_2 = \frac{1}{2}(1 - \|\vec{a}\|)$. L'entropie de von Neumann d'un qubit est donc

$$\begin{aligned} S(\rho) &= -\frac{1}{2}(1 + \|\vec{a}\|)\log\left(\frac{1}{2}(1 + \|\vec{a}\|)\right) - \frac{1}{2}(1 - \|\vec{a}\|)\log\left(\frac{1}{2}(1 - \|\vec{a}\|)\right) \\ &= \log(2) - \frac{1}{2}(1 + \|\vec{a}\|)\log(1 + \|\vec{a}\|) - \frac{1}{2}(1 - \|\vec{a}\|)\log(1 - \|\vec{a}\|) \end{aligned}$$

Figure 7.1 L'entropie de von Neumann d'un qubit est maximale au centre de la boule de Bloch et décroit pour s'annuler sur la surface, correspondant aux états purs (extrémaux).

On peut tenter de définir des quantités analogues à l'entropie conditionnelle et l'information mutuelle de Shannon par pure analogie formelle avec les expressions classiques. Par exemple:

$$\begin{aligned} S(\rho_A|\rho_B) &= S(\rho_{AB}) - S(\rho_B) \\ I(\rho_A;\rho_B) &= S(\rho_A) - S(\rho_A|\rho_B) \\ &= S(\rho_A) + S(\rho_B) - S(\rho_{AB}) \\ &= S(\rho_B) - S(\rho_B|\rho_A) \end{aligned}$$

Bien que ces définitions puissent s'avérer parfois utiles, leur sens opérationnel n'est pas clair à ce stade. En effet, la notion de "probabilité conditionnelle" classique n'a pas d'analogue immédiat dans le cas quantique car l'observation ou la mesure sur une partie du système change son état. Il existe en fait plusieurs définitions possibles de ce que l'on pourrait appeler des entropies conditionnelles et information mutuelle et nous ne

poursuivrons pas cette voie plus en détail ici.

Une quantité utile qui joue un rôle important est la divergence de Kullback-Leibler quantique entre deux matrices densité:

$$D(\rho\|\sigma) = \mathrm{Tr}(\rho\log\rho) - \mathrm{Tr}(\rho\log\sigma) \tag{7.6}$$

Nous verrons que $D(\rho\|\sigma) \geq 0$ et $D(\rho\|\sigma) = 0$ si et seulement si $\rho = 0$. Bien que non-symétrique, on peut penser à cette divergence comme une mesure de distance entre deux matrices densité.

7.4 Propriétés de l'entropie quantique

Nous discutons ici plusieurs inégalités analogues au cas classique, mais nous verrons aussi quelques propriétés sans analogue classique (notamment l'inégalité d'Araki-Lieb).

• Le maximum de $S(\rho)$ est atteint pour $\rho = \frac{1}{d}\mathbb{1}_d$ pour un espace de Hilbert de dimension d. En effet, prenons la décomposition spectrale de ρ. Alors, $S(\rho) = -\sum_i \lambda_i \log\lambda_i$ est maximisée pour $\lambda_i = \frac{1}{d}$ et donc pour la matrice densité $\rho = \frac{1}{d}\mathbb{1}_d$ dans la base où elle est diagonale. Mais la matrice identité est diagonale dans toutes les bases. Donc, $\rho = \frac{1}{d}\mathbb{1}_d$ maximise $S(\rho)$. De plus

$$0 \leq S(\rho) \leq \log d$$

• L'entropie de von Neumann est concave.

$$S(t\rho + (1-t)\sigma) \geq tS(\rho) + (1-t)S(\sigma) \qquad 0 \leq t \leq 1 \tag{7.7}$$

pour deux matrices densité ρ et σ. La preuve est basée sur l'inégalité de Klein:

Lemme: Soit A et B deux matrices hermitiennes et $f : \mathbb{R} \longrightarrow \mathbb{R}$ une fonction convexe. Alors

$$\mathrm{Tr}\left(f(A) - f(B) - (A-B)f'(B)\right) \geq 0$$

Preuve: Soit $A|\phi_i\rangle = a_i|\phi_i\rangle$ et $B|\psi_i\rangle = b_i|\psi_i\rangle$. La trace ci dessus vaut:

$$\begin{aligned}
&\sum_i \langle\phi_i|f(A) - f(B) - (A-B)f'(B)|\phi_i\rangle \\
&= \sum_i f(a_i) - \langle\phi_i|f(B)|\phi_i\rangle - a_i\langle\phi_i|f'(B)|\phi_i\rangle + \langle\phi_i|Bf'(B)|\phi_i\rangle \\
&= \sum_{i,j} |\langle\phi_i|\psi_j\rangle|^2\Big\{f(a_i) - f(b_j) - (a_i - b_j)f'(b_j)\Big\}
\end{aligned}$$

où on a utilisé la *relation de fermeture* $\mathbb{1} = \sum_j |\psi_j\rangle\langle\psi_j|$ et $\sum_j |\langle\phi_i|\psi_j\rangle|^2 = 1$ pour obtenir la deuxième égalité. Puisque f est convexe, $f(a_i) - f(b_j) \geq (a_i - b_j)f'(b_j)$ ce qui prouve le lemme.

Corollaire 1: Soient A et B deux matrices hermitiennes semi-définie positives. On a

$$\mathrm{Tr}(A\log A) - \mathrm{Tr}(A\log B) \geq \mathrm{Tr}(A-B)$$

Preuve: Appliquer l'inégalité de Klein à $f(x) = x\log x$, $x \geq 0$.

Corollaire 2: L'entropie relative est positive.

Preuve: Choisir $A = \rho$ et $B = \sigma$ et utiliser $\mathrm{Tr}(\rho - \sigma) = \mathrm{Tr}(\rho) - \mathrm{Tr}(\sigma) = 0$.

Pour achever de montrer que l'entropie de von Neumann est concave, on choisit $A = \rho$ et $B = t\rho + (1-t)\sigma$. Les corollaires entraine

$$\mathrm{Tr}(\rho\log\rho) - \mathrm{Tr}(\rho\log(t\rho + (1-t)\sigma)) \geq (1-t)\,\mathrm{Tr}(\rho-\sigma) = 0 \tag{7.8}$$

Ensuite, avec $A = \sigma$ et $B = t\rho + (1-t)\sigma$, on obtient

$$\mathrm{Tr}(\sigma\log\sigma) - \mathrm{Tr}(\sigma\log(t\rho + (1-t)\sigma)) \geq t\,\mathrm{Tr}(\sigma-\rho) = 0 \tag{7.9}$$

En multipliant (7.8) par t et (7.9) par $(1-t)$ puis en additionnant les inégalités, on obtient le résultat (7.7).

• Sous additivité de l'entropie quantique. La sous-additivité dans le cas classique peut se voir comme suit:

$$H(X) + H(Y) - H(X,Y) = D_{KL}(\mathbb{P}_{X,Y} \| \mathbb{P}_X \mathbb{P}_Y) \geq 0$$

Le cas quantique est parfaitement analogue.

$$\begin{aligned}
S(\rho_A) + S(\rho_B) - S(\rho_{AB}) &= -\mathrm{Tr}(\rho_A\log\rho_A) - \mathrm{Tr}(\rho_B\log\rho_B) + \mathrm{Tr}(\rho_{AB}\log\rho_{AB}) \\
&= -\mathrm{Tr}(\rho_{AB}\log\rho_A\otimes\mathbb{1}_B) - \mathrm{Tr}(\rho_{AB}\log\mathbb{1}_A\otimes\rho_B) + \mathrm{Tr}(\rho_{AB}\log\rho_{AB}) \\
&= \mathrm{Tr}(\rho_{AB}\log\rho_{AB}) - \mathrm{Tr}(\rho_{AB}(\log\rho_A\otimes\mathbb{1}_B + \log\mathbb{1}_A\otimes\rho_B)) \\
&= \mathrm{Tr}(\rho_{AB}\log\rho_{AB}) - \mathrm{Tr}(\rho_{AB}\log\rho_A\otimes\rho_B) \\
&= D(\rho_{AB}\|\rho_A\otimes\rho_B) \geq 0
\end{aligned}$$

Notons que l'identité $\log\rho_A\otimes\mathbb{1}_B + \log\mathbb{1}_A\otimes\rho_B = \log\rho_A\otimes\rho_B$ peut être vérifiée via la décomposition spectrale de ρ_A et ρ_B et que l'inégalité finale suit du corollaire 2.

• Sous-additivité forte. Considérons un système tripartite ρ_{ABC} pour l'espace de Hilbert $\mathcal{H}_A\otimes\mathcal{H}_B\otimes\mathcal{H}_C$. On considère les matrices densité réduites $\rho_{AB} = \mathrm{Tr}_{\mathcal{H}_C}(\rho_{ABC}$, $\rho_{BC} = \mathrm{Tr}_{\mathcal{H}_A}(\rho_{ABC}$ et $\rho_B = \mathrm{Tr}_{\mathcal{H}_A\otimes\mathcal{H}_C}(\rho_{ABC}$. L'inégalité de sous-additivité forte stipule

$$S(\rho_{ABC}) + S(\rho_B) \leq S(\rho_{AB}) + S(\rho_{BC})$$

Cette inégalité est analogue à la sous-additivité forte classique. Néanmoins, dans le cas quantique, la preuve est différente et aussi beaucoup plus compliquée. Il s'agit d'une des inégalités les plus profondes de la théorie de l'information quantique. Nous omettons complètement cette preuve ici et nous mentionnons juste qu'elle est basée sur la

concavité jointe de la fonctionnelle

$$f(A,B) = \mathrm{Tr}\Big(M^\dagger A^s M B^{1-s}\Big)$$

pour n'importe quelle matrice M et $0 \leq s \leq 1$. La concavité jointe de cette fonctionnelle était une conjecture de Wigner, Yanase et Dyson et fut finalement démontrée par Lieb. Plus tard, Lieb et Ruskai montrèrent qu'elle implique la sous-additivité forte.

• Inégalité de Araki-Lieb. Nous avons déjà mentionné que l'inégalité classique $H(X,Y) \geq H(X)$ n'est pas toujours vraie dans le cas quantique. Par exemple, l'entropie de l'état $|\psi\rangle_{AB} = \frac{1}{\sqrt{2}}\Big(|0\rangle_A \otimes |0\rangle_B + |1\rangle_A \otimes |1\rangle_B\Big)$ vaut $S(\rho_{AB}) = 0$ alors que $\rho_A = \frac{1}{2}\mathbb{1}_A$ et $\rho_B = \frac{1}{2}\mathbb{1}_B$ si bien que $S(\rho_A) = S(\rho_B) = \log 2$. Plus généralement, pour n'importe quel état pur $|\psi_{AB}\rangle$, le théorème de Schmidt implique que ρ_A et ρ_B possèdent les mêmes valeurs propres non-nulles et donc $S(\rho_A) = S(\rho_B)$. Ainsi, pour tout état pur, nous voyons que $S(\rho_{AB} \geq |S(\rho_A) - S(\rho_B)| = 0$. Remarquablement, cela n'est pas limité aux états purs mais est complètement général.

Soit ρ_{AB} un état mixte d'entropie $S(\rho_{AB})$. Considérons $S(\rho_A)$ et $S(\rho_B)$ les entropies des matrices densité réduites (si $S(\rho_{AB})$ n'est pas pur, en général $S(\rho_A) \neq S(\rho_B)$). L'inégalité d'Araki-Lieb stipule

$$S(\rho_{AB}) \geq |S(\rho_A) - S(\rho_B)|$$

La preuve utilise la méthode de la purification. Le système AB peut être purifié en introduisant un "environnement" C tel que $\rho_{ABC} = |\psi\rangle_{ABC}\langle\psi|_{ABC}$ est pur et $\rho_{AB} = \mathrm{Tr}_{\mathcal{H}_C}(\rho_{ABC})$. Par la sous-additivité. on a

$$S(\rho_{AC}) \leq S(\rho_A) + S(\rho_C) \tag{7.10}$$

Par le théorème de Schmidt, puisque ρ_{ABC} est pur, on a

$$S(\rho_{AB}) = S(\rho_C) \quad \text{et} \quad S(\rho_{AC}) = S(\rho_B) \tag{7.11}$$

De (7.10) et (7.11), il suit que

$$S(\rho_B) \leq S(\rho_A) + S(\rho_{AB})$$

c'est-à-dire $S(\rho_{AB}) \geq S(\rho_B) - S(\rho_A)$. Par symétrie, on a aussi $S(\rho_{AB}) \geq S(\rho_A) - S(\rho_B)$. Ceci achève la preuve de l'inégalité d'Araki-Lieb.

7.5 Principe de non-communication

Nous avons déjà vu deux impossibilités quantiques au Chapitre 3: le théorème de non-clonage et le théorème de non-discernement d'états non-orthogonaux. Ici, nous montrons brièvement une autre impossibilité: la propriété de non-communication (ou *non-signaling*).

Pour un système bipartite, une opération locale (unitaire ou de mesure) sur une partie ne peut pas être détectée par l'autre partie sans communication entre les deux parties.

Formalisons cela en considérant un système d'espace de Hilbert $\mathcal{H}_A \otimes \mathcal{H}_B$, une matrice densité ρ_{AB} et les matrices densité réduites ρ_A et ρ_B de A et B respectivement. Une opération unitaire locale dans le système A transforme l'état en

$$(U_A \otimes \mathbb{1}_B)\, \rho_{AB}\, (U_A^\dagger \otimes \mathbb{1}_B)$$

Une mesure projective locale en A projette l'état sur

$$\sigma_{AB}^i = \frac{(P_A^i \otimes \mathbb{1}_B)\, \rho_{AB}\, (P_A^i \otimes \mathbb{1}_B)}{\mathrm{Tr}\Big((P_A^i \otimes \mathbb{1}_B)\, \rho_{AB}\, (P_A^i \otimes \mathbb{1}_B)\Big)}$$

avec probabilité

$$p_i = \mathrm{Tr}\Big((P_A^i \otimes \mathbb{1}_B)\, \rho_{AB}\, (P_A^i \otimes \mathbb{1}_B)\Big) = \mathrm{Tr}\Big(P_A^i\, \rho_A\Big)$$

Les matrices densités locales de A et B après l'opération unitaire sont

$$\rho_A^{\text{après}} = \mathrm{Tr}_B \left((U_A \otimes \mathbb{1}_B)\, \rho_{AB}\, (U_A^\dagger \otimes \mathbb{1}_B)\right) = U_A \rho_A U_A^\dagger$$
$$\rho_B^{\text{après}} = \mathrm{Tr}_A \left((U_A \otimes \mathbb{1}_B)\, \rho_{AB}\, (U_A^\dagger \otimes \mathbb{1}_B)\right) = \mathrm{Tr}_A\, \rho_{AB} = \rho_B$$

Nous voyons qu'après l'opération unitaire locale chez A, la matrice densité de B est inchangée et donc l'opération chez A n'est pas détectée en B.

Après l'opération de mesure locale dans A, l'état de A est décrit par

$$\sigma_A^i = \mathrm{Tr}_B(\sigma_{AB}^i) = \frac{P_A^i \rho_A P_A^i}{\mathrm{Tr}\Big(P_A^i \rho_A\Big)}$$

avec probabilité $p_i = \mathrm{Tr}\Big(P_A^i \rho_A\Big)$. En B, le résultat de la mesure n'est pas observé donc la description est donné par le mélange statistique des états σ_B^i:

$$\sigma_B^i = \mathrm{Tr}_A(\sigma_{AB}^i) = \frac{\mathrm{Tr}_A(P_A^i \otimes \mathbb{1}_B\, \rho_{AB})}{\mathrm{Tr}\Big(P_A^i \rho_A\Big)}, \qquad p_i = \mathrm{Tr}\Big(P_A^i \rho_A\Big)$$

La matrice densité correspondant à ce mélange est

$$\sum_i p_i \sigma_B^i = \sum_i \mathrm{Tr}_A(P_A^i \otimes \mathbb{1}_B\, \rho_{AB}) = \mathrm{Tr}_A(\rho_{AB}) = \rho_B$$

car $\sum_i P_A^i = \mathbb{1}_A$. À nouveau, nous concluons que la mesure locale de A laisse la matrice densité de B inchangée et cette mesure est indétectable.

Notons que si on permet une communication entre A et B, alors A pourrait communiquer B que le résultat de la mesure locale est j (ou bien σ_A^j). Dans ce cas, B peut en déduire que son état est décrit par σ_B^j (qui est différent de ρ_B). La concavité de l'entropie implique

$$S(\rho_B) = S(\sum_i p_i \sigma_B^i) \geq \sum_i p_i S(\sigma_B^i) \geq S(\sigma_B^j)$$

Après communication entre A et B, l'entropie de B est donc réduite. Mais avant toute communication, il n'y a pas de changement d'entropie.

Exemple: Processus de téléportation. Supposons que A et B partagent une paire EPR dans l'état $\frac{1}{\sqrt{2}}\big(|00\rangle_{AB} + |11\rangle_{AB}\big)$ et A possède en plus un qubit dans l'état $|\varphi\rangle_A = \alpha|0\rangle_A + \beta|1\rangle_A$. Les matrices densité réduites de A et B sont

$$\rho_A = \frac{1}{2}|\varphi\rangle_A\langle\varphi|_A \otimes \mathbb{1}_A \qquad\qquad \rho_B = \frac{1}{2}\mathbb{1}_B$$

Pour les entropies, on a $S(\rho_{AB}) = 0$, $S(\rho_A) = S(\rho_B) = \log 2$. Lors de la mesure dans la base de Bell chez A, l'état total devient égal à

$$|B_{ij}\rangle_A\langle B_{ij}|_A \otimes |\varphi^{ij}\rangle_B\langle\varphi^{ij}|_B$$

avec probabilité $p_{ij} = \frac{1}{4}$ et $|\varphi^{ij}\rangle_B = X^iZ^j|\varphi\rangle_B$. Les matrices densité réduites deviennent

$$\rho_A^{\text{après}} = |B_{ij}\rangle_A\langle B_{ij}|_A$$

si (i, j) est observé chez A et

$$\rho_B^{\text{après}} = \frac{1}{4}\sum_{i,j=0,1}|\varphi^{ij}\rangle_B\langle\varphi^{ij}|_B = \frac{1}{2}\mathbb{1}_B$$

qui reste inchangé. Les entropies correspondantes sont $S(\rho_A^{\text{après}}) = 0$ (si (i, j) est observé) et $S(\rho_B^{\text{après}}) = \ln 2$. Pour achever le processus de téléportation, A communique (i, j) à B et l'état de B est donc décrit par $|\varphi^{ij}\rangle_B\langle\varphi^{ij}|_B$ qui possède une entropie nulle.

En résumé, avant la communication classique, lors du processus de mesure locale chez A, aucune information n'est transférée à B. Ce n'est qu'après la communication classique que l'information a été transférée et que l'entropie (≡ l'incertitude) devient nulle.

7.6 Entropie d'un mélange d'états mixtes

Dans ce paragraphe, nous prouvons le théorème suivant qui joue un rôle important en information quantique.

Théorème. Soit X une variable aléatoire de distribution $\mathbb{P}_X[X = x] = p_x$ et la matrice densité $\rho = \sum_x p_x\rho_x$ où les ρ_x sont eux-mêmes des états mixtes. On a

$$S(\rho) \leq \sum_x p_xS(\rho_x) + H(X)$$

L'interprétation de cette inégalité est claire. L'incertitude sur ρ ne peut pas excéder l'incertitude moyenne de chaque état mixte plus l'incertitude sur la préparation du mélange des ρ_x (cette préparation est décrite par X). En particulier, si $\rho_x = |\varphi_x\rangle\langle\varphi_x|$

sont des états purs, alors $S(\rho) \leq H(X)$ comme discuté dans la remarque (Section **7.3**).

Preuve. Tout d'abord, on considère le cas d'un mélange d'états purs

$$\rho_S = \sum_x p_x |\phi_x\rangle_S \langle\phi_x|_S$$

Purifions ce système en introduisant un environnement $\mathcal{H}_\mathcal{E}$ de dimension égal au nombre d'états purs $|\phi_x\rangle_S$. Soit $|x\rangle_\mathcal{E}$ les états d'une base orthonormée de $\mathcal{H}_\mathcal{E}$. L'état purifié est

$$|\Psi\rangle_{S\mathcal{E}} = \sum_x \sqrt{p_x} |\phi_x\rangle_S \otimes |x\rangle_\mathcal{E}$$

On vérifie facilement que $\mathrm{Tr}_\mathcal{E}(|\Psi\rangle\langle\Psi|_{S\mathcal{E}}) = \rho_S$. De plus,

$$\begin{aligned} \rho_\mathcal{E} &= \mathrm{Tr}_S(|\Psi\rangle\langle\Psi|_{S\mathcal{E}}) \\ &= \sum_{x,x'} \sqrt{p_x}\sqrt{p'_x}\,\langle\phi_x|\phi_{x'}\rangle_S\, |x\rangle\langle x'|_\mathcal{E} \end{aligned} \tag{7.12}$$

D'après le théorème de Schmidt, $S(\rho_S) = S(\rho_\mathcal{E})$. Considérons maintenant un état $\rho'_\mathcal{E} = \sum_x p_x |x\rangle\langle x'|_\mathcal{E}$. L'entropie relative

$$D(\rho_\mathcal{E}||\rho'_\mathcal{E}) = \mathrm{Tr}\Big(\rho_\mathcal{E} \log \rho_\mathcal{E}\Big) - \mathrm{Tr}\Big(\rho_\mathcal{E} \log \rho'_\mathcal{E}\Big) \geq 0$$

Donc

$$S(\rho_S) = S(\rho_\mathcal{E}) \leq -\mathrm{Tr}\Big(\rho_\mathcal{E} \log \rho'_\mathcal{E}\Big)$$

Puisque $\log \rho'_\mathcal{E} = \sum_x \log(p_x) |x\rangle\langle x|_\mathcal{E}$, cette inégalité devient

$$\begin{aligned} S(\rho_S) &\leq -\sum_x \log(p_x)\,\mathrm{Tr}(\rho_\mathcal{E}|x\rangle\langle x|_\mathcal{E}) \\ &= -\sum_x \log(p_x)\langle x|\rho_\mathcal{E}|x\rangle_\mathcal{E} \end{aligned}$$

Il suit de (7.12) que $\langle x|\rho_\mathcal{E}|x\rangle_\mathcal{E} = p_x$ et on trouve bien

$$S(\rho_S) \leq H(X)$$

Pour étendre cette inégalité aux mélanges d'états mixtes $\rho = \sum_x p_x \rho_x$, on utilise la décomposition spectrale $\rho_x = \sum_j \lambda_j^x |e_j^x\rangle\langle e_j^x|$ si bien que

$$\rho = \sum_{x,j} p_x \lambda_j^x |e_j^x\rangle\langle e_j^x|$$

Par le résultat précédent

$$\begin{aligned} S(\rho) &\leq -\sum_{x,j} p_x \lambda_j^x \log\left(p_x \lambda_j^x\right) \\ &= -\sum_{x,j} p_x \lambda_j^x \log(p_x) - \sum_{x,j} p_x \lambda_j^x \log\left(\lambda_j^x\right) \\ &= -\sum_{x} p_x \log(p_x) - \sum_{x} p_x \sum_{j} \lambda_j^x \log\left(\lambda_j^x\right) \\ &= H(X) + \sum_x p_x S(\rho_x) \end{aligned}$$

7.7 Information accessible et borne de Holevo

Dans ce paragraphe, nous considérons la question de l'information acquise lors d'une mesure d'un état de mélange $\rho = \sum_x p_x \rho_x$. Nous imaginons que cet état provient de la préparation d'un ensemble statistique d'états ρ_x chacun pris avec probabilité p_x. Soit X la variable aléatoire décrivant cette préparation $\mathbb{P}_{\mathrm{X}}[X = x] = p_x$. Lors d'une mesure projective dans une base décrite par un ensemble de projecteurs P_y, l'état est projeté sur $P_y \rho P_y$ avec probabilité $\mathrm{Tr}\left(P_y \rho\right) = q_y$. Cela définit une variable aléatoire Y avec $\mathbb{P}_{\mathrm{Y}}[Y = y] = q_y$. On peut aussi définir la probabilité conditionnelle $\mathbb{P}_{\mathrm{Y|X}}[Y = y|X = x] = p_{y|x} = \mathrm{Tr}\left(P_y \rho_x\right)$ et une probabilité jointe $\mathbb{P}_{\mathrm{X,Y}}[X = x, Y = y] = p_x \mathrm{Tr}\left(P_y \rho_x\right)$. On vérifie que ces définitions sont consistantes avec les contraintes de marginalisation.

L'information contenue dans les mesures Y à propos de la préparation X est donnée par $I(X;Y) = H(X,Y) - H(X) - H(Y)$ l'information mutuelle entre X et Y.

L'information accessible à propos de la préparation de ρ est définie par le maximum de $I(X;Y)$ sur l'ensemble de toutes les mesures possibles de l'état.

$$\mathrm{Acc}\left(\{p_x, \rho_x\}\right) = \sup_{\{P_y\}} I(X;Y)$$

C'est une fonctionnelle de la préparation $\{p_x, \rho_x\}$ car l'unique connaissance de ρ ne suffit pas à calculer cette quantité. Son calcul requiert la solution d'un problème d'optimisation qui peut être difficile à résoudre.

Il existe une inégalité importante, la borne de Holevo, qui est universelle et joue un rôle important en théorie de l'information.

Théorème (Borne de Holevo).

On définit la quantité de Holevo

$$\chi\left(\{p_x, \rho_x\}\right) \equiv S(\rho) - \sum_x p_x S(\rho_x)$$

Alors, l'information accessible satisfait

$$\mathrm{Acc}\Big(\{p_x, \rho_x\}\Big) \leq \chi\Big(\{p_x, \rho_x\}\Big)$$

Preuve. Soit $\mathcal{H}_S$ l'espace de Hilbert de la matrice densité ρ. On introduit un espace de Hilbert $\mathcal{H}_X \otimes \mathcal{H}_S \otimes \mathcal{H}_Y$ avec $\Big\{|x\rangle\Big\}$ une base orthonormale de $\mathcal{H}_X$, $\Big\{|y\rangle\Big\}$ une base orthonormale de $\mathcal{H}_Y$ et un état $\rho_{XSY} = \sum_x p_x |x\rangle\langle x| \otimes \rho \otimes |0\rangle\langle 0|$. Soit $U_{XSY} = \mathbb{1}_X \otimes U_{SY}$ l'opération unitaire définie par

$$U_{SY}|\phi\rangle_S \otimes |a\rangle_Y = \sum_y P_y|\phi\rangle_S \otimes |a \oplus y\rangle_Y$$

où $a \oplus y$ est calculé (modulo $\dim(\mathcal{H}_Y)$). On peut montrer que U_{SY} est bien unitaire car le produit scalaire est préservé. Soit également $\rho'_{XSY} = U_{XSY}\rho_{XSY}U^\dagger_{XSY}$. Par conséquent, les deux matrices densité ρ_{XSY} et ρ'_{XSY} possèdent les mêmes valeurs propres, donc $S(\rho_{XSY}) = S(\rho'_{XSY})$.

Considérons maintenant les matrices densité réduites ρ_{SY} et ρ'_{SY}.

$$\begin{aligned}
\rho_{SY} &= \mathrm{Tr}_{\mathcal{H}_X}\left(\sum_x p_x |x\rangle\langle x| \otimes \rho \otimes |0\rangle\langle 0|\right) \\
&= \rho \otimes |0\rangle\langle 0| \\
\rho'_{SY} &= \mathrm{Tr}_{\mathcal{H}_X}\left(\sum_{x,y,y'} p_x |x\rangle\langle x| \otimes P_y \rho P'_y \otimes |y\rangle\langle y'|\right) \\
&= P_y \rho P'_y |y\rangle\langle y'|
\end{aligned}$$

Ces deux matrices densité sont aussi unitairement reliées par U_{SY} et donc $S(\rho_{SY}) = S(\rho'_{SY})$.

Par sous-additivité forte,

$$S(\rho'_{XSY}) - S(\rho'_{SY}) \leq S(\rho'_{XY}) - S(\rho'_Y)$$

et donc

$$S(\rho_{XSY}) - S(\rho_{SY}) \leq S(\rho'_{XY}) - S(\rho'_Y) \tag{7.13}$$

Le reste de la preuve est un calcul de chaque entropie dans cette inégalité. Clairement

$$S(\rho_{XSY}) = H(X) + \sum_x p_x S(\rho_x)$$

car $|0\rangle\langle 0|$ ne contient pas d'entropie et les états $|x\rangle$ sont orthogonaux. D'autre part

$$S(\rho_{SY}) = S(\rho)$$

puisque $|0\rangle\langle 0|$ ne contient pas d'entropie. De plus,

$$\begin{aligned}\rho'_{XY} &= \sum_{x,y,y'} p_x |x\rangle\langle x| \otimes |y\rangle\langle y'| \operatorname{Tr}_{\mathcal{H}_S}(P_y \rho_x P_{y'}) \\ &= \sum_{x,y,y'} p_x |x\rangle\langle x| \otimes |y\rangle\langle y'| \delta_{yy'} p_{y|x} \\ &= \sum_{x,y} p_x p_{y|x} |x\rangle\langle x| \otimes |y\rangle\langle y|\end{aligned}$$

car $\operatorname{Tr}_{\mathcal{H}_S}(P_y \rho_x P_{y'}) = \operatorname{Tr}_{\mathcal{H}_S}(P_{y'} P_y \rho) = \delta_{yy'} \operatorname{Tr}_{\mathcal{H}_S}(P_y \rho)$. Donc,

$$S(\rho'_{XY}) = H(X, Y)$$

Finalement, $\rho'_Y = \sum_{y,x} p_{x,y} |y\rangle\langle y| = \sum_y p_y |y\rangle\langle y|$ si bien que

$$S(\rho'_Y) = H(Y)$$

En mettant tous ces résultats dans (7.13), on obtient

$$I(X;Y) \leq S(\rho) - \sum_x p_x S(\rho_x)$$

Ceci est valable pour toute mesure projective $\{P_y\}$ et on peut maximiser le membre de gauche sur l'ensemble des mesures projectives.

Exemple

Soit $\rho = \frac{1}{2}|0\rangle\langle 0| + \frac{1}{2}\left(\frac{|0\rangle+|1\rangle}{\sqrt{2}} \frac{\langle 0|+\langle 1|}{\sqrt{2}}\right) = \frac{1}{4}\begin{pmatrix} 3 & 1 \\ 1 & 1 \end{pmatrix}$.

Les valeurs propres de ρ sont $\frac{1}{2} \pm \frac{\sqrt{2}}{4}$ et la borne de Holevo est, dans ce cas, égale à $S(\rho) = -\left(\frac{1}{2} + \frac{\sqrt{2}}{4}\right)\log\left(\frac{1}{2} + \frac{\sqrt{2}}{4}\right) - \left(\frac{1}{2} - \frac{\sqrt{2}}{4}\right)\log\left(\frac{1}{2} - \frac{\sqrt{2}}{4}\right) = 0.59 \ln(2)$. Ainsi, aucune mesure ne permettra d'obtenir plus de 0.59 bits d'information. Considérons la mesure dans la base $\left\{|0\rangle\langle 0|, |1\rangle\langle 1|\right\}$. On obtient

$$p_y = \begin{pmatrix} \frac{3}{4} & \frac{1}{4} \end{pmatrix}, \qquad p_{y|x} = \begin{pmatrix} 1 & \frac{1}{2} \\ 0 & \frac{1}{2} \end{pmatrix}, \qquad p_{x,y} = \begin{pmatrix} \frac{1}{2} & \frac{1}{4} \\ 0 & \frac{1}{4} \end{pmatrix}$$

Un calcul donne $I(X;Y) = \frac{3}{2}\log(2) - \frac{3}{4}\log(3) = 0.31 \ln(2)$. Donc, on obtient 0.31 bits d'informations avec cette mesure.

8 Opérations locales avec Communication Classique

Le processus de téléportation étudié dans le Chapitre **6** est un exemple de protocole basé sur une classe d'opérations locales avec communication classique (LOCC). Alice et Bob partagent une paire de Bell et effectuent uniquement des opérations unitaires et/ou mesures locales et ne peuvent communiquer que via un canal classique. Il est naturel d'étudier cette classe d'opération pour des systèmes plus généraux, le but ultime étant d'établir des protocoles de communication quantique dans des réseaux multipartites partageant des ressources intriquées.

Dans ce chapitre, nous introduisons quelques bases de ce vaste sujet. Nous commençons par discuter un exemple simple qui permet de dégager certains aspects essentiels comme les notions de dilution et distillation d'états intriqués. Ensuite, nous formalisons la classe d'opérations locales avec communication classique (LOCC) ainsi que la version stochastique (SLOCC) puis discutons les protocoles optimaux de distillation et dilution en nous basant sur les notions d'entropie de von Neumann vues au Chapitre **7**. Finalement, nous discutons comment les opérations SLOCC permettent une classification des ressources intriquées en classes d'équivalence, notamment pour les états à deux et trois qubits.

8.1 Transformations de dilution

Un exemple simple

Considérons le problème suivant. Alice et Bob partagent deux qubits dans l'état intriqué de Bell $|B_{00}\rangle = \frac{1}{\sqrt{2}}\big(|00\rangle + |11\rangle\big)$. Ils veulent convertir cet état en un état "moins" intriqué de la forme $|\psi_\theta\rangle = \cos\theta\,|00\rangle + \sin\theta\,|11\rangle$ avec $\theta \neq \frac{\pi}{4}$. Cet état est moins intriqué au sens où l'entropie d'intrication des matrices densités réduites est inférieure à $\ln 2$ si $\theta \neq \frac{\pi}{4}$. Alice et Bob partagent uniquement un canal de communication classique et sont physiquement distants, si bien que les seules opérations possibles (unitaires ou mesures) sont locales.

Alice fait une mesure de type POVM[1] avec

$$M_0 = \begin{pmatrix} \cos\theta & 0 \\ 0 & \sin\theta \end{pmatrix} \quad \text{et} \quad M_1 = \begin{pmatrix} \sin\theta & 0 \\ 0 & \cos\theta \end{pmatrix}$$

On note que $\{M_0, M_1\}$ est un POVM car $M_0^\dagger M_0 + M_1^\dagger M_1 = \mathbb{1}$. L'état résultant après la mesure est proportionnel à l'un des deux états

$$M_0 \otimes \mathbb{1}|B_{00}\rangle = \frac{1}{\sqrt{2}}\Big(\cos\theta\,|00\rangle + \sin\theta\,|11\rangle\Big)$$
$$M_1 \otimes \mathbb{1}|B_{00}\rangle = \frac{1}{\sqrt{2}}\Big(\sin\theta\,|00\rangle + \cos\theta\,|11\rangle\Big)$$

avec probabilités

$$p_0 = \langle B_{00}|(M_0 \otimes \mathbb{1})^\dagger (M_0 \otimes \mathbb{1})|B_{00}\rangle = \frac{1}{2}$$
$$p_1 = \langle B_{00}|(M_1 \otimes \mathbb{1})^\dagger (M_1 \otimes \mathbb{1})|B_{00}\rangle = \frac{1}{2}$$

Si Alice obtient le résultat $i = 0, 1$, elle applique l'opération unitaire locale $X^i \otimes \mathbb{1}$ (donc pour $i = 0$, elle ne fait rien et pour $i = 1$, elle applique X à son qubit). Ensuite, elle communique le bit d'information classique $i = 0, 1$ à Bob et celui-ci applique l'opération unitaire locale $\mathbb{1} \otimes X^i$. En fin de compte, l'état résultant de ce protocole est

$$(\mathbb{1} \otimes X^i)(X^i \otimes \mathbb{1})(M_i \otimes \mathbb{1})|B_{00}\rangle \propto |\psi_\theta\rangle \quad \text{pour } i = 0, 1$$

Alice et Bob ont donc bien effectué la transformation $|B_{00}\rangle \longrightarrow |\psi_\theta\rangle$ grâce à un protocole LOCC.

Un théorème caractérisant les transformations de dilution

La transformation ci-dessus est parfaite et déterministe au sens où le résultat sera toujours $|\psi_\theta\rangle$ et ce n'est pas un hasard. En effet, cela suit d'un théorème général (énoncé ici sans démonstration).

[1] Les Positive Operator Valued Measurement (POVM) généralisent les mesures projectives usuelles. Un POVM décrit la situation physique la plus générale où la mesure est faite sur le système et son environnement. Mathématiquement, $\{M_0, \ldots, M_n\}$ est un POVM si $\sum_i M_i^\dagger M_i = \mathbb{1}$. L'état du système (sans son environnement) est $M_i|\psi\rangle$ (à la normalisation près) avec probabilité $\langle\psi|M_i^\dagger M_i|\psi\rangle$.

Théorème

Soit $|\phi\rangle$ et $|\psi\rangle$ deux états bipartites de l'espace de Hilbert $\mathcal{H}_A \otimes \mathcal{H}_B$ avec $\dim \mathcal{H}_A = \dim \mathcal{H}_B = d$. Soient $\vec{\lambda_\phi} = \left\{\lambda_\phi^1, \ldots, \lambda_\phi^d\right\}$ et $\vec{\lambda_\psi} = \left\{\lambda_\psi^1, \ldots, \lambda_\psi^d\right\}$ l'ensemble des valeurs propres des matrices densité réduites $\rho_\phi^A = \mathrm{Tr}_{\mathcal{H}_B}|\phi\rangle\langle\phi|$ et $\rho_\psi^B = \mathrm{Tr}_{\mathcal{H}_A}|\psi\rangle\langle\psi|$ respectivement.

Si le vecteur $\vec{\lambda_\phi}$ est "dégradé" par rapport au vecteur $\vec{\lambda_\psi}$, au sens où $\vec{\lambda_\phi} = P\vec{\lambda_\psi}$ avec P une matrice doublement stochastique, alors on peut effectuer la transformation $|\phi\rangle \longrightarrow |\psi\rangle$ par un protocole LOCC. La réciproque est aussi vraie.

Remarque. Une matrice doublement stochastique est une matrice telle que ses éléments sont dans $[0, 1]$, la somme des éléments de chaque ligne vaut 1 et la somme des éléments de chaque colonne vaut aussi 1. En particulier, P (et aussi P^T) est la matrice de transition d'une chaîne de Markov.

Il n'est pas très difficile de montrer que $\vec{\lambda_\phi} = P\vec{\lambda_\psi}$ implique l'inégalité au niveau des entropies des distributions de probabilités $S(\rho_\phi^A) \geq S(\rho_\psi^B)$. En d'autres termes, $\vec{\lambda_\phi}$ est une version "plus aléatoire" de $\vec{\lambda_\psi}$. Intuitivement, l'état $|\phi\rangle$ peut être transformé en $|\psi\rangle$ car $|\phi\rangle$ est plus intriqué que $|\psi\rangle$.

Entropie de dilution

Notons que pour tout état $|\psi\rangle \in \mathbb{C}^2 \otimes \mathbb{C}^2$, la transformation $|B_{00}\rangle \longrightarrow |\psi\rangle$ peut être effectuée de façon certaine par un protocole LOCC. En effet, d'après le théorème précédent, il suffit de vérifier qu'il existe une matrice doublement stochastique telle que $\lambda_{|B_{00}\rangle} = P\vec{\lambda_\psi}$. Or, $\lambda_{|B_{00}\rangle} = (\frac{1}{2}, \frac{1}{2})^T$, donc la matrice $P = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}$ fait toujours l'affaire[2].

La prochaine question naturelle est donc: étant donné n paires de Bell $|B_{00}\rangle$, combien d'états $|\psi\rangle$ peut-on produire? On peut clairement en produire au moins $m = n$. Mais peut-on en produire plus que n? En d'autres termes, si on pense à une paire de Bell comme à une "monnaie d'intrication", la question est combien "d'états $|\psi\rangle$ vaut-elle?" Une représentation de cette question est montré à la Figure **8.1**. Notez que pour se poser cette question, il est nécessaire qu'Alice et Bob aient des qubits auxiliaires.

Le processus de transformation $|B_{00}\rangle^{\otimes n}$ en $|\psi\rangle^{\otimes m}$ pour $m \geq n$ s'appelle la "dilution des paires de Bell". On définit "l'entropie de dilution" de $|\psi\rangle$ comme

$$E_{\text{dilution}}(|\psi\rangle) = \sup\left\{\frac{n}{m} \;\middle|\; |B_{00}\rangle^{\otimes n} \longrightarrow |\psi\rangle^{\otimes m} \quad \text{via un LOCC}\right\}$$

[2] Ceci est vrai car ρ_ψ^B est une matrice 2×2 dont les valeurs propres sont $\lambda_\psi^{1/2} \geq 0$ (car semi-définie positive) et $\mathrm{Tr}\,\rho_\psi^B = \lambda_\psi^1 + \lambda_\psi^2 = 1$.

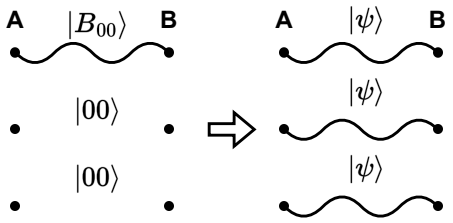


Figure 8.1 Représentation de la dilution de $|B_{00}\rangle$ en plusieurs états $|\psi\rangle$.

On peut démontrer que le supremum est atteint pour n et m qui tendent vers $+\infty$ et que $E_{\text{dilution}}(|\psi\rangle) = S(\rho_\psi^A) = S(\rho_\psi^B)$; l'entropie de dilution de $|\psi\rangle$ est égale à son entropie d'intrication.

8.2 Transformation de distillation

Un exemple simple

Nous discutons maintenant le processus inverse. Supposons qu'Alice et Bob partagent l'état bipartite

$$|\psi_\theta\rangle = \cos\theta\,|00\rangle + \sin\theta\,|11\rangle$$

Pour $\theta \neq \frac{\pi}{4}$, cet état n'est pas maximalement intriqué. Alice et Bob aimeraient améliorer cette intrication et même fabriquer une paire parfaite $|B_{00}\rangle$ en utilisant uniquement des opérations locales et de la communication classique. Nous allons voir que cette fois, la transformation $|\psi_\theta\rangle \longrightarrow |B_{00}\rangle$ ne peut qu'avoir lieu avec une probabilité inférieure à 1. Notons déjà que si $\theta = 0$ ou $\theta = \frac{\pi}{2}$, $|\psi_\theta\rangle$ est l'état produit $|00\rangle$ ou $|11\rangle$ respectivement et clairement des opérations locales ne peuvent pas produire d'intrication. Dans ce cas, la probabilité de succès du processus est nulle. En effet, les opérations locales sont de la forme d'un produit tensoriel et ne peuvent pas transformer un état produit en autre chose qu'état produit.

Alice effectue d'abord une mesure avec un POVM $\{M_0, M_1\}$ comme avant, mais cette fois-ci sur l'état $|\psi_\theta\rangle$. Le résultat est un des deux états proportionnel à

$$M_0 \otimes \mathbb{1}|\psi_\theta\rangle = \cos^2\theta\,|00\rangle + \sin^2\theta\,|11\rangle$$
$$M_1 \otimes \mathbb{1}|\psi_\theta\rangle = \sin\theta\,\cos\theta\,\big(|00\rangle + |11\rangle\big)$$

avec les probabilités

$$p_0 = \langle\psi_\theta|(M_0 \otimes \mathbb{1})^\dagger(M_0 \otimes \mathbb{1})|\psi_\theta\rangle = \cos^4\theta + \sin^4\theta$$
$$p_1 = \langle\psi_\theta|(M_1 \otimes \mathbb{1})^\dagger(M_1 \otimes \mathbb{1})|\psi_\theta\rangle = 2(\sin\theta\,\cos\theta)^2$$

Ainsi, l'état $|B_{00}\rangle$ survient avec probabilité $2(\sin\theta\ \cos\theta)^2$ qui est la probabilité de succès. Dans ce cas, Alice communique le bit classique d'information $i = 1$ à Bob. Si c'est l'état $\cos^2\theta\,|00\rangle + \sin^2\theta\,|11\rangle$ qui survient, Alice communique le bit d'information classique $i = 0$ à Bob et ils ignorent le résultat.

Pour $\theta = 0$ ou $\frac{\pi}{2}$, on voit bien que la probabilité de succès est nulle. Pour $\theta = \frac{\pi}{4}$, en revanche, l'état de départ était déjà $|B_{00}\rangle$ et la probabilité de succès est de 1 (i.e. les deux états obtenables après avoir mesuré avec le POVM restent $|B_{00}\rangle$).

Contrairement à la dilution, le protocole utilisé pour la distillation réussi avec une probabilité inférieure à un. Dans le cas de protocoles dont le résultat final n'est pas déterministe (par exemple la distillation), on parle de protocole SLOCC (stochastique LOCC).[3]

Entropie de distillation:

Étant donné un certain nombre m d'états $|\psi\rangle \in \mathbb{C}^2 \otimes \mathbb{C}^2$, combien de paires de Bell (disons n) peut-on fabriquer grâce aux opérations décrites ci-dessus? Comme nous l'avons vu, le protocole réussit avec une probabilité inférieure à un (en général). Par exemple, dans l'exemple précédent, la probabilité de réussite est $2(\sin\theta\ \cos\theta)^2$. Ainsi, pour m et n grands, ce protocole simple permet d'extraire environ n paires de Bell à partir de $m = \frac{n}{2(\sin\theta\ \cos\theta)^2}$ états $|\psi_\theta\rangle$.

L'opération de transformation $|\psi_\theta\rangle^{\otimes m} \longrightarrow |B_{00}\rangle^{\otimes n}$ s'appelle *la distillation de paires de Bell*. On définit *l'entropie de distillation* d'un état $|\psi\rangle$ par

$$E_{\text{distillation}}(|\psi\rangle) = \sup\left\{\frac{n}{m} \;\middle|\; |\psi\rangle^{\otimes m} \longrightarrow |B_{00}\rangle^{\otimes n} \quad \text{via un LOCC}\right\}$$

Nous ne formalisons pas la définition complète ici, mais il faut ajouter que l'on demande que la probabilité de succès soit $\geq 1 - \epsilon$, puis on prend $\epsilon \to 0$. On démontre que le supremum est atteint pour n et m qui tendent vers l'infini et que $E_{\text{distillation}}(|\psi\rangle) = S(\rho_\psi^A) = S(\rho_\psi^B)$. Ainsi, l'entropie de distillation est donnée par l'entropie d'intrication.

Pour conclure ce paragraphe, remarquons que les entropies de dilution et de distillation sont égales pour un état pur. Cela n'est pas évident à priori. En effet, nous avons vu qu'il y a une asymétrie dans la probabilité de succès de la transformation $|\psi_\theta\rangle \longrightarrow |B_{00}\rangle$ (distillation) et $|B_{00}\rangle \longrightarrow |\psi_\theta\rangle$ (dilution).

[3] Notez que les deux protocoles font intervenir une mesure quantique, qui n'est pas déterministe. Cependant, le processus de dilution en lui-même est déterministe, car il réussit avec probabilité 1 (c'est donc un LOCC) tandis que le processus de distillation n'est pas déterministe, car il a un taux d'échec (c'est donc un SLOCC).

8.3 Protocoles optimaux de dilution et distillation

Reprenons le problème des transformations $|B_{00}\rangle^{\otimes n} \longrightarrow |\psi\rangle^{\otimes m}$ et $|\psi\rangle^{\otimes m} \longrightarrow |B_{00}\rangle^{\otimes n}$ par des opérations de la classe SLOCC. Nous appliquons ici des méthodes usuelles de théorie de l'information pour montrer les affirmations précédentes et notamment que les entropies de dilution et de distillations sont égales à l'entropie d'intrication $S(\rho_\psi^A) = S(\rho_\psi^B)$.

Cela nécessite de faire appel à la notion de *typicalité* en théorie de l'information classique et nous faisons donc une digression à ce propos.

Typicalité en théorie de l'information classique

Soient $x_1, x_2, \ldots, x_n$ n copies indépendantes et identiquement distribuées d'une variable aléatoire de distribution $p(x)$, $x \in \mathcal{X}$ où $\mathcal{X}$ est un alphabet fini (par exemple $\mathcal{X} = \{0, 1\}$).

Définition

Étant donné $\epsilon > 0$, une suite $x_1, x_2, \ldots, x_n$ est dite ϵ-typique si

$$\left| \frac{1}{n} \sum_{k=1}^{n} \log p\left(\frac{1}{x_k}\right) - H(X) \right| \leq \theta \tag{8.1}$$

En d'autres termes, une suite de symboles émis par la source $X \sim \mathbb{P}_X$ est typique si son entropie empirique est proche de l'entropie de Shannon $H(X) = \frac{1}{n} \sum_{x \in \mathcal{X}} p(x) \log p\left(\frac{1}{x}\right)$.

Théorème des séquences typiques

Soit $\epsilon > 0$ et $x_1, \ldots, x_n \overset{\text{i.i.d}}{\sim} \mathbb{P}_X$. Soit $T_{n,\epsilon}$ l'ensemble des séquences ϵ-typiques. Pour tout $\delta > 0$, on peut prendre n assez grand tel que (a)

$$\mathbb{P}\left[T_{n,\epsilon}\right] \geq 1 - \delta$$

et

$$(1 - \delta) 2^{n(H(X) - \epsilon)} \leq \left| T_{n,\epsilon} \right| \leq 2^{n(H(X) + \epsilon)}$$

[a] Ici, $\mathbb{P}\left[T_{n,\epsilon}\right]$ veut dire $\mathbb{P}\left[x_1, \ldots x_n \in T_{n,\epsilon}\right] = \sum_{\substack{x_1, \ldots, x_n \\ \in T_{n,\epsilon}}} p(x_1) \ldots p(x_n)$.

Le théorème affirme que pour tout $\delta > 0$, on peut toujours prendre un n (suffisamment grand, selon le δ choisi initialement) tel qu'une proportion $1 - \delta$ des séquences est typiques et que le nombre de telles séquences est d'environ $2^{nH(X)}$. Ce théorème est à la base du fameux théorème de compression de l'information de Shannon qui affirme qu'il

suffit de $nH(X)$ bits pour représenter les suites typiques $x_1, \ldots, x_n$ émises par la source $X \sim \mathbb{P}_X$.

Preuve du Théorème

La preuve est une application de la loi des grands nombres. Comme $x_1, x_2, \ldots, x_n$ sont indépendants et identiquement distribués avec $\mathbb{P}_X$, la somme $-\frac{1}{n}\sum_{i=1}^{n} \log \mathbb{P}_X(x = x_i)$ se concentre sur sa moyenne qui n'est autre que $H(X)$. Plus formellement, la loi des grands nombres stipule

$$\mathbb{P}\left[\left|-\frac{1}{n}\sum_{i=1}^{n} \log \mathbb{P}_X(x = x_i) - H(X)\right| \leq \epsilon\right] \geq 1 - \epsilon$$

pour n assez grand.

Pour obtenir la borne sur $\left|T_{n,\epsilon}\right|$, on remarque

$$\begin{aligned}\mathbb{P}\left[T_{n,\epsilon}\right] &= \sum_{\substack{x_1,\ldots,x_n \\ \in T_{n,\epsilon}}} p(x_1)\ldots p(x_n) \\ &= \sum_{\substack{x_1,\ldots,x_n \\ \in T_{n,\epsilon}}} 2^{-\sum_{i=1}^{n} \log \frac{1}{p(x_i)}}\end{aligned}$$

Or, la définition (8.1) des séquences typiques implique que

$$\left|T_{n,\epsilon}\right|2^{-n(H(X)+\epsilon)} \leq \mathbb{P}\left[T_{n,\epsilon}\right] \leq \left|T_{n,\epsilon}\right|2^{-n(H(X)-\epsilon)} \tag{8.2}$$

De la première inégalité de (8.2), il suit

$$\left|T_{n,\epsilon}\right|2^{-n(H(X)+\epsilon)} \leq \mathbb{P}\left[T_{n,\epsilon}\right] \leq 1$$

et donc

$$\left|T_{n,\epsilon}\right| \leq 2^{n(H(X)+\epsilon)}$$

De la seconde inégalité de (8.2), il suit

$$1 - \delta \leq \mathbb{P}\left[T_{n,\epsilon}\right] \leq \left|T_{n,\epsilon}\right|2^{-n(H(X)-\epsilon)}$$

et donc

$$\left|T_{n,\epsilon}\right| \geq (1-\delta)2^{n(H(X)-\epsilon)}$$

Nous sommes maintenant prêt à prouver les affirmations énoncées plus haut sur les entropies de dilution et distillation.

Protocole optimal de dilution

Nous prenons n copies de $|B_{00}\rangle$ et voulons les transformer en m copies de $|\psi\rangle$ (si $m > n$, les qubits supplémentaires nécessaires sont fournis à Alice et Bob). Nous décrivons ici

un protocole SLOCC qui maximise $\frac{n}{m}$ pour $n, m \to \infty$.

Supposons que $|\psi\rangle$ possède la décomposition de Schmidt

$$|\psi\rangle = \sum_x \sqrt{p(x)}|x\rangle_A \otimes |x\rangle_B$$

Ici (dans le cadre le plus simple possible), $x \in \{0, 1\}$. Nous utilisons la notation $\sqrt{p(x)} \in \left\{\sqrt{p(0)}, \sqrt{p(1)}\right\}$ pour les coefficients de Schmidt car $p(x) \in \{p_0, p_1\}$ sera interprété comme une distribution de probabilité.

L'état que l'on veut fabriquer par des opérations SLOCC est

$$|\psi\rangle^{\otimes m} = \sum_{\substack{x_1,\ldots,x_n \\ \in \{0,1\}^n}} \sqrt{p(x_1)\ldots p(x_n)}\, |x_1 x_2 \ldots x_n\rangle_A \otimes |x_1 x_2 \ldots x_n\rangle_B \tag{8.3}$$

Nous allons considérer une approximation de cet état. C'est cette approximation qui sera en fait fabriquée. Soit

$$|\varphi_m\rangle = \sum_{\substack{x_1,\ldots,x_n \\ \in T_{n,\epsilon}}} \sqrt{p(x_1)\ldots p(x_n)}\, |x_1 x_2 \ldots x_n\rangle_A \otimes |x_1 x_2 \ldots x_n\rangle_B$$

où la somme porte sur les séquences typiques associées à la distribution $p(x)$. La norme de ce vecteur est

$$\langle\varphi_m|\varphi_m\rangle = \sum_{\substack{x_1,\ldots,x_n \\ \in T_{n,\epsilon}}} p(x_1)\ldots p(x_n) = \mathbb{P}\left[T_{n,\epsilon}\right]$$

Soit $|\varphi'_m\rangle = \frac{|\varphi_m\rangle}{\sqrt{\langle\varphi_m|\varphi_m\rangle}}$ le vecteur d'état normalisé. On remarque que

$$\begin{aligned}\langle\varphi'_m|\psi\rangle^{\otimes m} &= \frac{\sum_{\substack{x_1,\ldots,x_n \\ \in T_{n,\epsilon}}} p(x_1)\ldots p(x_n)}{\sqrt{\langle\varphi_m|\varphi_m\rangle}} \\ &= \sqrt{\mathbb{P}\left[T_{n,\epsilon}\right]} \geq \sqrt{1-\delta}\end{aligned}$$

Ainsi, on peut prendre m suffisamment grand pour que $|\varphi'_m\rangle$ soit une très bonne approximation de $|\psi\rangle^{\otimes m}$. Pour $m \to +\infty$, on peut choisir $\delta \to 0$ et $\left\||\varphi'_m\rangle - |\psi\rangle^{\otimes m}\right\|_2 \to 0$.

De plus, le nombre de termes de la base computationnelle dans la superposition $|\varphi'_m\rangle$ est $\left|T_{m,\epsilon}\right| \leq 2^{m(H(p(x))+\epsilon)}$. On remarque aussi que $H(p(x)) = S(\rho^A_\psi) = S(\rho^B_\psi)$. Ainsi, $|\varphi'_m\rangle$ est décrit par $m\, S(\rho^A_\psi)$ (ou bien $m\, S(\rho^B_\psi)$) qubits chez Alice (et Bob).

Nous décrivons maintenant comment Alice et Bob fabriquent $|\varphi'_m\rangle$ à partir de paires de Bell. Supposons qu'ils possèdent $n = m\, S(\rho^A_\psi) = m\, S(\rho^B_\psi)$ paires de Bell. Comme nous l'avons mentionné avant, ils peuvent utiliser des qubits supplémentaires (auxiliaires) dans chacun de leurs laboratoires. Alice utilise $2n$ qubits auxiliaires pour préparer "chaque partie" de $|\varphi'_m\rangle$ et utilise un protocole de téléportation (c.-à-d. des mesures de Bell dans son labo et l'envoi de $2n$ bits classiques à Bob) pour transférer à Bob sa partie de $|\varphi'_m\rangle$. Nous avons étudié le protocole de téléportation dans un cadre plus simple au

Chapitre **6** mais il est facile de voir qu'il est possible de le généraliser à la situation présente.

Imaginons d'abord un seul terme de la préparation $|\varphi'_m\rangle$ dans le labo d'Alice: $|x_1 \dots x_m\rangle_A \otimes |x_1 \dots x_m\rangle_{A\text{-copy}}$. Étant donné que $x_1 \dots x_m$ est une séquence typique, elle peut être décrite par $m\, S(\rho^A_\psi) = m\, S(\rho^B_\psi) = n$ bits et cet état est équivalent à $|y_1 \dots y_m\, 0 \dots 0\rangle_A \otimes |y_1 \dots y_m\, 0 \dots 0\rangle_{A\text{-copy}}$. Par exemple, on peut l'obtenir via une opération de rotation de base chez Alice. Il est maintenant clair qu'Alice et Bob peuvent effectuer le protocole de téléportation standard pour téléporter $|y_1 \dots y_m\, 0 \dots 0\rangle_{A\text{-copy}}$ chez Bob et obtenir $|y_1 \dots y_m\, 0 \dots 0\rangle_A \otimes |y_1 \dots y_m\, 0 \dots 0\rangle_B$. Ensuite, grâce à des rotations locales dans chaque labo, ils reconstituent $|x_1 \dots x_m\rangle_A \otimes |x_1 \dots x_m\rangle_B$.

Bien sûr, $|\varphi'_m\rangle$ est constitué de la somme sur les séquences typiques. Mais, puisque toutes les opérations de mesures et unitaires de la téléportation sont linéaires, on peut se convaincre que ces mêmes opérations permettent de fabriquer $|\varphi'_m\rangle$.

En résumé, nous avons décrit un protocole qui permet de fabriquer $|\psi\rangle^{\otimes m}$ à partir de $n = m\, S(\rho^A_\psi)$ paires de Bell. On a donc certainement $E_{\text{dilution}}(|\psi\rangle) \geq S(\rho^A_\psi)$.

Il reste à prouver que le protocole décrit ici est optimal, c.-à-d. $E_{\text{dilution}}(|\psi\rangle) \leq S(\rho^A_\psi)$. La démonstration de cette inégalité est omise ici.

Protocole optimal de distillation

Alice et Bob partagent m copies d'un état $|\psi\rangle$ et veulent transformer $|\psi\rangle^{\otimes m}$ en n paires EPR $|B_{00}\rangle^{\otimes n}$. On s'attend à ce que $n \leq m$ et donc, contrairement au protocole de dilution, aucun qubit auxiliaire n'est nécessaire.

L'ensemble des séquences typiques $T_{n,\epsilon}$ correspond à un ensemble "d'états typiques" de la base computationnelle $\left\{|x_1 \dots x_n\rangle_A \mid (x_1 \dots x_n) \in T_{m,\epsilon}\right\}$. Ces états engendrent un sous-espace de l'espace de Hilbert $\left(\mathbb{C}^2\right)^{\otimes m}$, appelé *sous-espace typique*, et noté $\tau_{n,\epsilon} \subset \left(\mathbb{C}^2\right)^{\otimes m}$. On a bien sûr la décomposition de l'espace de Hilbert $\left(\mathbb{C}^2\right)^{\otimes m} = \tau_{n,\epsilon} \oplus \tau^\perp_{n,\epsilon}$. Soient $\mathcal{P}_{n,\epsilon}$ et $\mathcal{P}^\perp_{n,\epsilon}$ les deux projecteurs orthogonaux associés à $\tau_{n,\epsilon}$ et $\tau^\perp_{n,\epsilon}$. Puisque $\mathcal{P}_{n,\epsilon} + \mathcal{P}^\perp_{n,\epsilon} = \mathbb{1}$, on peut considérer des mesures projectives avec $\left\{\mathcal{P}_{n,\epsilon}, \mathcal{P}^\perp_{n,\epsilon}\right\}$.

Alice effectue une mesure locale dans son laboratoire, ce qui transforme $|\psi\rangle^{\otimes m}$ en l'état proportionnel à

$$\mathcal{P}_{n,\epsilon} \otimes \mathbb{1} |\psi\rangle^{\otimes m} = \sum_{\substack{x_1,\dots,x_n \\ \in T_{n,\epsilon}}} \sqrt{p(x_1) \dots p(x_n)}\, |x_1 x_2 \dots x_n\rangle_A \otimes |x_1 x_2 \dots x_n\rangle_B$$

avec probabilité $\langle\psi|^{\otimes m}\mathcal{P}_{n,\epsilon}\otimes\mathbb{1}|\psi\rangle^{\otimes m}=\mathbb{P}\left[T_{n,\epsilon}\right]\geq 1-\epsilon$ ou bien en l'état

$$\mathcal{P}^{\perp}_{n,\epsilon}\otimes\mathbb{1}|\psi\rangle^{\otimes m}=\sum_{\substack{x_1,\dots,x_n\\ \in T^C_{n,\epsilon}}}\sqrt{p(x_1)\dots p(x_n)}\,|x_1x_2\dots x_n\rangle_A\otimes|x_1x_2\dots x_n\rangle_B$$

avec probabilité complémentaire $\mathbb{P}\left[T^C_{n,\epsilon}\right]\leq\epsilon$.

On note que le premier état n'est autre que $|\varphi'_m\rangle$. Le second état est perpendiculaire, appelons le $|\varphi'^{\perp}_m\rangle$. Alice envoie à Bob un bit d'information pour indiquer si elle a obtenu $|\varphi'_m\rangle$ ou bien $|\varphi'^{\perp}_m\rangle$. Ainsi, Alice et Bob ont transformé $|\psi\rangle^{\otimes m}$ en $|\varphi'_m\rangle$ avec une probabilité de succès supérieure à $1-\delta$. On peut prendre $\delta\to 0$ pour $m\to+\infty$.

La prochaine étape du protocole consiste à transformer $|\varphi'_m\rangle$ en $|B_{00}\rangle^{\otimes n}$. C'est l'opération de distillation. Tout d'abord, on remarque que $|\varphi'_m\rangle$ peut être décrit avec $n=mH(X)=m\,S(\rho^{A/B}_{\phi'_m})$ qubits puisque $\left|T_{n,\epsilon}\right|\approx 2^{mH(X)}$. Alice et Bob peuvent faire une rotation locale $U_A\otimes U_B$ telle que

$$U_A\otimes U_B\,|\varphi'_m\rangle=\frac{1}{\sqrt{\mathbb{P}\left[T_{n,\epsilon}\right]}}\sum_{\substack{y_1\dots y_n\\ \in\{0,1\}^n}}\sqrt{p(y_1)\dots p(y_n)}\,|y_1\dots y_n\,0\dots0\rangle_A\otimes|y_1\dots y_n\,0\dots0\rangle_B$$

Éventuellement, Alice et Bob peuvent laisser tomber les qubits $|0\dots0\rangle_A\otimes|0\dots0\rangle_B$. Le reste peut être dilué en n paires de Bell. En effet, les matrices densités associées à cette expression sont

$$\tilde{\rho}^{A,B}_{\varphi'_m}=\frac{1}{\sqrt{\mathbb{P}\left[T_{n,\epsilon}\right]}}\sum_{\substack{y_1\dots y_n\\ \in\{0,1\}^n}}p(y_1\dots y_n)\,|y_1\dots y_n\rangle\langle y_1\dots y_n|_{A,B}$$

et leurs valeurs propres

$$\lambda_{y_1\dots y_n}=\frac{p(y_1\dots y_n)}{\sqrt{\mathbb{P}\left[T_{n,\epsilon}\right]}}\leq\frac{2^{-n(H(X)-\epsilon)}}{1-\delta}$$

Pour n choisi tel que

$$\frac{2^{-n(H(X)-\epsilon)}}{1-\delta}\leq 2^{-n}\tag{8.4}$$

le vecteur $(2^{-n},\dots,2^{-n})$ est dégradé par rapport au vecteur $(\lambda_{y_1\dots y_n}\mid(y_1\dots y_n)\in\{0,1\}^n)$.

Le théorème de la Section **8.1** affirme alors qu'il est possible de transformer $U_A\otimes U_B\,|\varphi'_m\rangle$ en $|B_{00}\rangle^{\otimes n}$ tant que l'inégalité (8.4) est satisfaite. La plus grande valeur de n possible (avec ce protocole) et donc $n=m(H(X)-\epsilon)+\log(1-\delta)$. Pour $\epsilon,\delta\to 0$ et $n,m\to+\infty$, on obtient $\frac{n}{m}\to H(X)=S(\rho^{A.B}_{\psi})$.

Pour conclure, mentionnons sans démonstration que ce protocole donne une valeur optimale pour $\frac{n}{m}$.

8.4 États tripartites

Pour les systèmes bipartites, nous avons identifié deux classes naturelles d'états: les états produits et les états intriqués. Dans le cas multipartite général, la classification des ressources offertes par l'intrication est un problème beaucoup plus riche et encore largement ouvert. Le cas tripartite permet déjà de distinguer plusieurs classes d'états intriqués qui ne peuvent pas être obtenus les unes des autres par des transformations SLOCC. Celles-ci permettent de partitionner l'espace des états multipartites en classes d'équivalence (disjointes) naturelles. Un protocole avec des opérations locales et de la communication classique basé sur un état peut être aussi basé n'importe quel autre état de la même classe d'équivalence (en effet, il suffit d'adjoindre au protocole la transformation SLOCC qui fait passer d'un état à l'autre).

Dans le reste de ce chapitre, nous discutons en détail le cas tripartite qui est bien compris. Nous énumérons dans ce paragraphe 6 classes d'états tripartites et discutons quelques une de leurs propriétés. L'analyse mathématique détaillée et la preuve que ces 6 classes sont exhaustives est reportée à la Section **8.5**.

États tri-produits

Un *état tri-produit* est un état de la forme $|\psi\rangle_{ABC} = |\varphi_A\rangle \otimes |\varphi_B\rangle \otimes |\varphi_C\rangle$ où $|\varphi_i\rangle \in \mathbb{C}^2$ arbitraire pour $i = A, B, C$. On peut prendre comme représentant de cette classe d'équivalence l'état $|0\rangle \otimes |0\rangle \otimes |0\rangle$. Tous les états tri-produits peuvent être obtenus en appliquant des opérations unitaires locales de la forme $U_A \otimes U_B \otimes U_C$ au représentant $|0\rangle \otimes |0\rangle \otimes |0\rangle$.

États bi-produits

On peut considérer des états où A est produit alors que B et C sont intriqués: $|\psi\rangle_{ABC} = |\varphi_A\rangle \otimes |\phi_{BC}\rangle$ avec $|\varphi_A\rangle \in \mathbb{C}^2$ et $|\phi_{BC}\rangle \in \mathbb{C}^2 \otimes \mathbb{C}^2$ intriqué. Un représentant naturel est l'état $|0\rangle \otimes |B_{00}\rangle$ où $|B_{00}\rangle$ est l'état de Bell $|B_{00}\rangle = \frac{1}{\sqrt{2}}\big(|00\rangle + |11\rangle\big)$. Tous les états bi-produits peuvent être obtenus en appliquant une opération unitaire locale sur $U_A|0\rangle = |\varphi_A\rangle$ sur A et une transformation de dilution sur BC pour transformer $|B_{00}\rangle \to |\tilde{\phi}_{BC}\rangle = \cos\theta\,|00\rangle + \sin\theta\,|11\rangle$ (voir Section **8.1**). Pour obtenir $|\phi_{BC}\rangle$ général, on remarque par décomposition de Schmidt qu'on a toujours $|\phi_{BC}\rangle = \sqrt{\lambda_0}\,|\chi_0\rangle_B \otimes |\bar{\chi}_0\rangle_C + \sqrt{\lambda_1}\,|\chi_1\rangle_B \otimes |\bar{\chi}_1\rangle_C$ (ici, $\rho_B|\chi_i\rangle_B = \lambda_i|\chi_i\rangle_B$ et $\rho_C|\bar{\chi}_i\rangle_C = \lambda_i|\bar{\chi}_i\rangle_C$ avec $0 \le \lambda_i \le 1$ et $\lambda_0 + \lambda_1 = 1$). Il suffit donc d'appliquer des unitaires locales U_B et U_C qui transforment les bases $U_B|i\rangle_B = |\chi_i\rangle_B$ et $U_C|i\rangle_C = |\bar{\chi}_i\rangle_C$ si bien que $(U_B \otimes U_C)|\tilde{\phi}_{BC}\rangle = |\phi_{BC}\rangle$.

Par symétrie, on voit qu'il existe 3 classes d'équivalences d'états bi-produits correspondants aux bipartitions $A - BC$ (discutée ci-dessus), $B - AC$ et $C - AB$. Il est clair que des opérations locales ne peuvent pas faire passer d'une classe à l'autre. De plus, il est également clair que les opérations locales ne peuvent pas transformer les états tri-produits en bi-produits.

États avec intrication tripartite

Nous montrerons au paragraphe suivant qu'il existe deux classes disjointes d'états intriqués sous la relation d'équivalence SLOCC: les états obtenus à partir de l'état $|\mathrm{GHZ}\rangle = \frac{1}{\sqrt{2}}\big(|000\rangle + |111\rangle\big)$ et ceux obtenus à partir de l'état $|\mathrm{W}\rangle = \frac{1}{\sqrt{3}}\big(|001\rangle + |010\rangle + |100\rangle\big)$.

Discutons quelques propriétés qui distinguent ces deux états. Les matrices densités réduites à 1 qubits sont $\rho_A^{\mathrm{GHZ}} = \frac{1}{2}\mathbb{1}$ et $\rho_A^{\mathrm{W}} = \frac{2}{3}|0\rangle\langle 0| + \frac{1}{3}|1\rangle\langle 1|$ (et de même pour les sous-systèmes B et C). Localement, chaque partie possède une matrice densité diagonale. Autrement dit, chaque qubit se comporte localement comme une variable binaire aléatoire.

La matrice densité à 2 qubits, disons du système AB, est pour GHZ:

$$\rho_{AB}^{\mathrm{GHZ}} = \frac{1}{2}\big(|00\rangle\langle 00| + |11\rangle\langle 11|\big)$$

qui est aussi diagonale et équivalente à une variable aléatoire binaire. En particulier, dans l'état GHZ, il n'y a aucune intrication bipartite. Par exemple, A et B ne peuvent pas utiliser l'état GHZ pour de la téléportation ou du dense coding entre eux.

Pour W, la matrice densité à 2 qubits du système AB est

$$\begin{aligned}\rho_{AB}^{\mathrm{W}} &= \frac{1}{3}\big\{|00\rangle\langle 00| + |01\rangle\langle 01| + |01\rangle\langle 10| + |10\rangle\langle 01| + |10\rangle\langle 10|\big\}\\ &= \frac{1}{3}|00\rangle\langle 00| + \frac{2}{3}|B_{01}\rangle\langle B_{01}|\end{aligned}$$

avec $|B_{01}\rangle = \frac{1}{\sqrt{2}}\big(|01\rangle + |10\rangle\big)$. Ainsi, A et B partagent une ressource intriquée sous la forme d'un état mixte non-séparable. Notons que si C décohère localement. ρ_{AB}^{W} est inchangé et cette ressource intriquée est stable.

Hiérarchie sous les transformations LOCC

Une application du théorème de la Section **8.1** permet de montrer l'existence de transformations LOCC qui font passer des états $|\mathrm{GHZ}\rangle$ et $|\mathrm{W}\rangle$ aux états bi-produits $|0\rangle\otimes|B_{00}\rangle$ et des états bi-produits à l'état tri-produit $|0\rangle\otimes|0\rangle\otimes|0\rangle$. Pour les représentants des classes d'équivalence, ces transformations réussissent avec probabilité 1. Pour les autres états des classes d'équivalence, le passage est effectué via SLOCC avec probabilité non nulle. On note que ces transformations ne sont pas réversibles.

8.5 Analyse de la relation d'équivalence SLOCC

Caractérisation générale des transformations SLOCC

Nous formalisons la relation d'équivalence sous les transformations SLOCC dans le cas général multipartite et donnons un critère utile pour décrire les classes d'équivalence.

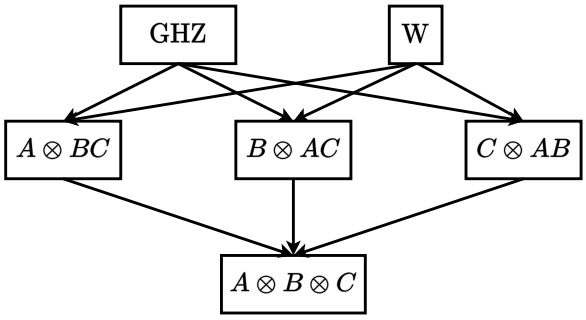


Figure 8.2 Passages entre classes d'équivalences via des LOCC. Le passage inverse n'est pas possible.

L'application au cas tripartite montre que les 6 classes discutées à la Section **8.4** sont les seules possibles.

Définition

Deux états $|\psi\rangle$ et $|\varphi\rangle$ entre les parties $A_1, A_2, \ldots, A_n$ sont dits SLOCC-équivalents s'il existe une suite de transformations locales avec communication classique (et résultat possiblement stochastique) qui opèrent les transformations $|\psi\rangle \leftrightarrow |\varphi\rangle$ (dans les deux sens). La classe d'équivalence d'un état $|\psi\rangle$ est constituée de l'ensemble des états obtenus à partir de $|\psi\rangle$ sous ces transformations.

Clairement, l'ensemble des transformations SLOCC est un groupe qui définit une relation réflexive, symétrique et transitive. C'est donc une relation d'équivalence qui partitionne l'ensemble des états multipartite en classes disjointes (en effet, si deux classes possédaient nu élément en commun, on pourrait relier tous les éléments de ces deux classes entre eux par transitivité, ce qui impliquent que les deux classes sont identiques).

On peut se représenter une transformation SLOCC comme un ensemble d'opérations du type $A_1 \otimes A_2 \otimes \ldots \otimes A_n$ avec probabilités $p_1 p_2 \ldots p_n$. En effet, on peut collecter toutes les opérations effectuées par A_1 (des unitaires ou des POVM) dans un produit de matrices inversibles $A_1 \otimes \mathbb{1} \otimes \ldots \otimes \mathbb{1}$ (avec probabilité p_1), celles de A_2 dans un produit $\mathbb{1} \otimes A_2 \otimes \ldots \otimes \mathbb{1}$ (avec probabilité p_2) et ainsi de suite. Les parties $A_1, A_2, \ldots A_n$ peuvent effectuer leurs opérations séquentiellement ou non et la transformation résultante sera du type $A_1 \otimes A_2 \otimes \ldots \otimes A_n$.

Deux états $|\psi\rangle$ et $|\varphi\rangle$ sont donc SLOCC-équivalents si on a

$$|\psi\rangle = A_1 \otimes A_2 \otimes \ldots \otimes A_n |\varphi\rangle$$

avec probabilité $p_1 p_2 \ldots p_n \neq 0$ et

$$|\varphi\rangle = A_1' \otimes A_2' \otimes \ldots \otimes A_n' |\psi\rangle$$

avec probabilité $p'_1 p'_2 \dots p'_n \neq 0$.

Théorème

Deux états $|\psi\rangle$ et $|\varphi\rangle$ sont SLOCC-équivalents si et seulement si on peut trouver des transformations locales inversibles telles que

$$|\psi\rangle = A_1 \otimes A_2 \otimes \dots \otimes A_n |\varphi\rangle$$
$$|\varphi\rangle = A_1^{-1} \otimes A_2^{-1} \otimes \dots \otimes A_n^{-1} |\psi\rangle$$

Démonstration

On considère d'abord le cas de transformations où c'est uniquement A_1 qui fait des transformations locales et $A_2 = \dots = A_n = \mathbb{1}$. Le cas général n'est pas beaucoup plus compliqué. Supposons d'abord qu'il existe une transformation locale inversible telle que

$$|\psi\rangle = A_1 \otimes \mathbb{1} |\varphi\rangle$$
$$|\varphi\rangle = A_1^{-1} \otimes \mathbb{1} |\psi\rangle$$

Puisque $\langle\psi|\psi\rangle = \langle\varphi|\varphi\rangle = 1$, il suit que $\langle\varphi|A_1^\dagger A_1|\varphi\rangle = \langle\psi|(A_1^{-1})^\dagger A_1^{-1}|\psi\rangle = 1$. Donc, $\left\{\sqrt{p_1}A_1,\ \sqrt{1-p_1}A_1^\dagger A_1\right\}$ et $\left\{\sqrt{q_1}A_1^{-1},\ \sqrt{1-q_1}(A_1^{-1})^\dagger A_1^{-1}\right\}$ sont des POVM tant que $0 \leq p_1, q_1 \leq 1$ sont des probabilités suffisamment petites pour que $p_1 A_1^\dagger A_1 < \mathbb{1}$ et $q_1 (A_1^{-1})^\dagger A_1^{-1} < \mathbb{1}$. Grâce à ces POVM, on effectue la transformation $|\psi\rangle \to |\varphi\rangle$ avec probabilité p_1 et $|\varphi\rangle \to |\psi\rangle$ avec probabilité q_1.

Pour démontrer la réciproque, on commence par supposer qu'il existe une transformation SLOCC avec probabilité de succès p_1 telle que $|\psi\rangle = A_1 \otimes \mathbb{1}|\varphi\rangle$. On veut montrer qu'il est toujours possible de choisir A_1 inversible. Soient λ_i^ψ et λ_i^φ les valeurs propres des matrices densités réduites $\rho_{A_1}^\psi$ et $\rho_{A_1}^\varphi$ et soient $|\chi_i\rangle$ et $|\tau_i\rangle$ les bases de la décomposition de Schmidt pour $|\varphi\rangle$:

$$|\varphi\rangle = \sqrt{\lambda_0^\varphi}\,|\chi_0\rangle \otimes |\tau_0\rangle + \sqrt{\lambda_1^\varphi}\,|\chi_1\rangle \otimes |\tau_1\rangle$$

Si U_{A_1} est l'unitaire pour passer de la base de Schmidt de la partie A_1 de $|\varphi\rangle$ à la partie A_1 de $|\psi\rangle$, on a:

$$|\psi\rangle = \sqrt{\lambda_0^\psi}\,U_{A_1}|\chi_0\rangle \otimes |\tau_0\rangle + \sqrt{\lambda_1^\psi}\,U_{A_1}|\chi_1\rangle \otimes |\tau_1\rangle$$

Soit maintenant

$$\tilde{A}_1 = \sqrt{\frac{\lambda_0^\psi}{\lambda_0^\varphi}}\,|\chi_0\rangle\langle\chi_0| + \sqrt{\frac{\lambda_1^\psi}{\lambda_1^\varphi}}\,|\chi_1\rangle\langle\chi_1|$$

On vérifie facilement que $A_1 = U_{A_1}\tilde{A}_1$.

Puisque

$$\tilde{A}_1^{-1} = \sqrt{\frac{\lambda_0^\varphi}{\lambda_0^\psi}}\,|\chi_0\rangle\langle\chi_0| + \sqrt{\frac{\lambda_1^\varphi}{\lambda_1^\psi}}\,|\chi_1\rangle\langle\chi_1|$$

on a bien que A_1 est inversible avec $A_1^{-1} = \tilde{A}_1^{-1}U_{A_1}^\dagger$.

Application aux états à deux qubits

Nous démontrons ici que les états produits et intriqués sont les deux seules classes d'équivalence pour la relation d'équivalence SLOCC.

Soit $|\psi\rangle_{AB}$ un état pur bipartite. Alors, d'après le théorème de Schmidt, les matrices densité réduites ont le même rang. Donc, $\text{rang}(\rho_A) = \text{rang}(\rho_B) = 1$ ou $\text{rang}(\rho_A) = \text{rang}(\rho_B) = 2$. Dans le premier cas, ρ_A et ρ_B sont des états purs $|\rho_A\rangle\langle\rho_A|$ et $|\rho_B\rangle\langle\rho_B|$ et donc $|\psi\rangle_{AB} = |\rho_A\rangle \otimes |\rho_B\rangle$ est un état produit. Puisque tous les états produits peuvent être obtenus à partir d'unitaires locales $U_A \otimes U_B$ agissant sur $|0\rangle \otimes |0\rangle$ et qui sont inversibles, le théorème précédent affirme que les états produits forment une classe d'équivalence avec représentant $|0\rangle \otimes |0\rangle$.

Dans le cas de rang 2, le théorème de Schmidt dit que

$$|\psi\rangle_{AB} = \sqrt{\lambda_0}\,|\chi_0\rangle_A \otimes |\chi_0\rangle_B + \sqrt{\lambda_1}\,|\chi_1\rangle_A \otimes |\chi_1\rangle_B$$

avec les λ_i et $|\chi_i\rangle$ les valeurs propres non nulles et bases orthonormées de ρ_A et ρ_B. Par transformation unitaire locale, cet état est équivalent à

$$|\tilde{\psi}\rangle_{AB} = \sqrt{\lambda_0}\,|0\rangle_A \otimes |0\rangle_B + \sqrt{\lambda_1}\,|1\rangle_A \otimes |1\rangle_B$$

Puisque $\lambda_0 + \lambda_1 = 1$, on peut poser $\sqrt{\lambda_0} = \cos\theta$ et $\sqrt{\lambda_1} = \sin\theta$. L'analyse faite à la Section **8.1** ou bien la technique de preuve du théorème précédent montrent que cet état est SLOCC-équivalent à l'état de Bell $|B_{00}\rangle = \frac{1}{\sqrt{2}}\big(|0\rangle \otimes |0\rangle + |1\rangle \otimes |1\rangle\big)$. L'état de Bell peut être choisi comme représentant de la classe d'équivalence des états intriqués.

Application aux états à trois qubits

Nous montrons dans ce paragraphe que les classes d'équivalences des états avec 3 qubits sont les 6 classes énumérées à la Section **8.4**. Tout d'abord, considérons la bipartition A et BC et les matrices densité réduites correspondantes ρ_A et ρ_{BC}. D'après le théorème de Schmidt, on a $\text{rang}(\rho_A) = \text{rang}(\rho_{BC}) = 1$ ou bien $\text{rang}(\rho_A) = \text{rang}(\rho_{BC}) = 2$.

Dans le premier cas, ρ_A et ρ_{BC} sont des états purs dans $|\psi\rangle_{ABC} = |\varphi_A\rangle \otimes |\varphi_{BC}\rangle$. Pour le système BC, d'après l'analyse précédente, $|\varphi_{BC}\rangle$ est SLOCC-équivalent à $|0\rangle_B \otimes |0\rangle_C$ ou bien à $|B_{00}\rangle = \frac{1}{\sqrt{2}}\big(|0\rangle_B \otimes |0\rangle_C + |1\rangle_B \otimes |1\rangle_C\big)$. De plus, $|\varphi_A\rangle$ est unitairement relié

à $|0\rangle_A$. Ainsi, trouvons dans ce cas deux classes de représentants $|0\rangle_A \otimes |0\rangle_B \otimes |0\rangle_C$ et $|0\rangle_A \otimes |B_{00}\rangle_{BC}$. Bien sûr, en considérant les bipartitions $B-AC$ et $C-AB$, le même raisonnement montre qu'il existe encore deux autres classes de représentants $|0\rangle_B \otimes |B_{00}\rangle_{AC}$ et $|0\rangle_C \otimes |B_{00}\rangle_{AB}$.

Passons maintenant au cas où $\text{rang}(\rho_A) = \text{rang}(\rho_{BC}) = 2$. Le théorème de Schmidt affirme que

$$|\psi\rangle_{ABC} = \sqrt{\lambda_0}\,|\chi_0\rangle_A \otimes |\phi_0\rangle_{BC} + \sqrt{\lambda_1}\,|\chi_1\rangle_A \otimes |\phi_1\rangle_{BC}$$

où $|\chi_i\rangle$ et $|\phi_i\rangle$ sont les bases orthonormales d'états propres de ρ_A et ρ_{BC} et $\lambda_i \neq 0$ (car le rang vaut 2). Par transformation unitaire locale sur A, cet état est équivalent à

$$\sqrt{\lambda_0}\,|0\rangle_A \otimes |\phi_0\rangle_{BC} + \sqrt{\lambda_1}\,|1\rangle_A \otimes |\phi_1\rangle_{BC}$$

Deux cas de figures se présentent: $|\phi_0\rangle_{BC}$ peut être un état produit ou intriqué. Dans les deux cas, des transformations SLOCC transforment $|\phi_0\rangle_{BC}$ en $|0\rangle_B \otimes |0\rangle_C$ ou en $\frac{1}{\sqrt{2}}\Big(|0\rangle_B \otimes |0\rangle_C + |1\rangle_B \otimes |1\rangle_C\Big)$. La partie $|\phi_1\rangle_{BC}$ subit également l'action de cette transformation, donc, l'état total est équivalent à

$$\sqrt{a_0}\,|0\rangle_A \otimes |0\rangle_B \otimes |0\rangle_C + \sqrt{a_1}\,|1\rangle_A \otimes |\tilde{\phi}_1\rangle_{BC} \tag{8.5}$$

où a_0 et a_1 sont réels non nuls (Notez que si $a_1 = 0$, on retrouve la classe d'états produits).

Supposons maintenant que $|\tilde{\phi}_1\rangle_{BC}$ est produit: $|\tilde{\phi}_1\rangle_{BC} = \Big(s_B\,|0\rangle + t_B\,|1\rangle\Big) \otimes \Big(s_C\,|0\rangle + t_C\,|1\rangle\Big)$. On peut appliquer la transformation locale inversible

$$\frac{1}{\sqrt{2}} \begin{pmatrix} \frac{1}{\sqrt{a_0}} & 0 \\ 0 & \sqrt{a_1} \end{pmatrix} \otimes \begin{pmatrix} 1 & -\frac{s_B}{t_B} \\ 0 & \frac{1}{t_B} \end{pmatrix} \otimes \begin{pmatrix} 1 & -\frac{s_C}{t_C} \\ 0 & \frac{1}{t_C} \end{pmatrix}$$

pour transformer l'état en

$$|\text{GHZ}\rangle = \frac{1}{\sqrt{2}}\Big(|000\rangle + |111\rangle\Big)$$

Il reste à considérer le cas où $|\tilde{\phi}_1\rangle_{BC}$ n'est pas un état produit. Nous montrons ci-dessous qu'il est possible, sans perte de généralité, de poser $|\tilde{\phi}_1\rangle_{BC} = \frac{1}{\sqrt{2}}\Big(|01\rangle + |10\rangle\Big)$. Dans ce cas, l'état $|\psi\rangle_{ABC}$ est équivalent à

$$\sqrt{a_0}\,|000\rangle + \sqrt{\frac{a_1}{2}}|101\rangle + \sqrt{\frac{a_1}{2}}|110\rangle$$

En appliquant la transformation inversible $X_A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, on obtient

$$\sqrt{a_0}\,|100\rangle + \sqrt{\frac{a_1}{2}}|001\rangle + \sqrt{\frac{a_1}{2}}|010\rangle$$

Finalement, en appliquant la transformation inversible

$$\frac{1}{\sqrt{3}}\begin{pmatrix}1 & 0\\ 0 & \frac{1}{\sqrt{a_0}}\end{pmatrix}\otimes\begin{pmatrix}\frac{2}{\sqrt{a_1}} & 0\\ 0 & 1\end{pmatrix}\otimes\begin{pmatrix}1 & 0\\ 0 & \frac{2}{\sqrt{a_1}}\end{pmatrix}$$

l'état obtenu est

$$|\mathrm{W}\rangle = \frac{1}{\sqrt{3}}\Big(|100\rangle + |010\rangle + |001\rangle\Big)$$

Il reste à démontrer l'affirmation proposée ci-dessus: dans le cas où $|\tilde{\phi}_1\rangle_{BC}$ est intriqué, on peut choisir $|\tilde{\phi}_1\rangle_{BC} = \frac{1}{\sqrt{2}}\Big(|01\rangle + |10\rangle\Big)$. En général, $|\tilde{\phi}_1\rangle_{BC}$ est une combinaison linéaire de $|00\rangle$, $|01\rangle$, $|10\rangle$ et $|11\rangle$ donc l'état (8.5) est

$$\sqrt{a_0}\,|0\rangle\otimes|00\rangle + \sqrt{a_1}\,|1\rangle\otimes\Big\{x\,|00\rangle + y\,|01\rangle + z\,|10\rangle + t\,|11\rangle\Big\} =$$
$$\Big(\sqrt{a_0}\,|0\rangle + x\sqrt{a_1}\,|1\rangle\Big)\otimes|00\rangle + \sqrt{a_1}\,|1\rangle\otimes\Big\{y\,|01\rangle + z\,|10\rangle + t\,|11\rangle\Big\}$$

Grâce à un changement de base inversible chez A avec la matrice $\begin{pmatrix}1 & 1\\ 0 & 1\end{pmatrix}$, cet état est équivalent à

$$\alpha\,|0\rangle\otimes|00\rangle + \beta\,|1\rangle\otimes\Big\{y\,|01\rangle + z\,|10\rangle + t\,|11\rangle\Big\}$$

ce qui montre que, sans perte de généralité, $|00\rangle$ peut être omis de $|\tilde{\phi}_1\rangle_{BC}$. Nous remarquons maintenant que

$$\rho_{BC} = \alpha^2\,|00\rangle\langle 00| + \beta^2\Big\{y\,|01\rangle + z\,|10\rangle + t\,|11\rangle\Big\}\Big\{y\,\langle 01| + z\,\langle 10| + t\,\langle 11|\Big\}$$

Donc, tous les vecteurs appartenant à l'espace image de ρ_{BC} sont des combinaisons linéaires de $|00\rangle$ et $\Big\{y\,|01\rangle + z\,|10\rangle + t\,|11\rangle\Big\}$. D'autre part, il faut exclure les combinaisons linéaires qui donnent des états produits indépendants de $|00\rangle$ sinon nous serions ramenés au cas où $|\tilde{\phi}_1\rangle_{BC}$ est produit. Nous devons donc poser la condition que

$$\lambda\,|00\rangle + \Big\{y\,|01\rangle + z\,|10\rangle + t\,|11\rangle\Big\}$$

ne soit pas produit pour tout λ. C'est-à-dire $\lambda t \neq yz\ \forall\lambda$[4]. Ceci implique $t = 0$ (sinon il existe $\lambda = \frac{yz}{t}$ qui donne un état produit) et nous concluons que $|11\rangle$ doit être omis dans $|\tilde{\phi}_1\rangle_{BC}$.

[4] Pour un état produit, $\Big(a_0\,|0\rangle + a_1\,|1\rangle\Big)\otimes\Big(b_0\,|0\rangle + b_1\,|1\rangle\Big) = \alpha\,|00\rangle + \beta\,|01\rangle + \gamma\,|10\rangle + \delta\,|11\rangle$, la condition $\alpha\delta = \beta\gamma$ doit être satisfaite.

9 Modèle des Circuits et Algorithmes Quantique

Ce chapitre est une première introduction au calcul quantique. Après une brève introduction historique (Section **9.1**) et une discussion de la notion de circuit classique (Section **9.2**), nous introduisons le modèle de Deutsch des "circuits quantiques" (1995, Section **9.3**). Comme nous allons le voir, ce modèle sert à définir ce qui sera pour nous un algorithme quantique, tout en fournissant une représentation très concrète de ces algorithmes. Enfin, nous illustrons ces notions grâce à un algorithme quantique simple, mais important: l'algorithme de Deutsch et Jozsa (Section **9.6**). Cet algorithme quantique contient déjà la plupart des ingrédients d'une classe plus large, qui traite le problème plus général de la recherche de symétries cachées. Le célèbre algorithme de Shor (1997), permettant la factorisation d'un nombre entier en temps polynomial (par rapport aux nombres de bits de l'entier) appartient à cette classe. Celui-ci fera l'objet du Chapitre **11**. Il est suspecté, mais pas démontré, que cette famille d'algorithmes quantiques permet une accélération exponentielle du temps de calcul, par rapport aux algorithmes classiques.

9.1 Brève introduction historique

Un calcul est de façon ultime réalisé par un dispositif physique. Il est donc naturel de se demander quelles sont les limites fondamentales que les lois de la physique imposent au calcul. Un travail précurseur fut celui de Landauer (dans les années 60-70) qui montra que l'effacement d'un bit - un processus irréversible - est toujours accompagné d'une dissipation de chaleur. Essentiellement, ceci provient de l'augmentation de l'entropie du système due à l'effacement du bit (perte d'information) et des lois de la thermodynamique reliant la variation d'entropie d'un système au flux de chaleur entre le système et son environnement. En conséquence, tout calcul utilisant des portes logiques *irréversibles* (par exemple AND, OR) dissipe de la chaleur. Mais existe-t-il un principe fondamental qui nécessite absolument une dissipation minimale de chaleur lors d'un calcul ? Ou bien pourrait-on, en théorie du moins, éliminer la dissipation de chaleur lors du calcul ? Une *réponse positive* à cette deuxième question a été présentée par Bennett, Benioff et d'autres. Plus précisément, *tout calcul irréversible peut être rendu réversible*, grâce à des portes élémentaires appropriées. Néanmoins, il y a un coût: l'espace de stockage doit être augmenté pour éliminer les effacements.

En l'absence de la chaleur générée par un calcul réversible, lorsque le support physique

des bits atteint les dimensions moléculaires ou atomiques et que la température du système est maintenue très basse, le comportement quantique de la matière et la cohérence des états quantiques (le principe de superposition) deviennent importants. Il est naturel de se poser la question suivante: quels sont les effets du comportement quantique de la matière sur le calcul ? Est-ce que les effets quantiques aident, ou bien au contraire apportent-ils de nouvelles limites ?

Ces questions ont été soulevées et discutées par Feynman, Benioff et Manin au début des années 1980. En principe la MQ n'apporte pas de nouvelles limitations. Au contraire ! Le principe de superposition appliqué à des systèmes à plusieurs particules (plusieurs qubits) nous permet d'effectuer des "calculs parallèles". Ce "parallélisme quantique" accélère les calculs classiques, et parfois même de façon exponentielle. Cela fut reconnu notamment par Feynman qui a fait valoir que les ordinateurs classiques ne peuvent simuler des processus quantiques "efficacement" (du point de vue du temps de calcul). Feynman a suggéré que nous devrions construire des "machines quantiques" pour simuler efficacement les processus quantiques.

La raison fondamentale pour laquelle le calcul classique ne peut pas simuler efficacement les processus quantiques est la suivante. Un état général quantique de N bits quantiques contient une superposition de 2^N "états classiques":

$$|\Psi\rangle = \sum_{b_1 \dots b_N \in \{0,1\}^N} C(b_1 \dots b_N)|b_1 \dots b_N\rangle$$

Ici $|b_1 \dots b_N\rangle = |b_1\rangle \otimes \dots \otimes |b_N\rangle$ (avec $b_i = 0, 1$) sont les états de la base dite "computationnelle". Une simulation classique de l'évolution du ket $|\Psi\rangle$ doit effectuer essentiellement 2^N calculs pour l'évolution de suite binaire classique $(b_1 \dots b_N)$. Au contraire, la dynamique quantique unitaire U agit sur $|\Psi\rangle$ dans son ensemble (ou en parallèle sur chaque ket $|b_1 \dots b_N\rangle$). Un dispositif physique réalisant la dynamique unitaire U sur $|\Psi\rangle$ peut être considéré comme un ordinateur quantique.

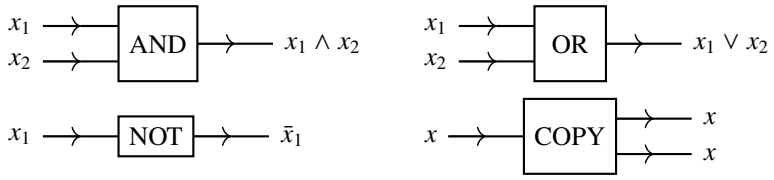
Un calcul classique peut être représenté par un modèle de circuit construit à partir d'un ensemble donné de portes élémentaires universelles agissant de manière récursive sur l'entrée du calcul. En 1985, David Deutsch a montré que la même chose est valable dans le cas quantique. Notamment, toute évolution unitaire peut être assez bien approchée par un ensemble universel de portes quantiques universelles.

Il existe aussi d'autres modèles d'ordinateur quantique, mais le modèle de Deutsch - un modèle de circuit quantique - est le plus populaire aujourd'hui. Un des buts de ce chapitre est d'expliquer ce modèle. Une des raisons de sa popularité est qu'il s'agit d'un modèle universel: en principe, tout calcul quantique peut être représenté comme un circuit quantique construit à partir d'un ensemble restreint de porte logiques quantiques. De plus, cette représentation est relativement concrète.

Il existe aussi une notion de machine de Turing quantique (analogue aux machines de Turing classiques) qui est le cadre naturel pour discuter des classes de complexité quantiques. Il a été démontré que le modèle de machine de Turing quantique est équivalent au modèle des circuits quantiques. Cet aspect ne sera pas abordé ici.

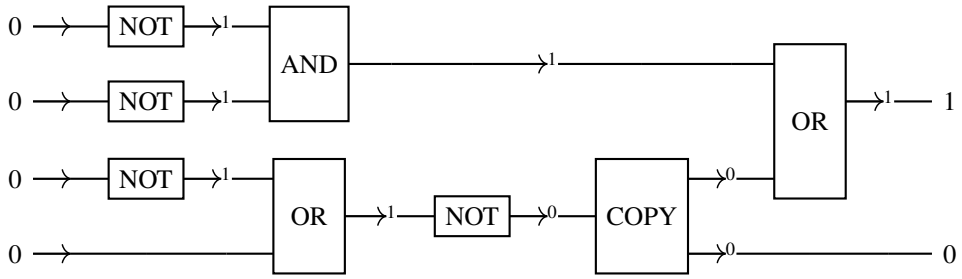
9.2 Modèle des circuits pour le calcul classique

Nous discutons brièvement le calcul classique basé sur les circuits booléens. Considérons les portes logiques classiques de base $x_i \in \mathbf{F}_2 = \{0, 1\}$.



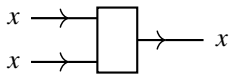
L'opération COPY s'appelle aussi parfois FANOUT. Cet ensemble de portes peut être utilisé pour définir les circuits Booléens.

Un circuit Booléen est un graphe acyclique (sans cycles) dirigé (les liens ont une direction) avec n bits d'entrée et m bits de sortie. L'entrée peut toujours être initialisée à $(0 \dots 0)$ car tout $(x_1 \dots x_n)$ est obtenu par une série de portes NOT. Le schéma suivant illustre cette définition avec un exemple de circuit Booléen.



Un célèbre théorème d'Emil Post (~1950) montre que toute fonction $f : \mathbf{F}_2^n \to \mathbf{F}_2^m$ peut être réalisée par un circuit Booléen. Plus précisément, pour toute fonction $f : \mathbf{F}_2^n \to \mathbf{F}_2^m$, il existe un circuit Booléen qui applique les entrées $(x_1 \dots x_n)$ sur les sorties $(y_1 \dots y_m) = f(x_1 \dots x_n)$. Le circuit est entièrement construit avec les portes NOT, AND, OR, COPY et est un graphe acyclique dirigé. On dit que l'ensemble des portes (NOT, AND, OR, COPY) est universel.

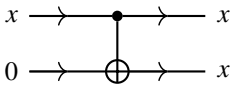
La porte NOT est logiquement réversible. Cela veut dire qu'à partir de la sortie, il est possible de récupérer l'entrée. Les portes AND et OR, elles, sont logiquement irréversibles. Il est impossible de reconstruire l'entrée à partir de la sortie. L'irréversibilité logique entraîne l'irréversibilité physique, c.-à-d. une dissipation de chaleur lors du calcul. En effet, la perte d'information et l'augmentation d'entropie est reliée à un flux de chaleur du système vers l'environnement. La porte COPY, quant à elle, est logiquement réversible, mais physiquement irréversible. En effet, l'opération inverse



efface un bit et, comme Landauer l'a montré, cela entraîne une dissipation de chaleur.

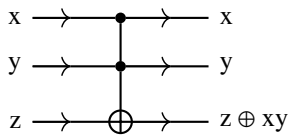
Bennett a montré que n'importe quel circuit Booléen peut être simulé ou remplacé par un circuit réversible. De plus, il existe un ensemble universel de portes logiques réversibles (logiquement et physiquement réversibles). Nous n'allons pas donner cette preuve ici, mais nous contenter de l'idée essentielle: on peut toujours remplacer les portes AND, OR et COPY par des portes réversibles à condition d'utiliser des bits auxiliaires.

Commençons par la porte COPY. Elle peut être remplacée par



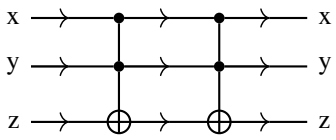
qui est une porte control-NOT (aussi notée CNOT) réversible utilisant deux bits. Le bit supérieur est appelé *bit de contrôle* et celui inférieur est le *bit de stockage*, égal à 0 dans l'entrée.

Pour les portes AND et OR, nous utilisons la *porte de Toffoli*. Celle-ci n'est rien d'autre qu'un CCNOT (control-control-NOT) et utilise trois bits.



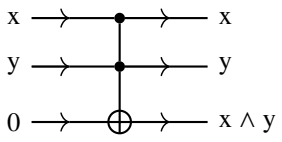
$\equiv \quad T(x, y, z) = (x, y, z \oplus xy)$

Cette porte flip le bit z si les deux bits de contrôle x et y sont égaux à 1. Sinon z est inchangé. La porte de Toffoli est effectivement réversible, car



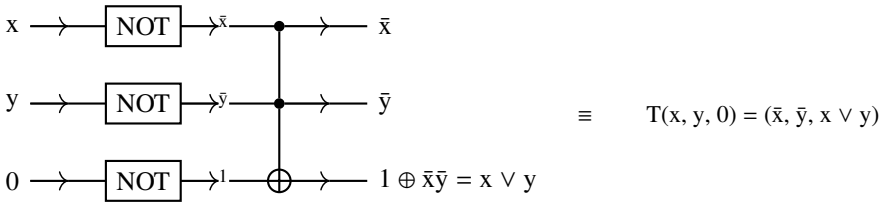
$\equiv \quad T^2(x, y, z) = (x, y, z)$

La porte AND correspond au cas où z = 0. Dans ce cas, T(x, y, 0) = (x, y, xy) retourne $x \wedge y$ pour le troisième bit.



$\equiv \quad T(x, y, 0) = (x, y, x \wedge y)$

Pour la porte OR, on utilise



Résumons: un circuit Booléen utilisant {AND, OR, COPY, NOT} peut être remplacé par un circuit utilisant les portes universelles {CNOT; Toffoli; NOT}. L'ensemble {AND, OR, COPY, NOT} contient uniquement des portes à un et deux bits, alors que {CNOT; Toffoli; NOT} fait intervenir une porte à trois bits (Toffoli). On peut montrer qu'il n'est pas possible de se passer de portes à tris bits si on veut la réversibilité d'un circuit classique. Nous verrons que dans le cas quantique, cela est possible !

9.3 Circuits quantiques.

Les circuits quantiques sont analogues aux circuits classiques. En particulier, ils sont construits à partir d'un petit ensemble de "portes quantiques" élémentaires. Une "porte quantique" n'est rien d'autre qu'une opération unitaire qui agit sur un petit nombre de qubits (typiquement un ou deux qubits). Ces portes sont réversibles, car une opération unitaire est inversible, en effet $UU^\dagger = U^\dagger U = \mathbb{1}$. Nous verrons dans des chapitres ultérieurs comment elles sont réalisées en pratique. Nous commençons par discuter certaines de ces portes élémentaires.

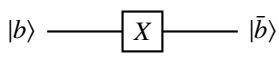
Portes à un qubit

Les portes à un qubit sont des matrices unitaires qui agissent sur les vecteurs d'état de l'espace de Hilbert $\mathbb{C}^2$. Certaines de ces matrices joueront un rôle particulièrement important.

- Les trois "matrices de Pauli" :

$$X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad iY = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \quad Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

  Ce sont bien des matrices unitaires. (Attention: Y n'est pas unitaire, mais hermitienne, néanmoins iY est unitaire; X et Z sont unitaires et hermitiennes.) Les circuits correspondants sont les suivants:



$$|b\rangle \text{ —— } \boxed{iY} \text{ —— } (-1)^{b+1}|\bar{b}\rangle$$

$$|b\rangle \text{ —— } \boxed{Z} \text{ —— } (-1)^{b}|b\rangle$$

La porte X n'est rien d'autre que la porte NOT quantique. La différence principale avec le cas classique est que le qubit en entrée peut être une superposition cohérente des états $\{|0\rangle, |1\rangle\}$ (contrairement à un bit classique qui vaut soit 0, soit 1). Par exemple,

$$X(\alpha|0\rangle + \beta|1\rangle) = \alpha|1\rangle + \beta|0\rangle.$$

Ainsi, l'action de la porte NOT est beaucoup plus générale dans le cas quantique. Cette remarque est simple, mais cruciale et profonde (et s'applique à toutes les portes quantiques discutées ci-dessous) !

- La porte de Hadamard $H = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$

$$|b\rangle \text{ —— } \boxed{H} \text{ —— } H|b\rangle = \tfrac{1}{\sqrt{2}}\left(|0\rangle + (-1)^{b}|1\rangle\right)$$

Cette porte n'a pas d'analogue classique.

- La porte $T = \begin{pmatrix} 1 & 0 \\ 0 & e^{i\pi/4} \end{pmatrix}$

$$|b\rangle \text{ —— } \boxed{T} \text{ —— } e^{ib\pi/4}|b\rangle = \begin{cases} |0\rangle \\ e^{i\pi/4}|1\rangle \end{cases}$$

Elle agit sur les superpositions comme

$$T(\alpha|0\rangle + \beta|1\rangle) = \alpha|0\rangle + \beta e^{i\pi/4}|1\rangle$$

- La porte $S = \begin{pmatrix} 1 & 0 \\ 0 & e^{i\pi/2} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix}$

$$|b\rangle \text{ —— } \boxed{S} \text{ —— } e^{ib\pi/2}|b\rangle = \begin{cases} |0\rangle \\ i|1\rangle \end{cases}$$

Elle agit sur les superpositions comme

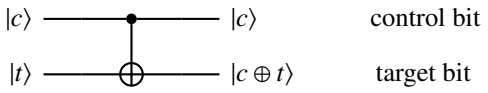
$$S(\alpha|0\rangle + \beta|1\rangle) = \alpha|0\rangle + i|1\rangle$$

Très généralement, on peut montrer que toute matrice unitaire U 2×2 peut être approximée avec une précision arbitraire par des produits des matrices élémentaires H et T. Notez que $S = T^2$.

Portes à deux qubits

Les portes à deux qubits sont des matrices unitaires qui agissent sur les vecteurs d'état de l'espace de Hilbert $\mathbb{C}^2 \otimes \mathbb{C}^2$. La plus importante est la porte control-NOT quantique.

- La porte CNOT (control-NOT) est définie par:



  Dans la base

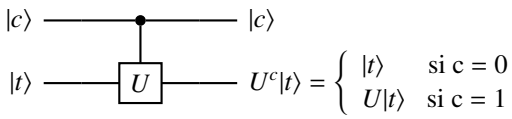
$$|0\rangle \otimes |0\rangle \equiv \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad |0\rangle \otimes |1\rangle \equiv \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad |1\rangle \otimes |0\rangle \equiv \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \quad |1\rangle \otimes |1\rangle \equiv \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

  la représentation matricielle est

$$\text{CNOT} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

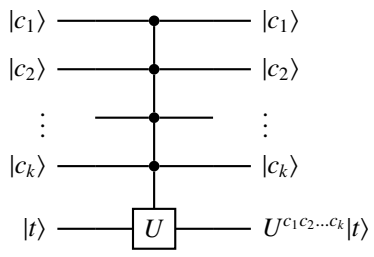
  Nous remarquons que cette porte agit en général sur des superpositions cohérentes d'états à deux qubits.

- Une généralisation importante est la porte control-U où U est une opération unitaire à un qubit:



Portes multi-control-U

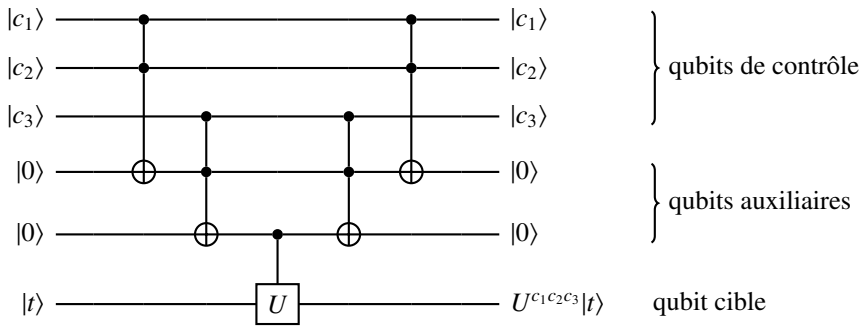
Une généralisation des portes précédentes est



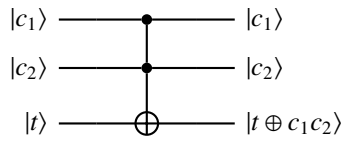
La porte U agit sur le dernier qubit si tous les qubits de contrôle sont égaux à 1. À nouveau, il est important de remarquer que ces portes agissent sur des superpositions

cohérentes d'états de base de l'espace de Hilbert $\mathbb{C}^2 \otimes \cdots \otimes \mathbb{C}^2$. Ce sont des matrices $2^n \times 2^n$.

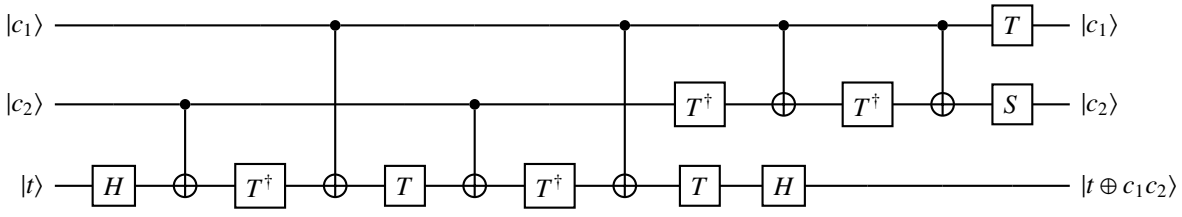
En augmentant le nombre de qubits auxiliaires la porte multi-control-U peut être réalisée grâce à une concaténation de control-control-NOT et un control-U. En effet, le circuit suivant réalise bien l'opération control-U avec trois qubits de contrôle:



La porte control-control-NOT s'appelle aussi porte de Toffoli quantique (la différence avec la porte classique est que les entrées peuvent être des superpositions d'états). Remarquablement, cette porte quantique peut être représentée par des portes à un & deux qubits {T, S, H, CNOT}. En effet, on peut vérifier que



est équivalent au circuit suivant :



Résumons cette discussion: Toute porte multi-control-U de dimension $2^n \times 2^n$ peut être réalisée grâce à l'ensemble {T, S, H, CNOT, U} où U est une porte à un qubit. De plus, nous avons aussi vu que U peut être approximée avec une précision arbitraire par un produit de matrices H et T.

Le modèle des circuits quantiques de Deutsch

Le dernier résultat du paragraphe précédent peut être généralisé. Un théorème important affirme que toute opération unitaire agissant sur l'espace à n qubits (c.-à-d. toute matrice $2^n \times 2^n$) peut être approximée avec une précision arbitraire par l'ensemble des portes {T, S, H, CNOT}. Ce théorème est à la base du modèle des circuits quantiques.

La complexité du circuit dépendra du nombre de matrices utilisées pour approximer l'opération unitaire à n qubits. Nous verrons par exemple que l'opération unitaire utilisée pour la factorisation des entiers requiert $O(\text{poly}N)$ portes quantiques élémentaires. Cela n'est pas possible avec un circuit classique. Néanmoins, on sait qu'il existe des matrices unitaires qui requièrent un nombre exponentiellement grand (en n) de portes élémentaires.

Définissons maintenant le modèle des circuits:

a). Un circuit quantique est graphe dirigé acyclique dont les vertex sont les portes {T, S, H, CNOT} et les arcs "portent" des qubits ($\alpha|0\rangle + \beta|1\rangle$).

b). L'entrée est donnée par l'état produit:

$$|0\rangle \otimes ... \otimes |0\rangle$$

c). La sortie est le résultat de l'évolution unitaire. Cette sortie prend la forme générale:

$$|\Psi\rangle = \sum_{c_1...c_n} A(c_1...c_n)|c_1c_2...c_n\rangle$$

où $A(c_1...c_n)$ sont des coefficients complexes et $|c_1c_2...c_n\rangle = |c_1\rangle \otimes \cdots \otimes |c_n\rangle$ sont les états de la base computationnelle.

d). Finalement une opération de mesure est effectuée sur $|\Psi\rangle$ avec un appareil mesurant dans la "base computationnelle" $\{|c_1c_2...c_n\rangle, c_i = 0, 1\}$. Le résultat de l'opération de mesure est l'état $|c_1...c_n\rangle$ avec probabilité $|A(c_1...c_n)|^2$ (règle de Born). Le résultat du calcul quantique est donc un résultat probabiliste. En pratique, l'algorithme est bon si la probabilité est concentrée sur le résultat cherché. La plupart du temps, on répète l'expérience (le calcul du circuit) pour amplifier cette probabilité (nous verrons cela en pratique).

Remarquons les points importants:

1. Des "qutrits" au lieu des "qubits" ne changeraient rien de fondamental (et la nature offre des qutrits).
2. Faire les opérations de mesure à des stades intermédiaires ou bien à la fin ne change rien.
3. Faire les opérations de mesure dans une autre base est équivalent à faire des opérations de changement de base - qui sont unitaires car toutes les bases de mesure sont orthonormées - et donc à changer le circuit et faire la mesure dans la base computationnelle. Néanmoins, cela peut modifier la complexité (la réduire ou l'agrandir).
4. Commencer avec une autre entrée est aussi équivalent à ajouter des opérations unitaires et donc à changer le circuit et initialiser avec l'entrée $|0, \cdots, 0\rangle$. Néanmoins, cela peut modifier la complexité (la réduire ou l'agrandir).

5. Il existe aussi d'autres ensembles de portes universelles.
6. Le calcul quantique est réversible (pas de perte d'information, pas d'augmentation d'entropie et pas de dissipation de chaleur) car le circuit est une opération unitaire, tant que l'opération de mesure n'a pas été effectuée.
7. Une opération classique réversible peut être représentée par une opération unitaire quantique. En effet:

   $$\tilde{f}(x_1...x_n, y) = (x_1...x_n, y \oplus f(x_1...x_n))$$

   induit l'unitaire

   $$U_f|x_1...x_n, y\rangle = |x_1...x_n, y \oplus f(x_1...x_n)\rangle.$$

   Si la sortie $f(x_1...x_n)$ possède m composantes, on prend y avec m composantes et l'addition modulo 2 (le XOR) est faite composante par composante. On vérifie aisément que U_f est unitaire, en vérifiant que la matrice préserve le produit scalaire.
8. Le point précédent montre que tout calcul réversible classique est contenu dans le modèle des circuits quantiques.
9. La puissance du calcul quantique vient de l'action simultanée ou parallèle de l'évolution unitaire sur toutes les "suites classiques" $c_1 \dots c_n$ d'un état de n qubits. C'est ce qu'on appelle parfois le parallélisme quantique: celui-ci provient des principes 1 (superposition) et 5 (produit tensoriel des systèmes composés) mis ensemble (Section **3.2**). La complexité du calcul est donné par la taille du circuit. Le résultat est aléatoire. En général, il faut répéter un certain nombre de fois que le calcul quantique pour obtenir le résultat voulu avec une probabilité proche de 1. Cette répétition peut augmenter la complexité.

9.4 Le problème de Deutsch-Jozsa

L'algorithme de Deutsch et Jozsa est probablement l'algorithme quantique le plus simple. Dans sa version initiale, il fut inventé par David Deutsch dans son article fondateur[1] en 1985. L'algorithme fut ensuite amélioré par Deutsch et Jozsa (1992) et finalement par Cleve-Ekert-Macchiavello-Mosca (1998). Cette version définitive fait clairement apparaître que l'algorithme est le prototype d'une classe plus vaste, étudiée dans les chapitres ultérieurs, basée sur la *transformée de Fourier quantique* et le *principe d'interférence des chemins quantiques*. De plus, il constitue une très bonne illustration d'un cas où l'on peut tirer parti du *parallélisme quantique* de façon assez spectaculaire.

Formulons d'abord le problème à résoudre. Soit

$$f : \mathbf{F}_2^n \to \mathbf{F}_2, \qquad (x_1 \dots x_n) \mapsto f(x_1 \dots x_n) \in \{0, 1\}$$

une fonction booléenne dont on sait a priori qu'elle est *constante* ou *balancée*. La fonction est constante si l'image prend toujours la même valeur quel que soit l'argument $(x_1 \dots x_n)$. Notez qu'il y a 2^n arguments possibles. *Balancée* signifie que pour une moitié

[1] *Quantum Theory, the Church-Turing Principle and the Universal Quantum Computer* Proc. Roy. Soc of London A **400** pp 97-117 (1985)

des arguments (c.-à-d. 2^{n-1}), elle prend la valeur 0 et pour l'autre moitié (c.-à-d. 2^{n-1}) elle prend la valeur 1.

Le problème de Deutsch-Jozsa est un *problème de décision avec Oracle*. Cela veut dire que l'on ne connaît pas la fonction f mais que l'on a à disposition un *Oracle* (Figure 9.1) qui donne la réponse $f(x_1 \dots x_n)$ pour toute entrée $x_1 \dots x_n$ qui lui est soumise.

Figure 9.1 Oracle classique retournant les valeurs de la fonction f.

Le problème est de décider si f est constante. Le nombre de questions nécessaires à l'Oracle détermine la complexité temporelle de la résolution. Le but est de prendre la décision correcte en posant le moins de questions possibles.

Discutons d'abord la solution classique. Si on se limite à utiliser un *algorithme déterministe*, il existe des fonctions f pour lesquelles la complexité temporelle est de $2^{n-1}+1$, c'est-à-dire exponentielle par rapport à la taille des entrées. En effet, supposons que f soit constante et prenne la valeur 0 si bien que l'Oracle retourne toujours la réponse 0. Si l'on pose strictement moins de $2^{n-1}+1$ questions à l'Oracle, on n'a aucun moyen de savoir si la prochaine réponse sera aussi 0. Par contre, si à la $2^{n-1}+1$-ième question la réponse est 0 alors on peut affirmer avec certitude que la fonction n'est pas balancée, car si elle l'était, cette réponse aurait été 1.

Notons que si la fonction est balancée le nombre minimum de questions à poser est deux et le maximum est $2^{n-1}+1$. Le point important est qu'il existe des situations défavorables qui nécessitent un nombre exponentiel de questions.

Nous allons voir qu'il existe un algorithme quantique qui permet de déterminer si f est balancée ou constante avec une et une seule utilisation de l'Oracle et ceci quelle que soit la fonction f. Cela est assez spectaculaire. Notons que par rapport à un algorithme classique déterministe, le gain est exponentiel.

9.5 L'Oracle quantique

Pour chaque n on construit un circuit qui constitue l'algorithme. Les constituants du circuit sont la porte quantique de Hadamard et un *Oracle quantique*. Ici, nous discutons la modélisation de l'Oracle.

L'Oracle quantique est une porte donnée (par la Nature par exemple; c'est-à-dire que cela pourrait être un système physique) qui effectue l'opération *unitaire* suivante :

$$U_f|x_1 \dots x_n, y\rangle = |x_1 \dots x_n, y \oplus f(x_1 \dots x_n)\rangle$$

Cet opérateur agit sur un ket de Dirac à plusieurs qubits appartenant à l'espace

$$\underbrace{\mathbb{C}^2 \otimes \mathbb{C}^2 \dots \mathbb{C}^2}_{n \text{ fois}} \otimes \mathbb{C}^2$$

et donne un autre ket appartenant au même espace. C'est donc une matrice de dimensions $2^{n+1} \times 2^{n+1}$.

L'Oracle agit comme une porte *multicontrôle* (c.-à-d. qu'il y a plusieurs qubits de contrôle). Son circuit quantique est représenté sur la Figure **9.2**.

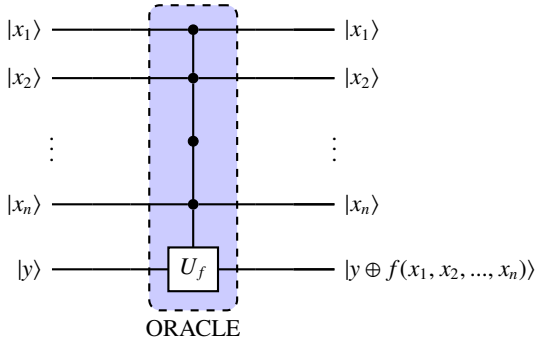


Figure 9.2 L'Oracle quantique retourne le résultat de f stocké dans les bits auxiliaires.

Vérifions que l'on a bien à faire à une matrice *unitaire*. Pour cela, il suffit de montrer qu'elle préserve le produit scalaire. Prenons deux vecteurs

$$U_f|x_1 \dots x_n, y\rangle \quad \text{et} \quad U_f|x'_1 \dots x'_n, y'\rangle$$

et effectuons les produits scalaires :

$$\begin{aligned}
\langle x'_1 \dots x'_n, y'|U_f^\dagger U_f|x_1 \dots x_n, y\rangle &= \langle x'_1 \dots x'_n, y' \oplus f(x'_1 \dots x'_n)|x_1 \dots x_n, y \oplus f(x_1 \dots x_n)\rangle \\
&= \langle x'_1|x_1\rangle \dots \langle x'_n|x_n\rangle \langle y' + f(x'_1 \dots x'_n)|y + f(x_1 \dots x_n)\rangle \\
&= \delta_{x_1 x'_1} \dots \delta_{x_n x'_n} \langle y' + f(x'_1 \dots x'_n)|y + f(x_1 \dots x_n)\rangle \\
&= \delta_{x_1 x'_1} \dots \delta_{x_n x'_n} \delta_{y' y}
\end{aligned}$$

D'autre part

$$\begin{aligned}
\langle x'_1 \dots x'_n, y'|x_1 \dots x_n, y\rangle &= \langle x'_1|x_1\rangle \dots \langle x'_n|x_n\rangle \langle y'|y\rangle \\
&= \delta_{x'_1 x_1} \dots \delta_{x'_n x_n} \delta_{y' y}
\end{aligned}$$

9.6 Algorithme quantique de Deutsch-Jozsa

Le circuit de l'algorithme quantique de Deutsch-Jozsa est donné à la Figure **9.3**. L'algorithme est initialisé dans l'état (instant t_0).

$$\underbrace{|0\rangle \otimes \dots \otimes |0\rangle}_{n \text{ fois}} \otimes |0\rangle = |\Psi_{in}\rangle$$

et se termine par une mesure dans la base computationnelle des n premiers qubits. Les projecteurs utilisés dans la mesure sont

$$P(b_1 \dots b_n) = |b_1 \dots b_n\rangle\langle b_1 \dots b_n|$$

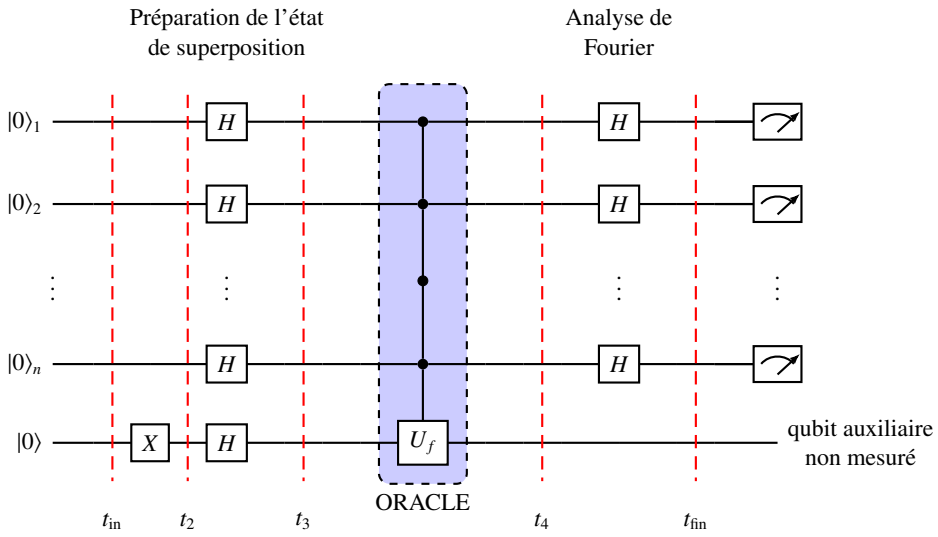


Figure 9.3 Circuit de l'algorithme de Deutsch-Jozsa.

Nous allons analyser l'évolution temporelle effectuée par un circuit aux instants t_{in}, t_2, t_3, t_4 et t_{fin}. L'état final est donné par

$$|\Psi_{\text{fin}}\rangle = U(t_{\text{fin}}, t_{\text{in}})|\Psi_{\text{in}}\rangle = U(t_{\text{fin}}, t_4)U(t_4, t_3)U(t_3, t_2)U(t_2, t_{\text{in}})|\Psi_{\text{in}}\rangle$$

Les opérations d'évolution de chaque tranche sont

$$\begin{aligned}
U(t_2, t_{\text{in}}) &= (\underbrace{\mathbb{1} \otimes \mathbb{1} \otimes \cdots \otimes \mathbb{1}}_{n \text{ fois}}) \otimes X \\
U(t_3, t_2) &= (\underbrace{H \otimes H \otimes \cdots \otimes H}_{n \text{ fois}}) \otimes H \\
U(t_4, t_3) &= U_f \\
U(t_{\text{fin}}, t_2) &= (\underbrace{H \otimes H \otimes \cdots \otimes H}_{n \text{ fois}}) \otimes \mathbb{1}
\end{aligned}$$

État à l'instant t_3.

$$\begin{aligned}
U(t_3,t_2)U(t_2,t_{\text{in}})|\Psi_{\text{in}}\rangle &= \underbrace{(H\otimes H\otimes\cdots\otimes H)}_{n\text{ fois}}\otimes HX\ \underbrace{|0\rangle\otimes\cdots\otimes|0\rangle}_{n\text{ fois}}\otimes|0\rangle \\
&= (\underbrace{H|0\rangle\otimes H|0\rangle\otimes\cdots\otimes H|0\rangle}_{n\text{ fois}})\otimes HX|0\rangle \\
&= \left(\frac{1}{\sqrt{2}}(|0\rangle+|1\rangle)\otimes\cdots\otimes\frac{1}{\sqrt{2}}(|0\rangle+|1\rangle)\right)\otimes H|1\rangle \\
&= \left(\frac{1}{\sqrt{2}}(|0\rangle+|1\rangle)\otimes\cdots\otimes\frac{1}{\sqrt{2}}(|0\rangle+|1\rangle)\right)\otimes\frac{1}{\sqrt{2}}(|0\rangle-|1\rangle) \\
&= \left(\frac{1}{2^{\frac{n}{2}}}\sum_{b_1\ldots b_n}|b_1\ldots b_n\rangle\right)\otimes\frac{1}{\sqrt{2}}(|0\rangle-|1\rangle)\equiv|\Psi_{t_3}\rangle
\end{aligned}$$

À cet instant, l'entrée est une *superposition cohérente de toutes les entrées classiques possibles*. Le dernier bit $\frac{1}{\sqrt{2}}(|0\rangle-|1\rangle)$ est un bit auxiliaire qui va servir à stocker le résultat de l'Oracle.

État à l'instant t_4.

$$\begin{aligned}
U(t_4,t_3)|\Psi_{t_3}\rangle &= U_f\frac{1}{2^{\frac{n}{2}}}\sum_{b_1\ldots b_n}|b_1\ldots b_n\rangle\otimes\left(\frac{1}{\sqrt{2}}|0\rangle-|1\rangle\right) \\
&= \frac{1}{2^{\frac{n}{2}}}\sum_{b_1\ldots b_n}\left(\frac{1}{\sqrt{2}}U_f|b_1\ldots b_n,0\rangle-\frac{1}{\sqrt{2}}U_f|b_1\ldots b_n,1\rangle\right) \\
&= \frac{1}{2^{\frac{n}{2}}}\sum_{b_1\ldots b_n}\left(\frac{1}{\sqrt{2}}|b_1\ldots b_n,f(b_1\ldots b_n)\rangle-\frac{1}{\sqrt{2}}|b_1\ldots b_n,\overline{f(b_1\ldots b_n)}\rangle\right)
\end{aligned}$$

Notons que si $f(b_1\ldots b_n)=0$ le terme entre parenthèses vaut

$$|b_1\ldots b_n\rangle\otimes\frac{|0\rangle-|1\rangle}{\sqrt{2}}$$

et que si $f(b_1\ldots b_n)=1$ le terme entre parenthèses vaut

$$|b_1\ldots b_n\rangle\otimes\frac{|1\rangle-|0\rangle}{\sqrt{2}}$$

On peut donc écrire l'état à l'instant t_4 comme suit:

$$\left(\frac{1}{2^{\frac{n}{2}}}\sum_{b_1\ldots b_n}(-1)^{f(b_1\ldots b_n)}|b_1\ldots b_n\rangle\right)\otimes\frac{|0\rangle-|1\rangle}{\sqrt{2}}$$

Cet état est à nouveau une superposition cohérente où l'Oracle a déphasé chaque entrée classique de 0 à π suivant que[2] l'image de f soit 0 ou 1.

État à l'instant t_{fin}. On applique finalement l'opérateur unitaire $\underbrace{H\otimes H\otimes\cdots\otimes H}_{n\text{ fois}}\otimes\mathbb{1}$,

[2] Car $e^{i0}=1$ et $e^{i\pi}=-1$.

ce qui donne par linéarité :

$$|\Psi_{\text{fin}}\rangle = \left(\frac{1}{2^{\frac{n}{2}}} \sum_{b_1 \dots b_n} (-1)^{f(b_1 \dots b_n)} H|b_1\rangle \otimes \cdots \otimes H|b_n\rangle \right) \otimes \frac{|0\rangle - |1\rangle}{\sqrt{2}}$$

Notons que

$$H|b\rangle = \frac{1}{\sqrt{2}} \left(|0\rangle + (-1)^b |1\rangle \right) = \frac{1}{\sqrt{2}} \sum_{c=0,1} (-1)^{cb} |c\rangle$$

si bien que

$$\begin{aligned} H|b_1\rangle \otimes \cdots \otimes H|b_n\rangle &= \frac{1}{\sqrt{2}} \left(|0\rangle + (-1)^{b_1} |1\rangle \right) \otimes \cdots \otimes \frac{1}{\sqrt{2}} \left(|0\rangle + (-1)^{b_n} |1\rangle \right) \\ &= \frac{1}{\sqrt{2}} \sum_{c_1=0,1} (-1)^{c_1 b_1} |c_1\rangle \otimes \cdots \otimes \sum_{c_n=0,1} (-1)^{c_n b_n} |c_n\rangle \\ &= \frac{1}{2^{\frac{n}{2}}} \sum_{c_1 \dots c_n} (-1)^{\sum_{i=1}^n b_i c_i} |c_1 \dots c_n\rangle \end{aligned}$$

Ainsi

$$|\Psi_{\text{fin}}\rangle = \sum_{c_1 \dots c_n} \left\{ \frac{1}{2^n} \sum_{b_1 \dots b_n} (-1)^{f(b_1 \dots b_n)} (-1)^{\sum_{i=1}^n b_1 c_i} \right\} |c_1 \dots c_n\rangle \otimes \frac{1}{\sqrt{2}} (|0\rangle - |1\rangle)$$

L'état final est à nouveau une superposition cohérente d'états classiques affectés d'amplitudes

$$\frac{1}{2^n} \sum_{b_1 \dots b_n} (-1)^{f(b_1 \dots b_n)} (-1)^{\sum_{i=1}^n b_i c_i}$$

Les amplitudes contiennent de l'information sur la fonction f. Si nous les connaissions toutes, nous pourrions en fait déterminer cette fonction. Mais la seule chose qui est à notre disposition est la totalité de l'état $|\Psi_{\text{fin}}\rangle$ (pris dans sa globalité) et *la seule chose que nous puissions faire, pour extraire de l'information, est une mesure*.

Dernière étape de l'algorithme: la mesure

Appliquons le postulat de la mesure. Si nous mesurons l'état des n premiers qubits dans la base computationnelle$\{|c_1 \dots c_n\rangle, c_i = 0, 1\}$, l'état est projeté (ou réduit) sur *un* des états $|c_1 \dots c_n\rangle$ avec probabilité (règle de Born ou postulat de la mesure)

$$\begin{aligned} \mathbb{P}\left[c_1 \dots c_n\right] &= [\text{carré de l'amplitude devant } |c_1 \dots c_n\rangle] \\ &= \frac{1}{2^{2n}} \left| \sum_{b_1 \dots b_n} (-1)^{f(b_1 \dots b_n)} (-1)^{\sum_{i=1}^n c_i b_i} \right|^2 \end{aligned}$$

La signification de cette assertion est la suivante : tant que la mesure est faite une fois sur un unique état $|\Psi_{\text{fin}}\rangle$ l'état final observé est un des états $|c_1 \dots c_n\rangle$ et il n'y a aucun moyen

de prédire lequel. Si l'on dispose d'un ensemble d'états $|\Psi_{\text{fin}}\rangle$, en répétant l'expérience plusieurs fois, la fréquence des observations $|c_1 \dots c_n\rangle$ est donnée par $\mathbb{P}[c_1 \dots c_n]$.

Calculons cette probabilité. Si f est constante on trouve

$$\begin{aligned}
\mathbb{P}[c_1 \dots c_n] &= \frac{1}{2^{2n}} \left| \sum_{b_1 \dots b_n} (-1)^{\sum_{i=1}^n c_i b_i} \right|^2 \\
&= \frac{1}{2^{2n}} \left| \sum_{b_1 \dots b_n} \prod_{i=1}^n (-1)^{c_i b_i} \right|^2 \\
&= \frac{1}{2^{2n}} \left| \prod_{i=1}^n \left((-1)^{c_i 0} + (-1)^{c_i 1} \right) \right|^2 \\
&= \begin{cases} 1 & \text{si } (c_1 \dots c_n) = (0 \dots 0) \\ 0 & \text{dans tous les autres cas} \end{cases}
\end{aligned}$$

Donc si f est constante nous observerons certainement $(0 \dots 0)$ (c.-à-d. avec probabilité 1) en faisant une seule expérience ! Par contre, si f est balancée on constate que

$$\mathbb{P}[0 \dots 0] = \frac{1}{2^{2n}} \left| \sum_{b_1 \dots b_n} (-1)^{f(b_1 \dots b_n)} \right|^2 = 0$$

et on n'observera certainement pas $(0 \dots 0)$.

En conclusion : après le processus de mesure si $(0 \dots 0)$ est observé on peut conclure "f constante" et si autre chose est observé on peut conclure "f balancée". Remarquablement, il suffit de faire l'expérience et d'utiliser l'Oracle quantique une seule fois ! Ceci n'est pas typique des autres algorithmes quantiques: nous verrons par exemple que pour l'algorithme de Shor il faut faire l'expérience plusieurs fois pour amplifier la probabilité de succès.

Note sur la complexité de l'algorithme. En général, nous devons distinguer la complexité du circuit et celle de l'algorithme proprement dit. Dans ce cours, la complexité des circuits sera mesurée en termes de deux grandeurs, la "profondeur" et la "largeur". La taille du circuit est alors définie comme le produit $(profondeur) \times (largeur)$. Ici le circuit contient 3 tranches de temps intermédiaires: on dit que ce circuit à une profondeur égale à 3 (ce qui est important ici, c'est qu'elle est $O(1)$). Si le temps élémentaire requis pour effectuer une porte quantique (par RMN par ex) est de τ alors le temps de calcul du circuit est de 3τ. La largeur du circuit est donnée par le nombre de bits d'entrée plus le nombre de bits auxiliaires, ici $n+1$, et représente le nombre de calculs que le circuit quantique effectue à chaque tranche temporelle. La taille du circuit de DJ est donc $3(n+1) = O(n)$. Finalement, quelle est la complexité de l'algorithme lui-même ? Puisqu'on peut résoudre le problème de DJ en posant une seule question à l'Oracle, l'algorithme possède un temps de calcul $O(1)$ avec un circuit de taille $O(n)$.

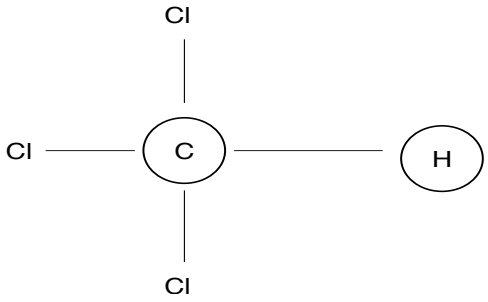


Figure 9.4 L'algorithme DJ a été réalisé par Résonance Magnétique Nucléaire sur les molécules de $CHCl_3$. On agit sur deux qubits associés aux spins nucléaires des atomes H et C entourés

9.7 Quelques remarques sur les réalisations expérimentales

Nous reviendrons sur les réalisations expérimentales à la fin du cours. Ce paragraphe peut être omis en première lecture.

L'algorithme de DJ fut un des premiers algorithmes quantique à être réalisé expérimentalement (2001) pour le cas $n = 1$ par la résonance magnétique nucléaire. Cette expérience utilise un liquide de $CHCl_3$ (le chloroforme, fig. 9.4). Pour $n = 1$ il faut deux bits quantiques, un pour l'entrée et un bit auxiliaire pour stocker le résultat de l'Oracle. Ceux-ci sont matérialisés par les spins nucléaires de l'atome d'hydrogène H et de carbone C. Les portes de Hadamard et l'Oracle peuvent être réalisés en manipulant ces deux spins nucléaires par des champs magnétiques radiofréquences. Un des problèmes principaux est de préparer toutes les molécules du liquide dans l'état initial $|00\rangle$. En effet, on ne peut pas se débarrasser complètement des fluctuations thermiques qui induisent des transitions vers les autres états pour une fraction des molécules du liquide. Pour augmenter n, il faut prendre de plus grosses molécules avec des atomes appropriés. Ceci pose un problème ("scalability problem") parce que plus la molécule est grosse, plus les niveaux d'énergie sont nombreux et rapprochés (le spectre devient continu en quelque sorte) et il devient de plus en plus difficile de manipuler sélectivement les bits quantiques grâce à des transitions de radiofréquence. Plus récemment, l'algorithme de DJ a aussi été réalisé grâce aux technologies des pièges à ions et des cavités résonantes (cavity QED). Toutes ces expériences sont limitées à un faible nombre de qubits (< 10 qubits).

10 Algorithme de Simon

Dans ce chapitre, nous étudions un algorithme qui contient déjà les ingrédients essentiels qui interviendront dans l'algorithme de Shor. (Chapitre **11**) En fait, ce dernier peut être vu comme l'une des généralisations possibles de celui de Simon. Le problème classique de Simon est un problème avec oracle qui ne peut pas être résolu par un algorithme classique aléatoire non-exponentiel (une preuve de cette assertion existe). Par contre, comme nous le verrons, il existe un algorithme quantique (aléatoire) qui résout le problème et qui possède des complexités temporelles et spatiales polynomiales. Ainsi, nous connaissons au moins un problème avec oracle pour lequel il est mathématiquement prouvé, que le calcul quantique offre une accélération exponentielle par rapport au temps de calcul classique. Cela n'implique pas que ce soit vraiment le cas aussi pour des problèmes sans oracle, bien qu'il soit assez naturel de le conjecturer (à cause notamment de l'algorithme de Shor).

10.1 Le problème de Simon

On se donne une fonction de n variables booléennes telle que

$$f : \mathbf{F}_2^n \to X, \qquad (x_1 \dots x_n) \mapsto f(x_1 \dots x_n) \in X$$

où X est un ensemble fini et

$$f(x_1 \dots x_n) = f(y_1 \dots y_n) \;\Leftrightarrow\; \vec{x} = \vec{y} \;\text{ ou }\; \vec{x} - \vec{y} = \vec{a}$$

pour un vecteur $\vec{a}$ fixé. Pour nous familiariser avec une telle fonction, on peut se faire l'image suivante. Il y a 2^n vecteurs d'entrée $\vec{x}$. Prenons un tel $\vec{x}$ et considérons son partenaire $\vec{x}+\vec{a}$. La fonction satisfait $f(\vec{x}) = f(\vec{x}+\vec{a})$. De plus, si $\vec{x}$ et $\vec{y}$ sont différents mais ne sont pas partenaires (c'est-à-dire $\vec{x}-\vec{y} \neq \vec{a}$) alors $f(\vec{x}) \neq f(\vec{y})$. Le nombre de valeurs que prend la fonction est égale au nombre de paires $(\vec{x}, \vec{x}+\vec{a})$[1], c'est-à-dire $\frac{2^n}{2} = 2^{n-1}$. Ainsi la cardinalité de X est nécessairement $|X| = 2^{n-1}$. Un exemple est donné à la Figure **10.1**.

[1] Mathématiquement, chacune de ces paires forme une classe d'équivalence pour la relation d'équivalence suivante: $\vec{x} \sim \vec{y} \;\Leftrightarrow\; f(\vec{x}) = f(\vec{y})$ et l'ensemble de ces classes d'équivalence forme une partition de $\mathbf{F}_2^n$.

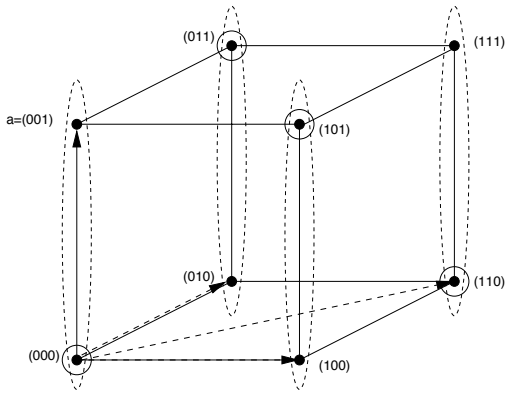


Figure 10.1 Problème de Simon dans le cas $n = 3$. On cherche le vecteur $\vec{a} = (0, 0, 1)$. Chaque ellipse représente une classe d'équivalence sur laquelle la fonction f est constante. Les points entourés sont des représentants arbitraires des classes d'équivalences. Le circuit quantique retourne les vecteurs du plan perpendiculaire à $\vec{a}$

Le problème classique à résoudre est le suivant. On dispose d'un oracle :

$$\vec{x} \longrightarrow \boxed{f} \longrightarrow f(\vec{x})$$

et on doit déterminer $\vec{a}$ en posant des questions à l'oracle. Avec un algorithme classique déterministe, on est assuré d'avoir la réponse en posant $2^{n-1} + 1$ questions à l'oracle (au pire), et en 2 questions (au mieux). Mais de façon générique, le "temps de calcul" de ce procédé est exponentiel. Considérons maintenant l'algorithme aléatoire suivant:

1. Soumettre q paires de questions $(\vec{x}_1^{(1)}, \vec{x}_2^{(1)}), \ldots, (\vec{x}_1^{(q)}, \vec{x}_2^{(q)})$ à l'oracle. On soumet les paires aléatoirement uniformément sur l'ensemble des paires (on s'assure que deux vecteurs d'une même paire sont distincts);
2. Si pour au moins une paire i la réponse est $f(\vec{x}_1^{(i)}) = f(\vec{x}_2^{(i)})$ en déduire $\vec{a} = \vec{x}_1^{(i)} - \vec{x}_2^{(i)}$ et déclarer: SUCCÈS;
3. Sinon (pour toutes les paires la réponse est $f(\vec{x}_1^{(i)}) \neq f(\vec{x}_2^{(i)})$) déclarer: ÉCHEC;

Calculons le temps de calcul nécessaire à une probabilité de succès d'au moins $1 - \epsilon$. D'après le principe d'exclusion,

$$\mathbb{P}\,[\text{SUCCES}] \leq \sum_i \mathbb{P}\left[f(\vec{x}_1^{(i)}) = f(\vec{x}_2^{(i)})\right]$$

Pour chaque terme on a $\mathbb{P}\left[f(\vec{x}_1^{(i)}) = f(\vec{x}_2^{(i)})\right] = \frac{2^{n-1}}{2^{n-1}(2^n-1)}$ puisque le nombre total de paires est $\frac{2^n(2^n-1)}{2}$ et le nombre total de paires avec égalité est 2^{n-1}. Donc

$$\mathbb{P}\,[\text{SUCCES}] \leq \frac{q}{2^n - 1}$$

Si on demande $\mathbb{P}\,[\text{SUCCES}] \geq 1 - \epsilon$, on trouve

$$q \geq (1 - \epsilon)(2^n - 1) \tag{10.1}$$

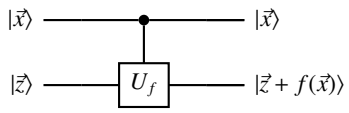
Tant que ϵ est fixé, il faut un nombre exponentiel de questions avec cet algorithme spécifique.

En fait, il est possible de montrer que tout algorithme classique aléatoire requiert un nombre exponentiel de questions[2] et qu'essentiellement on ne pas faire mieux que ci-dessus (avec le calcul classique).

Par contre, nous verrons qu'il existe un circuit quantique simple, fonctionnant avec un oracle quantique, qui permet de déterminer $\vec{a}$ avec probabilité $1-\epsilon$ en un temps polynomial $O(\ln(\epsilon)\,\mathrm{poly}(n))$ (et comme d'habitude, la "complexité interne de l'oracle" n'est pas prise en compte). L'oracle quantique est représenté par l'opérateur unitaire

$$U_f|\vec{x},\vec{z}\rangle = |\vec{x},\vec{z}+f(\vec{x})\rangle$$

Comme $f(\vec{x})$ prend 2^{n-1} valeurs, il faut $n-1$ bits pour représenter ses valeurs et donc pour $\vec{z}$ aussi.



Nous pouvons traiter en fait un problème un peu plus général par le même circuit quantique. Sa formulation est la suivante. On peut penser à $\mathbf{F}_2^n$ comme à l'espace vectoriel des vecteurs binaires à n composantes. Soit H un sous-espace vectoriel de dimension k (si on remplaçait $\mathbf{F}_2^n$ par $\mathbb{R}^n$ alors H serait un hyperplan à k dimensions passant par l'origine, comme montré à la Figure **10.2**).

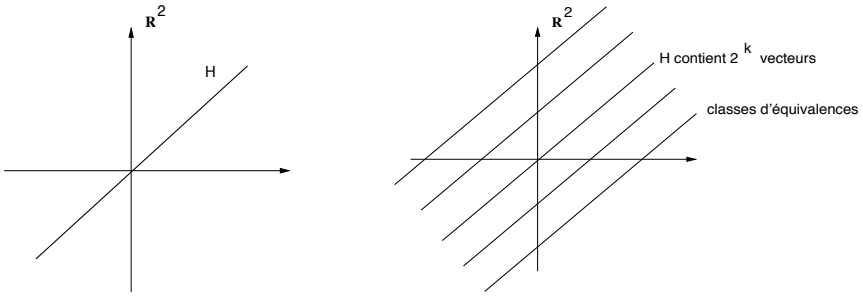


Figure 10.2 À gauche: Le sous-espace vectoriel recherché H (si on remplacait $\mathbf{F}_2^n$ par $\mathbb{R}^n$). À droite: les classes d'équivalence sur lesquelles f est constante (si on remplacait $\mathbf{F}_2^n$ par $\mathbb{R}^n$).

On se donne un oracle qui sait calculer

$$f : \mathbf{F}_2^n \to X, \qquad \vec{x} \mapsto f(\vec{x})$$

tel que $f(\vec{x}) = f(\vec{y}) \Leftrightarrow \vec{x}-\vec{y} \in H$. Dans l'exemple précédent $H = \{0,\vec{a}\}$ était le sous-espace vectoriel unidimensionnel généré par $\vec{a}$. Le but est de déterminer une base

[2] Voir A. Y. Kitaev, A. H. Shen, M. N. Vyalvi "Classical and Quantum Computation" Graduate Studies in Mathematics vol 47, American Mathematical Society (2002) pp 117.

de vecteurs de H en posant le moins de questions possible à l'oracle. Quelle est la cardinalité de X ? Si $\dim H = k$ alors H possède 2^k vecteurs binaires. En effet, tout vecteur de H peut s'écrire $b_1\vec{h_1} + b_2\vec{h_2} + \ldots + b_k\vec{h_k}$ avec $b_i = 0, 1$. De plus, $\mathbf{F}_2^n$ n'est rien d'autre que l'ensemble de tous les "hyperplans" translatés de H. Puisque $f(\vec{x}) \neq f(\vec{y})$ si $\vec{x}$ et $\vec{y}$ n'appartiennent pas au même hyperplan, on doit avoir $|X| = \frac{2^n}{2^k} = 2^{n-k}$.

10.2 Circuit quantique pour l'algorithme de Simon

Le circuit de l'algorithme est représenté à la Figure **10.3**. Sa largeur est $2n - k$ et sa profondeur est constante (égale à 3). En d'autres termes, la taille du circuit est $O(n)$. Ci-dessous nous analysons en détail l'évolution de l'état quantique au cours du temps.

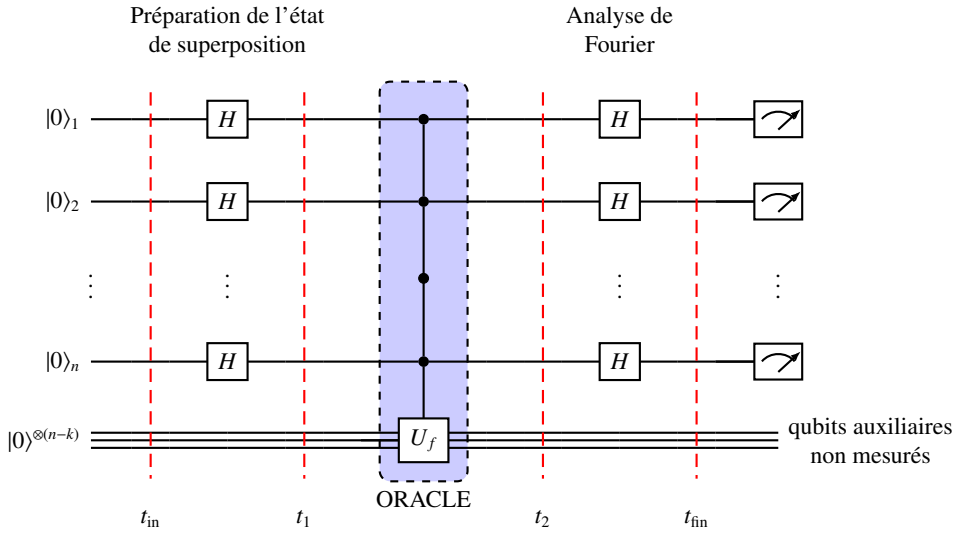


Figure 10.3 Circuit quantique pour l'algorithme de Simon. Au total il y a $n + (n - k)$ qubits. On ne fait pas d'opération de mesure sur les $n - k$ qubits auxiliaires.

L'état initial à l'instant t_{in} est

$$|\Psi_{\text{in}}\rangle = \underbrace{|0\rangle \otimes \ldots \otimes |0\rangle}_{n \text{ qubits}} \otimes \underbrace{|\vec{0}\rangle}_{n-k \text{ qubits}}$$

À l'instant t_1, on obtient l'état

$$\begin{aligned} U(t_1, t_{\text{in}})|\Psi_{\text{in}}\rangle &= \Big(\underbrace{H \otimes \ldots \otimes H}_{n \text{ fois}} \otimes \mathbb{1}_{n-k}\Big) \underbrace{|0\rangle \otimes \ldots \otimes |0\rangle}_{n \text{ qubits}} \otimes \underbrace{|\vec{0}\rangle}_{n-k \text{ qubits}} \\ &= H|0\rangle \otimes \ldots \otimes H|0\rangle \otimes |\vec{0}\rangle \\ &= \frac{1}{2^{\frac{n}{2}}} \sum_{x_1 \ldots x_n} |x_1 \ldots x_n, \vec{0}\rangle \end{aligned}$$

Pour obtenir l'état à l'instant t_2, on agit avec $U(t_2,t_1) = U_f$.

$$\begin{aligned} U(t_2,t_1)U(t_1,t_{\text{in}})|\Psi_{\text{in}}\rangle &= \frac{1}{2^{\frac{n}{2}}}\sum_{x_1\dots x_n} U_f|x_1\dots x_n,\vec{0}\rangle \\ &= \frac{1}{2^{\frac{n}{2}}}\sum_{x_1\dots x_n}|x_1\dots x_n, f(x_1\dots x_n)\rangle \end{aligned}$$

Avant de procéder plus avant, analysons la structure de cette somme. Tout vecteur $x_1\dots x_n \in \mathbf{F}_2^n$ peut se décomposer en

$$\vec{v}+\vec{h}$$

où $\vec{v}$ caractérise l'hyperplan (il y en a 2^{n-k}) et $\vec{h} \in H$ (voir Figure **10.2**). On peut toujours choisir un représentant pour chaque hyperplan. Dans la suite on se donne un ensemble de tels représentants $\vec{v}$, et les $\sum_{\vec{v}}$ portent sur cet ensemble de représentants. Donc

$$\begin{aligned} U(t_2,t_1)|\Psi_{\text{in}}\rangle &= \frac{1}{2^{\frac{n}{2}}}\sum_{\vec{v}}\sum_{\vec{h}\in H}|\vec{v}+\vec{h}, f(\vec{v}+\vec{h})\rangle \\ &= \frac{1}{2^{\frac{n}{2}}}\sum_{\vec{v}}\sum_{\vec{h}\in H}|\vec{v}+\vec{h}, f(\vec{v})\rangle \end{aligned}$$

où on a utilisé $f(\vec{v}+\vec{h}) = f(\vec{v})$. Il reste à calculer l'état à l'instant t_{fin} :

$$U(t_{\text{fin}},t_{\text{in}})|\Psi_{\text{in}}\rangle = \frac{1}{2^{\frac{n}{2}}}\sum_{\vec{v}}\sum_{\vec{h}\in H} H^{\otimes n}\otimes \mathbb{1}_{n-k}|\vec{v}+\vec{h}, f(\vec{v})\rangle$$

Calculons

$$\begin{aligned} H^{\otimes n}|\vec{v}+\vec{h}\rangle &= H|v_1+h_1\rangle\otimes\dots\otimes H|v_n+h_n\rangle \\ &= \frac{1}{2^{\frac{n}{2}}}\left(|0\rangle+(-1)^{v_1+h_1}|1\rangle\right)\otimes\dots\otimes\left(|0\rangle+(-1)^{v_n+h_n}|1\rangle\right) \\ &= \frac{1}{2^{\frac{n}{2}}}\sum_{y_1\dots y_n}(-1)^{\sum_{i=1}^n (v_i+h_i)y_i}|y_1\dots y_n\rangle \end{aligned}$$

L'état à l'instant final (avant la mesure) est :

$$\begin{aligned} |\Psi_{\text{fin}}\rangle &= \frac{1}{2^n}\sum_{\vec{v}}\sum_{\vec{h}\in H}\sum_{\vec{y}\in\mathbf{F}_2^n}(-1)^{(\vec{v}+\vec{h})\cdot\vec{y}}|\vec{y}\rangle\otimes|f(\vec{v})\rangle \\ &= \frac{1}{2^n}\sum_{\vec{v}}\sum_{\vec{y}\in\mathbf{F}_2^n}(-1)^{\vec{v}\cdot\vec{y}}|\vec{y}\rangle\otimes|f(\vec{v})\rangle\left\{\sum_{\vec{h}\in H}(-1)^{\vec{h}\cdot\vec{y}}\right\} \end{aligned}$$

Si $\vec{y} \in H^\perp$ (l'hyperplan perpendiculaire[3] à H) on a

$$\sum_{\vec{h}\in H}(-1)^{\vec{h}\cdot\vec{y}} = \sum_{\vec{h}\in H} 1 = |H| = 2^k$$

[3] Par définition $H^\perp = \{\vec{y} \mid \vec{y}\cdot\vec{h} \equiv 0 \mod 2 \quad \forall \vec{h}\in H\}$

Par contre si $\vec{y} \notin H^\perp$ il existe un vecteur non-nul de H, $\vec{h_0}$, tel que

$$\vec{y} \cdot \vec{h_0} \neq 0 \mod 2$$

Grâce au changement de variable $\vec{h} = \vec{h'} + \vec{h_0}$, on obtient :

$$\begin{aligned}\sum_{\vec{h} \in H} (-1)^{\vec{h} \cdot \vec{y}} &= \sum_{\vec{h'} \in H} (-1)^{(\vec{h'} + \vec{h_0}) \cdot \vec{y}} \\ &= (-1)^{\vec{h_0} \cdot \vec{y}} \sum_{\vec{h'} \in H} (-1)^{\vec{h'} \cdot \vec{y}}\end{aligned}$$

Notons que $(-1)^{\vec{h_0} \cdot \vec{y}} = -1$ (car $\vec{y} \cdot \vec{h_0} \neq 0 \Rightarrow \vec{y} \cdot \vec{h_0} = 1$ dans le cas binaire). Cela implique

$$\sum_{\vec{h} \in H} (-1)^{\vec{h} \cdot \vec{y}} = 0 \quad \text{pour} \quad \vec{y} \notin H^\perp$$

En conclusion, seuls les $\vec{y} \in H^\perp$ contribuent à l'état final qui peut s'écrire,

$$|\Psi_{\text{fin}}\rangle = \frac{1}{2^{n-k}} \sum_{\vec{y} \in H^\perp} \sum_{\vec{v}} (-1)^{\vec{v} \cdot \vec{y}} |\vec{y}\rangle \otimes |f(\vec{v})\rangle$$

À ce stade, le circuit quantique produit l'état $|\Psi_{\text{fin}}\rangle$. Pour en extraire une information, il faut encore faire une mesure dans la base computationnelle sur les n premiers bits. Les projecteurs associés à ce processus de mesure sont

$$P_{\vec{y}} = |\vec{y}\rangle\langle\vec{y}| \otimes \mathbb{1}_{n-k}$$

Où $\mathbb{1}$ représente le fait que les qubits auxiliaires ne sont pas mesurés. L'état après la mesure est

$$\frac{P_{\vec{y}} |\Psi_{\text{fin}}\rangle}{\| P_{\vec{y}} |\Psi_{\text{fin}}\rangle \|}$$

avec la probabilité

$$\mathbb{P}\left[\vec{y}\right] = \langle \Psi_{\text{fin}} | P_{\vec{y}} | \Psi_{\text{fin}} \rangle$$

- Si $\vec{y} \notin H^\perp$ on a,

$$P_y |\Psi_{\text{fin}}\rangle = 0$$

  Donc

$$\mathbb{P}\left[\vec{y}\right] = 0 \quad \text{si} \quad \vec{y} \notin H^\perp$$

- Si $\vec{y} \in H^\perp$ on a,

$$\begin{aligned}P_{\vec{y}} |\Psi_{\text{fin}}\rangle &= \frac{1}{2^{n-k}} \sum_{\vec{y} \in H^\perp} \sum_{\vec{v}} (-1)^{\vec{v} \cdot \vec{y}} |\vec{y}\rangle \langle\vec{y}|\vec{y}\rangle \otimes |f(\vec{v})\rangle \\ &= |\vec{y}\rangle \otimes \frac{1}{2^{n-k}} \sum_{\vec{v}} (-1)^{\vec{v} \cdot \vec{y}} |f(\vec{v})\rangle\end{aligned}$$

Il s'ensuit que,

$$\mathbb{P}\left[\vec{y}\right] = \frac{1}{2^{n-k}} \sum_{\vec{y} \in H^{\perp}} \sum_{\vec{v}} (-1)^{\vec{v}\cdot\vec{y}} \langle \vec{y} | \vec{y} \rangle \sum_{\vec{v}} (-1)^{\vec{v}\cdot\vec{y}} \langle f(\vec{v}) | f(\vec{v}) \rangle$$
$$= \begin{cases} \frac{1}{2^{n-k}} & \text{si} \quad \vec{y} \in H^{\perp} \\ 0 & \text{si} \quad \vec{y} \notin H^{\perp} \end{cases}$$

Donc après une mesure on obtient un vecteur uniformément aléatoire entièrement contenu dans $H^{\perp}$.

10.3 Analyse probabiliste de l'algorithme

Nous allons analyser le temps de calcul de l'algorithme suivant:

1. Initialiser le circuit avec $|0_n\rangle \otimes |0_{n-k}\rangle$ et faire une demande à l'oracle quantique.
2. Mesurer l'état final $|\Psi_{\text{fin}}\rangle$. la mesure fournit $\vec{y} \in H^{\perp}$.
3. Itérer $n-k$ fois les points 1 et 2 pour obtenir $\vec{y}_1, \ldots, \vec{y}_{n-k}$, des vecteurs de $H^{\perp}$.
4. Si $\vec{y}_1, \ldots, \vec{y}_{n-k}$ forme un ensemble de vecteurs indépendants, c'est une base de $H^{\perp}$. Calculer la base duale de H, $\vec{h}_1, \ldots, \vec{h}_k$ par les méthodes usuelles d'algèbre linéaire (p.ex. élimination gaussienne). Output : "SUCCÈS et donner la base de H;
5. Si $\vec{y}_1, \ldots, \vec{y}_{n-k}$ n'est pas un ensemble de vecteurs indépendants, donner "ÉCHEC".

Montrons que la probabilité de succès lors d'un round $1 \to 5$ est $\geq \frac{1}{4}$. Pour cela, il faut montrer que quand on tire uniformément des vecteurs aléatoires avec remise dans $H^{\perp}$, on a

$$\mathbb{P}\left[\vec{y}_1, \ldots, \vec{y}_{n-k} \text{ indépendants}\right] \geq \frac{1}{4}$$

.

Preuve : Supposons que nous avons tiré $\vec{y}_1$. Il faut que $\vec{y}_1 \neq \vec{0}$ et ceci a lieu avec probabilité $= \frac{2^{n-k}-1}{2^{n-k}}$. Maintenant tirons $\vec{y}_2$. Il faut que $\vec{y}_2 \neq \{\vec{0}, \vec{y}_1\}$ et ceci a lieu avec une probabilité $= \frac{2^{n-k}-2}{2^{n-k}}$. Ensuite tirons $\vec{y}_3$. Il faut que $\vec{y}_3 \notin \text{Vect}\{\vec{y}_1, \vec{y}_2\}$. Or, $\text{Vect}\{\vec{y}_1, \vec{y}_2\}$ contient 2^2 vecteurs. Donc $\vec{y}_3 \notin span\{\vec{y}_1, \vec{y}_2\}$ avec probabilité $= \frac{2^{n-k}-2^2}{2^{n-k}}$. En itérant cette réflexion, on trouve :

$$\mathbb{P}\left[\vec{y}_1, \ldots, \vec{y}_{n-k} \text{ indépendants}\right]$$
$$= \frac{2^{n-k}-2^0}{2^{n-k}} \cdot \frac{2^{n-k}-2^1}{2^{n-k}} \cdot \frac{2^{n-k}-2^2}{2^{n-k}} \cdot \cdots \cdot \frac{2^{n-k}-2^{m-k-1}}{2^{n-k}}$$
$$= \prod_{j=1}^{n-k} \left(1 - \frac{1}{2^j}\right) = \exp\left(\sum_{j=1}^{n-k} \ln\{1 - \frac{1}{2^j}\}\right)$$

En utilisant que $\ln(1-x) \geq -x(2\ln 2)$ quand $0 \leq x \leq \frac{1}{2}$ (Figure **10.4**), cette probabilité

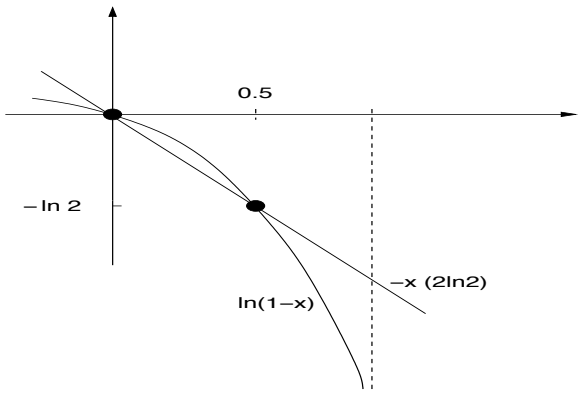


Figure 10.4 Une inégalité utile

est supérieure à :

$$\geq \exp\{-(2\ln 2)\sum_{j=1}^{n-k}\frac{1}{2^j}\}$$
$$\geq \exp\{-(2\ln 2)\sum_{j=1}^{\infty}\frac{1}{2^j}\}$$
$$= \exp(-2\ln 2) = \frac{1}{4}$$

Puisque la probabilité de succès lors d'un round est $\geq \frac{1}{4}$, la probabilité d'avoir au moins un succès lors de T rounds est

$$\mathbb{P}\,[\text{au moins un succès avec } T \text{ rounds}] = 1 - (1 - \mathbb{P}\,[\text{succès 1 round}])^T$$
$$\geq 1 - \left(1 - \frac{1}{4}\right)^T = 1 - \left(\frac{3}{4}\right)^T$$

Celle-ci est $\geq 1 - \epsilon$ si et seulement si :

$$1 - \left(\frac{3}{4}\right)^T \geq 1 - \epsilon$$

$$\Leftrightarrow \epsilon \leq \left(\frac{3}{4}\right)^T$$

$$\Leftrightarrow T \geq \frac{|\ln \epsilon|}{|\ln \frac{3}{4}|}$$

En résumé, on a une probabilité de succès $\geq 1 - \epsilon$ avec $O(|\ln \epsilon|)$ rounds. Le temps de calcul de chaque round est: $O(n)$ pour l'itération des points 1 et 2, plus $O\left(n^3\right)$ pour le calcul classique de la base duale. Donc nous avons un algorithme probabiliste polynomial avec temps de calcul $O\left(|\ln \epsilon|\, n^3\right)$. Rappelons que la taille du circuit est $O(n)$.

Note sur le calcul de la base duale

Soit $\vec{y}_1, \ldots, \vec{y}_{n-k}$ une base de $H^\perp$ donnée par l'algorithme. Les vecteurs de H sont orthogonaux, c'est-à-dire satisfont au système d'équations :

$$\begin{aligned} \vec{h}^T \cdot \vec{y}_1 &= 0 \\ &\vdots \\ \vec{h}^T \cdot \vec{y}_{n-k} &= 0 \end{aligned}$$

Ce système possède K solutions distinctes $\vec{h}_1, \ldots, \vec{h}_k$ formant une base de H. On peut écrire le système sous forme matricielle

$$\vec{h}^T \cdot \underbrace{[\vec{y}_1 \ldots \vec{y}_{n-k}]}_{\text{matrice des vecteurs colonne } n\times(n-k)} = 0$$

Donc $\vec{h}$ est dans le noyau d'une matrice de rang $n-k$. Puisque $\dim(\text{noyau}) + \underbrace{\dim(\text{image})}_{\text{rang}} = n$ on doit avoir $\dim(\text{noyau}) = k$. C'est-à-dire qu'il existe k solutions indépendantes $\vec{h}_1 \ldots \vec{h}_k$ pour le système d'équations ci-dessus. Celles-ci peuvent être effectivement trouvées en résolvant le système par élimination gaussienne, ce qui requiert en général $O\left(n^3\right)$ opérations.

11 Factorisation et Algorithme de Shor

L'un des développements les plus spectaculaires du calcul quantique est l'algorithme de Shor. Il s'agit d'un algorithme de factorisation pour les entiers de complexité polynomiale dans la taille de l'entier.

Le théorème fondamental de l'arithmétique nous assure que tout entier N peut être décomposé de façon unique en un produit de nombres premiers. Étant donné les facteurs de N, il est facile de vérifier que le produit de ces facteurs redonne N. Plus précisément, supposons que $N = p \cdot q$ avec p et q deux nombres premiers à $O(b)$ bits. Si p et q sont connus, la vérification $N = p \cdot q$ peut se faire facilement avec $O\left(b^2\right)$ opérations. Par contre, étant donné un entier général N avec $O(b)$ bits, on ne connaît pas d'algorithme polynomial permettant de calculer p et q (on ne sait même pas s'il en existe un ou non). On connaît plusieurs algorithmes plus rapides que $O\left((1+\epsilon)^b\right)$ pour tout $\epsilon > 0$. Le meilleur algorithme connu possède un temps de calcul $O\left(\exp\left(\frac{64}{9}b\right)^{\frac{1}{3}}(\log b)^{\frac{2}{3}}\right)$. Concrètement, un des records[1] relativement récents de factorisation atteint le 12 décembre 2009 pour un nombre à 232 décimales (b=768 bits) a utilisé des centaines de processeurs sur deux années de calcul.

Comme nous le verrons plus tard, l'algorithme de Shor permet une factorisation en temps polynomial avec un circuit quantique de taille polynomiale. L'algorithme de Shor résout en fait un autre problème de théorie des nombres appelé *la recherche de l'ordre*. Il est connu depuis 1976 (environ) que la factorisation peut se réduire à la recherche de l'ordre.

Pour analyser la complexité de l'algorithme de Shor nous aurons aussi besoin de quelques notions supplémentaires sur les fractions continuées et la fonction d'Euler. Celles-ci ainsi que la réduction de la factorisation à la recherche de l'ordre sont exposées dans la section suivante.

11.1 Une parenthèse de théorie de nombres

Factorisation basée sur la recherche de l'ordre

Soit $N = p_1^{e_1} p_2^{e_2} \dots p_k^{e_k}$ avec $p_i \neq 2$ et $k \geq 2$. En d'autres termes, N ne contient pas de puissance de 2 et N n'est pas la puissance d'un nombre premier unique ($N \neq p^e$). Notez que les puissances de 2 sont aisément extraites, car il est facile de voir si un entier et

[1] Voir http://en.wikipedia.org/wiki/RSA_numbers#RSA-768

pair et de diviser par 2. De plus, si $N = p^e$ il existe une méthode efficace pour trouver p et e.

Algorithme de factorisation basé sur la recherche de l'ordre.

a. Choisir aléatoirement uniformément $a \in \{2, \ldots, N-1\}$ et calculer

$$d = \text{PGCD}\,(a, N)$$

grâce à l'algorithme d'Euclide (de complexité $O\left((\log_2 N)^3\right)$).

b. Si $d > 1$ nous avons un facteur non-trivial de N car d divise N. On garde ce facteur et on retourne à l'étape **a**.

c. Si $d = 1$ (c'est-à-dire que a et N sont premiers entre eux) on calcule *le plus petit* entier r tel que

$$a^r = 1 \pmod N.$$

Cet entier s'appelle *l'ordre de a* (mod N) aussi noté $r = \text{Ord}_N(a)$. *Pour cette étape, on ne connaît pas d'algorithme classique polynomial. C'est l'étape qui sera traitée par l'algorithme de Shor.*

d. Supposons que r soit impair. Output `Fail` et retourner à l'étape **a**.

e. Si r est pair alors on sait que:

$$a^r - 1 = \left(a^{\frac{r}{2}} - 1\right)\left(a^{\frac{r}{2}} + 1\right)$$

Notez que N divise $a^r - 1$. Donc il y a a priori 3 possibilités:

e1. N divise $a^{\frac{r}{2}} - 1$. Ceci est en fait impossible, car alors on aurait $a^{\frac{r}{2}} = 1 \pmod N$ et $\frac{r}{2}$ serait l'ordre.

e2. N divise $a^{\frac{r}{2}} + 1$. C'est-à-dire $a^{\frac{r}{2}} = -1 \pmod N$. Output `Fail` et retourner à l'étape **a**.

e3. N partage des facteurs non-triviaux avec $a^{\frac{r}{2}} - 1$ et $a^{\frac{r}{2}} + 1$. Par exemple si $N = pq$ il faut que p divise $a^{\frac{r}{2}} - 1$ et que q divise $a^{\frac{r}{2}} + 1$ (ou vice-versa). En d'autres termes,

$$d_{\pm} = \text{PGCD}\left(a^{\frac{r}{2}} \pm 1, N\right)$$

sont non-triviaux $d_+ > 1$ et $d_- > 1$ et nous avons 2 facteurs non-triviaux de N. Ceux-ci sont calculés grâce à l'algorithme d'Euclide (Complexité $O\left((\log_2 N)^3\right)$).

f. Vérifier si le produit des facteurs trouvés dans les étapes précédentes vaut N. Si ce n'est pas encore le cas, retourner à l'étape **a**.

Nous voyons qu'en présence d'un oracle qui permettrait résoudre l'étape **c**, la complexité d'une expérience (un round) provient uniquement de l'algorithme d'Euclide qui est $O\left(b^3\right)$. Néanmoins, cette expérience est probabiliste (étape **a**) et nous devons nous assurer que la probabilité de succès est non-négligeable. En fait on peut prouver $\mathbb{P}$ [succès] $\geq \frac{3}{4}$. Cela implique qu'avec $T = \frac{|\ln \epsilon|}{|\ln 2|}$ expériences (rounds) on peut amplifier cette probabilité de succès à $\mathbb{P}$ [succès en T rounds] $\geq 1 - \epsilon$.

Lors d'un round, les seuls outputs `Fail` interviennent en **d** et **e2**. Cela correspond à l'évènement

$$(r \text{ est impair}) \text{ ou } (a^{\frac{r}{2}} = -1 \pmod N)$$

Ainsi

$$\mathbb{P}\,[\text{échec}] = \mathbb{P}\left[(r \text{ est impair}) \text{ ou } (a^{\frac{r}{2}} = -1 \pmod N)\right].$$

Théorème. Soit a pris uniformément aléatoirement dans $\{2, \ldots, N-1\}$. Soit r le plus petit entier satisfaisant à $a^r = 1 \pmod N$. Alors $\mathbb{P}\,[\text{échec}] \leq \frac{1}{4}$.

Nous ne donnons pas de preuve ici. Ce théorème assure que la probabilité de succès de l'algorithme de factorisation basé sur la recherche de l'ordre est d'au moins $\frac{3}{4}$ lors d'un seul round. En faisant des ronds successifs, il est possible d'amplifier cette probabilité à $1 - \epsilon$ ($\epsilon \ll 1$). Comme d'habitude, le nombre de rounds requis est de l'ordre de $O(|\ln \epsilon|)$.

Fractions continuées

Dans ce paragraphe, nous donnons, sans démonstration, quelques résultats de théorie des nombres, utiles pour le développement ultérieur de l'algorithme de Shor.

Tout nombre réel peut être développé en *fraction continuée*. Ici nous discutons ce processus uniquement pour les *fractions rationnelles*. Prenons un exemple. Soit la fraction $x = \frac{263}{189}$. Les étapes successives de l'algorithme d'Euclide du calcul du PGCD (263, 189) sont:

$$\begin{aligned} 263 &= 1 \cdot 189 + 74 \\ 189 &= 2 \cdot 74 + 41 \\ 74 &= 1 \cdot 41 + 33 \\ 41 &= 1 \cdot 33 + 8 \\ 33 &= 4 \cdot 8 + 1 \end{aligned}$$

On voit que PGCD (263, 189) = 1. Étant donné qu'à chaque fois on divise par un nombre ≥ 2, le nombre d'étapes (de lignes ci-dessus) est de l'ordre de $\log_2(263)$, c'est-à-dire $\log_2(N)$ en général. De plus, chaque division requiert $O\left((\log_2 N)^2\right)$ opérations. Donc le nombre total d'opérations pour l'algorithme d'Euclide est $O\left((\log_2 N)^3\right)$. Maintenant, pour obtenir le *développement en fraction continuée*, on procède de la sorte:

$$\frac{263}{189} = 1 + \frac{74}{189} = 1 + \frac{1}{\frac{189}{74}} = 1 + \cfrac{1}{2 + \frac{41}{74}}$$

$$= 1 + \cfrac{1}{2 + \cfrac{1}{\cfrac{74}{41}}} = 1 + \cfrac{1}{2 + \cfrac{1}{1 + \cfrac{33}{41}}}$$

$$= 1 + \cfrac{1}{2 + \cfrac{1}{1 + \cfrac{1}{\cfrac{41}{33}}}}$$

$$= 1 + \cfrac{1}{2 + \cfrac{1}{1 + \cfrac{1}{1 + \cfrac{8}{33}}}}$$

$$= 1 + \cfrac{1}{2 + \cfrac{1}{1 + \cfrac{1}{1 + \cfrac{1}{\cfrac{33}{8}}}}}$$

$$= 1 + \cfrac{1}{2 + \cfrac{1}{1 + \cfrac{1}{1 + \cfrac{1}{4 + \cfrac{1}{8}}}}}$$

On dit que le développement en fraction continue est

$$\frac{263}{189} = [1; 2; 1; 1; 4; 8]$$

La forme générale du développement est :

$$x = [a_0; a_1; \ldots; a_n]$$

$$x = a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \ldots}}}$$

Si $x > 1$ on a $a_0 \geq 1$ et si $x < 1$ on a $a_0 = 0$. De plus, ce développement n'est pas unique, car on peut toujours écrire le dernier terme $\frac{1}{a_n}$ comme $\frac{1}{(a_n-1)+\frac{1}{1}}$. Ainsi

$$x = [a_0; a_1; \ldots; a_{n-1}; a_n] = [a_0; a_1; \ldots; a_{n-1}; a_n - 1; 1]$$

Mais à cette ambigüité près, le développement est unique. De plus, on peut le rendre unique en déclarant que l'on choisit toujours le développement le plus court possible. Le nombre d'opérations requises est le même que pour l'algorithme d'Euclide, $O\big((\log_2 N)^3\big)$ ou $N = \max$(numérateur, dénominateur). La longueur du développement est $O((\log_2 N))$.

Définition: Notion de Convergent. Soit $x = [a_0; a_1; a_2; \ldots; a_n]$ un développement en fraction continuée de x. On appelle *convergents* les séries tronquées $[a_0; a_1; a_2; \ldots; a_m]$, $1 \leq m \leq n$. Ces convergents sont des nombres rationnels

$$[a_0; a_1; a_2; \ldots; a_m] = \frac{p_m}{q_m}.$$

Théorème: Propriétés des Convergents. Soit p_m/q_m l'ensemble des convergents d'une fraction x. Alors

a) PGCD $(p_m, q_m) = 1$ (p_m et q_m sont premiers entre eux) et
$\left|x - \frac{p_m}{q_m}\right| \leq \frac{1}{q_m^2}$ (les convergents forment de bonnes approximations).
b) Réciproquement, toutes les approximations de la forme p/q avec p et q premiers entre eux, telles que $\left|x - \frac{p}{q}\right| < \frac{1}{q^2}$ sont données par l'ensemble des convergents de x. On peut donc calculer ces approximations de facon systématique.

Pour nous, c'est la propriété b) qui sera utile dans l'analyse de l'algorithme de Shor.

Fonction d'Euler

Pour un entier $r > 2$, on dit que $a \in \{1, 2, 3, \ldots, r-1\}$ est premier avec r (a is coprime with r) si PGCD $(a, r) = 1$. Le nombre de tels entiers a est donné par $\varphi(r)$, la *fonction d'Euler*. Si N est premier, $N = p$ on a bien sûr : $a = 1, 2, 3, \ldots, p-1$ et $\varphi(p) = p - 1$. En général, on montre que si

$$N = p_1^{e_1} p_2^{e_2} \ldots p_k^{e_k}$$

est la décomposition (unique) en facteurs premiers de N,

$$\begin{aligned}\varphi(N) &= \varphi(p_1^{e_1})\varphi(p_2^{e_2}) \ldots \varphi(p_k^{e_k}) \\ &= p_1^{e_1-1}(p_1 - 1)p_2^{e_2-1}(p_2 - 1) \ldots p_k^{e_k-1}(p_k - 1)\end{aligned}$$

Exemple. Si $N = 9$ les a premiers avec N sont $a = 1, 2, 4, 5, 7, 8$ et donc $\varphi(9) = 6$. On vérifie $\varphi(9) = 3^{2-1}(3-1) = 6$

L'inégalité suivante (valable pour r assez grand) sera utile pour nous:

$$\varphi(r) \geq \frac{r}{4 \ln \ln r}$$

Le dénominateur $\ln \ln r$ croît extrêmement lentement. Par exemple pour $r = 10^{1000}$ (ce qui représente un nombre à 1000 décimales) on a $\ln \ln r = \ln(1000 \ln 10) = 3 \ln 10 + \ln \ln 10 \leq 8$. En d'autres termes, étant donné r, une fraction appréciable des $r-1$ nombres inférieurs sont premiers avec r (dans l'exemple cette fraction est supérieure à $1/32$). Cette propriété peut être ré-exprimée comme suit. Fixons r et tirons $k \in \{1, 2, \ldots, r\}$ au hasard, uniformément, (c.-à-d. avec probabilité $\frac{1}{r}$). Alors:

$$\mathbb{P}\left[\text{PGCD}\,(k, r) = 1\right] = \frac{\varphi(r)}{r} \geq \frac{1}{4(\ln \ln r)} \tag{11.1}$$

Le membre de droite de l'inégalité décroît très lentement: on pourra y penser comme étant $O(1)$ (même si cela n'est pas vrai bien sûr!).

11.2 Recherche de la période d'une fonction arithmétique

Comme nous l'avons expliqué, on peut ramener la factorisation d'un entier N à la recherche de l'ordre d'un nombre a pris au hasard dans $\{2, 3, \ldots, N-1\}$. L'ordre $\text{Ord}_N(a)$ est le plus petit entier r tel que

$$a^r = 1 \pmod{N}.$$

En d'autres termes, nous cherchons la période de la fonction arithmétique

$$\begin{aligned} f_{a,N} : \mathbb{Z} &\to \mathbb{Z} \\ x &\to f_{a,N}(x) = a^x \pmod{N}. \end{aligned}$$

Cette période (ou l'ordre) est le plus petit entier r t.q.

$$f_{a,N}(x) = f_{a,N}(x + r), \qquad \forall x \in \mathbb{Z}.$$

Nous commençons donc par étudier un algorithme général de "recherche de la période d'une fonction arithmétique".

Soit $f : \mathbb{Z} \to \mathbb{Z}$ de période inconnue

$$f(x) = f(x + r), \qquad \forall x \in \mathbb{Z}.$$

Comme nous serons obligés de travailler avec un nombre fini de bits, nous allons tronquer $\mathbb{Z}$ à $\frac{\mathbb{Z}}{M\mathbb{Z}} = \{0, 1, 2, \ldots, M-1\}$ où M est choisi bien plus grand que r : $M >> r$. Ici, $\frac{\mathbb{Z}}{M\mathbb{Z}}$ est le groupe additif des entiers pris $\pmod{M}$. En fait r est inconnu, mais nous supposons que l'on connait une borne supérieure, et qu'il est donc possible de choisir

$M >> r$. Par exemple, pour la recherche de l'ordre, nous savons que $r < N$. Nous verrons dans ce cas que $M = O(N^2)$ est suffisant.

Tout d'abord, il nous faut représenter les entiers $x \in \{0, \ldots, M-1\}$ par des états quantiques. Nous prenons (sans perte de généralité) $M = 2^m$ et notons que x peut être représenté grâce à son expansion binaire

$$x = 2^{m-1}x_{m-1} + 2^{m-2}x_{m-2} + \ldots + 2^2 x_2 + 2x_1 + x_0,$$

avec m bits

$$x = \underbrace{(x_{m-1} \ldots x_0)}_{\text{dev binaire de x}}.$$

En particulier $(0 \ldots 0) = 0$ et $(1 \ldots 1) = 2^m - 1$. Il est donc naturel de prendre comme espace de Hilbert

$$\mathcal{H} = \underbrace{\mathbb{C}^2 \otimes \mathbb{C}^2 \otimes \ldots \otimes \mathbb{C}^2}_{\text{m fois}},$$

et de stocker l'entier x dans un état quantique $|x\rangle \in \mathcal{H}$ construit à partir de m qubits (m systèmes à 2 niveaux: spins nucléaires, polarisation des photons ...)

$$|x\rangle = |x_{m-1}\rangle \otimes \ldots \otimes |x_0\rangle = |x_{m-1} \ldots x_0\rangle.$$

La fonction f est comme d'habitude représentée par l'opération unitaire

$$U_f : |x\rangle \otimes |0\rangle \rightarrow U_f(|x\rangle \otimes |0\rangle) = |x\rangle \otimes |f(x)\rangle$$

où $|0\rangle$ et $|f(x)\rangle$ sont des états à m qubits (dans l'algorithme de Shor, on calcule $f(x) \pmod N$ et donc m bits suffisent certainement). Nous aurons aussi besoin de la "Transformée de Fourrier Quantique" (Section **11.6**) définie par :

$$QFT|x\rangle = \frac{1}{\sqrt{M}} \sum_{y=0}^{M-1} e^{2\pi i \frac{xy}{M}} |y\rangle = \frac{1}{2^{m/2}} \sum_{\substack{y_0 \ldots y_{m-1} \\ \in \{0,1\}^m}} e^{2\pi i \frac{xy}{M}} |y_{m-1} \ldots y_0\rangle \tag{11.2}$$

Cette opération est linéaire, c.-à-d. que si $|\Psi\rangle = \sum_{x=0}^{M-1} c_x |x\rangle$, alors

$$QFT|\Psi\rangle = \sum_{x=0}^{M-1} c_x QFT|x\rangle.$$

On peut aussi montrer que l'opération est unitaire : ceci est un prérequis important pour pouvoir la réaliser grâce à un circuit quantique.

11.3 Circuit pour la recherche de la période

Le circuit de l'algorithme de recherche de la période est représenté sur la Figure **11.1**.

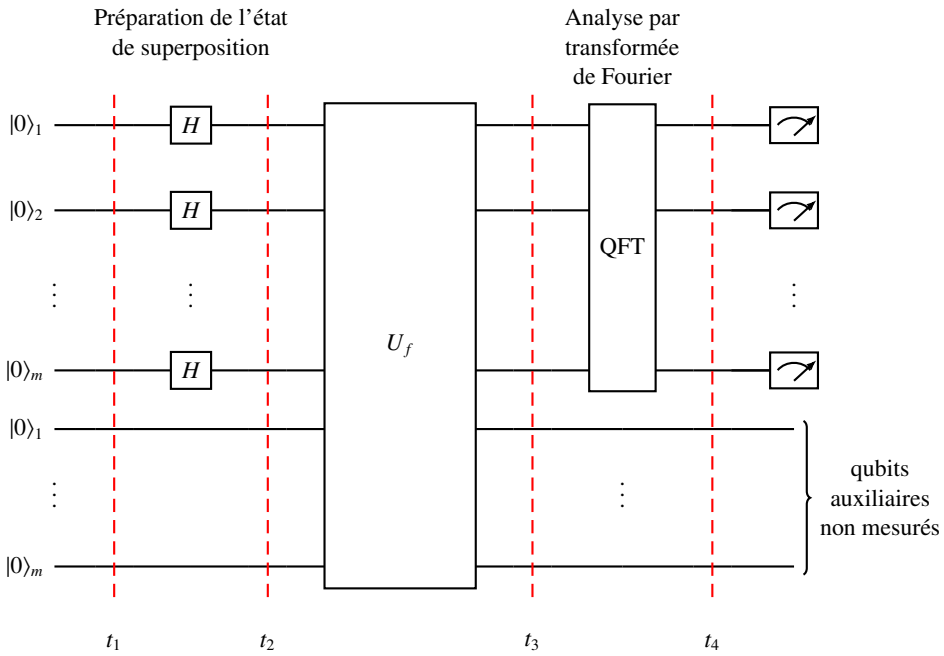


Figure 11.1 Circuit quantique pour la recherche de la période d'une fonction arithmétique.

Le circuit pour U_f dépend de la fonction spécifique. Pour la recherche de l'ordre, nous prendrons la fonction $f(x) = a^x \pmod N$ et verrons comment réaliser son circuit à la Section **11.7**. Le circuit pour QFT est réalisé à la Section **11.6**.

Calculons maintenant l'évolution de l'état initial (t_1):

$$|0\rangle \otimes |0\rangle = \underbrace{|0\ldots0\rangle}_{\text{m fois}} \otimes \underbrace{|0\ldots0\rangle}_{\text{m fois}}.$$

Juste après les portes de Hadamard (t_2):

$$H^{\otimes m} \underbrace{|0\ldots0\rangle}_{\text{m fois}} \otimes |0\rangle^{\otimes m} = \left(\frac{1}{2^{\frac{m}{2}}} \sum_{\substack{x_0 \ldots x_{m-1} \\ \in \{0,1\}^m}} |x_{m-1} \ldots x_0\rangle \right) \otimes |0\rangle^{\otimes m}.$$

C'est un état de superposition cohérente sur toutes les entrées classiques. Il peut aussi s'écrire de façon plus compacte:

$$\left(\frac{1}{\sqrt{M}} \sum_{x=0}^{M-1} |x\rangle \right) \otimes |0\rangle^{\otimes m}.$$

Après U_f (t_3), nous obtenons l'état:

$$\frac{1}{\sqrt{M}} \sum_{x=0}^{M-1} |x\rangle \otimes |f(x)\rangle.$$

Exploitons le fait que f est périodique pour réorganiser cette somme. L'intervalle $[0, M-1]$ est décomposé en morceaux de longueur r, sauf pour le dernier qui sera plus court (Figure **11.2**). Les entiers dans la première période sont $x_0 \in \{0, 1, \ldots, r-1\}$. Si M était

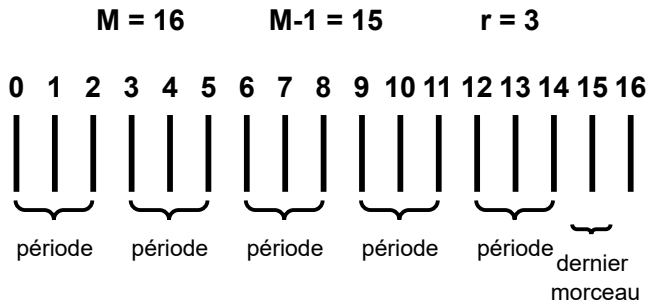


Figure 11.2 Exemple de décomposition de $\{0, 1, \ldots, M-1\}$ pour $r = 3$ et $M = 16$.

un multiple de r, on pourrait représenter chaque x comme

$$x = x_0 + jr \text{ avec } 0 \leq j \leq \frac{M}{r} - 1.$$

Dans le cas général (voir Figure **11.2**) on aura

$$x = x_0 + jr \text{ avec } 0 \leq j \leq A(x_0) - 1,$$

et $A(x_0)$ un entier dépendant de x_0 qui doit satisfaire

$$M - r \leq x_0 + (A(x_0) - 1)r \leq M - 1.$$

Nous avons :

$$\begin{aligned}\frac{1}{\sqrt{M}} \sum_{x=0}^{M-1} |x\rangle \otimes |f(x)\rangle &= \frac{1}{\sqrt{M}} \sum_{x_0}^{r-1} \sum_{j=0}^{A(x_0)-1} |x_0 + jr\rangle \otimes |f(x_0 + jr)\rangle \\ &= \frac{1}{\sqrt{M}} \sum_{x_0}^{r-1} \sum_{j=0}^{A(x_0)-1} |x_0 + jr\rangle \otimes |f(x_0)\rangle.\end{aligned}$$

Finalement, nous agissons sur cet état avec QFT. L'état obtenu (t_4) est :

$$\begin{aligned}|\Psi_{\text{fin}}\rangle &= \frac{1}{\sqrt{M}} \sum_{x_0}^{r-1} \sum_{j=0}^{A(x_0)-1} QFT|x_0 + jr\rangle \otimes |f(x_0)\rangle \\ &= \frac{1}{\sqrt{M}} \sum_{x_0}^{r-1} \sum_{j=0}^{A(x_0)-1} \frac{1}{\sqrt{M}} \sum_{y=0}^{M-1} e^{2\pi i \frac{(x_0+jr)y}{M}} |y\rangle \otimes |f(x_0)\rangle \\ &= \frac{1}{M} \sum_{x_0=0}^{r-1} \left(\sum_{y=0}^{M-1} \left(e^{2\pi i \frac{x_0 y}{M}} \sum_{j=0}^{A(x_0)-1} e^{2\pi i \frac{jy}{M/r}} \right) |y\rangle \right) \otimes |f(x_0)\rangle.\end{aligned}$$

Cette dernière expression est l'état final $|\Psi_{\text{fin}}\rangle$ juste avant la mesure.

11.4 Le Processus de Mesure

Il reste maintenant à analyser l'opération de mesure. Tout d'abord, il nous faut choisir une "base représentant l'appareil de mesure". Celle-ci est formée par l'ensemble des projecteurs.

$$P_y = |y\rangle\langle y| \otimes \mathbb{1}_{m\times m}, \qquad y \in \{0, 1, 2, \ldots, M-1\}.$$

L'état quantique résultant juste après la mesure est (à une normalisation près)

$$P_y|\Psi\rangle_{\text{fin}}$$

avec la probabilité

$$\mathbb{P}\,[y] = \langle\Psi_{\text{fin}}|P_y|\Psi_{\text{fin}}\rangle.$$

D'abord, on calcule $P_y|\Psi_{\text{fin}}\rangle$,

$$P_y|\Psi_{\text{fin}}\rangle = \frac{1}{M}\sum_{x_0=0}^{r-1}\left(e^{2\pi i\frac{x_0 y}{M}}\sum_{j=0}^{A(x_0)-1}e^{2\pi i\frac{jy}{M/r}}\right)|y\rangle\otimes|f(x_0)\rangle.$$

Puis $\langle\Psi_{\text{fin}}|P_y|\Psi_{\text{fin}}\rangle = \langle\Psi_{\text{fin}}|P_yP_y|\Psi_{\text{fin}}\rangle$. Cela donne

$$\mathbb{P}\,[y] = \frac{1}{M^2}\sum_{x_0=0}^{r-1}\left|\sum_{j=0}^{A(x_0)=1}e^{2\pi i\frac{jy}{M/r}}\right|^2.$$

Remarquons que les différents termes de la somme sur x_0 n'interfèrent pas, car les kets $|f(x_0)\rangle$ sont orthogonaux entre eux.

11.5 Analyse de la probabilité $\mathbb{P}\,[y]$

Traitons d'abord le cas (irréaliste) simple où M serait multiple de r.

Dans ce cas, $A(x_0) = \frac{M}{r}$ et donc

$$\mathbb{P}\,[y] = \frac{r}{M^2}\left|\sum_{j=0}^{\frac{M}{r}-1}e^{2\pi i\frac{jy}{M/r}}\right|^2.$$

Si $y = k\frac{M}{r}$ avec $k = \{0, 1, \ldots, r-1\}$, on a

$$e^{2\pi i\frac{jy}{M/r}} = e^{2\pi ijk} = 1.$$

Si bien que $\mathbb{P}\,[y] = \frac{r}{M^2}\left|\frac{M}{r}\right|^2 = \frac{1}{r}$. Puisque cette probabilité doit se sommer à 1, nous en déduisons qu'elle est nulle pour toutes les autres valeurs de $y \neq k\frac{M}{r}$. Cette distribution est représentée sur la Figure **11.3**.

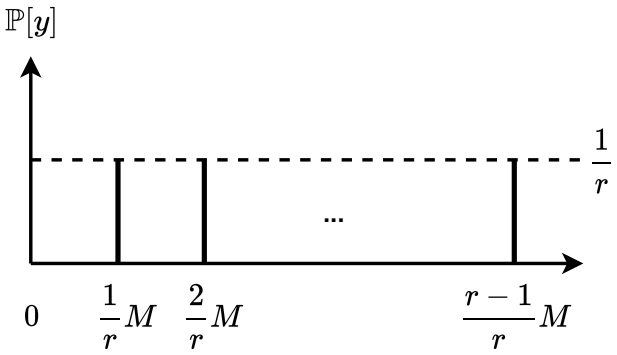


Figure 11.3 Distribution de probabilité des résultats de mesures pour M multiple de r.

La mesure donne avec probabilité 1 une valeur de y de la forme

$$y = k\frac{M}{r} \quad \text{avec} \quad k \in \{0, 1, \ldots, r-1\}.$$

Puisque M est connu, grâce à la valeur de y donnée par la mesure, nous calculons $\frac{y}{M}$. Deux cas de figure se présentent à nous:

- $\frac{y}{M} = \frac{k}{r}$ et PGCD $(k, r) = 1$. Alors nous pouvons trouver k et r en simplifiant la fraction $\frac{y}{M}$ "au maximum" jusqu'à ce que les numérateurs et dénominateurs n'aient plus de facteurs communs. Nous trouvons ainsi r.
- $\frac{y}{M} = \frac{k}{r}$ et PGCD $(k, r) \neq 1$. Alors nous ne savons pas jusqu'où simplifier la fraction (et n'avons pas de façon systématique de trouver k et r).

En "pratique" nous ne savons pas à priori si k et r sont premiers entre eux ou non. Ainsi nous adoptons la procédure suivante: dans tous les cas, simplifier la fraction $\frac{y}{M}$ au maximum, et tester si le r trouvé est une période de $f(x)$ ou non.

La probabilité de succès est la probabilité d'avoir PGCD $(k, r) = 1$ quand k est tiré uniformément dans $\{0, 1, \ldots, r-1\}$. D'après ce que nous avons appris dans le chapitre précédent (équation (11.1)):

$$\mathbb{P}\left[\text{PGCD}\,(k, r) = 1, k \in \{0, 1, \ldots, r-1\}\right] = \frac{\varphi(r)}{r} \geq \frac{1}{4(\ln \ln r)}.$$

Puisque $r < M$, nous avons *une probabilité de succès pour une expérience*:

$$\mathbb{P}\left[\text{succes}\right] \geq \frac{1}{4 \ln \ln M} \quad \left(= \frac{1}{4 \ln 2 \ln m}\right).$$

Bien que cette probabilité soit faible, nous pouvons l'amplifier en faisant tourner le

circuit plusieurs fois. Au bout de T expériences (ou "rounds"):

$$\mathbb{P}\left[\text{au moins 1 succes au bout de T rounds}\right] \geq 1 - \left(1 - \frac{1}{4\ln\ln\ M}\right)^T,$$

ce qui peut être rendu proche de $1-\epsilon$ si on prend

$$T = O(|\ln\ \epsilon|\ \ln\ln\ M) = O(|\ln\ m|\,|\ln\ \epsilon|\,).$$

En effet:

$$\begin{aligned}
&1 - \left(1 - \frac{1}{4\ln\ M}\right)^T \geq 1 - \epsilon \\
&\leftrightarrow \epsilon \geq \left(1 - \frac{1}{4\ln\ M}\right)^T \leftrightarrow \ln\ \epsilon \geq T\ln\left(1 - \frac{1}{4\ln\ M}\right) \\
&\leftrightarrow \ln\ \epsilon \geq -T\frac{1}{4\ln\ M}\ (M \text{ grand}) \\
&\leftrightarrow T \geq 4(\ln\ln\ M)|\ln\ \epsilon|
\end{aligned}$$

Passons maintenant au cas général ou M n'est pas un multiple de r.

Nous allons utiliser un Lemme technique (la démonstration n'est pas donnée ici). Une illustration graphique de son contenu est fournie par la Figure **11.4**. En gros, le Lemme affirme que la distribution de probabilité $\mathbb{P}\left[y\right]$ est concentrée sur les entiers proches des fractions kM/r.

Lemme. Soit $I = \cup_{k=0}^{r-1}\left[k\frac{M}{r} - \frac{1}{2}, k\frac{M}{r} + \frac{1}{2}\right] = \cup_{k=0}^{r-1} I_k$ une union d'intervalles disjoints I_k. Alors,

$$\mathbb{P}\left[y \in I\right] \geq \frac{2}{5}.$$

Ainsi, les mesures fournissent des entiers y proches de $k\frac{M}{r}$ avec $k \in \{0, 1, \ldots, r-1\}$ avec probabilité au moins $\frac{2}{5}$.

Lorsqu'une mesure donne $y \in I$, cela signifie qu'il existe k entier tel que

$$k\frac{M}{r} - \frac{1}{2} \leq y \leq k\frac{M}{r} + \frac{1}{2},$$

ce qui est équivalent à

$$\left|\frac{y}{M} - \frac{k}{r}\right| \leq \frac{1}{2M}. \tag{11.3}$$

Maintenant supposons que nous prenions $M > r^2$. Alors cette inégalité entraine,

$$\left|\frac{y}{M} - \frac{k}{r}\right| \leq \frac{1}{2r^2} \ \text{ pour } \ k \in \{0, 1, \ldots, r-1\}.$$

Comment pouvons-nous déterminer k et r à partir de y et M ? D'après ce que nous avons vu dans la théorie des fractions continuées, si le PGCD $(k, r) = 1$, alors $\frac{k}{r}$ est nécessairement un "convergent" du développement en fractions continuées de $\frac{y}{M}$. Il y a

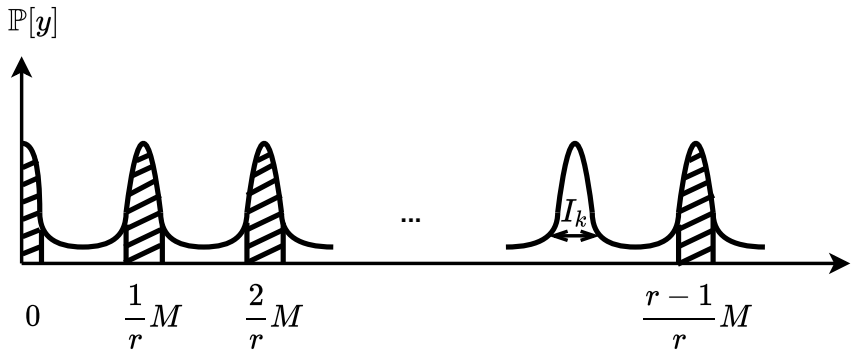


Figure 11.4 Distribution de probabilité fournie par les mesures. L'aire hachurée est superieure a $\frac{2}{5}$. Notez que les intervalles I_k ont une longueur 1 et sont distants d'environ $M/r >> 1$.

un nombre fini de "convergents" car $\frac{y}{M}$ est rationnel, et ceux-ci peuvent être systématiquement calculés grâce à l'algorithme d'Euclide (en temps $O((\ln M)^3)$. Par contre, si PGCD $(k, r) \neq 1$ on ne peut pas affirmer que $\frac{k}{r}$ est un convergent de $\frac{y}{M}$ et n'avons, dans ce cas, pas de moyen systématique de calculer k et r.

Nous adoptons donc la procédure suivante. Nous calculons tous les convergents de $\frac{y}{M}$ (grâce à l'algorithme d'Euclide) et examinons leurs dénominateurs r. Pour chacun de ces dénominateurs, nous testons si c'est une période de $f(x)$. Le succès est assuré si PGCD $(k, r) = 1$, ce qui a lieu avec probabilité $O(\frac{1}{4 \ln \ln r})$.

Récapitulons. Quel est la probabilité de succès lors d'une expérience avec le circuit quantique ? Le circuit quantique est initialisé dans l'état $|0\rangle \otimes |0\rangle$. L'évolution unitaire conduit à l'état $|\Psi_{\text{final}}\rangle$, après quoi on effectue une mesure. Cette mesure donne l'entier y. Pour en déduire r avec succès, il faut remplir deux conditions:

- $y \in I$ pour un certain $k \in \{0, 1, \ldots, r-1\}$.
- Étant donné $y \in I_k$, il faut PGCD $(k, r) = 1$.

Donc:

$$\mathbb{P}\ [\text{succes}] \geq \frac{2}{5} \times \frac{1}{4 \ln \ln r}.$$

En itérant l'expérience $T \approx O(|\ln \epsilon|\ \ln \ln r)$ fois, on peut amplifier la probabilité de succès à $1 - \epsilon$.

11.6 Le circuit de la QFT

Dans ce paragraphe, nous montrons comment réaliser le circuit de la QFT. Rappelons l'équation que doit implémenter le circuit (11.2):

$$QFT|x\rangle = \frac{1}{\sqrt{M}} \sum_{y=0}^{M-1} e^{2\pi i \frac{xy}{M}} |y\rangle = \frac{1}{2^{m/2}} \sum_{\substack{y_0 \dots y_{m-1} \\ \in \{0,1\}^m}} e^{2\pi i \frac{xy}{M}} |y_{m-1} \dots y_0\rangle$$

On peut se convaincre que pour $M = 2$, on a $QFT = H$ (la porte de Hadamard):

$$(QFT)_{M=2}|x\rangle = \frac{1}{\sqrt{2}}(|0\rangle + (-1)^x|1\rangle).$$

Pour $M = 4$,

$$\begin{aligned}(QFT)_{M=4}|x\rangle &= \frac{1}{\sqrt{4}}\left(|0\rangle + e^{i\frac{\pi}{2}x}|1\rangle + e^{i\pi x}|2\rangle + e^{3i\frac{\pi}{2}x}|3\rangle\right) \\ &= \frac{1}{\sqrt{4}}\left(|00\rangle + e^{i\frac{\pi}{2}x}|01\rangle + e^{i\pi x}|10\rangle + e^{3i\frac{\pi}{2}x}|11\rangle\right) \\ &= \frac{1}{\sqrt{2}}\left(|0\rangle + e^{i\pi x}|1\rangle\right) \otimes \frac{1}{\sqrt{2}}\left(|0\rangle + e^{i\frac{\pi}{2}x}|1\rangle\right).\end{aligned}$$

En notation binaire $x \in \{0, 1, 2, 3\}$ est représenté par

$$x = 2x_1 + x_0 \quad , \quad x_0, x_1 \in \{0, 1\}$$

si bien que $e^{i\pi x} = e^{2\pi i x_1} e^{i\pi x_0} = (-1)^{x_0}$ et $e^{i\frac{\pi}{2}x} = e^{i\pi x_1} e^{i\frac{\pi}{2}x_0} = (-1)^{x_1} e^{i\frac{\pi}{2}x_0}$. On trouve alors

$$(QFT)_{M=4}|x\rangle = \frac{1}{\sqrt{2}}\Big(|0\rangle + (-1)^{x_0}|1\rangle\Big) \otimes \frac{1}{\sqrt{2}}\Big(|0\rangle + (-1)^{x_1} e^{i\frac{\pi}{2}x_0}|1\rangle\Big).$$

Cette factorisation est à la base de la réalisation du circuit de la QFT. La factorisation suggère le circuit de la Figure **11.5**.

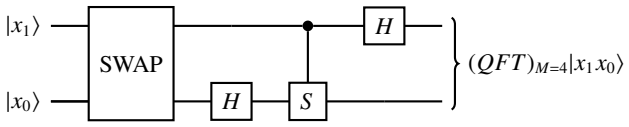


Figure 11.5 Circuit de la $(QFT)_{M=4}$.

La première opération SWAP échange les deux qubits. Elle peut être réalisée par trois portes $CNOT$ (Figure **11.6**).

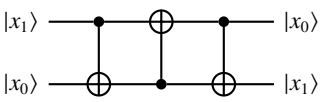


Figure 11.6 Circuit pour un SWAP.

La seconde opération de la Figure **11.5** est une porte de Hadamard agissant sur $|x_1\rangle$ pour produire $\frac{1}{\sqrt{2}}(|0\rangle + (-1)^{x_1}|1\rangle)$. La troisième opération est un "phase shift" contrôlé

par le premier bit x_0: si $x_0 = 0$ il n'y a pas de phase shift et le second bit reste dans l'état $\frac{1}{\sqrt{2}}(|0\rangle + (-1)^{x_1}|1\rangle)$; par contre si $x_0 = 1$, il y a un phase shift et le second bit est transformé en $\frac{1}{\sqrt{2}}(|0\rangle + (-1)^{x_1} e^{i\frac{\pi}{2}}|1\rangle)$. Enfin, la dernière porte de Hadamard agit sur $|x_0\rangle$ pour produire $\frac{1}{\sqrt{2}}(|0\rangle + (-1)^{x_0}|1\rangle)$.

Le circuit général de la QFT est obtenu par une généralisation des remarques ci-dessus.

QFT générale

Lemme. Pour $x \in \{0, 1, \ldots, M-1\}$ et $M = 2^m$

$$QFT|x\rangle = \prod_{l-1}^{m} \frac{(|0\rangle + e^{i\frac{\pi}{2^{l-1}}x}|1\rangle)}{\sqrt{2}}.$$

Démonstration. Rappelons que la formule de la QFT est donnée par (11.2):

$$QFT|x\rangle = \frac{1}{\sqrt{M}} \sum_{y=0}^{M-1} e^{2\pi i \frac{xy}{M}}|y\rangle = \frac{1}{2^{\frac{m}{2}}} \sum_{y=0}^{2^m-1} e^{2\pi i \frac{xy}{2^m}}|y\rangle.$$

Chaque $y \in \{0, 1, \ldots, 2^m - 1\}$ possède un développement binaire

$$\begin{aligned} y &= 2^{m-1}y_{m-1} + 2^{m-2}y_{m-2} + \ldots + 2y_1 + y_0 \\ &= 2y' + y_0 \end{aligned}$$

où $y' = 2^{m-2}y_{m-1} + \ldots + y_1$. On décompose la somme sur y en une somme avec $y_0 = 0$ et une somme avec $y_0 = 1$ (cela revient à séparer les y pairs et impairs.)

$$\begin{aligned} QFT|x\rangle &= \frac{1}{2^{\frac{m}{2}}} \sum_{y'=0}^{2^{m-1}-1} e^{2\pi i \frac{x2y'}{2^m}}|y'\rangle \otimes |0\rangle + \frac{1}{2^{\frac{m}{2}}} \sum_{y'=0}^{2^{m-1}-1} e^{2\pi i \frac{x(2y'+1)}{2^m}}|y'\rangle \otimes |1\rangle \\ &= \left(\frac{1}{2^{\frac{m-1}{2}}} \sum_{y'=0}^{2^{m-1}-1} e^{2\pi i \frac{xy'}{2^{m-1}}}|y'\rangle\right) \otimes (|0\rangle + e^{\frac{\pi i x}{2^{m-1}}}|1\rangle). \end{aligned}$$

Cette factorisation peut maintenant être répétée sur la première parenthèse. La seule différence est que $m \to m-1$. On obtient

$$QFT|x\rangle = \left(\frac{1}{2^{\frac{m-2}{2}}} \sum_{y''=0}^{2^{m-2}-1} e^{2\pi i \frac{xy''}{2^{m-2}-1}}|y''\rangle\right) \otimes \frac{|0\rangle + e^{\frac{i\pi x}{2^{m-2}}}|1\rangle}{\sqrt{2}} \otimes \frac{|0\rangle + e^{\frac{i\pi x}{2^{m-1}}}|1\rangle}{\sqrt{2}}.$$

En itérant ce procédé, on obtient le résultat du lemme.

La dernière étape consiste à remplacer x par son développement binaire (comme nous l'avons fait pour $M = 4$)

$$x = 2^{m-1}x_{m-1} + \ldots + 2^2 x_2 + 2x_1 + x_0,$$

ce qui implique pour tout $1 \leq l \leq m$

$$e^{i\frac{\pi}{2^{l-1}}x} = e^{i\pi x_{l-1}} e^{i\frac{\pi}{2}x_{l-2}} \dots e^{i\frac{\pi}{2^{l-2}}x_1} e^{i\frac{\pi}{2^{l-1}}x_0}.$$

Ici le point crucial est que les bits x_i avec $i \geq l$ ne contribuent pas. Remplaçant cette expression dans la formule du lemme, on trouve la décomposition finale qui permet de construire un circuit:

$$QFT|x\rangle = \prod_{l=1}^{m} \left(\frac{|0\rangle + e^{i\pi x_{l-1}} e^{i\frac{\pi}{2}x_{l-2}} \dots e^{i\frac{\pi}{2^{l-1}}x_0}|1\rangle}{\sqrt{2}} \right). \tag{11.4}$$

La Figure **11.7** représente le circuit correspondant à cette dernière formule pour $m = 4$, c.-à-d. $M = 16$.

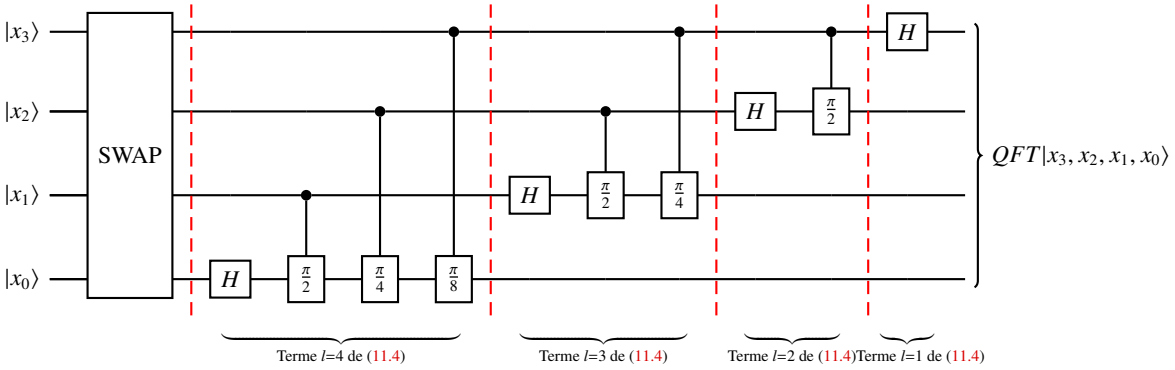


Figure 11.7 Circuit de la QFT pour 4 qubits.

On peut se convaincre que l'opération de SWAP requiert $O(3m)$ portes $CNOT$. D'autre part, le nombre de portes H et déphasages contrôlés est

$$m + (m-1) + \ldots + 1 = \frac{m(m+1)}{2}.$$

La profondeur du circuit est donc de l'ordre de $O(m^2)$. Cette profondeur indique comment le temps de calcul pour la QFT augmente avec la taille des entrées. D'autre part, la largeur du circuit est m. Ainsi la taille totale est profondeur $\times$ largeur $= O(m^3)$.

11.7 Circuit pour $U_{f_{a,N}}$

Dans la Section **11.3**, nous avons donné un algorithme aléatoire de factorisation d'entiers N basé sur la recherche de la période de l'exponentielle modulaire. Plus précisément, pour a t.q. PGCD $(a, N) = 1$, on cherche la période de la fonction $f_{a,N}(x) = a^x \pmod N$. Ceci est équivalent à la recherche de $\mathrm{Ord}_N(a) = r$ c.-à-d. le plus petit entier r t.q $a^r = 1 \pmod N$.

Nous devons trouver un circuit qui réalise l'opérateur unitaire correspondant $U_{f_{a,N}}$ (Figure **11.8**).

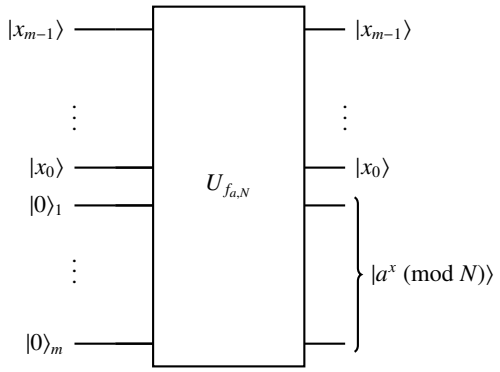


Figure 11.8 Représentation unitaire de l'exponentielle modulaire

Notons d'abord que

$$\begin{aligned} a^x &= a^{2^{m-1}x_{m-1}} a^{2^{m-2}x_{m-2}} \dots a^{2x_1} a^{x_0} \\ &= \left(a^{2^{m-1}}\right)^{x_{m-1}} \left(a^{2^{m-2}}\right)^{x_{m-2}} \dots \left(a^2\right)^{x_1} a^{x_0}. \end{aligned}$$

Il est possible de pré-calculer les puissances $\{a, a^2, a^4, a^8, \dots, a^{2^{m-1}}\}$ en un nombre polynomial d'opérations. En effet, on part de a qui possède m bits (au plus). Son carré a^2 se calcule en m^2 opérations. Puisque a^2 est pris (mod M), a^2 possède aussi m bits au plus. Le carré de ce dernier $a^4 = (a^2)^2$ se calcule en m^2 opérations, et ainsi de suite. En itérant ce procédé m fois, on va jusqu'au calcul de $a^{2^{m-1}}$. Ainsi on peut pré-calculer toutes ces puissances en $O(m^3)$ opérations. Il existe des circuits classiques réversibles pour faire ce calcul, et puisqu'ils sont réversibles, ils peuvent aussi être rendus quantiques (c.-à-d. unitaires). Finalement, pour calculer a^x, en vertu de l'identité ci-dessus, il suffit de prendre le circuit de la Figure **11.9**.

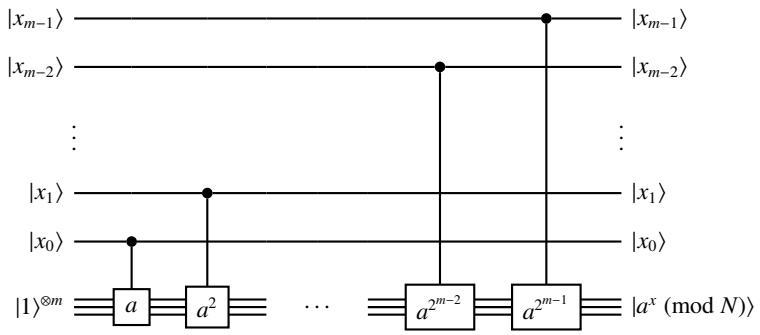


Figure 11.9 Circuit pour l'exponentielle modulaire.

La profondeur de ce circuit est $O(m^3)$, sa largeur $O(m)$ et sa taille $O(m^4)$.

11.8 Résumé de l'algorithme de Shor

Nous sommes maintenant en mesure de résumer la totalité de l'algorithme quantique de Shor pour la factorisation d'un entier N.

input: N impair et avec au moins deux facteurs premiers distincts.

output: facteur non trivial de N.

temps de calcul: $O((\ln N)^3(\ln\ln N)|\ln\ \epsilon|\,)$ pour une probabilité de succès supérieure à $1-\epsilon$.

taille du circuit: $O((\ln N)^3)$.

Algorithme:

1. Choisir uniformément aléatoirement $a \in \{2,\ldots,N-1\}$.
2. Calculer PGCD $(a,N) = d$ par l'algorithme d'Euclide:

- si $d > 1 \rightarrow$ SUCCÈS; on a un facteur,
- sinon $d = 1 \rightarrow$ aller à l'étape **3**.

3. Calculer $\mathrm{Ord}_N(a)$ (i.e trouver le plus petit r tel que $a^r = 1 \pmod N$). Pour cela, utiliser le circuit quantique avec m qubits et $2^m = M \approx N^2$. Faire une mesure quantique et considérer le résultat y. Calculer les convergents de $\frac{y}{M}$ (grâce à l'algorithme d'Euclide). Trouver si r se trouve parmi les dénominateurs de ces convergents en testant $a^r = 1 \pmod N$.
   - si oui (la théorie assure que c'est le plus petit possible) $\rightarrow$ aller à l'étape **4**,
   - sinon $\rightarrow$ ÉCHEC.

4. Vérifier si r est pair et $a^r \neq -1 \pmod N$
   - si oui $\rightarrow$ aller à l'étape **5**,
   - sinon $\rightarrow$ ÉCHEC.

5. Calculer PGCD $\left(a^{\frac{r}{2}}+1,N\right)$ et PGCD $\left(a^{\frac{r}{2}}-1,N\right)$. Cela donne deux facteurs non triviaux de N (grâce à l'algorithme d'Euclide).

La probabilité de succès d'un tel "round" est $O\left(\frac{1}{\ln\ln N}\right)$ et sa complexité (temps de calcul) $O\left((\ln N)^3\right)$. On peut amplifier (comme d'habitude) la probabilité de succès à $1-\epsilon$ en faisant $O(\ln\ln N)$ rounds. Le temps de calcul total sera alors $O\left(|\ln\ \epsilon|(\ln\ln N)(\ln N)^3\right)$.

12 Algorithme de Grover

Dans ce chapitre, nous étudions un algorithme entièrement différent des précédents, qui s'applique à des objets sans structure ou sans symétrie (il n'y a pas de sous groupe caché, pas de période, etc). Il s'agit d'un algorithme avec oracle qui effectue la recherche d'un *élément marqué* dans un *ensemble sans structure* ou une *base de donnée sans structure*. Donnons un exemple concret.

Prenons un annuaire (la base de donnée) et supposons que les numéros de téléphone jouent le rôle des entrées, les personnes correspondantes jouent le rôle des sorties. Étant donné une entrée (le numéro de téléphone) il est difficile de retrouver la sortie (le nom de la personne correspondante). La seule façon de procéder est de faire une recherche exhaustive ou bien de procéder à une recherche aléatoire. On montre facilement que ces deux méthodes requièrent de soumettre $O(N)$ questions à l'annuaire où N est le nombre d'entrées. Donc si $N = 2^n$, c'est-à-dire que l'on peut décrire tout l'annuaire avec $O(n)$ bits, le temps de recherche est exponentiel $O(2^n)$. La raison fondamentale étant que la liste des numéros de téléphone n'est pas ordonnée et ne possède aucune structure. Bien sûr, le problème consistant à rechercher le téléphone d'une personne connaissant son nom est facile car la liste des personnes est ordonnée par ordre alphabétique.

Nous verrons que l'algorithme de Grover permet de résoudre le problème en un temps $O\left(\sqrt{N}\right)$. Le temps est toujours exponentiel mais (pour N assez grand) au moins nous obtenons, grâce au calcul quantique, une accélération quadratique. Il est possible de montrer que si le problème n'a pas de structure spéciale sous-jacente, il n'est pas possible de faire mieux que $O\left(\sqrt{N}\right)$ dans le cadre du modèle de Deutsch (circuits quantiques).

Pour illustrer ce dernier point, reconsidérons le problème de la factorisation d'un entier $N = p{\cdot}q$ avec deux facteurs premiers. Il n'est pas possible d'avoir p et q supérieurs à $\sqrt{N}$, donc l'un des deux est $\leq \sqrt{N}$. En cherchant à travers la liste $\{1, 2, \ldots, \sqrt{N}\}$, nous trouverons sûrement un des deux facteurs premiers. Avec une recherche exhaustive il faut un temps $\sqrt{N}$, mais avec l'algorithme de Grover (ici l'oracle serait une boîte noire testant la division de N par $x \in \{1, 2, \ldots, \sqrt{N}\}$) il faut un temps $O\left(N^{1/4}\right)$ ce qui est déjà intéressant. Nous savons qu'il est possible de faire mieux (même classiquement) et d'obtenir une accélération exponentielle grâce à l'algorithme de Shor, mais cela est

possible en exploitant la symétrie (la périodicité) d'une fonction arithmétique reliée au problème de la factorisation.

12.1 Formulation Mathématique du problème

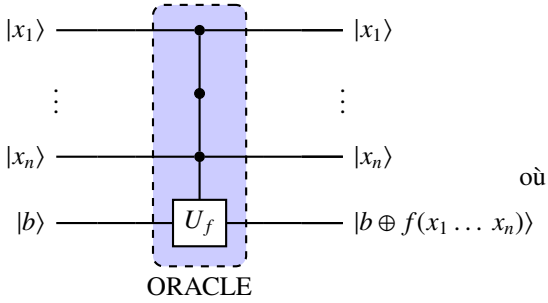
Soit $f : \mathbf{F}_2^n \to \mathbf{F}_2 = \{0, 1\}$ et soit l'équation

$$f(x_1 \dots x_n) = 1$$

Nous voulons déterminer parmi les 2^n entrées possibles pour $f(\cdot)$, au moins une solution $\overline{x_1} \dots \overline{x_n}$. Nous ne voulons pas déterminer toutes les solutions possibles mais simplement au moins une. L'application typique qui nous intéresse est celle où il y a une solution unique.

Comme nous le verrons dans l'algorithme de base, il faut connaître à priori le nombre total de solutions, mais cette restriction peut ensuite être supprimée grâce à une extension facile.

Nous supposons avoir à disposition un oracle quantique:



$U_f(|x_1 \dots x_n\rangle \otimes |b\rangle) =$
$|x_1 \dots x_n\rangle \otimes |b \oplus f(x_1 \dots x_n)\rangle$
où $x_1 \dots x_n \in \mathbf{F}_2^n$ et $b = 0$ ou 1.

Le lecteur constatera la généralité du problème formulé ci-dessus.

12.2 Dérivation de l'algorithme

Dans les chapitres précédents, nous avons toujours commencé par donner le circuit de l'algorithme. Cette fois nous procédons de façon plus inductive et verrons que l'algorithme de Grover peut être construit de façon assez naturelle. L'état initial entrant dans le circuit est :

$$\underbrace{|0 \dots 0\rangle}_{n \text{ fois}} \otimes |1\rangle$$

où les n premiers qubits stockent les entrées et le dernier bit est un bit auxiliaire. Pour exploiter le parallélisme quantique, nous formons une superposition cohérente sur toutes

les entrées possibles grâce aux portes de Hadamard agissant sur tous les qubits du premier registre (t_1) :

$$(H \otimes \ldots \otimes H \otimes \mathbb{1})|0 \ldots 0\rangle \otimes |1\rangle = \frac{1}{2^{n/2}} \sum_{x \in \mathbf{F}_2^n} |x\rangle \otimes |1\rangle$$

Ensuite, l'opération de Hadamard sur le second registre se révèle être utile pour obtenir le phénomène *kick-back phase* (t_2). L'état devient alors :

$$\frac{1}{2^{n/2}} \sum_{x \in \mathbf{F}_2^n} |x\rangle \otimes \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$$

Appliquons maintenant U_f (t_3). L'état devient

$$\frac{1}{2^{n/2}} \sum_{x \in \mathbf{F}_2^n} |x\rangle \otimes \frac{1}{\sqrt{2}}(|f(x)\rangle - |\overline{f(x)}\rangle)$$

On note que

$$|f(x)\rangle - |\overline{f(x)}\rangle = (-1)^{f(x)}(|0\rangle - |1\rangle)$$

ce qui donne l'état

$$\frac{1}{2^{n/2}} \sum_{x \in \mathbf{F}_2^n} (-1)^{f(x)}|x\rangle \otimes \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$$

Les opérations ci-dessus sont résumées sur la Figure **12.1**.

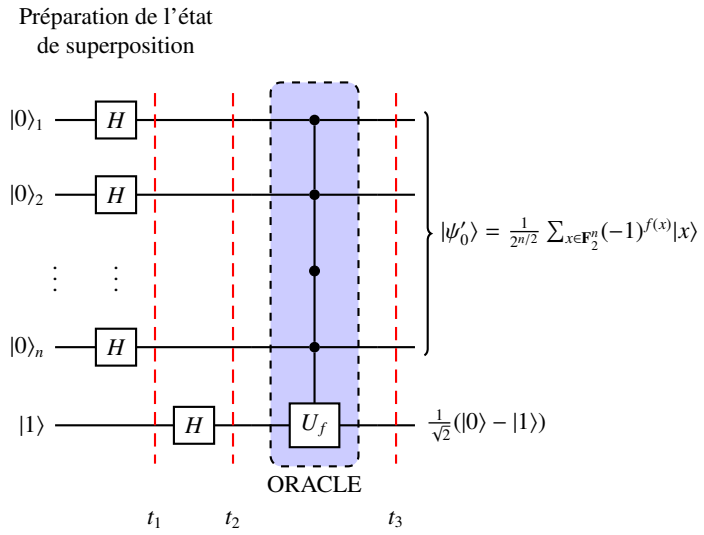


Figure 12.1 Début du circuit de l'algorithme de Grover.

Dans l'état résultant, nous voyons que les états *ket* $* x$ qui sont solution de l'équation $f(x) = 1$ ont été *marquée* d'une phase égale à -1 (ou π).

Supposons que le nombre de solutions est M et définissons deux états quantiques

$$|S\rangle \equiv \frac{1}{\sqrt{M}} \sum_{\text{x sol}} |x\rangle$$
$$|P\rangle \equiv \frac{1}{\sqrt{N-M}} \sum_{\text{x pas sol}} |x\rangle$$

où $N = 2^n$ (le nombre total d'états $|x\rangle$). On peut alors voir que :

$$|\psi_0\rangle = \frac{1}{\sqrt{N}} \sum_x |x\rangle = \sqrt{\frac{M}{N}}|S\rangle + \sqrt{\frac{N-M}{N}}|P\rangle$$
$$|\psi_0'\rangle = \frac{1}{\sqrt{N}} \sum_x (-1)^{f(x)}|x\rangle = -\sqrt{\frac{M}{N}}|S\rangle + \sqrt{\frac{N-M}{N}}|P\rangle$$

Puisque $\sqrt{\frac{M}{N}}$ et $\sqrt{\frac{N-M}{N}}$ sont ≤ 1 et $\left(\sqrt{\frac{M}{N}}\right)^2 + \left(\sqrt{\frac{N-M}{N}}\right)^2 = 1$ on peut définir un angle θ_0 tel que

$$\sin\theta_0 = \sqrt{\frac{M}{N}} \qquad \text{et } \cos\theta_0 = \sqrt{\frac{N-M}{N}}$$

Nous avons maintenant une interprétation géométrique du circuit ci-dessus. L'entrée du premier registre (les n premiers qubits juste avant l'oracle, dont l'état est $|\psi_0\rangle$) est un vecteur dans le plan $\{|P\rangle, |S\rangle\}$ formant un angle θ_0 avec l'axe $|P\rangle$. La sortie (i.e. après application de l'oracle) est l'état $|\psi_0'\rangle$ obtenu par réflexion par rapport à $|P\rangle$ et formant l'angle $-\theta_0$ avec ce dernier (Figure **12.2**).

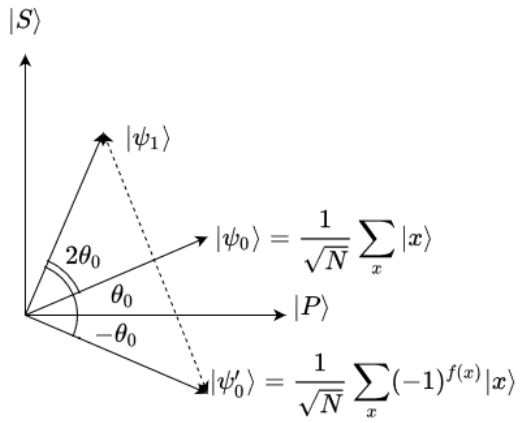


Figure 12.2

L'état réfléchi contient les solutions marquées par une phase -1. Nous devons maintenant agir sur cet état de façon à se rapprocher de l'axe $|S\rangle$ (qui est le sous-espace des solutions). Une idée naturelle est de faire une réflexion par rapport à $|\psi_0\rangle$, ce qui nous amène sur $|\psi_1\rangle$ (Figure **12.2**). On notera deux points importants :

- En obtenant, $|\psi_1\rangle$, on s'est rapproché de $|S\rangle$ (tout du moins si θ_0 était faible au départ, donc si $M << N$).
- La réflexion est faite par rapport à $|\psi_0\rangle$ et n'utilise aucune information sur les solutions (inconnues).

Nous avons

$$|\psi_1\rangle = \cos 3\theta_0|P\rangle + \sin 3\theta_0|S\rangle$$

Le troisième point crucial qu'il faut remarquer est que

- $|\psi_1\rangle$ est une rotation d'angle $2\theta_0$ de $|\psi_0\rangle$ dans le plan $\{|P\rangle, |S\rangle\}$.

En itérant ce procédé — réflexion autour de $|P\rangle$, puis réflexion autour de $|\psi_0\rangle$ — on obtient la suite d'états obtenus par rotations successives d'angle $2\theta_0$:

$$|\psi_k\rangle = \cos((2k+1)\theta_0)|P\rangle + \sin((2k+1)\theta_0)|S\rangle$$

Si θ_0 est assez petit (ce qui sera le cas en pratique car $M << N$) et si M est connu on peut choisir le nombre d'itérations k tel que : $(2k+1)\theta_0 \approx \frac{\pi}{2}$ pour rendre l'état presque parallèle à $|S\rangle$. Une mesure donnera alors une solution avec probabilité proche de 1. Cette analyse est présentée dans le prochain paragraphe.

Revenons maintenant à l'implémentation de la réflexion autour de $|\psi_0\rangle$ dans le circuit quantique. La réflexion d'un vecteur $\vec{v}$ par rapport à $\vec{\psi_0}$ correspond à l'opération suivante (prouvez le avec un dessin!)

$$\vec{v} \mapsto 2(\vec{\psi_0}^T \cdot \vec{v})\vec{\psi_0} - \vec{v}$$

En notation de Dirac :

$$\begin{aligned}|v\rangle \mapsto & 2|\psi_0\rangle\langle\psi_0|v\rangle - |v\rangle \\ = & (2|\psi_0\rangle\langle\psi_0| - \mathbb{1})|v\rangle \\ = & H^{\otimes n}(2|0\dots0\rangle\langle 0\dots0| - \mathbb{1})H^{\otimes n}|v\rangle\end{aligned}$$

L'opération dans les parenthèses (une matrice égale à 2 (projecteur sur $|0\dots0\rangle$) - Identité) a pour effet de changer la phase des entrées non nulles :

$$(2|0\dots0\rangle\langle 0\dots0| - \mathbb{1})|x\rangle = \begin{cases}|0\dots0\rangle & \text{si } x = (0\dots0) \\ -|x\rangle & \text{sinon}\end{cases} \tag{12.1}$$

Nous appelons cet opérateur *Phase Cond* pour *phase conditionnelle*. Le circuit implémentant la réflexion par rapport à $|\psi_0\rangle$ est montré à la Figure **12.3**.

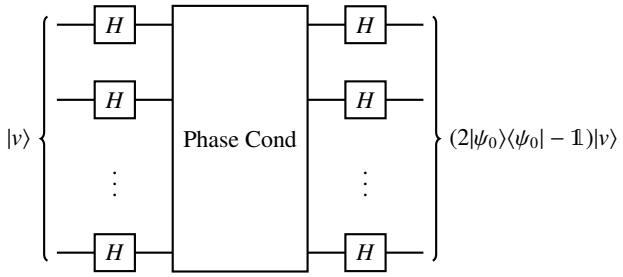


Figure 12.3 Circuit implémentant la réflexion d'un vecteur $|v\rangle$ par rapport à l'axe $|\psi_0\rangle$.

En combinant ce circuit avec celui de la Figure **12.1** nous obtenons le circuit correspondant à une itération du circuit de Grover (Figure **12.4**).

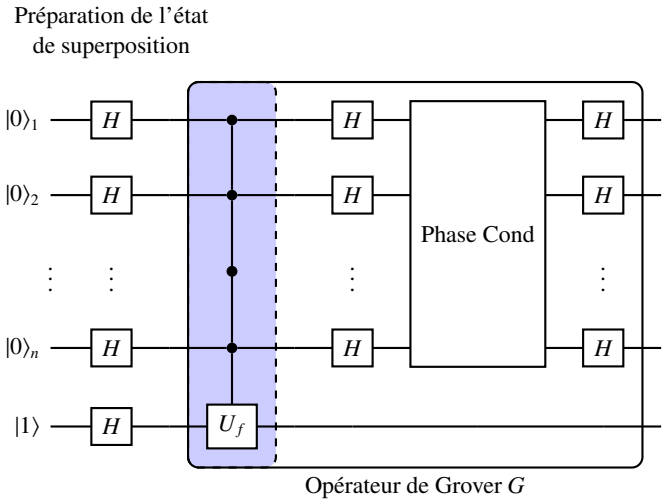


Figure 12.4 Définition de l'opérateur de Grover, qui utilise le circuit donné à la Figure **12.2**.

La boîte encadrée représente l'*opérateur de Grover* appelé G. Son effet sur

$$\frac{1}{\sqrt{N}}\sum_x |x\rangle \otimes \frac{|0\rangle - |1\rangle}{\sqrt{2}} = (\cos\theta_0|P\rangle + \sin\theta_0|S\rangle) \otimes \frac{|0\rangle - |1\rangle}{\sqrt{2}}$$

est d'effectuer une rotation d'angle $2\theta_0$ dans le plan $\{|P\rangle, |S\rangle\}$. La sortie est

$$(\cos 3\theta_0|P\rangle + \sin 3\theta_0|S\rangle) \otimes \frac{|0\rangle - |1\rangle}{\sqrt{2}}$$

L'algorithme de Grover consiste à itérer cette opération G un certain nombre de fois (k fois). Son circuit est donné à la Figure **12.5**.

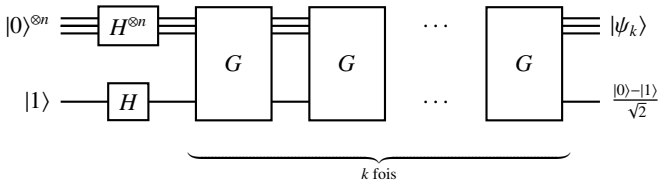


Figure 12.5 Circuit de l'algorithme de Grover. Chaque porte G correspond au circuit donné à la Figure **12.4**.

La profondeur (temps de calcul) du circuit est $O(k)$ et sa longueur $n+1$.

12.3 Analyse Probabiliste

Nous faisons une mesure sur les bits du premier registre dans la base computationnelle. La probabilité d'obtenir un état $|x\rangle$ qui est solution est :

$$|\langle S|\psi_k\rangle|^2 = \sin^2\Big((2k+1)\theta_0\Big)$$

Rappelons que $\sin(\theta_0) = \sqrt{\frac{M}{N}}$ et $\cos(\theta_0) = \sqrt{1-\frac{M}{N}}$.

Discutons d'abord le cas facile $M = \frac{N}{4}$.
Dans ce cas $\sin(\theta_0) = \frac{1}{2}$ et $\cos(\theta_0) = \frac{\sqrt{3}}{2}$ ce qui signifie $\theta_0 = \frac{\pi}{6}$. Au bout d'une itération ($k = 1$) l'état est aligné avec $|S\rangle$ car $3\theta_0 = \frac{\pi}{2}$. Nous trouvons donc 1 solution avec probabilité égale à 1 en une mesure et l'oracle a été consulté 1 fois. Bien sûr, ce cas est modérément intéressant car quand il y a $\frac{N}{4}$ solutions, on peut en trouver une avec probabilité $\frac{1}{4}$ en posant une question aléatoire à un oracle classique. Ensuite, on peut amplifier cette probabilité à $1-t$ en répétant l'expérience $O(|\ln e|)$ fois.

Prenons maintenant le cas $M = 1$
qui est intuitivement le plus difficile. Dans ce cas $\sin(\theta_0) = \frac{1}{\sqrt{N}}$ et donc l'angle θ_0 est très petit, $\theta_0 \approx \frac{1}{\sqrt{N}}$. Donc

$$\sin^2\Big((2k+1)\theta_0\Big) = \sin^2\left(\frac{2k+1}{\sqrt{N}}\right)$$

Pour $\frac{2k+1}{\sqrt{N}} \approx \frac{\pi}{2}$ cette probabilité est proche de 1. C'est-à-dire qu'avec $k = [\frac{\pi}{4}\sqrt{N}]$ (partie entière supérieure ou inférieure, c'est égal) la probabilité de succès est proche de 1. On peut voir que cette dernière est $1 - O\left(\frac{1}{N}\right)$.

Cas général avec M général connu.

- Si $M < \frac{3}{4}N$ alors $\sin(\theta_0) < \frac{\sqrt{3}}{2} \Rightarrow \theta_0 < \frac{\pi}{3}$.

  Itérons $\lfloor \frac{\pi}{4\theta_0} \rfloor = k$ fois.
  Puisque $\lfloor \frac{\pi}{4\theta_0} \rfloor = \frac{\pi}{4\theta_0} - \frac{1}{2} + \delta$ avec $|\delta| < \frac{1}{2}$ on a :

  $$(2k+1)\theta_0 = \frac{\pi}{2} + 2\delta\theta_0$$

  avec $2|\delta|\,\theta_0 < 2|\delta|\,\frac{\pi}{3} < \frac{\pi}{3}$. Donc $(2k+1)\theta_0 > \frac{\pi}{2} - \frac{\pi}{3}$ et $\sin^2\Big((2k+1)\theta_0\Big) > \sin^2\left(\frac{\pi}{2} - \frac{\pi}{3}\right) = \sin^2\left(\frac{\pi}{6}\right) = (\frac{1}{2})^2 = \frac{1}{4}$.
  La probabilité de succès est au moins $\frac{1}{4}$ et peut ensuite être amplifiée comme avant.
- Si $M > \frac{3}{4}N$ on a une probabilité de succès de $\frac{3}{4}$ grâce à une question classique.

Cas général avec M général inconnu.

Si M est inconnu, nous ne savons pas combien de fois itérer et le problème serait que nous n'itérions pas assez ou trop. L'algorithme suivant à une probabilité de succès de $\frac{1}{4}$ (ce qui nous suffit car celle-ci peut ensuite être amplifiée).

1 Prendre une entrée x aléatoirement uniformément. Si $f(x) = 1 \rightarrow$ SUCCES
2 Sinon, choisir $R \in \{0, 1, 2, \ldots, \sqrt{N} - 1\}$ uniformément aléatoirement et appliquer Grover avec R itérations. Mesurer la sortie.

Montrons que la probabilité de succès est toujours $\geq \frac{1}{4}$. Si $M \geq \frac{3}{4}N$ le point 1 à une probabilité de succès $\frac{3}{4}$ qui est plus grand que $\frac{1}{4}$.
Si $M < \frac{3}{4}N$, on a surement :

$$\mathbb{P}\,[\text{succès}] \geq \mathbb{P}\,[\text{succès au point 2}]$$

Mais

$$\mathbb{P}\,[\text{succès au point 2}] = \sum_{R=1}^{\sqrt{N}-1} \mathbb{P}\,[\text{succès au point 2}\,|\,R]\mathbb{P}\,[R]$$

$$= \frac{1}{\sqrt{N}} \sum_{R=0}^{\sqrt{N}-1} \sin^2\Big((2R+1)\theta_0\Big) = \frac{1}{2} - \frac{\sin\Big(4\theta_0\sqrt{N}\Big)}{4\sqrt{N}\sin\Big(2\theta_0\Big)}$$

Mais $\sin\Big(4\theta_0\sqrt{N}\Big) < 1$ et

$$\sin\Big(2\theta_0\Big) = 2\sin(\theta_0)\cos(\theta_0) = 2\sqrt{\frac{M}{N}}\sqrt{\frac{N-M}{M}} > 2\sqrt{\frac{M}{N}}\sqrt{\frac{N}{4N}} > \frac{1}{\sqrt{N}}$$

$$\Rightarrow \mathbb{P}\,[\text{succès au point 2}] \geq \frac{1}{2} - \frac{1}{4} = \frac{1}{4}$$

13 Calcul distribué - Problème de Deutsch-Jozsa distribué

Considérons la situation suivante; Alice et Bob possèdent chacun une machine quantique et veulent se partager une tâche. Par exemple, cette tâche pourrait être le calcul de la fonction $f(\vec{x}, \vec{y})$ où $\vec{x} \in \mathbf{F}_2^n$ est donné à Alice et $\vec{y} \in \mathbf{F}_2^n$ est donné à Bob. Selon le modèle de calcul considéré, on peut imaginer qu'Alice et Bob ne peuvent communiquer que via un canal de communication classique, ou alors un canal quantique, ou bien encore partagent des qubits intriqués. On s'intéressera uniquement à la complexité de la phase de communication, c.-à-d. par exemple le nombre de bits, qubits ou e-bits[1] échangés.

13.1 Modèles de calcul distribué

Tout d'abord, discutons un modèle classique, montré à la Figure **13.1**.

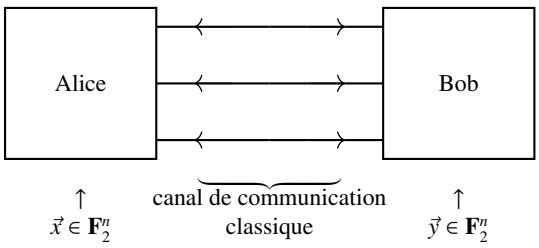


Figure 13.1 Modèle classique de calcul distribué.

Alice et Bob doivent calculer $f(\vec{x}, \vec{y})$ et on permet T échanges de bits classiques. La complexité du protocole est le minimum possible pour T sans compter la complexité du calcul de f elle-même (par exemple, Alice et Bob disposent de deux oracles). Un théorème classique affirme qu'il n'est pas possible de faire mieux que $T = \alpha n$ avec $\alpha = 0.007$.

Nous allons voir que dans deux modèles quantiques différents, on peut faire mieux avec $T = O(\log_2(n))$. Les modèles principaux introduits ici sont dus à Yao et Cleve-Buhrman.

[1] entangled-bits

Modèle de Yao:

Alice et Bob partagent un canal de communication quantique et échangent T qubits, comme montré à la Figure **13.2**. Les opérations qu'ils effectuent sont des opérations locales et unitaires et leurs mesures sont locales.

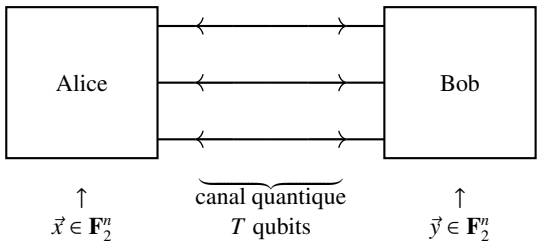


Figure 13.2 Modèle quantique de Yao de calcul distribué.

Modèle de Cleve-Buhrman:

Alice et Bob partagent des paires intriquées (ou de l'intrication générale). Ils peuvent aussi communiquer par un canal classique, comme montré à la Figure **13.3**. Les opérations qu'ils effectuent sont des opérations locales et unitaires et leurs mesures sont locales.

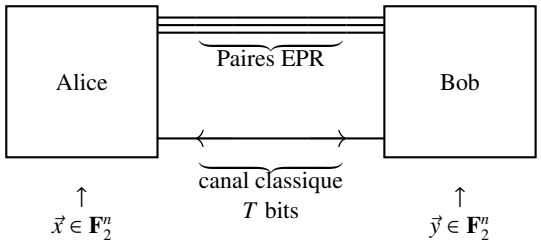


Figure 13.3 Modèle quantique de Cleve-Burhman de calcul distribué.

Notons qu'en fait, on peut montrer que les deux modèles sont équivalents grâce au protocole de téléportation quantique (Section **6.3**). En effet, dans le second modèle, en partageant T paires intriquées et en échangeant $2T$ bits classiques, on peut se ramener au premier modèle où on échangerait $2T$ bits quantiques.

Nous allons traiter les deux modèles séparément. Pour le second, l'analyse qui sera présentée n'utilisera pas le protocole de téléportation et nous verrons que l'on peut même se dispenser du facteur 2.

Problème de Deutsch-Jozsa distribué:

Nous n'allons pas travailler dans un contexte général mais plutôt regarder le problème particulier suivant. Soient deux vecteurs $\vec{x} \in \mathbf{F}_2^n$ et $\vec{y} \in \mathbf{F}_2^n$ (et n une puissance de 2)[2]. De plus, supposons la promesse suivante vraie: Soit $\vec{x} = \vec{y}$, soit ces deux vecteurs sont différents mais possèdent $\frac{n}{2}$ composantes opposées et $\frac{n}{2}$ composantes égales. En d'autres termes, si $d_H(\vec{x}, \vec{y})$ est la distance de Hamming[3] entre $\vec{x}$ et $\vec{y}$, on a :

$$d_H(\vec{x}, \vec{y}) = 0 \qquad \text{ou bien} \qquad d_H(\vec{x}, \vec{y}) = \frac{n}{2} \tag{13.1}$$

Il s'agit donc d'un problème de décision et le but d'Alice et de Bob est de décider dans lequel de ces deux cas se trouvent leurs vecteurs $\vec{x}$ et $\vec{y}$ tout en minimisant T.

Quelle est la fonction $f(\vec{x}, \vec{y})$ qu'ils doivent calculer dans ce problème particulier? Remarquons que (13.1) est équivalent à

$$\frac{1}{n}\sum_{i=0}^{n-1}(-1)^{x_i+y_i} = \begin{cases} 1 & \text{si} \quad d_H(\vec{x}, \vec{y}) = 0 \\ 0 & \text{si} \quad d_H(\vec{x}, \vec{y}) = \frac{n}{2} \end{cases} \tag{13.2}$$

La fonction que nous voulons calculer ici est $\vec{f} : \mathbf{F}_2^n \times \mathbf{F}_2^n \to \{-1, 1\}^n$ définie par

$$\vec{f}(\vec{x}, \vec{y}) = \begin{pmatrix} f_0(\vec{x}, \vec{y}) \\ f_1(\vec{x}, \vec{y}) \\ \vdots \\ f_{n-1}(\vec{x}, \vec{y}) \end{pmatrix} \tag{13.3}$$

$$f_i(\vec{x}, \vec{y}) \equiv (-1)^{x_i+y_i} = (-1)^{x_i}(-1)^{y_i} \tag{13.4}$$

En d'autres termes, Alice possède $\vec{x}$ et calcule $(-1)^{x_i}$ pour chaque indice $i = 0 \ldots n-1$ et Bob possède $\vec{y}$ et calcule $(-1)^{y_i}$ pour chaque indice $i = 0 \ldots n-1$. Ensuite, ils doivent parvenir (par exemple en communiquant) à calculer tous les produits $(-1)^{x_i}(-1)^{y_i}$, $i = 0 \ldots n-1$ pour décider si $\frac{1}{n}\sum_{i=0}^{n-1}(-1)^{x_i+y_i}$ vaut 0 ou 1.

Bien sûr, on peut vérifier cela avec $T = O(n)$ en échangeant n bits classiques. Nous verrons que pour les deux modèles quantiques proposés, on peut se débrouiller avec $T = O(\log_2(n))$.

13.2 Analyse du modèle de Yao

Alice et Bob possèdent chacun un circuit qui ressemble à celui du problème standard de Deutsch-Jozsa. Commençons par le circuit d'Alice, donné à la Figure **13.4**. On pose $n = 2^k$, $k = \log_2(n)$ par souci de simplicité.

[2] Notez que dans ce chapitre, nous utilisons une convention pour les composantes de $\vec{x}$ et $\vec{y}$, car nous commençons à l'indice 0 jusqu'à l'indice $n-1$, c.-à-d. $\vec{x} = \begin{pmatrix} x_0 & \ldots & x_{n-1} \end{pmatrix}$, au lieu de compter de 1 à n.

[3] La distance de Hamming entre deux bits est égale au nombre de composantes différentes entre les deux bits: $d_H(\vec{x}, \vec{y}) = \#\{i \mid x_i \neq y_i\}$.

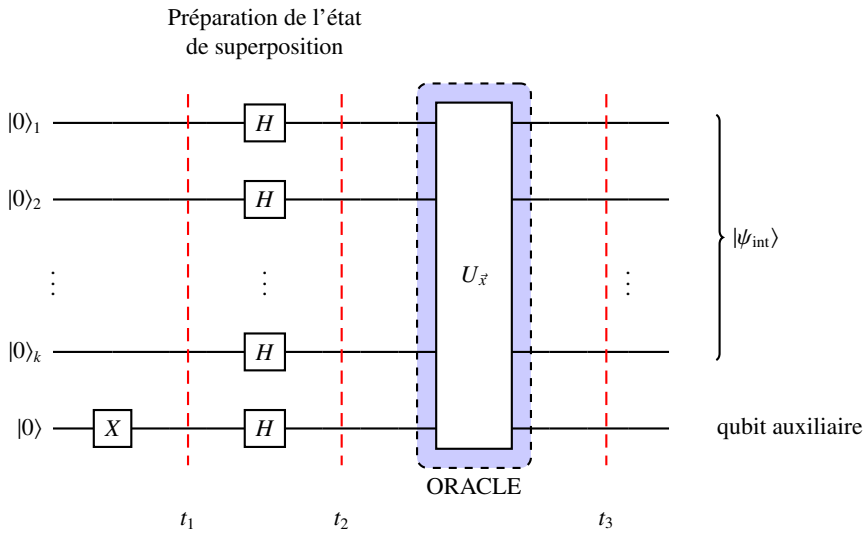


Figure 13.4 Circuit d'Alice dans le modèle de calcul distribué quantique de Yao.

À l'instant t_1, l'état est

$$|\psi_1\rangle = |0\rangle^{\otimes k} \otimes |1\rangle \tag{13.5}$$

À l'instant t_2, l'état est

$$\begin{aligned} |\psi_2\rangle &= \frac{1}{2^{\frac{k}{2}}} \sum_{\substack{b_1 \dots b_k \\ \in \{0,1\}^k}} |b_1 \dots b_k\rangle \otimes \left(\frac{|0\rangle - |1\rangle}{\sqrt{2}} \right) \\ &= \frac{1}{\sqrt{n}} \sum_{i=0}^{n-1} |i\rangle \otimes \left(\frac{|0\rangle - |1\rangle}{\sqrt{2}} \right) \end{aligned} \tag{13.6}$$

où $|i\rangle = |b_1 \dots b_k\rangle = |b_1\rangle \otimes \dots \otimes |b_k\rangle \in \left(\mathbb{C}^2\right)^{\otimes k}$ avec $b_1 \dots b_k$ la représentation binaire de i.

Pour l'opération unitaire $U_{\vec{x}}$ (l'oracle chez Alice), on a par définition

$$U_{\vec{x}}(|i\rangle \otimes |z\rangle) = |i\rangle \otimes |x_i \oplus z\rangle \qquad z \in \{0,1\}, i = 0 \dots n-1 \tag{13.7}$$

L'état à l'instant t_3, après application de l'oracle, est donc

$$\begin{aligned} |\psi_3\rangle &= \frac{1}{\sqrt{n}} \sum_{i=0}^{n-1} |i\rangle \otimes \left(\frac{|x_i\rangle - |1 \oplus x_i\rangle}{\sqrt{2}} \right) \\ &= \underbrace{\left\{ \frac{1}{\sqrt{n}} \sum_{i=0}^{n-1} (-1)^{x_i} |i\rangle \right\}}_{|\psi_{\text{int}}\rangle} \otimes \left(\frac{|0\rangle - |1\rangle}{\sqrt{2}} \right) \end{aligned} \tag{13.8}$$

À ce point, notons qu'Alice possède un état intermédiaire $|\psi_{\text{int}}\rangle = \frac{1}{\sqrt{n}} \sum_{i=0}^{n-1} (-1)^{x_i} |i\rangle$ avec $k = \log_2(n)$ qubits. Cet état stocke en parallèle toutes les informations $(-1)^{x_i}$ sur "sa partie" de la fonction f.

Alice envoie cet état à Bob. Qu'est-ce que cela veut dire physiquement? Cet état est incarné dans $k = \log_2(n)$ "particules" (photons, spins nucléaires, etc) et Alice envoie donc k "particules" à Bob via un canal de communication quantique (p.ex. une fibre optique pour des photons). Bob reçoit cet état et va l'utiliser comme entrée dans son circuit, montré à la Figure **13.5**.

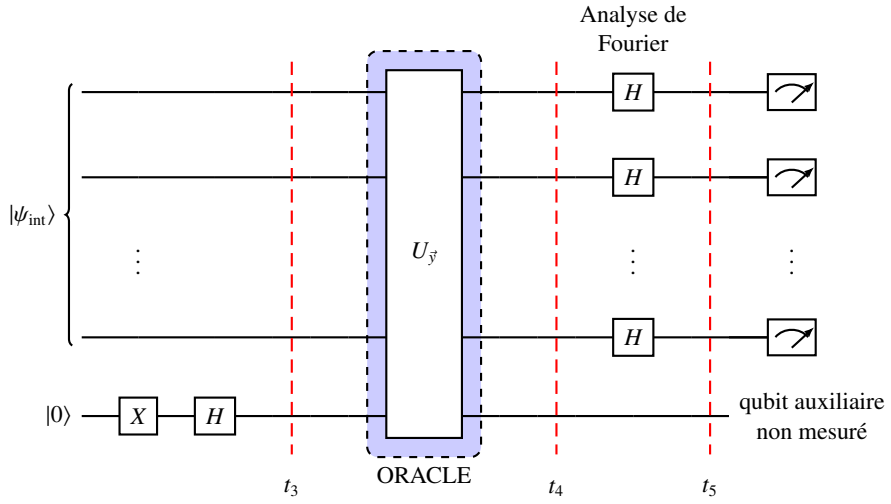


Figure 13.5 Circuit de Bob dans le modèle de calcul distribué quantique de Yao.

Après avoir reçu l'état d'Alice et en prenant en compte le qubit auxiliaire de son circuit, l'état de Bob à l'instant t_3 est donné par

$$|\psi_3\rangle = |\psi_{\text{int}}\rangle \otimes \left(\frac{|0\rangle - |1\rangle}{\sqrt{2}} \right) \tag{13.9}$$

L'état à l'instant t_4, après la porte unitaire $U_{\vec{y}}$ (l'oracle chez Bob), est donc

$$\begin{aligned} |\psi_4\rangle &= U_{\vec{y}} \left(\left\{ \frac{1}{\sqrt{n}} \sum_{i=0}^{n-1} (-1)^{x_i} |i\rangle \right\} \otimes \left(\frac{|0\rangle - |1\rangle}{\sqrt{2}} \right) \right) \\ &= \frac{1}{\sqrt{n}} \sum_{i=0}^{n-1} (-1)^{x_i} U_{\vec{y}} \left(|i\rangle \otimes \left(\frac{|0\rangle - |1\rangle}{\sqrt{2}} \right) \right) \end{aligned} \tag{13.10}$$

À nouveau, par définition, on a

$$U_{\vec{y}} (|i\rangle \otimes |z\rangle) = |i\rangle \otimes |y_i \oplus z\rangle \qquad z \in \{0, 1\}, i = 0 \ldots n-1 \tag{13.11}$$

On trouve chez Bob:

$$\begin{aligned} |\psi_4\rangle &= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} (-1)^{x_i} |i\rangle \otimes \left(\frac{|y_i\rangle - |1 \oplus y_i\rangle}{\sqrt{2}} \right) \\ &= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} (-1)^{x_i + y_i} |i\rangle \otimes \left(\frac{|0\rangle - |1\rangle}{\sqrt{2}} \right) \end{aligned} \tag{13.12}$$

Il s'agit maintenant d'extraire l'information chez Bob. Pour cela, on procède comme dans l'algorithme standard de Deutsch-Josza, c'est-à-dire par une "analyse de Fourier" suivie d'une mesure dans la base computationnelle.

À l'instant t_5, après les portes de Hadamard, l'état est:

$$|\psi_5\rangle = \frac{1}{\sqrt{n}} \sum_{i=0}^{n-1} (-1)^{x_i + y_i} H^{\otimes k} |i\rangle \otimes \left(\frac{|0\rangle - |1\rangle}{\sqrt{2}} \right) \tag{13.13}$$

Or, on a que

$$\begin{aligned} H^{\otimes k} |i\rangle &= H^{\otimes k} |b_1 \dots b_k\rangle \\ &= H|b_1\rangle \otimes \dots \otimes H|b_k\rangle \\ &= \frac{1}{2^{\frac{k}{2}}} \sum_{\substack{c_1 \dots c_k \\ \in \{0,1\}^k}} (-1)^{b_1 c_1 + \dots + b_k c_k} |c_1 \dots c_k\rangle \end{aligned} \tag{13.14}$$

Donc, l'état juste avant la mesure est donné par:

$$\begin{aligned} |\psi_5\rangle &= \frac{1}{\sqrt{n}} \sum_{i=0}^{n-1} (-1)^{x_i + y_i} \left\{ \frac{1}{\sqrt{n}} \sum_{\substack{c_1 \dots c_k \\ \in \{0,1\}^k}} (-1)^{b_1 c_1 + \dots + b_k c_k} |c_1 \dots c_k\rangle \right\} \otimes \left(\frac{|0\rangle - |1\rangle}{\sqrt{2}} \right) \\ &= \sum_{j=0}^{n-1} \left\{ \frac{1}{n} \sum_{i=0}^{n-1} (-1)^{x_i + y_i} (-1)^{\langle i,j \rangle} \right\} |j\rangle \otimes \left(\frac{|0\rangle - |1\rangle}{\sqrt{2}} \right) \end{aligned} \tag{13.15}$$

où $\langle i, j \rangle$ est le produit scalaire entre les deux représentations binaires des entiers i et j. Nous voyons que Bob possède un état de superposition de la base computationnelle et qui contient toutes les valeurs $(-1)^{x_i + y_i}$ $i = 0 \dots n-1$ stockées dans $|\psi_5\rangle$.

La mesure dans la base computationnelle donne un état $|j\rangle$, $j = 0 \dots n-1$ avec probabilité:

$$\mathbb{P}\left[j\right] = \left| \frac{1}{n} \sum_{i=0}^{n-1} (-1)^{x_i + y_i} (-1)^{\langle i,j \rangle} \right|^2 \tag{13.16}$$

En particulier, la probabilité d'obtenir l'entier $j = 0$ est:

$$\mathbb{P}\left[j = 0\right] = \frac{1}{n^2} \left| \sum_{i=0}^{n-1} (-1)^{x_i + y_i} \right|^2 = \begin{cases} 1 & \text{si} \quad \vec{x} = \vec{y} \\ 0 & \text{si} \quad d_H(\vec{x}, \vec{y}) = \frac{n}{2} \end{cases} \tag{13.17}$$

En conclusion, l'observation de l'entier 0 ou l'absence de cette observation permet de décider si $d_H(\vec{x}, \vec{y}) = 0$ ou si $d_H(\vec{x}, \vec{y}) = \frac{n}{2}$. Rappelons que dans cet algorithme, Alice a envoyé $k = \log_2(n)$ qubits à Bob.

13.3 Analyse du modèle de Cleve-Buhrman

Nous supposons qu'Alice et Bob partagent l'état intriqué suivant:

$$|\psi\rangle = \frac{1}{\sqrt{n}} \sum_{i=0}^{n-1} |i\rangle_A \otimes |i\rangle_B = \frac{1}{2^{\frac{k}{2}}} \sum_{\substack{b_1 \dots b_k \\ \in \{0,1\}^k}} |b_1 \dots b_k\rangle_A \otimes |b_1 \dots b_k\rangle_B \tag{13.18}$$

Notez qu'Alice et Bob possèdent chacun k qubits. Cet état n'est rien d'autres que le produit tensoriel de k paires EPR. Alice et Bob utilisent maintenant chacun le circuit de la Figure **13.6**

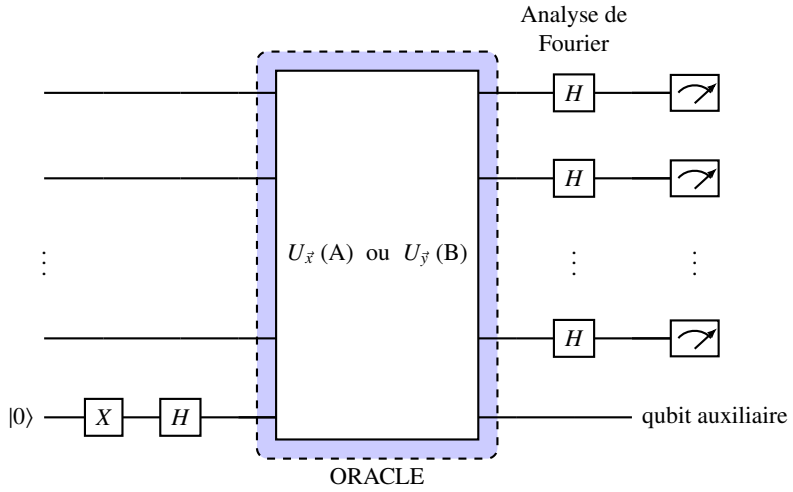


Figure 13.6 Circuits de Alice ($U_{\vec{x}}$) et Bob ($U_{\vec{y}}$) dans le modèle de calcul distribué quantique de Cleve-Buhrman.

dont l'état d'entrée pour chacun est la partie appropriée de l'état intriqué (13.18). Après les opérations $U_{\vec{x}}$ et $U_{\vec{y}}$, l'état (distribué et intriqué) est:

$$|\psi'\rangle = \frac{1}{\sqrt{n}} \sum_{i=0}^{n-1} (-1)^{x_i + y_i} |i\rangle_A \otimes \left(\frac{|0\rangle - |1\rangle}{\sqrt{2}}\right) \otimes |i\rangle_B \otimes \left(\frac{|0\rangle - |1\rangle}{\sqrt{2}}\right) \tag{13.19}$$

La prochaine étape avant la mesure est l'analyse de Fourier chez Alice et Bob. En laissant tomber les deux qubits auxiliaires $\left(\frac{|0\rangle - |1\rangle}{\sqrt{2}}\right)^{\otimes 2}$ qui sont maintenant inutiles, on a:

$$
\begin{aligned}
|\psi''\rangle &= \frac{1}{\sqrt{n}} \sum_{i=0}^{n-1} (-1)^{x_i+y_i} H^{\otimes k}|i\rangle_A \otimes H^{\otimes k}|i\rangle_B \\
&= \frac{1}{n^{\frac{3}{2}}} \sum_{i=0}^{n-1} (-1)^{x_i+y_i} \sum_{j_1=0}^{n-1} (-1)^{\langle i,j_1\rangle}|j_1\rangle_A \otimes \sum_{j_2=0}^{n-1} (-1)^{\langle i,j_2\rangle}|j_2\rangle_B \\
&= \frac{1}{n^{\frac{3}{2}}} \sum_{j_1=0}^{n-1}\sum_{j_2=0}^{n-1} \left\{ \sum_{i=0}^{n-1} (-1)^{x_i+y_i}(-1)^{\langle i,j_1\rangle}(-1)^{\langle i,j_2\rangle} \right\} |j_1\rangle_A \otimes |j_2\rangle_B
\end{aligned}
\tag{13.20}
$$

Ici, comme avant, $\langle i, j_1\rangle$ et $\langle i, j_2\rangle$ sont les produits scalaires des représentations binaires des entiers $i, j_1, j_2 \in \{0 \dots n-1\}$.

La mesure donne un résultat (aléatoire) $|j_1\rangle_A \otimes |j_2\rangle_B$ (donc $|j_1\rangle$ chez Alice et $|j_2\rangle$ chez Bob) avec probabilité:

$$
\mathbb{P}\left[j_1, j_2\right] = \frac{1}{n^3} \left| \sum_{i=0}^{n-1} (-1)^{x_i+y_i}(-1)^{\langle i,j_1\rangle+\langle i,j_2\rangle} \right|^2 \tag{13.21}
$$

En particulier:

$$
\mathbb{P}\left[j_1 = j, j_2 = j\right] = \frac{1}{n^3} \left| \sum_{i=0}^{n-1} (-1)^{x_i+y_i}(-1)^{2\langle i,j\rangle} \right|^2 = \frac{1}{n^3} \left| \sum_{i=0}^{n-1} (-1)^{x_i+y_i} \right|^2 \tag{13.22}
$$

Ainsi, la probabilité qu'Alice et Bob obtiennent le même entier est:

$$
\mathbb{P}\left[j_1 = j_2\right] = \sum_{j=0}^{n-1} \mathbb{P}\left[j_1 = j, j_2 = j\right] = \frac{1}{n^2} \left| \sum_{i=0}^{n-1} (-1)^{x_i+y_i} \right|^2 = \begin{cases} 1 & \text{si} \quad d_H(\vec{x},\vec{y}) = 0 \\ 0 & \text{si} \quad d_H(\vec{x},\vec{y}) = \frac{n}{2} \end{cases} \tag{13.23}
$$

Dans cet algorithme, Alice et Bob doivent encore échanger les k bits de j_1 et j_2 obtenus pour comparer si $j_1 = j_2$ ou si $j_1 \neq j_2$. Puisque $k = \log_2(n)$, on a la complexité d'échange $T = \log_2(n)$ bits classiques.

14 Correction d'erreur en MQ

Dans la théorie classique, la façon usuelle de compenser les effets du bruit sur le stockage ou la communication des données est d'introduire de la redondance de façon suffisamment structurée. C'est ce que l'on appelle la correction d'erreurs. La correction d'erreurs est efficace dans la mesure où l'information est digitale. En effet, pour des signaux analogues, les perturbations dues aux bruits sont continues ce qui rend la correction d'erreurs impraticable. On pourrait maintenant croire que dans le cas quantique, la correction d'erreurs est également impossible (ou néanmoins difficile) car les états de l'espace de Hilbert forment un continuum. D'autre part, le processus de mesure tend à détruire l'état quantique du système, ce qui n'est pas le cas dans le monde classique. Parfois, des arguments ont été avancés pour dire que la correction 'd'erreurs n'est pas possible pour des états quantiques, mais il s'est avéré (suite aux travaux de Shor) qu'il n'en est rien. La nature discrète de la MQ se manifeste dans la classification des perturbations possibles et on peut encore construire des codes correcteurs d'erreurs.

Dans ce chapitre, nous donnons quelques constructions élémentaires de codes quantiques. Les idées générales seront développées en partant des codes les plus simples: les codes de répétitions quantiques. En concaténant deux codes de répétitions quantiques - relatifs à des bases différentes de l'espace de Hilbert - on obtient déjà un code intéressant! Comme nous le verrons, on peut développer des constructions assez générales en élaborant cette idée.

14.1 Bref rappel sur les codes linéaires classiques

Le code correcteur le plus simple que l'on puisse imaginer est le code de répétition. Supposons que l'on veuille transmettre un bit 0 ou 1 à travers un canal BSC[1] qui renverse le bit avec probabilité $0 < p < \frac{1}{2}$. On pourrait utiliser le code de répétition

$$0 \longrightarrow 000$$
$$1 \longrightarrow 111$$

Pour aucun ou un seul renversement, on peut facilement détecter puis corriger l'erreur

[1] Canal Binaire Symétrique

grâce à la règle de la majorité. Par exemple, si 010 est reçu, on en conclu que très probablement le deuxième bit a été renversé et on décode en faisant l'opération $010 \to 000$. Ce faisant, peut-être que l'on fait une *erreur de décodage*; il se pourrait qu'en fait 111 était effectivement transmis et que les premier et troisième bits aient été renversés par le canal. Mais la probabilité de cet événement est faible si p est assez petit. Sans répétition, la probabilité d'erreur de décodage sera toujours égale à p alors qu'avec répétition elle est égale à $p^3 + 3p^2(1-p) = 3p^2 - 2p^3$ et est inférieure à p pour tout $0 < p < \frac{1}{2}$.

Définition. Un code linéaire est un sous-espace vectoriel de dimension k de $\mathbf{F}_2^n$. Le nombre de bits d'information est k, le nombre de mots de code est 2^k, leur longueur est n. Le rendement du code est défini par le rapport entre le nombre de bits d'information contenus dans un mot et sa longueur. On dit que les paramètres du code sont (n, k). Il existe deux représentations fondamentales d'un code linéaire. L'une utilise la matrice génératrice et l'autre la matrice de parité.

Matrice génératrice. Un code binaire linéaire peut être décrit par une application linéaire

$$\begin{aligned} \mathbf{F}_2^k &\to \mathbf{F}_2^n \\ \vec{u} &\mapsto \vec{x} = G\vec{u} \end{aligned}$$

où $\vec{u} = \begin{pmatrix} u_1 \\ \vdots \\ u_k \end{pmatrix}$ est un vecteur colonne contenant k bits d'information et G une matrice $n \times k$ (d'éléments 0 ou 1) . Le vecteur $G\vec{u}$ est le mot de code, vecteur binaire à n composantes. Toutes les opérations sont faites (mod 2). Pour que l'image de $\mathbf{F}_2^k$ dans $\mathbf{F}_2^n$ soit un sous-espace vectoriel de dimension k, on prend une matrice G de rang k. Autrement dit, les k colonnes de G sont linéairement indépendantes. La matrice G s'appelle la matrice génératrice.

Par exemple, le code de répétition (3,1) correspond à la matrice 3×1:

$$G = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

Matrice de Parité. Puisqu'un code linéaire est un sous-espace vectoriel de $\mathbf{F}_2^n$, il peut être vu comme le noyau d'une matrice H, c'est-à-dire l'ensemble des vecteurs solutions de

$$H\vec{x} = 0, \qquad H \text{ de dimension } (n-k) \times n$$

Puisque l'on veut $\dim(\ker(H)) = k$ et que $\dim(\ker(H)) + \dim(\mathrm{Im}(H)) = n$, il faut avoir $\dim(\mathrm{Im}(H)) = n-k$. Autrement dit, H doit posséder au moins $n-k$ lignes indépendantes. En général, on choisira H de dimension $(n-k) \times n$, mais ce choix n'est pas unique.

Relation entre les deux matrices. On observe que pour un code donné, on a $HG = 0$; H et G sont duales. Pour construire H à partir de G, on peut écrire $G = [\vec{g}_1 \dots \vec{g}_k]$ où $\vec{g}_1 \dots \vec{g}_k$ sont indépendants et prendre $n - k$ vecteurs orthogonaux $\vec{h}_1 \dots \vec{h}_{n-k}$ au sous-espace engendré par $\vec{g}_1 \dots \vec{g}_k$. Alors

$$H = \begin{bmatrix} \vec{h}_1^T \\ \vdots \\ \vec{h}_{n-k}^T \end{bmatrix}$$

Réciproquement, pour construire G à partir de H on prend les $n - k$ vecteurs lignes indépendants de H et on construit k vecteurs orthogonaux (au sous-espace engendré par les lignes de H) qui forment les k colonnes de G.

Par exemple, une matrice de parité pour le code de répétition (3,1) est la matrice 2×3:

$$H = \begin{bmatrix} 1 & -1 & 0 \\ 0 & -1 & 1 \end{bmatrix}$$

Détection, correction et distance minimale. La matrice de parité permet de détecter certaines erreurs (mais pas toutes). Supposons que le mot $\vec{x}$ soit transmis à travers un canal et $\vec{x}'$ reçu : $\vec{x}' = \vec{x} + \vec{e}$ où $\vec{e}$ est un vecteur contenant des 1 pour les bits erronés et 0 ailleurs. On a :

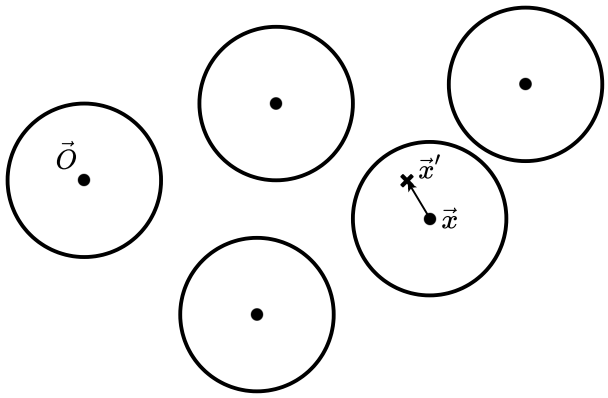
$$H\vec{x}' = H\vec{x} + H\vec{e} = H\vec{e}$$

On appelle $H\vec{e}$ *"le syndrome"*. Si celui-ci est non-nul, il indique que le mot reçu est erroné. Notez que s'il est nul, il peut ou non y avoir une erreur de transmission. Pour le canal BSC, la probabilité d'avoir un vecteur d'erreur $\vec{e}$ est $p^{|\vec{e}|}(1-p)^{n-|\vec{e}|}$ et il est naturel (pour p petit) de décoder $\vec{x}'$ en déclarant que $\vec{x}$ est le vecteur qui minimise $d(\vec{x}, \vec{x}') = |\vec{e}|$, la distance de Hamming entre $\vec{x}$ et $\vec{x}'$. En effet, $\vec{x}$ est le vecteur transmis le plus probable (qui maximise $p^{|\vec{e}|}(1-p)^{n-|\vec{e}|}$) étant donné que $\vec{x}'$ a été reçu (les vecteurs transmis sont choisis uniformément). C'est le décodeur dit de *"distance minimale"*.

Un paramètre important qui caractérise la qualité d'un code est sa distance minimale $d = \min_{x,x' \in C} d(\vec{x}, \vec{x}')$. Pour un code linéaire, on a aussi $d = \min_{x \neq 0 \in C} d(\vec{x}, \vec{0})$. On parle alors de code (n, k, d) : longueur n, dimension k et distance minimale d.

Théorème : *Soit un code (n, k, d). On peut toujours corriger jusqu'à t erreurs avec $d \geq 2t + 1$ grâce au décodeur de la distance minimale.*

Si on considère les boules de Hamming de rayon $\frac{d-1}{2}$ autour des mots de codes, celles-ci ne s'intersectent pas. De plus, si le nombre d'erreurs est $\leq \frac{d-1}{2}$ alors le mot reçu $\vec{x}'$ est certainement dans une et une seule des boules. On déclare que le mot transmis est l'unique mot de code dans la même boule que $\vec{x}'$.



Théorème : *Pour tout code linéaire, le nombre minimal de colonnes de H qui sont linéairement indépendantes donne exactement d. En particulier, toute collection de d-1 colonnes sont linéairement indépendantes.*

Pour s'en convaincre, il suffit de remarquer que si $\vec{x}$ est un mot de code, $H\vec{x} = 0$, ce qui signifie que les colonnes de H indexées par les composantes non nulles de $\vec{x}$ sont linéairement indépendantes.

Rappel sur les codes de Hamming. Les codes de Hamming sont commodément décrits par leur matrice de parité. Soit $r \geq 2$ et H la matrice de parité dont les colonnes sont tous les $2^r - 1$ vecteurs binaires non nuls de longueur r. On a $n = 2^r - 1$ et $n - k = r \Rightarrow k = 2^r - r - 1$. On obtient donc un code $(n, k) = (2^r - 1, 2^r - r - 1)$ de longueur $2^r - 1$ et dimension $2^r - r - 1$. Il faut vérifier que le rang de cette matrice est bien r : cela est vrai car parmi les $2^r - 1$ vecteurs binaires non nuls, il y en a r indépendants qui engendrent tout $\mathbf{F}_2^r$. Pour les codes de Hamming, on voit facilement que $d = 3$. En effet, l'ensemble des colonnes est formé de *tous* les vecteurs binaires non-nuls: il y en a nécessairement trois qui sont linéairement dépendants, et de plus toute paire de colonnes distinctes sont différentes. Ainsi, ces codes corrigent une erreur. Par exemple, le code de Hamming $(7, 4, 3)$ possède la matrice (ici $r = 3$)

$$H = \begin{bmatrix} 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{bmatrix}$$

Supposons qu'une erreur soit produite dans le $3^{\text{ième}}$ bit. Alors c'est $\vec{e} = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$ et le

syndrome vaut $H\vec{e} = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$ qui est égal à la $3^{\text{ième}}$ colonne de H. Le syndrome nous montre automatiquement quel bit est erroné et il suffit de le renverser pour effectuer la correction d'erreur. On est assuré que l'erreur est correctement détectée et corrigée *seulement* si elle est unique.

14.2 Codes de Répétition Quantique

Par analogie avec le cas classique, on peut considérer le code de répétition

$$|0\rangle \rightarrow |000\rangle$$
$$|1\rangle \rightarrow |111\rangle.$$

que nous allons utiliser sur un *canal bit-flip*.

Quel est le procédé de détection et correction dans le cas quantique? Il faut choisir correctement la base de mesure pour ne pas détruire irrémédiablement l'information contenue dans l'état sortant! Ici l'espace de Hilbert à la sortie du canal est de dimension $2^3 = 8$ et il faut donc une base qui contient 8 états. Considérons les 4 projecteurs de dimension 2 chacun (8 donc au total)

$$P_0 = |000\rangle\langle 000| + |111\rangle\langle 111|$$
$$P_1 = |100\rangle\langle 100| + |011\rangle\langle 011|$$
$$P_2 = |010\rangle\langle 010| + |101\rangle\langle 101|$$
$$P_3 = |001\rangle\langle 001| + |110\rangle\langle 110|.$$

P_0 projette dans le sous-espace à 0 erreurs, P_1 dans le sous-espace à une erreur sur le premier bit, P_2 dans le sous-espace à une erreur dans le deuxième bit, P_3 dans le sous-espace à une erreur sur le $3^{\text{ième}}$ bit. L'entrée du canal est le ket $\alpha |0\rangle + \beta |1\rangle$ codé :

$$\alpha |000\rangle + \beta |111\rangle$$

Notez que ce code de répétition ne viole pas le "no-cloning theorem" car on répète uniquement des états de base orthonormés. Ici, on ne répète pas l'état $\alpha |0\rangle + \beta |1\rangle$, ce qui violerait le théorème. Si une seule erreur est produite, mettons pour le deuxième bit (avec probabilité p) la sortie est

$$\alpha |010\rangle + \beta |101\rangle$$

Une mesure dans la base $\{P_0, P_1, P_2, P_3\}$ préserve l'état avec probabilité 1 car

$$P_2\Big(\alpha\,|010\rangle + \beta\,|101\rangle\Big) = \alpha\,|010\rangle + \beta\,|101\rangle$$
$$P_i\Big(\alpha\,|010\rangle + \beta\,|101\rangle\Big) = 0, \qquad \mathrm{i} \neq 2.$$

Ainsi, on détecte que l'erreur est dans le deuxième bit sans détruire l'état. La correction d'erreur se fait en appliquant l'opérateur unitaire $\mathbb{1}\otimes X\otimes\mathbb{1}$ qui renverse le deuxième bit.

$$\mathbb{1}\otimes X\otimes\mathbb{1}\Big(\alpha\,|010\rangle + \beta\,|101\rangle\Big) = \alpha\,|000\rangle + \beta\,|111\rangle.$$

Bien sûr (comme dans le cas classique), on ne détecte pas correctement des renversements de deux ou trois bits.

Il est utile de décrire le procédé de correction d'erreur d'un autre point de vue. Considérons un appareil de mesure qui mesure les observables

$$Z_1\otimes Z_2\otimes\mathbb{1} \qquad \text{et} \qquad \mathbb{1}\otimes Z_2\otimes Z_3.$$

Notons que la mesure simultanée de ces observables (avec un seul appareil de mesure) est possible car elles commutent. La mesure de chaque observable donne les résultats (±1) et (±1). Ces mesures constituent l'analogue du *"syndrome classique"* :

$$\begin{array}{lllll} +1 & \text{et} & +1 & \longrightarrow & \text{pas d'erreur} \\ -1 & \text{et} & +1 & \longrightarrow & \text{erreur dans le } 1^{\text{er}} \text{ bit} \\ -1 & \text{et} & -1 & \longrightarrow & \text{erreur dans le } 2^{\text{éme}} \text{ bit} \\ +1 & \text{et} & -1 & \longrightarrow & \text{erreur dans le } 3^{\text{éme}} \text{ bit.} \end{array}$$

Une fois l'erreur détectée on la corrige en appliquant un opérateur unitaire :

$$\begin{array}{rcl} \text{pas d'erreur} & \longrightarrow & \mathbb{1}\otimes\mathbb{1}\otimes\mathbb{1} \\ 1^{\text{er}} \text{ bit} & \longrightarrow & X_1\otimes\mathbb{1}\otimes\mathbb{1} \\ 2^{\text{éme}} \text{ bit} & \longrightarrow & \mathbb{1}\otimes X_2\otimes\mathbb{1} \\ 3^{\text{éme}} \text{ bit} & \longrightarrow & \mathbb{1}\otimes\mathbb{1}\otimes X_3. \end{array}$$

Toutes les opérations – encodage, transmission, mesure, correction – sont permises par le MQ (opérateurs unitaires ou mesures projectives).

Ce code n'est par contre pas efficace pour corriger les erreurs (disons uniques) produites par un canal *"phase-flip"*. En effet, un phase-flip (sur n'importe quel bit) produit l'état

$$\alpha\,|000\rangle - \beta\,|111\rangle$$

à la sortie du canal. On voit facilement que le syndrome est $(+1, +1)$, interprété comme une absence d'erreur. Un code de répétition utile pour le canal phase-flip (mais

inutile pour le bit-flip) est obtenu en travaillant dans la base $|+\rangle = \frac{|0\rangle+|1\rangle}{\sqrt{2}}; |-\rangle = \frac{|0\rangle-|1\rangle}{\sqrt{2}}$.

Encodage :

$$|0\rangle \to |+\rangle \to |+++\rangle$$
$$|1\rangle \to |-\rangle \to |---\rangle.$$

Détection d'erreur. Mesurer les observables (qui commutent)

$$X_1 \otimes X_2 \otimes \mathbb{1} \quad \text{et} \quad \mathbb{1} \otimes X_2 \otimes X_3$$

La mesure donne les syndromes:

$+1$	et	$+1$	$\longrightarrow$	pas d'erreur
-1	et	$+1$	$\longrightarrow$	erreur dans le 1^{er} bit
-1	et	-1	$\longrightarrow$	erreur dans le $2^{\text{éme}}$ bit
$+1$	et	-1	$\longrightarrow$	erreur dans le $3^{\text{éme}}$ bit.

Correction d'erreur. Appliquer les opérateurs unitaires :

pas d'erreur	$\longrightarrow$	$\mathbb{1} \otimes \mathbb{1} \otimes \mathbb{1}$
1^{er} bit	$\longrightarrow$	$Z_1 \otimes \mathbb{1} \otimes \mathbb{1}$
$2^{\text{éme}}$ bit	$\longrightarrow$	$\mathbb{1} \otimes Z_2 \otimes \mathbb{1}$
$3^{\text{éme}}$ bit	$\longrightarrow$	$\mathbb{1} \otimes \mathbb{1} \otimes Z_3$.

Par exemple, si on transmet $|0\rangle$, on code par $|+++\rangle$ et si le canal produit $|++-\rangle$, on mesurera $(+1,-1)$ ce qui indique que le $3^{\text{ième}}$ bit a subi un phase-flip. Donc, on applique

$$\mathbb{1} \otimes \mathbb{1} \otimes Z_3 |++-\rangle = |+++\rangle \to |0\rangle.$$

14.3 Le code de Shor

La situation du paragraphe précédent n'est pas encore satisfaisante car en MQ les deux types d'erreurs peuvent essentiellement survenir en même temps et nous aimerions un procédé qui les corrige toutes. La solution de ce problème a été présentée pour la première fois par Shor et elle se révèle être à la base des autres constructions plus élaborées qui ont suivies.

Si nous pouvons corriger à la fois les bits-flips et les phases-flips, alors nous pouvons corriger du même coup une très large classe d'erreurs à 1 qubit. En effet, nous pourrons corriger toutes les erreurs produites par des canaux du type

$$\mathcal{E}(\varrho) = p_0\varrho + p_1 X\varrho X + p_2 Y\varrho Y + p_3 Z\varrho Z$$

avec $p_0 + p_1 + p_2 + p_3 = 1$. Cette classe contient les canaux bit-flip, phase-flip, bit-phase-flip, dépolarisant, etc. Notons que cela n'inclut pas des erreurs sur l'amplitude des coefficients α et β.

L'idée principale du code de Shor est de *"concaténer"* les deux codes de répétition. La concaténation est en fait une méthode de la théorie classique du codage. Elle permet de construire des nouveaux codes en assemblant des codes élémentaires. Les mots de code sont construits comme suit :

$$
\begin{aligned}
|0\rangle \to |+\rangle \longrightarrow |+++\rangle &= \frac{|0\rangle + |1\rangle}{\sqrt{2}} \otimes \frac{|0\rangle + |1\rangle}{\sqrt{2}} \otimes \frac{|0\rangle + |1\rangle}{\sqrt{2}} \\
&\longrightarrow \frac{|000\rangle + |111\rangle}{\sqrt{2}} \otimes \frac{|000\rangle + |111\rangle}{\sqrt{2}} \otimes \frac{|000\rangle + |111\rangle}{\sqrt{2}} \equiv |0\rangle_S \\
|1\rangle \to |-\rangle \longrightarrow |---\rangle &= \frac{|0\rangle - |1\rangle}{\sqrt{2}} \otimes \frac{|0\rangle - |1\rangle}{\sqrt{2}} \otimes \frac{|0\rangle - |1\rangle}{\sqrt{2}} \\
&\longrightarrow \frac{|000\rangle - |111\rangle}{\sqrt{2}} \otimes \frac{|000\rangle - |111\rangle}{\sqrt{2}} \otimes \frac{|000\rangle - |111\rangle}{\sqrt{2}} \equiv |1\rangle_S
\end{aligned}
$$

Ainsi chaque qubit est encodé par un état à 9 qubits. Le *"rendement"* de ce code est 1/9. Si nous appelons $|0\rangle_S$ et $|1\rangle_S$ les deux états à 9 qubits, un mot de code général est

$$|\psi\rangle = \alpha\,|0\rangle + \beta\,|1\rangle \to |\psi_S\rangle = \alpha\,|0\rangle_S + \beta\,|1\rangle_S.$$

Le circuit de la figure Figure **14.1** réalise l'opération de codage de façon unitaire.

Montrons maintenant que ce code est capable de protéger l'état contre n'importe quelle erreur sur 1 qubit engendrée par les opérateurs $\mathbb{1}$ (pas d'erreur), X (bit-flip), Z (phase-flip) et Y (bit & phase flip). En d'autres termes, nous voulons montrer que si la sortie du canal est $|\psi_S\rangle, X_1|\psi_S\rangle, Y_1|\psi_S\rangle, Z_1|\psi_S\rangle$ agit sur 1 des 9 qubits, il est possible de détecter l'erreur et de reconstruire $|\psi\rangle_S$ (et donc $|\psi\rangle$) grâce à des opérations permises en MQ.

Erreur de type bit-flip. Considérons d'abord l'action de X tout seul sur l'un des trois premiers qubits. Comme avant, la mesure simultanée des observables Z_1Z_2 et Z_2Z_3 permet de décider si un des trois premiers qubits est renversé ou non. Ensuite, il est facile de le corriger en appliquant X_i (si c'est $i = 1, 2, 3$ qui est renversé). Pour pouvoir traiter tous les qubits de cette façon, il faut faire une mesure des opérateurs suivants

$$Z_1Z_2 \text{ et } Z_2Z_3, \qquad Z_4Z_5 \text{ et } Z_5Z_6, \qquad Z_7Z_8 \text{ et } Z_8Z_9.$$

Tous ces opérateurs commutent et la mesure simultanée est possible.

Erreur de type phase-flip. Considérons maintenant l'action de Z sur un seul des

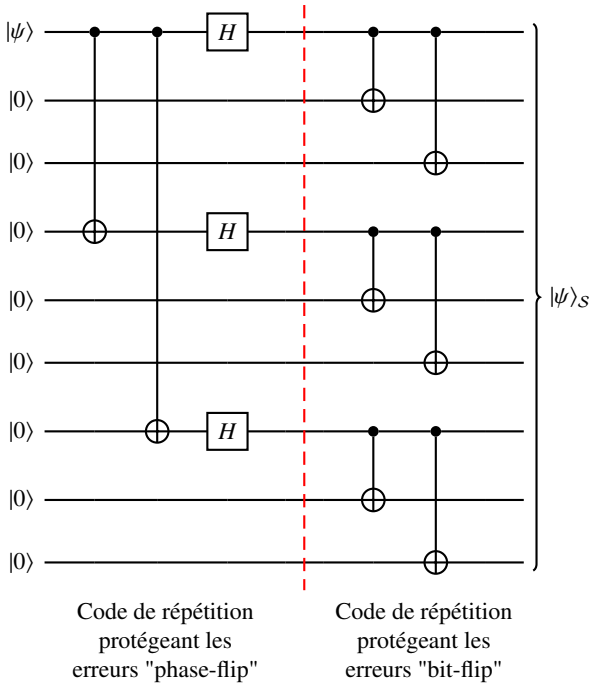


Figure 14.1 Le circuit du code de Shor.

trois premiers qubits. Un phase-flip sur le deuxième qubit change les états de base en $Z_2|\psi\rangle_S$ (l'état à la sortie du canal):

$$Z_2|0\rangle_S = \frac{|000\rangle - |111\rangle}{\sqrt{2}} \otimes \frac{|000\rangle + |111\rangle}{\sqrt{2}} \otimes \frac{|000\rangle + |111\rangle}{\sqrt{2}}$$

et

$$Z_2|1\rangle_S = \frac{|000\rangle + |111\rangle}{\sqrt{2}} \otimes \frac{|000\rangle - |111\rangle}{\sqrt{2}} \otimes \frac{|000\rangle - |111\rangle}{\sqrt{2}}$$

Cette erreur peut être détectée (sans détruire ces états) grâce à une mesure de (opérateurs qui commutent entre eux)

$$(X_1X_2X_3)(X_4X_5X_6) \qquad \text{et} \qquad (X_4X_5X_6)(X_7X_8X_9)$$

Par exemple, pour l'erreur ci-dessus

$$(X_1X_2X_3)(X_4X_5X_6)(Z_2|\psi_S\rangle) = -(Z_2|\psi_S\rangle)$$

et

$$(X_4X_5X_6)(X_7X_8X_9)(Z_2|\psi_S\rangle) = +(Z_2|\psi_S\rangle)$$

On conclut que l'erreur phase-flip est dans un des trois premiers qubits. Notez que

l'on ne peut pas affirmer lequel des trois est erroné. Mais cela ne nous empêche pas d'effectuer la correction en appliquant l'opérateur $Z_1Z_2Z_3$ (vérifiez!)

$$Z_1Z_2Z_3(Z_2|\psi_S\rangle) = |\psi_S\rangle.$$

Erreur de type bit-phase-flip. Illustrons maintenant la correction d'une erreur bit & phase-flip sur le 4$^{\text{ième}}$ bit. L'état sortant du canal est

$$X_4Z_4|\psi_S\rangle.$$

D'abord, on détecte les bit-flip comme suit :

$$\begin{aligned}
Z_1Z_2(X_4Z_4|\psi_S\rangle) &= X_4Z_4|\psi_S\rangle \\
Z_2Z_3(X_4Z_4|\psi_S\rangle) &= X_4Z_4|\psi_S\rangle \\
Z_4Z_5(X_4Z_4|\psi_S\rangle) &= -(X_4Z_4|\psi_S\rangle) \qquad (\text{car } Z_4X_4 = -X_4Z_4) \\
Z_5Z_6(X_4Z_4|\psi_S\rangle) &= X_4Z_4|\psi_S\rangle \\
Z_7Z_8(X_4Z_4|\psi_S\rangle) &= X_4Z_4|\psi_S\rangle \\
Z_8Z_9(X_4Z_4|\psi_S\rangle) &= X_4Z_4|\psi_S\rangle
\end{aligned}$$

On conclut qu'il y a un bit-flip sur le 4$^{\text{ième}}$ bit. Ensuite, on détecte les phases-flips comme suit

$$\begin{aligned}
X_1X_2X_3\ X_4X_5X_6(X_4Z_4|\psi_S\rangle) &= -X_4Z_4|\psi_S\rangle \qquad (\text{car } Z_4X_4 = -X_4Z_4) \\
X_4X_5X_6\ X_7X_8X_9(X_4Z_4|\psi_S\rangle) &= -X_4Z_4|\psi_S\rangle
\end{aligned}$$

et on conclut qu'il y a aussi un phase-flip sur le quatrième qubit. L'état est corrigé en appliquant (X_4Z_4) sur la sortie du canal : $X_4Z_4\ (X_4Z_4)|\psi_S\rangle = |\psi_S\rangle$.

Correction d'erreurs générales à 1 qubit. Les résultats obtenus jusqu'ici peuvent être obtenus aussi d'une façon plus unifiée qui montre leur généralité. Considérons la matrice densité décrivant la sortie d'un canal quelconque :

$$|\psi_S\rangle\langle\psi_S| \to \mathcal{E}(|\psi_S\rangle\langle\psi_S|)$$

Une mesure des observables Z_1Z_2, Z_2Z_3, Z_4Z_5, Z_5Z_6, Z_7Z_8, Z_8Z_9 et $X_1X_2X_3\ X_4X_5X_6$, $X_4X_5X_6\ X_7X_8X_9$ va projeter l'état $\mathcal{E}(|\psi_S\rangle\langle\psi_S|)$ sur un des sous-espaces propres de ces observables. Les valeurs propres sont toutes égales à ± 1. Si le bruit est faible et agit essentiellement sur un qubit au plus, le syndrome sera avec une grande probabilité du type discuté dans les paragraphes précédents, c'est-à-dire correspondant à un bit-flip, phase-flip ou bit & phase flip. On peut l'identifier puis corriger l'erreur. Il existe une faible probabilité (si le bruit est faible) que l'erreur affecte plus qu'un qubit (comme dans le cas classique) et alors la correction d'erreur est erronée.

Pour résumer, le code de Shor est de longueur 9, sa dimension est 2 et il corrige 1 erreur. Plus précisément, les mots de code sont de type $\alpha\ |0\rangle_S + \beta\ |1\rangle_S$ donc vivent dans un sous-espace de Hilbert de dimension 2^1, de l'espace à 9 qubits qui est lui de

dimension 2^9. On dit que les paramètres de ce code quantique sont [9, 1, 1] (longueur 9 et $\dim(\mathcal{H}) = 2^9$; $\dim(\mathbb{C}) = 2^1$; corrige une erreur).

14.4 Formalisme des stabilisateurs

Ce paragraphe peut être sauté en première lecture.

La construction du code de Shor permet de dégager une structure algébrique qui est à la base de la construction des codes correcteurs d'erreurs plus généraux. Nous avons vu dans le paragraphe précédent que:

1. Les états (mots) $|\Psi\rangle$ du code de Shor sont les états propres associés à la valeur propre égale à 1 des observables Z_1Z_2, Z_2Z_3, Z_4Z_5, Z_5Z_6, Z_7Z_8, Z_8Z_9 et $X_1X_2X_3$ $X_4X_5X_6$, $X_4X_5X_6$ $X_7X_8X_9$. Ces opérateurs sont appelés les *stabilisateurs* du code.
2. Ces opérateurs commutent entre eux, ce qui permet d'en faire une mesure simultanée.
3. $|\Psi\rangle \mapsto |\Psi'\rangle$ est corrompu par un bit, phase ou bit-phase flip, le nouvel état est encore un vecteur propre des opérateurs stabilisateurs[2].
4. Le syndrome est donné par la liste des valeurs propres des opérateurs stabilisateurs dans l'état $|\Psi'\rangle$.

Remarque : Les stabilisateurs sont les analogues des lignes de la matrice de parité. Chaque équation aux valeurs propres satisfaite par un état du code est l'analogue d'une contrainte de parité. Les valeurs propres des stabilisateurs sont l'analogue du syndrome. Notons qu'il y a en quelque sorte deux types de contraintes de parité: une contrainte de type "z" et une contrainte de type "x". Il existe aussi deux "parties" dans le syndrome: l'une ayant trait aux bit-flips et l'autre aux phase-flips.

Pour généraliser cette construction, il convient d'introduire le groupe de Pauli $\mathcal{P}^n$. Ce groupe est formé par l'ensemble de tous les produits (tensoriels) des opérateurs de Pauli X_j, Y_j, Z_j avec ± 1 et $\pm i$, $\forall j = 1, \ldots, n$. Il est facile de voir que deux éléments de ce groupe commutent ou anti-commutent. Un sous-groupe $\mathcal{S} \subset \mathcal{P}^n$ dont tous les éléments commutent est appelé *groupe stabilisateur*. On peut toujours trouver un ensemble de générateurs indépendants pour $\mathcal{S}$: $\{O_1, \ldots, O_m\}$. Le sous-espace de Hilbert $\mathcal{H} \subset (\mathbb{C}^2)^{\otimes n}$ des états $|\Psi\rangle$ tels que

$$O_j|\Psi\rangle = 1 \cdot |\Psi\rangle, \qquad j = 1, \ldots, m$$

est un code de longueur n et de dimension 2^{n-m}. Notez que $|\Psi\rangle$ est stable sous l'action de $\mathcal{S}$. Les observables $\{O_1, \ldots, O_m\}$ sont simultanément mesurables puisqu'elles commutent entre elles. Toute erreur (multiple) de type bit, phase, bit-phase ou combinaison linéaire de telles erreurs suivie par une mesure donne un état corrompu du type

$$E_k|\Psi\rangle = |\Psi'\rangle$$

[2] S'il est corrompu (toujours sur un seul bit) par une combinaison linéaire de ces erreurs, alors la mesure des observables projette sur un état propre et nous ramène à un de ces types d'erreurs discrètes.

où E_k est un élément de $\mathcal{P}^n$. Cet état est encore vecteur propre des stabilisateurs et donc une mesure ne le détruira pas. De plus, le résultat de la mesure donne les valeurs propres $\lambda_1 = \pm 1, \ldots, \lambda_m = \pm 1$. Cette liste de v.p. est le "syndrome". Si l'erreur E_k anti-commute avec un générateur O_j, alors on a

$$O_j|\Psi'\rangle = O_j E_k|\Psi\rangle = -E_k O_j|\Psi\rangle = -E_k|\Psi\rangle = -|\Psi'\rangle$$

c'est-à-dire $\lambda_j = -1$. Pour un code donné, on peut établir une table de "correction" avec les syndromes associés à chaque type d'erreur. Une fois l'erreur détectée et distinguée, la correction se fait par l'opération (unitaire)

$$|\Psi'\rangle \mapsto E_k|\Psi'\rangle = E_k^2|\Psi\rangle = |\Psi\rangle$$

Notons que dans les produits d'opérateurs de Pauli, il est toujours possible d'éliminer Y_j au profit de produits $X_j Z_j$. Ensuite, on peut toujours écrire un opérateur de Pauli comme un produit de Z multiplié par un produit de X. Par exemple:

$$X_1 \mathbb{1}_2 Y_3 Z_4 Y_5 \mathbb{1}_6 = -(X_1 \mathbb{1}_2 X_3 \mathbb{1}_4 X_5 \mathbb{1}_6)(\mathbb{1}_1 \mathbb{1}_2 Z_3 Z_4 Z_5 \mathbb{1}_6)$$

Discutons pour finir une représentation "semi-classique" du formalisme des stabilisaterus. On peut "coder" les produits de la droite de l'égalité par la "ligne"

$$\begin{pmatrix} 1 & 0 & 1 & 0 & 1 & 0 & | & 0 & 0 & 1 & 1 & 1 & 0 \end{pmatrix}$$

En appliquant ce procédé à tous les stabilisateurs O_j du code, on obtient une matrice de la forme

$$(A_1 \mid A_2)$$

appelée parfois "matrice de parité quantique". Chaque ligne de la matrice permet de reconstruire une contrainte $O_j|\Psi\rangle = |\Psi\rangle$ satisfaite par les états du code. On peut montrer que la condition de commutation des $\{O_1, \ldots, O_m\}$ est équivalente à

$$A_1 A_2^T + A_2 A_1^T = 0$$

Cette équation décrit le fait que chaque ligne doit être orthogonale par rapport à un produit scalaire symplectique. Un opérateur d'erreur E peut s'écrire de façon similaire comme un "vecteur d'erreur" $(\vec{e}_1 \mid \vec{e}_2)$. Le syndrome devient

$$A_1 \vec{e}_1 + A_2 \vec{e}_2 = \vec{s}$$

La correction d'erreur quantique est alors ramenée à la correction d'erreur classique (c'est-à-dire à la résolution d'in système d'équations linéaires). Ce dernier point est bi-envenu dans la mesure où la théorie classique du codage est très bien développée.

Dans le cas du code de Shor, on a une situation un peu particulière:

$$(A_1 \mid A_2) = \left[\begin{array}{cccccccccc|cccccccc}
1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1
\end{array}\right]$$

Les codes dont la matrice de parité quantique possède une structure par bloc s'appellent des codes de Calderbank-Shor-Steane. En général, un tel code s'écrit

$$(A_1 \mid A_2) = \left[\begin{array}{c|c} H_1 & 0 \\ 0 & G_2^T \end{array}\right]$$

avec la contrainte d'orthogonalité

$$H_1 G_2 = 0$$

On peut interpréter H_1 comme la matrice de parité d'un code classique C_1 et G_2 la matrice génératrice d'un code classique C_2. Ici, G_2^T est la matrice de parité[3] de $C_2^\perp$. La contrainte d'orthogonalité signifie que $C_2^\perp \subset C_1$. Pour le décodage, on doit résoudre des équations du type $H_1 \vec{e}_1 = \vec{s}_1$ et $G_2^T \vec{e}_2 = \vec{s}_2$ (ou bien $H_2^T = 0$).

Nous étudions cette construction plus en détail dans la Section **14.5**. Il deviendra clair que le code C_1 protège contre les erreurs de type bit-flip et $C_2^\perp$ protège contre les erreurs de type phase-flip. L'utilisation simultanée des deux codes permet de protéger contre toutes combinaisons de ces erreurs.

14.5 Codes de Calderbank-Shor-Steane

Ce paragraphe peut être lu indépendamment du précédent.

Les codes de Calderbank-Shor-Steane sont une classe particulière de codes de type stabilisateurs, dont les propriétés principales sont aisées à analyser. Cette classe de code utilise deux codes classiques: l'un pour corriger les bit-flips et l'autre pour corriger les phase-flips.

Soit C_1 et C_2 deux codes linéaires classiques tels que $C_2 \subset C_1$ de paramètres (n, k_1)

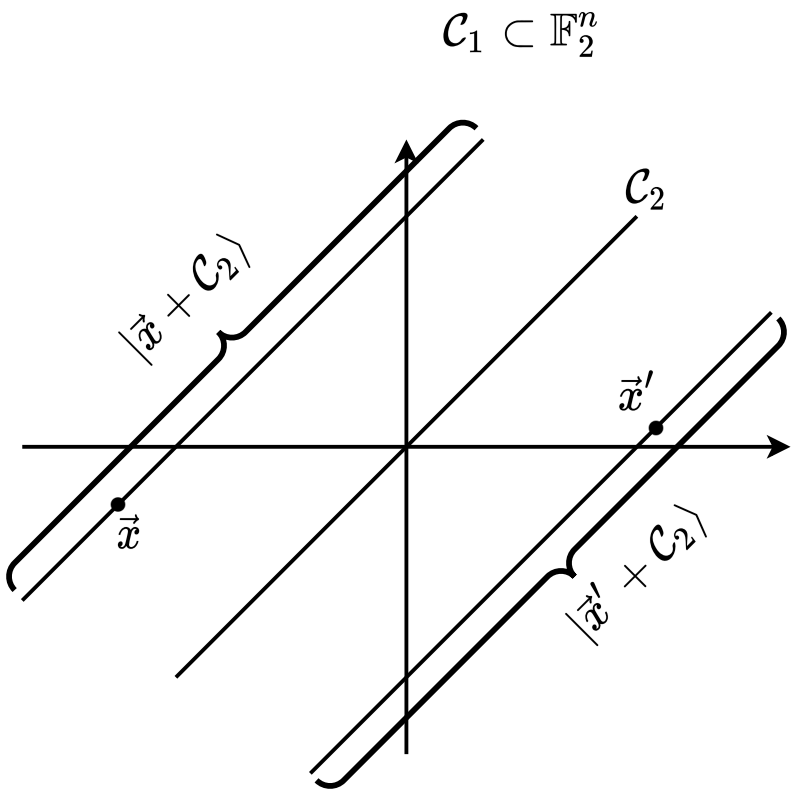
[3] Cette matrice est appelée $H_2^\perp$ dans la Section **14.5**.

et (n, k_2). Nous rappelons que C_1 est un sous-espace vectoriel de $\mathbf{F}_2^n$ de dimension k_1 et C_2 est un sous-espace vectoriel de C_1 de dimension $k_2 \leq k_1$. On suppose aussi que C_1 et $C_2^\perp$ corrigent t erreurs (par exemple, leur distance minimale est $d = 2t + 1$). Le code *dual* $C_2^\perp$ est le sous-espace vectoriel orthogonal à C_2 dans $\mathbf{F}_2^n$ et sa dimension est donc $n - k_2$.

En fait C_2 est aussi un sous-groupe de C_1 et on peut considérer l'ensemble des classes d'équivalence C_1/C_2. Soit $\vec{x}$ un mot de C_1, la classe d'équivalence de $\vec{x}$ est l'ensemble des mots de la forme $\{\vec{x} + \vec{y}; \vec{y} \in C_2\}$. C'est donc l'hyperplan parallèle à C_2 passant par $\vec{x} \in C_1$. Posons

$$|\vec{x} + C_2\rangle = \frac{1}{\sqrt{|C_2|}} \sum_{\vec{y} \in C_2} |\vec{x} + \vec{y}\rangle$$

Ce ket ne dépend que de la classe d'équivalence associée à $\vec{x}$ et non pas du représentant spécial choisi.



Le ket $|\vec{x}+C_2\rangle$ est la superposition cohérente de tous les états dans la classe d'équivalence $\vec{x}+C_2$. Pour deux classes d'équivalence différentes $\vec{x}+C_2$ et $\vec{x}'+C_2$ (c'est-à-dire que $\vec{x}-\vec{x}' \notin C_2$) les kets associés sont perpendiculaires :

$$\langle \vec{x}+C_2|\vec{x}'+C_2\rangle = 0$$

Il y a $\frac{|C_1|}{|C_2|} = 2^{k_1-k_2}$ classes d'équivalence et donc $2^{k_1-k_2}$ vecteurs orthogonaux.

Le code $\mathrm{CSS}(C_1,C_2)$ est le sous-espace de Hilbert engendré par ces $2^{k_1-k_2}$ vecteurs orthogonaux. C'est un sous-espace de l'espace de Hilbert à n qubits $\underbrace{\mathbb{C}^2\otimes\mathbb{C}^2\otimes\cdots\otimes\mathbb{C}^2}_{n \text{ fois}}$ lequel est de dimension 2^n. Les paramètres du code sont $[n, k_1-k_2]$; longueur n, $\dim(\mathcal{H}) = n$ et $\dim(\mathrm{CSS}(C_1,C_2)) = 2^{k_1-k_2}$. Nous allons montrer que si C_1 et $C_2^\perp$ corrigent t erreurs alors $\mathrm{CSS}(C_1,C_2)$ corrige t bit et phase flips.

Il suffit de montrer ces propriétés pour les vecteurs de base du code. Les mêmes arguments entraînent qu'elles sont sont vraies pour tout état de superposition du code. Soit $\vec{e}_1$ et $\vec{e}_2$ les vecteurs possédants des 1 et 0 avec la coordonnée 1 aux t endroits des bit et/ou phase flips. Le vecteur corrompu par le bruit est :

$$|\psi\rangle = \frac{1}{\sqrt{|C_2|}}\sum_{\vec{y}\in C_2}(-1)^{(\vec{x}+\vec{y})\cdot\vec{e}_2}|\vec{x}+\vec{y}+\vec{e}_1\rangle$$

Pour effectuer la correction d'erreur, il faut faire des opérations permises par la MQ et qui reconstituent $|\vec{x}+C_2\rangle$. Comme il faudra calculer et stocker les *"syndromes"* du bit et phase flip, on ajoute des bits auxiliaires de stockage et on travaille avec l'état

$$|\psi\rangle\otimes|0\rangle^{\otimes n}\otimes|0\rangle^{\otimes n} = \frac{1}{\sqrt{|C_2|}}\sum_{\vec{y}\in C_2}(-1)^{(\vec{x}+\vec{y})\cdot\vec{e}_2}|\vec{x}+\vec{y}+\vec{e}_1\rangle\otimes|0\rangle^{\otimes n}\otimes|0\rangle^{\otimes n}$$

Correction du bit-flip. On calcule les syndromes $H_1(\vec{x}+\vec{y}+\vec{e}_1) = H_1\vec{e}_1$ (grâce à la matrice de parité du code C_1). Ce calcul peut être implémenté de façon unitaire :

$$U_1|\vec{x}+\vec{y}+\vec{e}_1\rangle\otimes|0\rangle^{\otimes n}\otimes|0\rangle^{\otimes n} = |\vec{x}+\vec{y}+\vec{e}_1\rangle\otimes|H_1\vec{e}_1\rangle\otimes|0\rangle^{\otimes n}$$

si bien que

$$U_1|\psi\rangle\otimes|0\rangle^{\otimes n}\otimes|0\rangle^{\otimes n} = \frac{1}{\sqrt{|C_2|}}\sum_{\vec{y}\in C_2}(-1)^{(\vec{x}+\vec{y})\cdot\vec{e}_2}|\vec{x}+\vec{y}+\vec{e}_1\rangle\otimes|H_1\vec{e}_1\rangle\otimes|0\rangle^{\otimes n}$$

Une mesure du second registre dans la base computationnelle donne $|H_1\vec{e}_1\rangle$ avec probabilité 1 (car le ket $|H_1\vec{e}_1\rangle$ peut être factorisé dans la dernière expression). Donc, on connait le syndrome $H_1\vec{e}_1$ et si $|\vec{e}_1| \leq t$, on peut corriger les erreurs grâce à un algorithme classique (pour le code C_1). Cela permet d'appliquer l'opérateur unitaire

$$\prod_{\substack{\text{composantes de}\\ \vec{e}_1 \text{ égales à 1}}} X_i \equiv U_{1,\text{corr}}$$

pour corriger l'état quantique. On obtient

$$U_{1,\text{corr}} U_1 |\psi\rangle \otimes |0\rangle^{\otimes n} \otimes |0\rangle^{\otimes n} = \frac{1}{\sqrt{|C_2|}} \sum_{\vec{y}\in C_2} (-1)^{(\vec{x}+\vec{y})\cdot\vec{e}_2} |\vec{x}+\vec{y}\rangle \otimes |H_1\vec{e}_1\rangle \otimes |0\rangle^{\otimes n}$$

Correction du phase-flip. D'abord, pour passer dans une base plus naturelle, on applique les portes de Hadamard $H^{\otimes n}$, ce qui donne:

$$\begin{aligned}
&H^{\otimes n} U_{1,\text{corr}} U_1 |\psi\rangle \otimes |0\rangle^{\otimes n} \otimes |0\rangle^{\otimes n} \\
=&\frac{1}{2^{n/2}} \frac{1}{\sqrt{|C_2|}} \sum_{\vec{y}\in C_2} (-1)^{(\vec{x}+\vec{y})\cdot\vec{e}_2} \sum_{\vec{z}} (-1)^{\vec{z}\cdot(\vec{x}+\vec{y})} |\vec{z}\rangle \otimes |H_1\vec{e}_1\rangle \otimes |0\rangle^{\otimes n} \\
=&\frac{1}{2^{n/2}} \frac{1}{\sqrt{|C_2|}} \sum_{\vec{z}} \sum_{\vec{y}\in C_2} (-1)^{(\vec{z}+\vec{e}_2)\cdot(\vec{x}+\vec{y})} |\vec{z}\rangle \otimes |H_1\vec{e}_1\rangle \otimes |0\rangle^{\otimes n} \\
=&\frac{1}{2^{n/2}} \frac{1}{\sqrt{|C_2|}} \sum_{\vec{z}} \sum_{\vec{y}\in C_2} (-1)^{\vec{z}\cdot(\vec{x}+\vec{y})} |\vec{z}+\vec{e}_2\rangle \otimes |H_1\vec{e}_1\rangle \otimes |0\rangle^{\otimes n} = \dots
\end{aligned}$$

On note que

$$\sum_{\vec{y}\in C_2} (-1)^{\vec{z}\cdot\vec{y}} = \begin{cases} |C_2| & \text{si } \vec{z} \in C_2^\perp \\ 0 & \text{sinon} \end{cases}$$

ce qui donne l'état :

$$\dots = \frac{1}{|C_2^\perp|} \sum_{\vec{z}\in C_2^\perp} (-1)^{\vec{z}\cdot\vec{x}} |\vec{z}+\vec{e}_2\rangle \otimes |H_1\vec{e}_1\rangle \otimes |0\rangle^{\otimes n}.$$

On implémente maintenant le calcul du syndrome de $C_2^\perp$ de façon unitaire en stockant le résultat dans le dernier registre

$$\begin{aligned}
U_2 |\vec{z}+\vec{e}_2\rangle \otimes |H_1\vec{e}_1\rangle \otimes |0\rangle^{\otimes n} \\
&= |\vec{z}+\vec{e}_2\rangle \otimes |H_1\vec{e}_1\rangle \otimes |H_2^\perp(\vec{z}+\vec{e}_2)\rangle \\
&= |\vec{z}+\vec{e}_2\rangle \otimes |H_1\vec{e}_1\rangle \otimes |H_2^\perp\vec{e}_2\rangle.
\end{aligned}$$

où ici $H_2^\perp$ est la matrice de parité de $C_2^\perp$. L'état obtenu est :

$$\begin{aligned}
&U_2 H^{\otimes n} U_{1,\text{corr}} U_1 |\psi\rangle \otimes |0\rangle^{\otimes n} \otimes |0\rangle^{\otimes n} \\
=&\frac{1}{\sqrt{|C_2^\perp|}} \sum_{\vec{z}\in C_2^\perp} (-1)^{\vec{z}\cdot\vec{x}} |\vec{z}+\vec{e}_2\rangle \otimes |H_1\vec{e}_1\rangle \otimes |H_2^\perp\vec{e}_2\rangle.
\end{aligned}$$

La mesure du dernier registre (dans la base computationnelle) projette sur $|H_2^\perp\vec{e}_2\rangle$ lui-même, avec probabilité 1. Donc, $H_2^\perp\vec{e}_2$ est connue et $\vec{e}_2$ peut être calculé si $C_2^\perp$ corrige t

erreurs (on suppose $|\vec{e}_2| \leq t$). Cela permet alors d'appliquer l'opérateur unitaire

$$\prod_{\substack{\text{composantes de} \\ \vec{e}_2 \text{ égales à 1}}} X_i \equiv U_{2,\text{corr}}$$

pour trouver l'état :

$$\begin{aligned} & U_{2,\text{corr}} U_2 H^{\otimes n} U_{1,\text{corr}} U_1 |\psi\rangle \otimes |0\rangle^{\otimes n} \otimes |0\rangle^{\otimes n} \\ = & \frac{1}{\sqrt{|C_2^\perp|}} \sum_{\vec{z} \in C_2^\perp} (-1)^{\vec{z}\cdot\vec{x}} |\vec{z}\rangle \otimes |H_1 \vec{e}_1\rangle \otimes |H_2^\perp \vec{e}_2\rangle. \end{aligned}$$

La dernière étape consiste maintenant à appliquer une fois de plus $H^{\otimes n}$ pour revenir dans la base initiale. Cela donne

$$\begin{aligned} & H^{\otimes n} U_{2,\text{corr}} U_2 H^{\otimes n} U_{1,\text{corr}} U_1 |\psi\rangle \otimes |0\rangle^{\otimes n} \otimes |0\rangle^{\otimes n} \\ = & \frac{1}{\sqrt{|C_2^\perp|}} \frac{1}{2^{n/2}} \sum_{\vec{z} \in C_2^\perp} \sum_{\vec{y}} (-1)^{\vec{z}\cdot(\vec{x}+\vec{y})} |\vec{y}\rangle \otimes |H_1 \vec{e}_1\rangle \otimes |H_2^\perp \vec{e}_2\rangle \\ = & \frac{1}{\sqrt{|C_2^\perp|}} \frac{1}{2^{n/2}} \sum_{\vec{z} \in C_2^\perp} \sum_{\vec{y}} (-1)^{\vec{z}\cdot\vec{y}} |\vec{y} + \vec{x}\rangle \otimes |H_1 \vec{e}_1\rangle \otimes |H_2^\perp \vec{e}_2\rangle. \end{aligned}$$

Maintenant on note (comme avant) :

$$\sum_{\vec{z} \in C_2^\perp} (-1)^{\vec{z}\cdot\vec{y}} = \begin{cases} |C_2^\perp| & \text{si } \vec{y} \in (C_2^\perp)^\perp = C_2 \\ 0 & \text{sinon} \end{cases}$$

L'état final obtenu est donc :

$$\left(\frac{1}{\sqrt{|C_2|}} \sum_{\vec{y} \in C_2} |\vec{x} + \vec{y}\rangle \right) \otimes |H_1 \vec{e}_1\rangle \otimes |H_2^\perp \vec{e}_2\rangle.$$

Nous voyons que le mot de code initial a été reconstruit dans le premier registre!

Part III

Réalisations Expérimentales

15 La Dynamique du Spin

Dans ce chapitre, nous allons étudier la dynamique du spin 1/2 (moment magnétique de certains noyaux atomiques notamment). Celui-ci constitue l'une des réalisations naturelles les plus importantes du bit quantique. En effet, celui-ci est aisément manipulable grâce à des champs magnétiques dépendant du temps. La manipulation et le contrôle des moments magnétiques par des champs magnétiques dépendant du temps est à la base de la Résonance Magnétique Nucléaire (utilisée pour l'imagerie IRM, etc). Ce contrôle des spins est aussi à la base de la réalisation des portes quantiques. Nous verrons dans ce chapitre comment réaliser les portes de Hadamard et NOT. Pour ce qui concerne l'importante porte CNOT, il faut d'abord comprendre comment interagissent les paires de moments magnétiques. Ce sera l'objet du Chapitre **16**.

15.1 La sphère de Bloch

Nous commençons par un petit rappel sur la sphère de Bloch déjà introduite à la Section **2.10**. Un qubit appartenant à l'espace de Hilbert $\mathbb{C}^2$ peut toujours être paramétré comme suit

$$|\psi\rangle = \cos\left(\frac{\theta}{2}\right)|\uparrow\rangle + \sin\left(\frac{\theta}{2}\right)e^{i\phi}|\downarrow\rangle \tag{15.1}$$

Ce vecteur complexe à deux composantes, peut être représenté sur une sphère, appelée sphère de Bloch, où θ est l'angle par rapport à la direction z et ϕ est l'angle dans le plan xy par rapport à x (Figure **15.1**).

Il est utile de se remémorer (voir Chapitre **2**) la représentation des bases X, Y et Z c.-à-d. $\{|+\rangle, |-\rangle\}$, $\{|\circlearrowleft\rangle, |\circlearrowright\rangle\}$ et $\{|\uparrow\rangle, |\downarrow\rangle\}$ sur cette sphère. La base Z consiste en deux vecteurs opposés le long de l'axe z, la base X de deux vecteurs opposés le long de l'axe x et la base Y de deux vecteurs opposés le long de y.

Considérons la matrice

$$\vec{\sigma} \cdot \vec{n} = \sigma_x n_x + \sigma_y n_y + \sigma_z n_z \tag{15.2}$$

où $(\sigma_x, \sigma_y, \sigma_z)$ sont les trois matrices de Pauli et (n_x, n_y, n_z) est un vecteur unité tel que

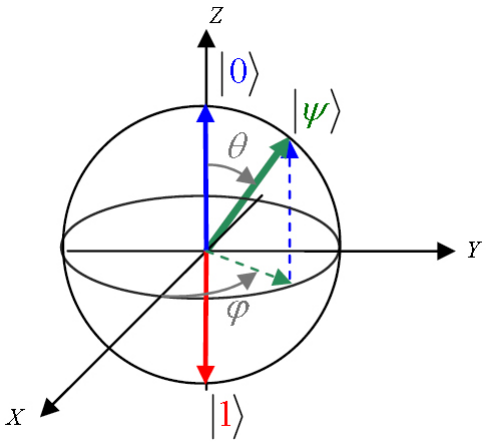


Figure 15.1 La sphère de Bloch.

$n_x^2 + n_y^2 + n_z^2 = 1$. Soit aussi $n_x = \sin(\theta)\cos(\phi)$, $n_y = \sin(\theta)\sin(\phi)$ et $n_z = \cos(\theta)$. C'est-à-dire que $\vec{n}$ est "identique" à la représentation de $|\psi\rangle$ sur la sphère de Bloch. On peut vérifier que

$$\vec{\sigma} \cdot \vec{n}|\psi\rangle = (+1)|\psi\rangle \tag{15.3}$$

C'est-à-dire que $|\psi\rangle$ est un vecteur propre de $\vec{\sigma} \cdot \vec{n}$ avec valeur propre $+1$. On peut donc se faire l'image que "$|\psi\rangle$ sur la sphère de Bloch est la projection du vecteur $\vec{\sigma}$ sur l'axe $\vec{n}$".

La formule pour l'exponentiation de $\vec{\sigma} \cdot \vec{n}$ (analogue à la formule d'Euler) est donné par

$$\exp\left(i\frac{\alpha}{2}\vec{\sigma} \cdot \vec{n}\right) = \cos\left(\frac{\alpha}{2}\right)\mathbb{1} + i(\vec{\sigma} \cdot \vec{n})\sin\left(\frac{\alpha}{2}\right). \tag{15.4}$$

Pour comprendre la signification de cette matrice 2Œ2 considérons un cas particulier. Prenons $\vec{n} = (0, 0, 1)$ (c.-à-d. l'axe z) et appliquons la matrice sur $|\psi\rangle$.

$$\exp\left(i\frac{\alpha}{2}\sigma_z\right) = \exp\left(i\frac{\alpha}{2}\sigma_z\right)\left\{\cos\left(\frac{\theta}{2}\right)|\uparrow\rangle + e^{i\phi}\sin\left(\frac{\theta}{2}\right)|\downarrow\rangle\right\} \tag{15.5}$$

$$= e^{i\frac{\alpha}{2}}\cos\left(\frac{\theta}{2}\right)|\uparrow\rangle + e^{i\phi}e^{-i\frac{\alpha}{2}}\sin\left(\frac{\theta}{2}\right)|\downarrow\rangle \tag{15.6}$$

$$= e^{i\frac{\alpha}{2}}\left\{\cos\left(\frac{\theta}{2}\right)|\uparrow\rangle + e^{i(\phi-\alpha)}\sin\left(\frac{\theta}{2}\right)|\downarrow\rangle\right\} \tag{15.7}$$

Le préfacteur $e^{i\frac{\alpha}{2}}$ n'a pas de signification physique puisque c'est une phase globale. Le

nouveau vecteur sur la sphère de Bloch fait toujours un angle θ avec z et un angle $(\phi-\alpha)$ dans le plan xy avec l'axe x. Ainsi l'opérateur $\exp\left(i\frac{\alpha}{2}\sigma_z\right)$ représente une rotation d'angle $(-\alpha)$ autour de l'axe z. De même, la matrice $\exp\left(-i\frac{\alpha}{2}\sigma_z\right)$ représente une rotation d'angle $(+\alpha)$ autour de l'axe z.

Plus généralement, la matrice $\exp\left(-i\frac{\alpha}{2}\vec{\sigma}\cdot\vec{n}\right)$ représente une rotation d'angle $(+\alpha)$ et d'axe $\vec{n}$ sur la sphère de Bloch.

Nous allons voir que la dynamique du spin dans un champ $\vec{B}$ constant fait intervenir de telles relations (autour de $\vec{B}$ qui joue le rôle de $\vec{n}$ essentiellement).

15.2 L'Hamiltonien du spin dans un champ magnétique

Nous avons vu à la Section **2.7** que l'énergie d'interaction d'un moment magnétique $\vec{M}$ avec un champ magnétique $\vec{B}$ est donnée par (comme en physique classique):

$$E = -\vec{M}\cdot\vec{B} \tag{15.8}$$

En physique quantique, l'observable $\vec{M}$ devient une matrice. Pour les moments magnétiques de spin 1/2 (comme celui de l'électron, du proton, de certains noyaux atomiques, etc) on a

$$\vec{M} = \frac{g\hbar}{2}\vec{\sigma} \tag{15.9}$$

où $\vec{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$ est le vecteur dont les composantes sont constituées par les trois matrices de Pauli. Ainsi l'énergie d'interaction d'un spin 1/2 dans un champ magnétique est donné par une matrice 2Œ2 (hermitienne, car c'est une observable) appelée Hamiltonien

$$H = -\frac{g\hbar}{2}\vec{B}\cdot\vec{\sigma} \tag{15.10}$$

En composantes, cette matrice est explicitement

$$H = -\frac{g\hbar}{2}\begin{pmatrix} B_z & B_x - iB_y \\ B_x + iB_y & -B_z \end{pmatrix} \tag{15.11}$$

Considérons maintenant un champ magnétique constant (qui ne varie pas avec le temps). On peut toujours choisir l'axe z le long du vecteur $\vec{B} = (0, 0, B)$. Par conséquent:

$$H = -\frac{g\hbar}{2}B\sigma_z \equiv -\frac{\hbar\omega_0}{2}\sigma_z = -\frac{\hbar\omega_0}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{15.12}$$

où $\hbar\omega_0 = g\hbar B$ par définition. Ici ω_0 à l'unité d'une fréquence [s^{-1}] et s'appelle la fréquence de Larmor. Les deux valeurs propres de H sont $-\frac{\hbar\omega_0}{2}$ et $+\frac{\hbar\omega_0}{2}$ et les vecteurs

propres correspondants sont $|\uparrow\rangle$ et $|\downarrow\rangle$. On peut représenter ces valeurs propres sur un diagramme donnant les "niveaux d'énergie" du système (Figure **15.2**).

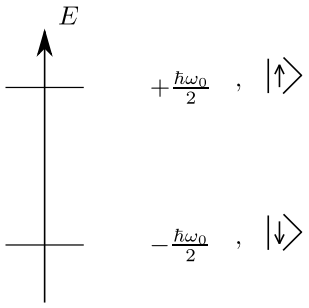


Figure 15.2 Niveaux d'énergie du système

Les systèmes dont l'Hamiltonien possède un spectre d'énergie de ce type s'appellent des "systèmes à deux niveaux". Il existe plusieurs types de systèmes à deux niveaux (exacts ou approximatifs) dans la nature. Ils peuvent tous être mathématiquement modélisés par "l'Hamiltonien d'un spin dans un champ magnétique", et donc le formalisme décrit dans ce chapitre dépasse de loin le cadre de la dynamique des moments magnétiques dans un champ magnétique.

Nous allons considérer aussi des champs magnétiques variables dans le temps du type

$$\vec{B} = (0, 0, B_0) + (B_1 \cos(\omega t), -B_1 \sin(\omega t), 0) \tag{15.13}$$

Il s'agit d'un champ auquel on a ajouté une partie tournante dans le plan xy. L'Hamiltonien du spin dans ce champ est donné par la matrice

$$H = -\frac{g\hbar}{2}\begin{pmatrix} B_0 & B_1\Big(\cos(\omega t) + i\sin(\omega t)\Big) \\ B_1\Big(\cos(\omega t) - i\sin(\omega t)\Big) & B_0 \end{pmatrix} \tag{15.14}$$

En introduisant les matrices

$$\sigma_+ = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \qquad \text{et} \qquad \sigma_- = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \tag{15.15}$$

On peut réécrire cet Hamiltonien sous la forme

$$H = -\frac{\hbar\omega_0}{2}\sigma_z - \frac{\hbar\omega_1}{2}\Big(\sigma_+ e^{i\omega t} + \sigma_- e^{-i\omega t}\Big) \tag{15.16}$$

Notons que $|\uparrow\rangle$ et $|\downarrow\rangle$ ne sont plus des états propres à cause du terme dépendant du temps. En fait, ce terme est responsable de transitions quantiques entre ces états, car

$$\sigma_+|\downarrow\rangle = |\uparrow\rangle \quad \text{et} \quad \sigma_+|\uparrow\rangle = 0 \tag{15.17}$$

$$\sigma_-|\uparrow\rangle = |\downarrow\rangle \quad \text{et} \quad \sigma_-|\downarrow\rangle = 0 \tag{15.18}$$

Nous allons voir que ces transitions entre les niveaux d'énergie de la partie $-\frac{\hbar\omega_0}{2}\sigma_z$ sont exploitées dans la RMN et la réalisation des portes logiques.

15.3 La précession de Larmor

Dans ce paragraphe, nous étudions l'évolution temporelle d'un état $|\psi\rangle = \cos\left(\frac{\theta}{2}\right)|\uparrow\rangle + e^{i\phi}\sin\left(\frac{\theta}{2}\right)|\downarrow\rangle$ dans un champ magnétique constant $\vec{B} = (0, 0, B_0)$ orienté dans la direction z.

D'après les postulats de la mécanique quantique, cette évolution temporelle est donnée par une matrice unitaire qui dépend du temps

$$U(t, 0)|\psi\rangle \equiv |\psi(t)\rangle \tag{15.19}$$

telle que $U(t_3, t_2)U(t_2, t_1) = U(t_3, t_1)$. (Ici $U(t, s)$ signifie l'évolution de l'instant s à l'instant t). Cette matrice est la solution de l'équation de Schrödinger:

$$i\hbar\frac{d}{dt}U(t, 0) = HU(t, 0) \tag{15.20}$$

Il s'agit donc de résoudre cette équation de Schrödinger.

La résolution est aisée si H ne dépend pas du temps t. En effet, il est facile de vérifier qu'alors

$$U(t, 0) = \exp\left(-i\frac{t}{\hbar}H\right) \tag{15.21}$$

est solution. Par contre, si H dépend du temps, la solution est plus compliquée et en particulier cette formule simple n'est plus valable.

Pour le cas du champ $\vec{B} = (0, 0, B_0)$ constant où $H = -\frac{\hbar\omega_0}{2}\sigma_z$ indépendant du temps, on a:

$$U(t, 0) = \exp\left(it\frac{\omega_0}{2}\sigma_z\right) \tag{15.22}$$

Nous reconnaissons ici la matrice de rotation autour de l'axe z et d'angle $(-t\omega_0)$. Ainsi l'état à l'instant t est

$$|\psi(t)\rangle = \exp\left(i\frac{t\omega_0}{2}\sigma_z\right)|\psi(0)\rangle \tag{15.23}$$

$$= \exp\left(i\frac{t\omega_0}{2}\sigma_z\right)\left\{\cos\left(\frac{\theta}{2}\right)|\uparrow\rangle + e^{i\phi}\sin\left(\frac{\theta}{2}\right)|\downarrow\rangle\right\} \tag{15.24}$$

$$= \exp\left(i\frac{t\omega_0}{2}\right)\cos\left(\frac{\theta}{2}\right)|\uparrow\rangle + e^{i\phi}\exp\left(-i\frac{t\omega_0}{2}\right)\sin\left(\frac{\theta}{2}\right)|\downarrow\rangle \tag{15.25}$$

$$= \exp\left(i\frac{t\omega_0}{2}\right)\left\{\cos\left(\frac{\theta}{2}\right)|\uparrow\rangle + \exp\Big(i(\phi - t\omega_0)\Big)\sin\left(\frac{\theta}{2}\right)|\downarrow\rangle\right\} \tag{15.26}$$

L'angle ϕ évolue comme $\phi - t\omega_0$ avec le temps. Sur la sphère de Bloch on a un mouvement appelé précession de Larmor autour de z (l'axe de $\vec{B}$) et la fréquence de la précession de Larmor est ω_0 lui-même (la périodicité temporelle étant $\dfrac{2\pi}{\omega_0} = T_0$).

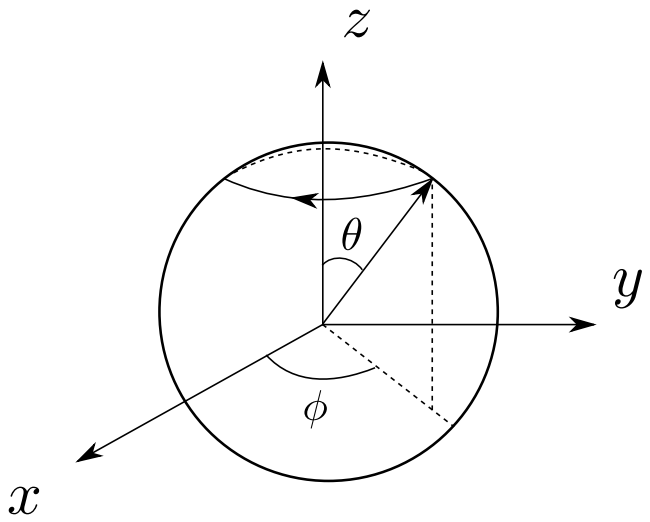


Figure 15.3 Précession de Larmor sur la sphère de Bloch. Le vecteur paramétrisé par θ et ϕ tourne autour de l'axe z à la fréquence ω_0.

15.4 Oscillations de Rabi

Nous allons maintenant nous attaquer à la dynamique du spin dans le champ $\vec{B} = (0, 0, B_0) + B_1\Big(\cos(\omega t), -\sin(\omega t), 0\Big)$ tournant.

L'équation de Schrödinger donnant l'opérateur d'évolution

$$i\hbar\frac{d}{dt}U(t,0) = H(t)U(t,0) \tag{15.27}$$

est toujours valable avec ici avec

$$H(t) = -\frac{\hbar\omega_0}{2}\sigma_z - \frac{\hbar\omega_1}{2}\left(\sigma_+ e^{i\omega t} + \sigma_- e^{-i\omega t}\right) \tag{15.28}$$

Néanmoins, l'équation de Schrödinger est moins aisée à résoudre car H(t) dépend du temps. Pour nous affranchir de cette difficulté, nous faisons un changement de référentiel. Dans le nouveau référentiel, l'Hamiltonien est indépendant du temps et il est facile de calculer l'opérateur d'évolution. Soit:

$$|\tilde{\psi}(t)\rangle = \exp\left(i\frac{t}{\hbar}K\right)|\psi(t)\rangle \tag{15.29}$$

avec $K = \begin{pmatrix} -\frac{\hbar\omega}{2} & 0 \\ 0 & \frac{\hbar\omega}{2} \end{pmatrix}$.
Ici l'exponentielle est la matrice de rotation d'angle $(t\omega)$ autour de z. Puisque $|\psi(t)\rangle = U(t,0)|\psi(0)\rangle$ on trouve

$$|\tilde{\psi}(t)\rangle = \exp\left(i\frac{t}{\hbar}K\right)U(t,0)|\psi(0)\rangle \tag{15.30}$$

$$\equiv \tilde{U}(t,0)|\psi(0)\rangle \tag{15.31}$$

Ainsi, le nouvel opérateur d'évolution dans le nouveau référentiel est $\tilde{U}(t,0) = e^{i\frac{t}{\hbar}K}U(t,0)$. Pour obtenir l'équation de Schrödinger, on calcule:

$$i\hbar\frac{d}{dt}\tilde{U} = i\hbar\frac{i}{\hbar}Ke^{i\frac{t}{\hbar}K}U + e^{i\frac{t}{\hbar}K}H(t)U \tag{15.32}$$

$$= \left\{-K + e^{i\frac{tK}{\hbar}}H(t)e^{-i\frac{tK}{\hbar}}\right\}\tilde{U} \tag{15.33}$$

$$\equiv \tilde{H}(t)\tilde{U} \tag{15.34}$$

Le nouvel Hamiltonien est $\tilde{H}$. On le calcule facilement. En effet:

$$e^{i\frac{tK}{\hbar}} = \begin{pmatrix} e^{-i\frac{t\omega}{2}} & 0 \\ 0 & e^{i\frac{t\omega}{2}} \end{pmatrix} \tag{15.35}$$

ce qui donne finalement,

$$\tilde{H} = \frac{\hbar\delta}{2}\sigma_z - \frac{\hbar\omega_1}{2}\sigma_x \tag{15.36}$$

avec $\delta = \omega - \omega_0$. L'opérateur d'évolution $\tilde{U}$ est donc

$$\tilde{U} = \exp\left(-\frac{it}{\hbar}\tilde{H}\right) = \exp\left\{-i\frac{t}{2}(\delta\sigma_z - \omega_1\sigma_x)\right\} \tag{15.37}$$

On peut calculer cette matrice à partir de la formule "d'Euler généralisée" (15.4). À partir de là, on déduit

$$U = e^{-i\frac{t}{\hbar}K}\tilde{U} \tag{15.38}$$

L'opérateur d'évolution final obtenu sans aucune approximation est:

$$U(t,0) = \begin{pmatrix} u_{\uparrow\uparrow} & u_{\uparrow\downarrow} \\ u_{\downarrow\uparrow} & u_{\downarrow\downarrow} \end{pmatrix} \tag{15.39}$$

avec les 4 éléments de matrice:

$$u_{\uparrow\uparrow} = e^{i\frac{t\omega}{2}}\left\{\cos\left(\frac{t}{2}\sqrt{\delta^2+\omega_1^2}\right) + i\frac{\delta}{\sqrt{\delta^2+\omega_1^2}}\sin\left(\frac{t}{2}\sqrt{\delta^2+\omega_1^2}\right)\right\} \tag{15.40}$$

$$u_{\uparrow\downarrow} = -i\frac{i\omega_1}{\sqrt{\delta^2+\omega_1^2}}e^{i\frac{t\omega}{2}}\sin\left(\frac{t}{2}\sqrt{\delta^2+\omega_1^2}\right) \tag{15.41}$$

$$u_{\downarrow\uparrow} = -i\frac{i\omega_1}{\sqrt{\delta^2+\omega_1^2}}e^{-i\frac{t\omega}{2}}\sin\left(\frac{t}{2}\sqrt{\delta^2+\omega_1^2}\right) \tag{15.42}$$

$$u_{\downarrow\downarrow} = e^{-i\frac{t\omega}{2}}\left\{\cos\left(\frac{t}{2}\sqrt{\delta^2+\omega_1^2}\right) - i\frac{\delta}{\sqrt{\delta^2+\omega_1^2}}\sin\left(\frac{t}{2}\sqrt{\delta^2+\omega_1^2}\right)\right\} \tag{15.43}$$

Nous avons complètement résolu le problème de la dynamique du vecteur d'état $|\psi\rangle$ dans un champ magnétique du type $(B_1\cos(\omega t), B_1\sin(\omega t), B_0)$. À partir de l'opérateur d'évolution, on peut calculer les probabilités de transitions suivantes

$$P_{|\uparrow\rangle\to|\downarrow\rangle}(t) = |\langle\downarrow|U(t)|\uparrow\rangle|^2 = |u_{\uparrow\downarrow}|^2 \tag{15.44}$$

$$P_{|\uparrow\rangle\to|\uparrow\rangle}(t) = |\langle\uparrow|U(t)|\uparrow\rangle|^2 = |u_{\uparrow\uparrow}|^2 \tag{15.45}$$

La première probabilité représente la probabilité d'observer l'état $|\downarrow\rangle$ à l'instant t lorsque l'état initial est $|\uparrow\rangle$. C'est donc la probabilité que le spin soit "retourné". La seconde est la probabilité que le spin reste inchangé. On trouve:

$$\begin{cases} P_{|\uparrow\rangle\to|\downarrow\rangle}(t) &= \dfrac{\omega_1^2}{\delta^2+\omega_1^2}\left(\sin\left(\dfrac{t}{2}\sqrt{\delta^2+\omega_1^2}\right)\right)^2 \\ P_{|\uparrow\rangle\to|\uparrow\rangle}(t) &= \left(\cos\left(\dfrac{t}{2}\sqrt{\delta^2+\omega_1^2}\right)\right)^2 + \dfrac{\delta^2}{\delta^2+\omega_1^2}\left(\sin\left(\dfrac{t}{2}\sqrt{\delta^2+\omega_1^2}\right)\right)^2 \end{cases} \tag{15.46}$$

On vérifie bien que ces deux probabilités se somment à 1, comme il se doit.

La Figure **15.4** représente la probabilité de transition $P_{|\uparrow\rangle\to|\downarrow\rangle}(t)$ en fonction du temps. C'est une fonction périodique de période $T_{Rabi} = \frac{2\pi}{\sqrt{\delta^2+\omega_1^2}}$ et de hauteur $\frac{\omega_1^2}{\sqrt{\delta^2+\omega_1^2}}$. Nous

voyons que l'amplitude est maximale lorsque $\delta = 0$ c.-à-d. $\omega = \omega_0$, lorsque la fréquence du champ tournant est égale à la fréquence de Larmor. Lorsque $\delta \gg \omega_1$ (on parle de detuning) la probabilité de transition est faible.

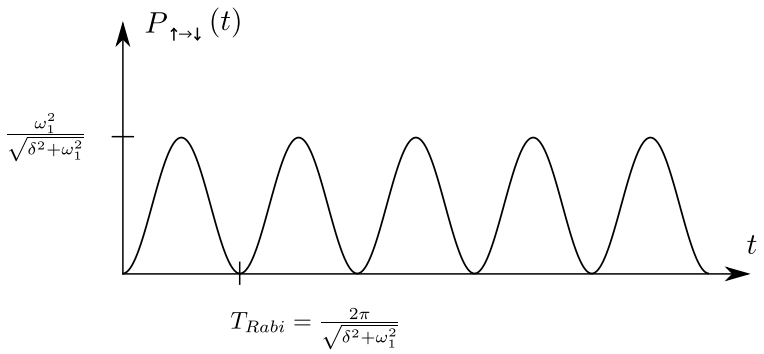


Figure 15.4 Probabilité de transition en fonction du temps.

15.5 Réalisations des portes quantiques

Les oscillations de Rabi, appliquées au cas $\delta = 0$ ($\omega = \omega_0$ tuning parfait entre la fréquence de rotation du champ tournant et la fréquence de Larmor) permettent de réaliser certaines portes à un qubit. Ici nous discutons les portes NOT et de Hadamard.

La porte NOT

En prenant $\delta = 0$ et $t = \frac{T_{Rabi}}{2} = \frac{\pi}{\omega_1}$ on voit que $P_{|\uparrow\rangle\to|\downarrow\rangle} = 1$. Ainsi le spin est retourné avec probabilité 1 par un champ tournant tel que $\omega = \omega_0$ enclenché pendant un temps $t = \frac{\pi}{\omega_1}$. Un tel champ s'appelle un "π-pulse" dans le langage de la RMN. Le temps $t = \frac{\pi}{\omega_1}$ est en gros la durée de basculement du spin.

Pour vérifier que l'opérateur d'évolution unitaire correspond bien (est équivalent) à la porte NOT, il est commode d'examiner l'évolution dans le référentiel tournant.

$$\tilde{U} = \exp\Big(- i\frac{t}{2}(\delta\sigma_z - \omega_1\sigma_x)\Big). \tag{15.47}$$

Pour $\delta = 0$ et $t = \frac{\pi}{\omega_1}$ on trouve

$$\tilde{U} = \exp\Big(i\frac{\pi}{2}\sigma_x\Big) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \tag{15.48}$$

Il s'agit d'une matrice de rotation d'angle π autour de l'axe x. Le vecteur $|\uparrow\rangle$ est bien transformé en $|\downarrow\rangle$ et $|\downarrow\rangle$ est bien transformé en $|\uparrow\rangle$.

La porte de Hadamard

Cette fois on fixe $\delta = 0$ (tuning parfait $\omega = \omega_0$) et on enclenche le champ tournant $(B_1 \cos(\omega t), -B_1 \sin(\omega t), B_0)$ pendant un temps $t = \frac{T_{Rabi}}{4} = \frac{\pi}{2\omega_1}$. Notez que cette durée est la moitié de celle de la porte NOT. On trouve alors $P_{|\uparrow\rangle\to|\downarrow\rangle} = \frac{1}{2}$ et $P_{|\downarrow\rangle\to|\uparrow\rangle} = \frac{1}{2}$. Plus précisément dans le référentiel tournant

$$\tilde{U} = \exp\Big(- i\frac{t}{2}(\delta\sigma_z - \omega_1\sigma_x)\Big) \tag{15.49}$$

$$= \exp\left(i\frac{\pi}{4}\sigma_x\right) \tag{15.50}$$

$$= \cos\left(\frac{\pi}{4}\right)\mathbb{1} + i\sin\left(\frac{\pi}{4}\right)\sigma_x \tag{15.51}$$

$$= \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & i \\ i & 1 \end{pmatrix} \tag{15.52}$$

Cette matrice effectue le basculement de $|\uparrow\rangle$ vers $\frac{1}{\sqrt{2}}(|\uparrow\rangle + i|\downarrow\rangle)$ et le basculement de $|\downarrow\rangle$ vers $\frac{1}{\sqrt{2}}(i|\uparrow\rangle + |\downarrow\rangle) = \frac{i}{\sqrt{2}}(|\uparrow\rangle - i|\downarrow\rangle)$. Cette opération est physiquement équivalente à une porte de Hadamard, $H = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$.

16 Hamiltonien d'Heisenberg et Portes à deux qubits

Dans le Chapitre **15**, nous avons discuté la dynamique du spin dans un champ magnétique dépendant du temps. Cela permet comme nous l'avons vu de réaliser des portes à un qubit. Par exemple, nous avons vu comment réaliser les portes NOT et Hadamard (Section **15.5**). Pour implémenter le calcul quantique, il faut encore être capable de réaliser des portes à deux qubits. Nous avons vu que l'ensemble des portes universelles contient CNOT et en principe cette porte à deux qubits suffit à fabriquer n'importe quel circuit. Cette porte fait intervenir deux qubits et ne peut être réalisée qu'à partir de leur interaction. Cela est généralement vrai pour toute porte à deux qubits (qui est non triviale). Pour cette raison, nous allons tout d'abord étudier l'interaction entre deux moments magnétiques. Nous verrons ensuite qu'avec certaines formes de cette interaction magnétique il est possible de fabriquer les portes voulues. De telles interactions magnétiques abondent dans la nature. Ce point sera discuté au Chapitre **17**.

16.1 Hamiltonien d'Heisenberg

Nous avons vu que l'énergie d'interaction entre un moment magnétique $\vec{M}$ et un champ magnétique $\vec{B}$ est donné par $-\vec{B}\cdot\vec{M}$ (Section **2.7**). Considérons maintenant deux moments magnétiques $\vec{M}_1$ et $\vec{M}_2$. On peut penser à ceux-ci comme à deux petits aimants. Pour minimiser leur énergie, ceux-ci vont avoir tendance à s'éloigner de façon "antiparallèle", comme l'illustre la Figure **16.1**.

Ici les boucles représentent les lignes du champ magnétique. Si nous demandons que l'énergie d'interaction soit indépendante de l'orientation globale du système, c'est-à-dire qu'il n'y a pas de direction privilégiée; on peut prendre comme Hamiltonien simple (ici " $\approx$ proportionnel "):

$$H \approx \vec{M}_1 \cdot \vec{M}_2 \tag{16.1}$$

Notons aussi que cette expression est la seule possible qui soit à la fois invariante sous les rotations du référentiel (pas de direction privilégiée) et du premier ordre dans $\vec{M}_1$ et $\vec{M}_2$.

Pour les moments magnétiques quantiques intrinsèques de spin 1/2 (par exemple proton, noyaux atomiques ^{13}C, ^{19}C, etc.) nous savons que $\vec{M}_1 = g_1\frac{\hbar}{2}\vec{\sigma_1}$ et $\vec{M}_2 = g_2\frac{\hbar}{2}\vec{\sigma_2}$ où

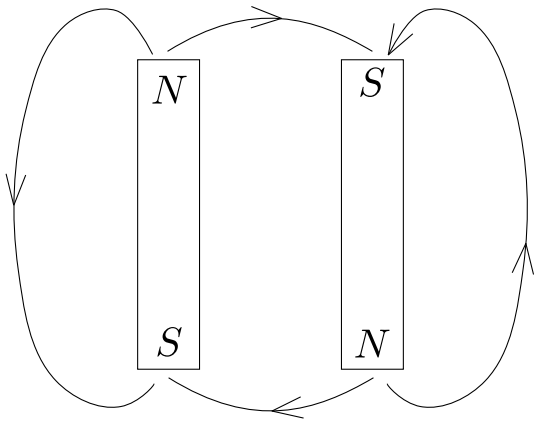


Figure 16.1 Lignes de champs de deux dipôles magnétiques

$\vec{\sigma_1}$ et $\vec{\sigma_2}$ sont les vecteurs des matrices de Pauli et g_1, g_2 dépendent du type précis de noyaux. L'Hamiltonien d'interaction entre deux moments magnétiques quantiques de spin 1/2 est donc donné par (système invariant de rotation):

$$H = \hbar J \vec{\sigma_1} \cdot \vec{\sigma_2} \tag{16.2}$$

où $\hbar J$ à l'unité d'énergie et J l'unité d'une fréquence [s^{-1}].

Faisons quelques remarques importantes sur l'interprétation du produit scalaire ci-dessus. Tout d'abord $\vec{\sigma_1} \cdot \vec{\sigma_2}$ doit être une matrice 4×4 puisqu'il s'agit d'une observable (matrice hermitienne) agissant sur les états de deux qubits (sur l'espace de Hilbert $\mathbb{C}^2 \otimes \mathbb{C}^2$). Ainsi

$$\vec{\sigma_1} \cdot \vec{\sigma_2} = \sigma_1^x \otimes \sigma_2^x + \sigma_1^y \otimes \sigma_2^y + \sigma_1^z \otimes \sigma_2^z. \tag{16.3}$$

Explicitement, dans la base canonique

$$\sigma_1^x \otimes \sigma_2^x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \otimes \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \tag{16.4}$$

$$\sigma_1^y \otimes \sigma_2^y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \otimes \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \tag{16.5}$$

$$\sigma_1^z \otimes \sigma_2^z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \otimes \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{16.6}$$

Et donc

$$H = \hbar J \vec{\sigma_1} \cdot \vec{\sigma_2} = \hbar J \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 2 & 0 \\ 0 & 2 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{16.7}$$

Il est facile de calculer les valeurs propres (niveaux d'énergie) et vecteurs propres de cette matrice. Néanmoins, il est encore plus instructif de le faire en utilisant l'algèbre de Pauli.

Introduisons les matrices

$$\sigma^+ = \frac{1}{2}(\sigma^x + i\sigma^y) = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \tag{16.8}$$

$$\sigma^- = \frac{1}{2}(\sigma^x - i\sigma^y) = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \tag{16.9}$$

Le lecteur vérifiera que $\sigma^+|\uparrow\rangle = 0$, $\sigma^+|\downarrow\rangle = |\uparrow\rangle$, $\sigma^-|\downarrow\rangle = 0$, $\sigma^-|\uparrow\rangle = |\downarrow\rangle$. En exprimant σ^x et σ^y en fonction de σ^+ et σ^- il est facile de montrer que

$$H = \hbar J \left\{ \sigma_1^z \otimes \sigma_2^z + 2(\sigma_1^+ \otimes \sigma_2^- + \sigma_1^- \otimes \sigma_2^+) \right\} \tag{16.10}$$

On peut vérifier que les éléments de matrice $\langle s_1' s_2' | H | s_1 s_2 \rangle$ où $s_1, s_2, s_1', s_2' = \uparrow, \downarrow$ donnent la matrice écrite précédemment. Par exemple

$$H|\uparrow\downarrow\rangle = 2\hbar J(\sigma_1^- \otimes \sigma_2^+)|\uparrow\downarrow\rangle \tag{16.11}$$

$$= 2\hbar J|\downarrow\uparrow\rangle \tag{16.12}$$

Si bien que $\langle\downarrow\uparrow|H|\uparrow\downarrow\rangle = 2\hbar J$.

Considérons maintenant l'action de H sur l'état $\frac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle)$ (Notez que cet état est l'un des états de Bell).

$$\sigma_1^z \otimes \sigma_2^z(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle) = -(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle) \tag{16.13}$$
$$\sigma_1^+ \otimes \sigma_2^-(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle) = -|\uparrow\downarrow\rangle \tag{16.14}$$
$$\sigma_1^- \otimes \sigma_2^+(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle) = |\downarrow\uparrow\rangle. \tag{16.15}$$

Ces trois équations impliquent

$$H(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle) = -3\hbar J(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle). \tag{16.16}$$

L'état $\frac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle)$ est donc un état propre de H d'énergie $-3\hbar J$.

Considérons l'action de H sur les trois états $|\uparrow\uparrow\rangle$; $\frac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle)$ et $|\downarrow\downarrow\rangle$. On voit facilement que $(\sigma_1^+ \otimes \sigma_2^- + \sigma_1^- \otimes \sigma_2^+)$ s'annule contre ces trois états. Il reste donc l'action de $\sigma_1^z \otimes \sigma_2^z$ qui donne

$$H|\uparrow\uparrow\rangle = \hbar J|\uparrow\uparrow\rangle \tag{16.17}$$
$$H(|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle) = \hbar J(|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle) \tag{16.18}$$
$$H|\downarrow\downarrow\rangle = \hbar J|\downarrow\downarrow\rangle \tag{16.19}$$

Ainsi, ces trois états sont états propres de H avec niveaux d'énergie (valeur propre) $+\hbar J$.

En résumé, l'état fondamental de H est l'état dit "singulet" d'énergie $-3\hbar J$ et les autres états, dits "triplets", possédant tous la même énergie $+\hbar J$.

À titre d'exercice, regardons l'effet d'un champ magnétique extérieur $\vec{B}$ sur les niveaux d'énergie du système. L'Hamiltonien devient:

$$H = -\frac{\hbar g_1}{2}\vec{B}\cdot\vec{\sigma_1} - \frac{\hbar g_2}{2}\vec{B}\cdot\vec{\sigma_2} + \hbar J\vec{\sigma_1}\cdot\vec{\sigma_2} \tag{16.20}$$

Notez que le terme $\vec{B}\cdot\vec{\sigma_1}$ doit être interprété comme $(\vec{B}\cdot\vec{\sigma_1} \otimes \mathbb{1}_2)$ et le terme $\vec{B}\cdot\vec{\sigma_2}$ doit être interprété comme $(\mathbb{1}_1 \otimes \vec{B}\cdot\vec{\sigma_2})$. Il est clair que nous pouvons orienter $\vec{B}$ le long de l'axe z. Puisque l'interaction magnétique $\hbar J\vec{\sigma_1}\cdot\vec{\sigma_2}$ est invariante sous les rotations cela ne fait pas de différence sur les niveaux d'énergie.

Considérons le cas le plus simple de deux noyaux identique $g_1 = g_2$. Alors en posant aussi $\hbar gB = \hbar\omega_0$ (la fréquence de Larmor, Section **15.3**), on est amené à étudier les niveaux d'énergie de

$$H = -\frac{\hbar\omega_0}{2}(\sigma_1^z \otimes \mathbb{1}_2 + \mathbb{1}_1 \otimes \sigma_2^z) + \hbar J\vec{\sigma_1}\cdot\vec{\sigma_2} \tag{16.21}$$

On a

$$H|\uparrow\uparrow\rangle = (-\hbar\omega_0 + \hbar J)|\uparrow\uparrow\rangle \tag{16.22}$$
$$H|\downarrow\downarrow\rangle = (+\hbar\omega_0 + \hbar J)|\downarrow\downarrow\rangle \tag{16.23}$$
$$H(|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle) = \hbar J(|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle) \tag{16.24}$$
$$H(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle) = -3\hbar J(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle) \tag{16.25}$$

Nous voyons que l'énergie de l'état singulet est inchangée. Cela n'est pas surprenant puisque dans cet état les composantes z du spin sont opposées. La même remarque vaut pour l'état $(|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle)$. Les états $|\uparrow\uparrow\rangle$ et $|\downarrow\downarrow\rangle$ voient leurs énergies descendre (orientation parallèle à $\vec{B}$) et monter (orientation anti-parallèle à $\vec{B}$). Le point important ici et que la dégénérescence de l'état triplet est "levée" par la perturbation additionnelle: ceci est une caractéristique assez générique en MQ. Le graphe de l'énergie des états en fonction de B où ω_0 est donné à la Figure **16.2**.

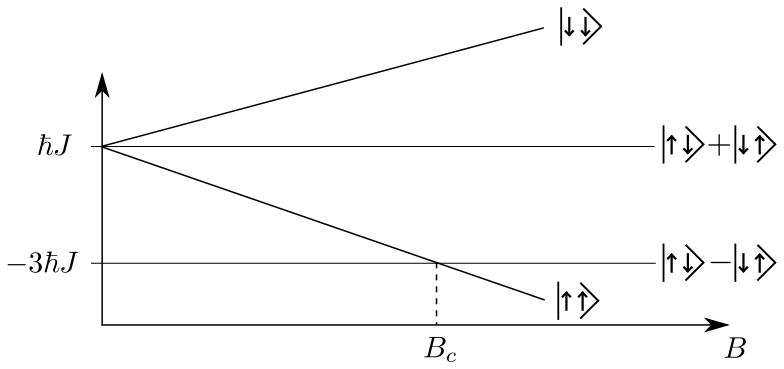


Figure 16.2 Graphe de l'énergie des états triplets et singulets en fonction de l'intensité du champ magnétique extérieur $\vec{B}$.

Il existe un champ critique $\vec{B}_c$ au-delà duquel c'est l'état $|\uparrow\uparrow\rangle$ qui devient état de plus basse énergie (état fondamental).

16.2 Porte SWAP et Hamiltonien de Heisenberg

La porte SWAP est donnée par la définition suivante sur les états de la base computationnelle:

$$\text{SWAP}\,|x, y\rangle = |y, x\rangle \tag{16.26}$$

Par linéarité, cette définition s'étend sur tout l'espace de Hilbert. Cette porte est importante car elle permet d'échanger les qubits d'un circuit. De plus, on peut réaliser une porte SWAP à partir de 3 portes CNOT: Figure **16.3**.

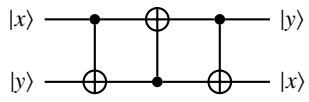


Figure 16.3 Circuit pour la réalisation du SWAP.

De plus, on peut montrer qu'il est possible de réaliser CNOT à partir de $\sqrt{\text{SWAP}}$ et de portes à un qubit. Ainsi $\sqrt{\text{SWAP}}$ joue aussi le rôle de porte universelle au même titre que CNOT.

Dans ce paragraphe, nous montrons que la porte SWAP peut être réalisée grâce à l'Hamiltonien d'Heisenberg isotrope. De même $\sqrt{\text{SWAP}}$ peut aussi être fabriqué avec ce même Hamiltonien.

Calculons tout d'abord l'opérateur d'évolution pour l'Hamiltonien d'Heisenberg.

$$U_{Heis}(t) = \exp\Big(-\frac{it}{\hbar}H_{Heis}\Big). \tag{16.27}$$

Nous allons le faire en notation de Dirac et dans la base des états propres de cet Hamiltonien. Sa décomposition spectrale est:

$$H = -3\hbar J\Big\{\frac{|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle}{\sqrt{2}}\cdot\frac{\langle\uparrow\downarrow| - \langle\downarrow\uparrow|}{\sqrt{2}}\Big\} \tag{16.28}$$

$$+\hbar J\Big\{|\uparrow\uparrow\rangle\langle\uparrow\uparrow| + \frac{|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle}{\sqrt{2}}\cdot\frac{\langle\uparrow\downarrow| + \langle\downarrow\uparrow|}{\sqrt{2}} + |\downarrow\downarrow\rangle\langle\downarrow\downarrow|\Big\} \tag{16.29}$$

Dans la base des états propres, il suffit de calculer l'exponentielle des valeurs propres:

$$U_{Heis}(t) = e^{i3Jt}\Big\{\frac{|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle}{\sqrt{2}}\cdot\frac{\langle\uparrow\downarrow| - \langle\downarrow\uparrow|}{\sqrt{2}}\Big\} \tag{16.30}$$

$$+e^{-itJ}\Big\{|\uparrow\uparrow\rangle\langle\uparrow\uparrow| + \frac{|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle}{\sqrt{2}}\cdot\frac{\langle\uparrow\downarrow| + \langle\downarrow\uparrow|}{\sqrt{2}} + |\downarrow\downarrow\rangle\langle\downarrow\downarrow|\Big\} \tag{16.31}$$

Un bon exercice que nous laissons au lecteur est d'écrire à partir de cette expression la matrice en composantes dans la base canonique $|\uparrow\uparrow\rangle, |\uparrow\downarrow\rangle, |\downarrow\uparrow\rangle, |\downarrow\downarrow\rangle$.

Prenons maintenant $t = \frac{\pi}{4J}$. C'est-à-dire que nous supposons que l'interaction magnétique est enclenchée pendant un intervalle de temps $\frac{\pi}{4J}$ seulement. Bien sûr, en pratique cela pose un problème car on ne peut pas enclencher et déclencher à volonté les interactions magnétiques entre moments magnétiques. Mais nous verrons au Chapitre **17**

comment "la technique de refocalisation" (Section **17.4**) permet de remédier à ce problème.

Pour $t = \frac{\pi}{4J}$ on a $e^{i3Jt} = e^{i\frac{3\pi}{4}} = -e^{i\frac{\pi}{4}}$ et $e^{-tiJ} = e^{-i\frac{\pi}{4}}$. Ainsi

$$U_{Heis}(\frac{\pi}{4J}) = e^{-i\frac{\pi}{4}}\Big\{-\frac{|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle}{\sqrt{2}} \cdot \frac{\langle\uparrow\downarrow| - \langle\downarrow\uparrow|}{\sqrt{2}} \tag{16.32}$$

$$+ |\uparrow\uparrow\rangle\langle\uparrow\uparrow| + \frac{|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle}{\sqrt{2}} \cdot \frac{\langle\uparrow\downarrow| + \langle\downarrow\uparrow|}{\sqrt{2}} + |\downarrow\downarrow\rangle\langle\downarrow\downarrow|\Big\} \tag{16.33}$$

Il est facile de voir sur cette formule que

$$U_{Heis}(\frac{\pi}{4J})|\uparrow\uparrow\rangle = e^{-i\frac{\pi}{4}}|\uparrow\uparrow\rangle \tag{16.34}$$

$$U_{Heis}(\frac{\pi}{4J})|\downarrow\downarrow\rangle = e^{-i\frac{\pi}{4}}|\downarrow\downarrow\rangle \tag{16.35}$$

$$U_{Heis}(\frac{\pi}{4J})(|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle) = e^{-i\frac{\pi}{4}}(|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle) \tag{16.36}$$

$$U_{Heis}(\frac{\pi}{4J})(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle) = -e^{-i\frac{\pi}{4}}(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle) \tag{16.37}$$

$$= e^{-i\frac{\pi}{4}}(|\downarrow\uparrow\rangle - |\uparrow\downarrow\rangle) \tag{16.38}$$

À une phase globale près (qui n'est pas importante physiquement), nous voyons que $U_{Heis}(\frac{\pi}{4J})$ échange bien les deux qubits. Sous cet échange, l'état triplet reste invariant et l'état singulet change de signe. Notez que l'addition et la soustraction des deux dernières équations donnent bien:

$$U_{Heis}(\frac{\pi}{4J})|\uparrow\downarrow\rangle = e^{-i\frac{\pi}{4}}|\downarrow\uparrow\rangle \tag{16.39}$$

$$U_{Heis}(\frac{\pi}{4J})|\downarrow\uparrow\rangle = e^{-i\frac{\pi}{4}}|\uparrow\downarrow\rangle \tag{16.40}$$

qui n'est rien d'autre que le SWAP.

Pour obtenir $\sqrt{\text{SWAP}}$, il suffit de choisir $t = \frac{\pi}{8j}$ au lieu de $\frac{\pi}{4J}$!

16.3 Porte CNOT et interaction magnétique anisotrope

Si les moments magnétiques sont dans un environnement anisotrope alors l'Hamiltonien de Heisenberg devient anisotrope lui aussi. Un cas extrême, mais qui est une très bonne approximation dans certaines expériences de RMN (Chapitre **17**) est celui d'une anisotropie où la direction z est dominante (par exemple à cause du champ externe $\vec{B_0} = (0, 0, B_0)$ très puissant par rapport aux interactions magnétiques). Dans ces cas, un Hamiltonien qui décrit le système raisonnablement bien est

$$H = \hbar J \sigma_1^z \otimes \sigma_2^z. \tag{16.41}$$

Cet Hamiltonien est purement diagonal:

$$H = \hbar J \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \otimes \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \hbar J \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{16.42}$$

$$= \hbar J \Big\{ |\uparrow\uparrow\rangle\langle\uparrow\uparrow| - |\uparrow\downarrow\rangle\langle\uparrow\downarrow| - |\downarrow\uparrow\rangle\langle\downarrow\uparrow| + |\downarrow\downarrow\rangle\langle\downarrow\downarrow| \Big\} \tag{16.43}$$

L'opérateur d'évolution est donc

$$e^{-i\frac{t}{\hbar}H} = e^{-itJ}|\uparrow\uparrow\rangle\langle\uparrow\uparrow| + e^{itJ}|\uparrow\downarrow\rangle\langle\uparrow\downarrow| + e^{itJ}|\downarrow\uparrow\rangle\langle\downarrow\uparrow| + e^{-itJ}|\downarrow\downarrow\rangle\langle\downarrow\downarrow| \tag{16.44}$$

$$= \begin{pmatrix} e^{-itJ} & 0 & 0 & 0 \\ 0 & e^{itJ} & 0 & 0 \\ 0 & 0 & e^{itJ} & 0 \\ 0 & 0 & 0 & e^{-itJ} \end{pmatrix} \tag{16.45}$$

L'identité suivante peut facilement être démontrée:

$$CNOT = (\mathbb{1}_1 \otimes H_2)(R_1 \otimes R_2) e^{-\frac{i}{\hbar}\frac{\pi}{4J}H} (\mathbb{1}_1 \otimes H_2) \tag{16.46}$$

où $R_1 = \exp\left(-i\frac{\pi}{4}\sigma_1^z\right)$ et $R_2 = \exp\left(-i\frac{\pi}{4}\sigma_2^z\right) = \begin{pmatrix} e^{-i\frac{\pi}{4}} & 0 \\ 0 & e^{i\frac{\pi}{4}} \end{pmatrix}$.

Ainsi la porte CNOT peut être réalisée en "enclenchant" l'interaction magnétique pendant un temps $\frac{\pi}{4J}$ et en combinant cela avec des manipulations à un qubit. À nouveau, en pratique, il faut utiliser la "technique de refocalisation" (Section **17.4**) pour "enclencher" et "déclencher" l'interaction magnétique.

Le circuit associé à la réalisation de CNOT correspond à la formule (16.46):

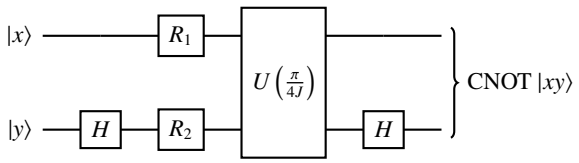


Figure 16.4 Circuit pour la réalisation du CNOT.

17 Réalisations expérimentales

Nous exposons dans ce chapitre les principes de base des réalisations expérimentales des algorithmes quantiques. L'illustration des principes sera faite dans le cadre de la résonance magnétique nucléaire (RMN), qui, comme nous le verrons, a permis de réaliser certains algorithmes, tels que celui de Shor, en laboratoire. Même s'il n'est probablement pas possible d'aller au-delà d'une dizaine de qubits avec la RMN, celle-ci a l'avantage d'illustrer de façon concrète et simple les principes universels qui s'appliquent aussi bien à d'autres technologies plus prometteuses. En effet, quelle que soit la technologie sous-jacente pour la fabrication et la manipulation de qubits effectifs, les hamiltoniens régissant leur dynamique ne peuvent être que des "polynômes" exprimés à partir des matrices de Pauli, contenant des termes linéaires et quadratiques à l'ordre le plus bas.

17.1 Les systèmes en jeu

Dans le cas de la RMN, les qubits sont les moments magnétiques de noyaux atomiques de spin 1/2, de molécules appropriées. Ces molécules sont naturelles ou éventuellement synthétiques, et forment un fluide ou solide macroscopique. La Figure **17.1** montre deux molécules naturelles - le chloroforme et le trichloroéthylène - qui ont déjà été utilisées pour le calcul quantique. Comme nous le verrons, chaque molécule correspond en quelque sorte à un circuit quantique. Dans ces molécules, les qubits en jeu sont les spins $1/2$ des noyaux d'hydrogène ^1H, de carbone ^{13}C et de fluor ^{19}F. Le noyau de l'atome hydrogène est un proton qui possède un spin $1/2$. Le ^{13}C contient 6 protons et 7 neutrons[1]. Dans l'état fondamental du noyau, les spins des protons et neutrons se compensent entre eux, si bien que le spin total est $1/2$. La situation est similaire pour l'isotope ^{19}F du fluor qui possède 9 protons et 10 neutrons.

Le fluide (ou le solide) à température ambiante, est placé dans une région ou règne un champ magnétique statique $\vec{B}_0 = (0, 0, B_0)$ orienté dans la direction z, d'intensité $B_0 \approx 1 - 10$T. Il s'agit d'un champ très intense: 1T = 10^5 Gauss, le Gauss étant l'ordre de grandeur du champ magnétique terrestre. Les fréquences de Larmor des moments

[1] Il s'agit de l'isotope du ^{12}C contenant 6 protons et 6 neutrons et possède un spin total nul; le ^{14}C connu pour être utilisé pour la datation, est un autre isotope contenant 6 protons et 8 neutrons, et son spin total est aussi nul. Les atomes de carbone sont électriquement neutres et contiennent tous 6 électrons; les propriétés chimiques sont régies par le nuage électronique et sont donc identiques pour des isotopes.

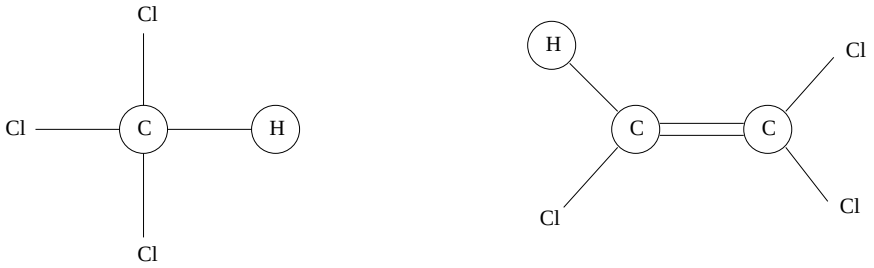


Figure 17.1 Molécules de chloroforme (gauche) et trichloroéthylène (droite). Les atomes d'hydrogène ^1H, carbone ^{13}C et fluor ^{19}F dont le noyau de spin $1/2$ est un qubit, sont entourés. Les noyaux des atomes de chlore possèdent un spin nul et ne sont pas des qubits.

magnétiques des noyaux ^1H, ^{13}C, ^{19}F sont de l'ordre de grandeur $\omega_L \approx 10 - 100$MHz (1MHz = 10^6Hz).

Les portes à un qubit sont réalisées par des pulses de radiofréquence $\vec{B}_1 e^{i\omega t}$ d'intensité $\omega_1 \approx 0.1 - 1$MHz. Ces pulses sont générés par une ou des bobines parcourues par un courant alternatif de radiofréquence ω. Il est possible d'agir sélectivement sur un type de qubit ^1H, ^{13}C, ^{19}F en réglant ω sur la fréquence de Larmor du qubit (cas résonant). L'influence sur les autres qubits ayant une fréquence de Larmor différente (non résonnante) peut être négligé. Nous verrons que l'environnement chimique d'un qubit dans la molécule a un effet correctif sur les fréquences de Larmor. Ce point est important, d'une part car il faut régler les fréquences de résonance suffisamment précisément, et d'autre part car cela permet d'agir sélectivement et distinguer les qubits de noyaux identiques. L'intensité ω_1 donne l'ordre de grandeur de la durée de la porte $\tau_1 \approx \frac{2\pi}{\omega_1} \approx 10^{-6}$ sec.

Les portes à deux qubits sont réalisées par la nature elle-même. Comme nous le verrons, il est possible de prendre avantage de l'interaction spin-spin de type Heisenberg pour réaliser des portes à deux qubits. La durée de ces portes est déterminée par l'intensité du couplage spin-spin et est d'environ 10^{-3} sec. Celle-ci est donc mille fois plus longue que la durée des portes à un qubit. Cette dernière pourra en première approximation être négligée.

À ce point, il est instructif de se faire une image des différentes échelles d'énergies en jeu. Le "circuit quantique moléculaire" est contrôlé grâce aux pulses de radiofréquence. L'énergie mise en jeu dans ces pulses de radio-fréquence est $\hbar\omega_1 \approx 10^{-6}$ eV. Celle-ci est au moins mille fois plus faible que toutes les autres énergies en jeu, si bien que ces pulses n'affectent en rien la structure moléculaire. En effet, les énergies d'excitations du nuage électronique sont de l'ordre de l'eV[2]; l'énergie du splitting Zeeman des spins électroniques est de l'ordre de 10^{-3} eV[3].

Le calcul quantique par RMN porte simultanément sur un nombre macroscopique de molécules et la mesure collective finale produit directement un résultat statistique. Pour pouvoir manipuler les moments magnétiques des noyaux actifs dans la molécule, il faut pouvoir caractériser et séparer de façon suffisamment précise leurs fréquences

[2] 1 eV = $1,6\,10^{-19}$ Joule, J = kg m^2/s^2, et $\hbar = 6,58211928(15).10^{-16}$ eV.

[3] L'énergie d'excitation interne des noyaux est 10^6 eV, et pour les nucléons $\sim 10^9$ eV (l'échelle de la chromodynamique quantique).

de Larmor. Ceci n'est possible que pour des molécules de tailles relativement, faibles contenant jusqu'à une dizaine de qubits. Même s'il était possible de synthétiser des molécules avec une centaine ou un millier de noyaux représentant les qubits actifs, les fréquences de Larmor formeraient un quasi-continuum et il deviendrait difficile de les manipuler et/ou de maintenir la cohérence des états quantiques. En effet, dans ce cas, d'une part les énergies d'excitations thermiques, et d'autre part la largeur des pulses de radiofréquence sont plus grandes que les différences entre fréquences de Larmor. C'est une des raisons pour lesquelles il n'a pas été encore possible d'utiliser la RMN pour fabriquer un ordinateur quantique fonctionnant sur plus d'une dizaine de qubits. Une autre raison importante provient du fait que les spins de l'état initial ne sont pas suffisamment bien polarisés et leur état est loin d'être un état pur. C'est un état équilibre thermique. L'une des innovations des expériences par RMN fut de transformer l'état thermique en état "pseudo-pur". Ce dernier point dépasse le cadre de ce cours ne sera pas abordé ici[4].

17.2 Oscillations de Rabi et portes à un qubit

Ce paragraphe résume l'essentiel de la théorie des oscillations de Rabi, qui sont à la base de la réalisation des portes à un qubit telles que les portes de Hadamard, les rotations autour de x, y ou z, ou encore les $\pi/2^k$ shifts.

L'hamiltonien d'un moment magnétique dans un champ $\vec{B}_0 = (0, 0, B_0)$ est

$$H = -\frac{g\hbar}{2m}\vec{B}_0 \cdot \vec{\sigma} = -\frac{\hbar\omega_L}{2}\sigma_z.$$

Pour le noyau d'hydrogène ^1H, $g \approx 5.59$, $m = 10^{-27}$ kg et $q = 1,6\,10^{-19}$ C, ce qui donne une fréquence de Larmor $\omega_L \approx 100$ MHz. L'évolution temporelle d'un état est donnée par[5]

$$|\psi\rangle = e^{i\frac{\varphi+\omega_L t}{2}} \cos\frac{\theta}{2}|\uparrow\rangle + e^{-i\frac{\varphi+\omega_L t}{2}} \sin\frac{\theta}{2}|\downarrow\rangle.$$

On peut se faire une image de cette dynamique en représentant l'état par un vecteur sur la sphère de Bloch. Ce vecteur possède un angle θ constant avec z, et effectue un mouvement de précession de fréquence ω_L autour de l'axe z.

Lors d'une expérience de RMN, le champ $B_0 \approx 1 - 10$ Tesla est fixé une fois pour toute et tous les moments magnétiques effectuent un mouvement de précession autour de z, avec une fréquence de Larmor qui leur est propre.

Pour manipuler les moments magnétiques on utilise un champ de radio-fréquences, de fréquence ω et d'intensité $B_1 \approx 10^{-2}$ Tesla. Pour que ce champ influence notablement un moment magnétique, il faut régler sa fréquence sur la fréquence de résonance $\omega \approx \omega_L$. Dans ce cas, le spin peut absorber ou émettre un quantum d'énergie $\hbar\omega$ précisément

[4] À titre d'information, indiquons qu'un état pseudo-pur est une matrice densité de la forme $\rho = \alpha\mathbb{1} + \delta|\Psi\rangle\langle\Psi|$. La moyenne de toute observable A à trace nulle, satisfait $\mathrm{Tr}\rho A = \delta\langle\Psi|A|\Psi\rangle$. Lorsque l'on transforme l'état thermique en état pseudo-pur, δ diminue avec le nombre de qubits et le signal résultant de la mesure devient trop faible

[5] L'état initial est obtenu en posant $t = 0$.

égal à la différence entre les niveaux d'énergie du moment magnétique dans le champ $\vec{B}_0$. Un basculement du spin autour des axes x et/ou y est réalisé en enclenchant le champ de radio-fréquences $\vec{B}_1(t) = (B_1 \cos \omega t, B_1 \sin \omega t, 0)$. L'hamiltonien est

$$H = -\frac{\hbar\omega_L}{2}\sigma_z - \frac{\hbar\omega_1}{2}(\sigma_x \cos \omega t + \sigma_y \sin \omega t).$$

Le calcul de l'opérateur d'évolution donne alors (si $\omega = \omega_L$)

$$\text{prob}(|\uparrow\rangle \to |\downarrow\rangle) = |\langle\downarrow|U(t)|\uparrow\rangle|^2 = \sin^2 \frac{\omega_1 t}{2}.$$

Nous voyons que la probabilité de transition oscille entre 0 et 1 avec une période égale à $\frac{2\pi}{\omega_1} \approx 10^{-6}$ sec. Ce sont les oscillations dites de Rabi.

Porte NOT. Si l'on pose $t = \frac{\pi}{\omega_1}$ (moitié d'une période de Rabi) on trouve

$$U(\frac{\pi}{\omega_1}) = \sigma_x = \text{NOT}.$$

Ainsi, pour réaliser une porte NOT, il suffit d'enclencher le champ de radio-fréquences pendant une durée $\frac{\pi}{\omega_1}$. Cette opération, s'appelle aussi un π-pulse et retourne le spin $|\uparrow\rangle \to |\downarrow\rangle$ et $|\downarrow\rangle \to |\uparrow\rangle$.

Porte de Hadamard. Pour $t = \frac{\pi}{2\omega_1}$ (quart d'une période de Rabi) on trouve

$$U(\frac{\pi}{2\omega_1}) = H.$$

Pour réaliser une porte H il suffit d'enclencher le champ de radio-fréquences pendant une durée $\frac{\pi}{2\omega_1}$. Cette opération, s'appelle aussi un $\pi/2$-pulse et fait basculer un spin, initialement le long de z, dans le plan xy. $|\uparrow\rangle \to \frac{1}{\sqrt{2}}(|\downarrow\rangle + |\uparrow\rangle)$ et $|\downarrow\rangle \to \frac{1}{\sqrt{2}}(|\downarrow\rangle - |\uparrow\rangle)$.

Autres portes. En théorie, on peut procéder de façon similaire pour réaliser d'autres opérations à un qubit, par exemple les rotations autour de z. Nous noterons qu'en pratique, vu la configuration des bobines, on ne peut pas créer des pulses de radiofréquences le long de z. Pour réaliser la rotation d'un qubit autour de z on utilise plutôt un déphasage des signaux ultérieurs sur ce qubit, ce qui simule une rotation du référentiel autour de z. Ce déphasage dépend de la fréquence de Larmor du qubit en question.

Finalement, il est important de noter que les pulses de radio-fréquences agissent sélectivement sur les qubits associés aux noyaux ^1H, ^{13}C, ^{19}F car les fréquences de Larmor sont bien séparées. Nous verrons que cela est aussi possible pour des noyaux identiques car leur environnement chimique dans la molécule modifie légèrement leurs fréquences de Larmor.

17.3 Couplage spin-spin et portes à deux qubits

Les portes à deux qubits exploitent l'interaction naturelle entre moments magnétiques. L'évolution temporelle correspondant à cette interaction est un opérateur unitaire qui ne peut pas se ramener au produit tensoriel de deux unitaires à un qubit. En d'autres

termes, cette interaction crée de l'intrication et permet en particulier de réaliser la porte $CNOT$. Notons immédiatement que cela ne va pas sans poser de problèmes car cette interaction n'est pas déclenchable à souhait. En effet, elle est *naturelle* ! Ce problème peut être contourné grâce à la technique de *refocalisation* (Section **17.4**).

Pour deux moments magnétiques nucléaires dans la molécule, reliés par un axe, $\vec{n}$ l'interaction dipolaire magnétique est de la forme

$$\hbar J_{\rm dip}(3(\vec{\sigma}_1\cdot\vec{n})(\vec{\sigma}_2\cdot\vec{n})-\vec{\sigma}_1\cdot\vec{\sigma}_2).$$

La constante de couplage décroit comme l'inverse du cube de la distance entre moments magnétiques. Notez qu'il y a dans cette expression deux structures algébriques: le produit scalaire dans l'espace Euclidien et le produit tensoriel entre matrices de Pauli dans l'espace de Hilbert des spins. Ce premier terme ne joue pas de rôle dans un fluide car les rotations thermiques des molécules induisent un effet de moyenne sur le vecteur $\vec{n}$. En effet, $\langle n_i\rangle = 0$ et $\langle n_i^2\rangle = \frac{1}{3}$ $(i = x, y, z)$ entrainent

$$\langle(\vec{\sigma}_1\cdot\vec{n})(\vec{\sigma}_2\cdot\vec{n})\rangle = \frac{1}{3}\langle\vec{\sigma}_1\cdot\vec{\sigma}_2\rangle$$

et donc l'interaction dipolaire magnétique s'annule en moyenne dans un fluide à l'équilibre thermique. Les moments magnétiques interagissent aussi avec le nuage électronique qui est dans leur environnement, ce qui entraine une interaction effective, à travers les liaisons chimiques, entre spin nucléaires. L'analyse[6] de ces effets est non-triviale et conduit à un hamiltonien de type Heisenberg (anisotrope)

$$\sum_{i,j=x,y,z}\hbar J_{ij}\sigma_{1i}\otimes\sigma_{2j}.$$

En fait $J_{ij} << \hbar\omega_L$ et cet hamiltonien peut être considéré comme une petite perturbation de l'hamiltonien de départ régissant la précession de Larmor, qui lui est diagonal dans la base computationnelle. Un calcul de perturbation à l'ordre le plus bas montre alors qu'il suffit de retenir la composante z de cette interaction.

Tout compte fait, le couplage entre spins sera modélisé par (on posera $J_{zz} = J$)

$$\mathcal{H}_{int} = \hbar J\sigma_{1z}\otimes\sigma_{2z}.$$

L'évolution unitaire associée à cette interaction est

$$U_{int}(t) = \exp\left(-\frac{it}{\hbar}\mathcal{H}_{int}\right) = \exp(-itJ\sigma_{1z}\otimes\sigma_{2z}).$$

C'est une matrice 4×4 diagonale dans la base computationnelle. Ses éléments diagonaux sont $e^{-itJ}; e^{itJ}; e^{itJ}; e^{-itJ}$. De plus $\hbar J << \hbar\omega_1$ $(<< \hbar\omega_L)$. Cette séparation entre échelles d'énergies est très importante car elle signifie que *l'échelle de temps des interactions est mille fois plus grande que la durée des portes à un qubit*. En effet, $\tau_{\rm int} \approx \frac{2\pi}{J} \approx 10^{-2} - 10^{-3}$ sec, alors que $\tau_{\rm pulse} \approx \frac{2\pi}{\omega_1} \approx 10^{-6}$ sec. Ainsi le pas de temps élémentaire des calculs quantiques par RMN est de l'ordre de la millième de seconde. Cette échelle de temps n'est pas particulièrement faible, comparée aux échelles de temps

[6] Celle-ci fait appel à l'interaction de contact de Fermi entre le nuage électronique et le noyau, le principe de Pauli et la règle de Hund.

des ordinateurs classiques; mais le gain potentiel du calcul quantique provient de la complexité potentiellement polynomiale des algorithmes, comparée à une complexité potentiellement exponentielle dans le cas classique.

L'identité suivante, et son circuit correspondant (Figure **17.3**), est à la base de la réalisation de la porte $CNOT$.

$$CNOT = (\mathbb{1} \otimes H)(R_1 \otimes R_2)U_{int}(t = \frac{\pi}{4J_z})(\mathbb{1} \otimes H)$$

ou

$$R_1 = \exp\left(i\frac{\pi}{2}\frac{\sigma_{1z}}{2}\right), \qquad R_2 = \exp\left(i\frac{\pi}{2}\frac{\sigma_{2z}}{2}\right).$$

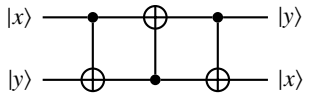


Figure 17.2 Circuit correspondant à l'identité pour la porte $CNOT$. Ici U_{int} est l'évolution unitaire due au couplage spin-spin. Les portes à un qubit sont réalisées grâce à des pulses de radiofréquences.

Prenons l'exemple de la molécule de chloroforme, où les qubits sont les moments magnétiques des noyaux ^1H et ^{13}C. Supposons que ^1H est le qubit 1 et ^{13}C le qubit 2. Le protocole expérimental pour réaliser la porte $CNOT$ est le suivant:

- A l'instant initial, envoyer un $\frac{\pi}{2}$-pulse de fréquence $\omega = \omega_L^C$, dans le plan xy: cela réalise la porte de Hadamard sur le deuxième qubit. L'hamiltonien correspondant est
$$-\frac{\hbar\omega_1}{2}(\sigma_x \cos\omega_L^C t + \sigma_y \sin\omega_L^C t).$$
- Puis envoyer deux $\frac{\pi}{2}$-pulses de fréquences $\omega = \omega_L^H$ et $\omega = \omega_L^C$, dans la direction z: cela réalise les deux rotations R_1 et R_2 d'angle $\pi/2$ autour de z. L'hamiltonien correspondant est
$$-\frac{\hbar\omega_1}{2}\sigma_z(\cos\omega_L^H t + \cos\omega_L^C t).$$
- Attendre un temps $\tau = \frac{\pi}{4J}$. Pendant ce temps, les deux qubits évoluent selon leur interaction naturelle d'hamiltonien
$$\hbar J\sigma_{1z} \otimes \sigma_{2z}.$$
- A l'instant final, envoyer un $\frac{\pi}{2}$-pulse, dans le plan xy, de fréquence $\omega = \omega_L^C$: cela réalise la dernière porte de Hadamard sur le deuxième qubit. L'hamiltonien correspondant est le même que ci-dessus.

Bien sûr, l'interaction naturelle $\hbar J\sigma_{1z} \otimes \sigma_{2z}$ agit aussi pendant les opérations sur les qubits individuels (les $\pi/2$-pulses). Donc ce protocole expérimental ne réalise la porte $CNOT$ qu'approximativement. Mais le point important est que cette approximation est

excellente, dans la mesure où la durée des $\pi/2$ pulses est négligeable - et peuvent donc être considérés instantanés - par rapport à $\tau_{\text{int}} = \frac{\pi}{4J}$. Finalement, n'oublions pas qu'au cours de toutes ces opérations, les spins effectuent leur précession de Larmor autour de la direction de $\vec{B}_0$.

17.4 Refocalisation

Considérons pour fixer les idées l'état obtenu après avoir effectué le protocole du paragraphe précédent pour la porte $CNOT$. À priori, cet état continue à évoluer avec l'interaction naturelle entre les moments magnétiques. Comment pouvons-nous le préserver pendant un temps appréciable, disons de l'ordre de la milliseconde ? Une autre situation d'intérêt serait celle où nous voudrions effectuer des opérations sur certains qubits, tout en maintenant d'autres qubits dans un état donné pendant un temps de l'ordre de la milliseconde. À priori, cela peut sembler difficile dans la mesure où les interactions naturelles ne peuvent pas être déclenchées à volonté. La technique de *refocalisation* permet de contourner le problème. Celle-ci joue un rôle capital dans la réalisation expérimentale des algorithmes: en fait, comme nous l'illustrons dans le dernier le paragraphe, la plupart des opérations effectuées sont relatives à la refocalisation.

Soit $t_{\text{in}} < t_{\text{fin}}$ deux instants initiaux et finaux, tels que $t_{\text{fin}} - t_{\text{in}}$ soit de l'ordre de quelques millisecondes. Soit $|\psi_{\text{in}}\rangle$ l'état initial, d'un système de deux qubits, que l'on veut préserver. Son évolution naturelle le conduirait à l'état final

$$|\psi_{\text{fin}}\rangle = \exp(-i(t_{\text{fin}} - t_{\text{in}})J\sigma_{1z} \otimes \sigma_{2z})|\psi_{\text{in}}\rangle.$$

Supposons maintenant qu'aux instants $\frac{t_{\text{fin}}+t_{\text{in}}}{2}$ et t_{fin} nous agissions avec un π-pulse correspondant à la rotation

$$R_{1x} = \exp\left(i\pi\frac{\sigma_{1x}}{2}\right) \otimes \mathbb{1}_2$$

d'angle π et d'axe x agissant sur le premier qubit (par exemple ${}^1H$). Comme ces pulses ont une durée négligeable par rapport à la durée totale de l'évolution, le nouvel état final - appelons le $|\psi_{\text{refoc}}\rangle$ - sera donné par

$$|\psi_{\text{refoc}}\rangle \approx R_{1x}\exp(-\frac{i(t_{\text{fin}} - t_{\text{in}})}{2}J\sigma_{1z} \otimes \sigma_{2z})R_{1x}\exp(-\frac{i(t_{\text{fin}} - t_{\text{in}})}{2}J\sigma_{1z} \otimes \sigma_{2z})|\psi_{\text{in}}\rangle.$$

Il n'est pas difficile de vérifier l'identité mathématique exacte

$$\mathbb{1}_1 \otimes \mathbb{1}_2 = R_{1x}\exp(-i\frac{t}{2}J\sigma_{1z} \otimes \sigma_{2z})R_{1x}\exp(-i\frac{t}{2}J\sigma_{1z} \otimes \sigma_{2z}),$$

valable pour tout t et J. Donc

$$|\psi_{\text{refoc}}\rangle \approx |\psi_{\text{in}}\rangle,$$

l'état obtenu par l'opération de refocalisation est une très bonne approximation de l'état initial (idéalement voulu).

Le principe général exposé ci-dessus peut être appliqué à des situations plus compliquées. En un mot, l'idée principale est de "corriger" la dynamique naturelle en agissant avec de courts pulses sur les qubits individuels.

17.5 Déplacements chimiques et effets de couplage

Il existe un effet important dont n'avons pas encore tenu compte: le déplacement chimique (ou chemical shift). Dans l'environnement de la molécule, la fréquence de Larmor nue d'un noyau est modifiée par l'environnement chimique du noyau. En effet, le champ local ressenti par un noyau est égal à B_0 plus des corrections diamagnétiques et paramagnétiques provenant du nuage électronique. Cela est important car des noyaux identiques auront des fréquences de Larmor légèrement différentes, ce qui permet de les adresser individuellement par les pulses de radiofréquences. Le décalage de ces fréquences de Larmor est appelé *déplacement chimique*.

Il y a encore un autre effet important sur les fréquences de Larmor dont il faut tenir compte, qui provient du couplage spin-spin. Celui-ci entraine un "éclatement" de chaque fréquence de Larmor en plusieurs fréquences satellites. Pour fixer les idées, considérons à nouveau la molécule de chloroforme. Supposons que l'état du ^{13}C soit $|\phi\rangle$. Dans ce cas l'hamiltonien effectif de ^{1}H sera

$$\mathcal{H}_{\text{eff}} = -\frac{\hbar\omega_L^H}{2}\sigma_z^H + \hbar J_{HC}\sigma_z^H\langle\phi|\sigma_z^C|\phi\rangle.$$

Selon que $|\phi\rangle = |\uparrow\rangle, |\downarrow\rangle$ on obtient

$$\mathcal{H}_{\text{eff}} = -\frac{\hbar(\omega_L^H \pm 2J_{HC})}{2}\sigma_z^H$$

et on trouve deux fréquences de Larmor effectives pour ^{1}H. La fréquence de précession est $\omega_L^H - 2J_{HC}$ quand ^{13}C est dans l'état $|\uparrow\rangle$, et $\omega_L^H + 2J_{HC}$ quand ^{13}C est dans l'état $|\downarrow\rangle$. De façon similaire, le ^{13}C possède deux fréquences de Larmor $\omega_L^C - 2J_{HC}$ et $\omega_L^C + 2J_{HC}$ selon que ^{1}H est dans l'état $|\uparrow\rangle$ ou $|\uparrow\rangle$.

Considérons la molécule de trichloroéthylène (Figure **17.1**). Les deux atomes de ^{13}C n'ont pas un environnement chimique identique, si bien que les déplacements chimiques de leur fréquence de Larmor diffèrent. Cela est bénéfique car nous pouvons distinguer les deux qubits ! De plus, chacune de ces fréquences est éclatée à cause du couplage spin-spin. Il n'est pas difficile de voir que pour la molécule de trichloroéthylène il y a *trois groupes de quatre fréquences*. En effet, ω_L^H est éclatée en quatre fréquences satellites correspondantes aux quatre états possibles des deux C^{13}, $|\uparrow\uparrow\rangle, |\uparrow\downarrow\rangle, |\downarrow\uparrow\rangle, |\downarrow\downarrow\rangle$. De même ω_L^C est éclatée en deux groupes (un pour chaque carbone) de quatre fréquences chacun correspondantes aux quatre états des noyaux voisins. La Figure **17.5** donne le spectre des deux ^{13}C. Les quatre fréquences de ^{1}H sont plus élevées et ne figurent pas sur cette échelle.

Il est facile de généraliser. Pour une molécule avec N noyaux de spin $1/2$ (ou N qubits) il y a à priori N fréquences de Larmor (à cause du déplacement chimique, ceci est le cas même si les noyaux sont identiques). À cause du couplage spin-spin, on aura en réalité N groupes de 2^{N-1} fréquences satellites[7]. Ces effets sont très important en spectroscopie par RMN. Les spectres moléculaires représentent une signature

[7] Ceci est vrai si tous les déplacements chimiques sont différents. Sinon certaines de ces fréquences satellites sont confondues.

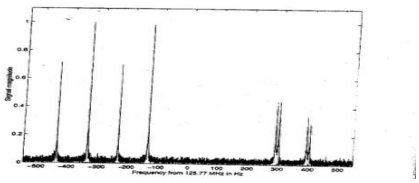


Figure 17.3 Spectre des deux ^{13}C dans la molécule de trichloroéthylène. L'origine des fréquences est placée 125.77 MHz et l'échelle horizontale est en Hz. [Source: Nielsen and Chuang, Quantum Computation and Information, CUP].

de la structure moléculaire, qu'il est souvent possible d'inférer à partir des différents groupes de fréquences. Le nombre, la position et l'intensité des pics donnent de précieuses informations sur la molécule. En ce qui concerne les réalisations expérimentales des algorithmes quantiques, il faut tenir compte de cet effet pour régler avec assez de précision la fréquence des pulses. Il permet aussi comme indiqué plus haut de distinguer les qubits associés à des noyaux identiques. Nous allons voir (Section **17.6**) qu'il permet aussi de lire les états de la base computationnelle.

17.6 Lecture des qubits

Dans les expériences de RMN, l'observation de l'état moléculaire est réalisé grâce à une mesure sur un ensemble statistique de qubits. Le signal obtenu est donc directement une valeur moyenne. Ainsi, il ne s'agit pas vraiment de la mesure projective d'un état quantique individuel, mais plutôt d'une mesure collective. Bien sûr, la valeur moyenne obtenue est conforme au postulat de la mesure.

Imaginons une expérience de RMN qui consiste à agir sur une solution moléculaire par une suite de pulses créés grâce à une bobine traversée par un courant alternatif. Cette même bobine va servir pour la mesure. L'état final des qubits $|\Psi_{\text{fin}}\rangle$ possède un moment magnétique total (dans ce paragraphe nous laissons tomber les facteurs de proportion-

nalité et les constantes)

$$\vec{M}_{\text{fin}} \sim \langle \Psi_{\text{fin}} | \sum_{i=1}^{N} \vec{\sigma}_i \, | \Psi_{\text{fin}} \rangle.$$

Cette aimantation macroscopique précesse dans le champ $\vec{B}_0$ (autour de z) ce qui induit, en conformité avec la loi de Faraday, une tension et un courant dans la bobine (ici nous supposons que la bobine est orientée selon x et que sa résistance est R)

$$i(t) = RV(t) \sim -\frac{d}{dt} M_x(t).$$

Ce signal est une combinaison linéaire de signaux sinusoïdaux de fréquences égales aux différentes fréquences de Larmor constituant la précession de l'aimantation. Une analyse de Fourier donne un spectre de fréquences de Larmor (corrigées par les déplacements chimiques). Notons que le signal est exponentiellement décroissant sur une échelle $O(T)$, dans le domaine temporel, à cause des effets de relaxation de l'aimantation vers une valeur d'équilibre (décohérence). Dans le domaine fréquentiel, cela se traduit par un élargissement $O(\frac{1}{T})$ des raies spectrales. Si cet élargissement est trop grand, c'est-à-dire la décroissance exponentielle du signal trop rapide, la mesure perd en précision.

Pour une seule fréquence on a

$$i(t) \sim e^{-t/T} \cos \omega_L t.$$

La raie spectrale correspondante est une fonction Lorentzienne

$$\hat{i}(\omega) \sim \frac{T}{1 + (\omega - \omega_L)^2 T^2}.$$

Appliquons ces principes à la lecture des états de la base computationnelle. Nous n'allons pas tenir compte des effets de relaxation, dont la description dépasse le formalisme élémentaire utilisé ici. Commençons par le cas le plus simple d'un état à N qubits ayant tous la même fréquence de Larmor, et de la forme

$$|\uparrow\uparrow \cdots \uparrow\rangle.$$

Un $\pi/2$-pulse renverse cet état dans le plan xy,

$$|\Psi_{\text{fin}}\rangle = (|\uparrow\rangle + e^{-it\omega_L}|\downarrow\rangle) \otimes \cdots (|\uparrow\rangle + e^{-it\omega_L}|\downarrow\rangle)$$

(à une phase globale près). Cet état possède une aimantation macroscopique $(M_x(t), M_y(t), 0)$ précessant autour de z. Un calcul simple donne

$$M_x(t) \sim N \sin \omega_L t,$$

ce qui implique pour le courant sinusoïdal traversant la bobine

$$i(t) \sim N\omega \cos \omega_L t, \qquad \text{et} \qquad \hat{i}(\omega) \sim N\delta(\omega - \omega_L), \omega > 0.$$

En commençant avec l'état initial $|\downarrow\downarrow \cdots \downarrow\rangle$, le $\pi/2$-pulse donne l'état $|\Psi_{\text{fin}}\rangle = (|\uparrow\rangle - e^{-it\omega_L}|\downarrow\rangle) \otimes \cdots (|\uparrow\rangle - e^{-it\omega_L}|\downarrow\rangle)$, et mène finalement à

$$M_x(t) \sim -N \sin \omega_L t$$

et

$$i(t) \sim -N\omega \cos \omega_L t, \qquad \text{et} \qquad \hat{i}(\omega) \sim -N\delta(\omega - \omega_L), \omega > 0.$$

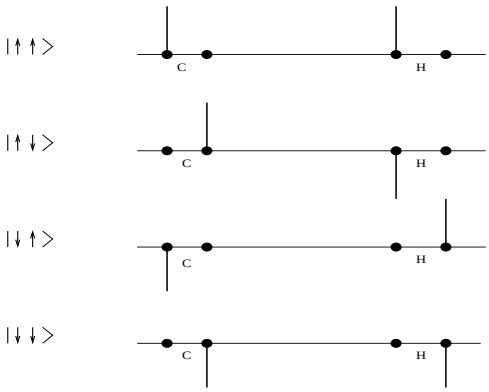


Figure 17.4 Représentation schématique des spectres de RMN correspondants aux états de la base computationnelle du chloroforme (axe horizontal = fréquences de Larmor).

Ces calculs simples permettent de tirer un enseignement intéressant.

- Un qubit dans l'état $|\uparrow\rangle$ donne des pics *positifs*[8].
- Un qubit dans l'état $|\downarrow\rangle$ donne des pics *négatifs*.
- Un qubit dans un état de superposition $\alpha|\uparrow\rangle + \beta|\downarrow\rangle$ donne un spectre avec des pics *positifs et négatifs*.

La Figure **17.6** donne une représentation schématique des pics associés aux états de la base computationnelle pour le cas de la molécule de chloroforme.

17.7 Réalisation de l'algorithme de Shor

Nous exposons ici les grandes lignes d'une expérience remarquable de Vandersypen-Steffen-Breyta-Yannoni-Sherwood-Chuang *"experimental realization of quantum factoring using nuclear magnetic resonance"*, Nature vol 414 pp. 883-887 (2001). Dans cette expérience, le nombre 15 est factorisé expérimentalement en suivant les principes de l'algorithme de Shor. En fait, il s'agit de calculer la période de la fonction $f(x) = a^x \bmod 15$ pour a premier avec 15. Ces auteurs ont réalisé l'expérience dans les deux cas $a = 11$ et $a = 7$. Ici nous allons illustrer cette expérience pour $a = 7$. Cette expérience, bien que ridicule du point de vue mathématique, est importante car elle prouve que les principes du calcul quantique peuvent être réalisés en laboratoire. D'autre part elle représente un tour de force de toute beauté.

Tout d'abord, nous résumons la théorie de l'algorithme de Shor dans le cas concret qui nous intéresse ici.

[8] En pratique on verra plusieurs pics a cause de l'effet du couplage avec les autres qubits.

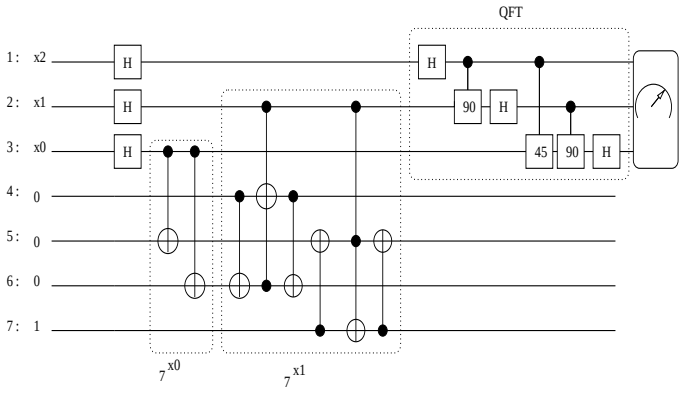


Figure 17.5 Circuit de l'algorithme de Shor pour $N = 15$ et $a = 7$. Les qubits numerotés de 1 à 7 correspondent aux spins des noyaux de la molécule de la Figure **17.6**. Les bits $x_2x_1x_0$ représentent les entiers $x \in \{0, 1, \cdots, 7\}$. L'état entrant est $|000\rangle \otimes |0001\rangle = |0\rangle \otimes |1\rangle$.

La fonction $f(x) = 7^x \bmod 15$ est représentée sur la figure pour les entiers $x \in \{0, 1, 2, 3, 4, 5, 6, 7, 8, \dots\}$. Nous voyons que la période est $r = 4$ En d'autres termes, l'ordre de 7 modulo 15 est égal a 4, un nombre pair. La factorisation est donc obtenue en calculant PGCD$(7^{4/2} \pm 1; 15) = 3$ et 5.

Pour voir la périodicité de cette fonction il suffit de travailler avec les entiers $\{0, \dots, 7\}$. Il faut trois qubits pour représenter les kets associés $|x\rangle$. Ainsi nous prendrons $M = 2^3 = 8$. Nous avons aussi besoin de stocker les images $7^x \bmod 15$. Pour cela, il faut au plus 4 qubits. Il nous suffit donc de travailler avec 7 qubits au total. Notons aussi que si $x = x_0 + 2x_1 + 4x_2 + 8x_3$,

$$a^x = a^{x_0} a^{2x_1} a^{4x_2} a^{8x_3}$$

et donc pour $a = 7$ (puisque $7^4 = 1$)

$$7^x = 7^{x_0} 7^{2x_1}.$$

L'exponentielle modulaire est donc contrôlée seulement par les bits x_0 et x_1. Le circuit de Shor est donné sur la figure 17.5. Cette figure illustre aussi la réalisation des deux multiplications contrôlées par 7^{x_0} et 7^{x_1}. Celle-ci fait intervenir uniquement des portes *CNOT* et nous laissons au lecteur le soin de vérifier que ce sont les circuits voulus.

Nous distinguons quatre phases dans l'évolution unitaire du circuit quantique. Cela sera pratique pour discuter l'expérience. Dans la phase (0), l'algorithme est initialisé dans l'état $|0000001\rangle$, qui correspond au produit tensoriel des "entiers" $|0\rangle \otimes |1\rangle$. La phase (1) correspond à l'action des portes de Hadamard qui prépare la superposition cohérente

$$\frac{1}{2^4} \sum_{x=0}^{7} |x\rangle \otimes |1\rangle$$

La troisième phase (2) correspond aux opérations d'exponentiation modulaire. L'état

Figure 17.6 Molécule de perfluobutadienyl utilisée dans la réalisation expérimentale de l'algorithme de Shor. Les moments magnétiques des noyaux numérotés de 1 a 7 correspondent aux qubits du circuit de la Figure **17.5**. Pour un champ magnetique de 11.7 T, les fréquences de Larmor decalées $\omega_i/2\pi$ (mesurées à 470 MHz pour F^{19} et 125 MHz pour C^{13}) ainsi que les couplages J_{ij} sont données en Hz. Les temps de relaxation T_{i1} et T_{i2} sont donnés en sec. [source: [VSBYSC] Nature vol 414 pp.883-887 (2001)].

devient

$$(|0\rangle + |4\rangle) \otimes |1\rangle + (|1\rangle + |5\rangle) \otimes |7\rangle + (|2\rangle + |6\rangle) \otimes |4\rangle + (|3\rangle + |7\rangle) \otimes |3\rangle.$$

Enfin dans la quatrième phase (3) on applique la transformée de Fourier quantique qui donne l'état sortant

$$\textit{une superposition des etats } |0\rangle, |2\rangle, |4\rangle, |6\rangle \textit{ faites le calcul } !$$

Si nous appliquons maintenant le postulat de la mesure, nous trouvons $\text{Prob}(y) = 1/4$ pour $y = 0, 2, 4, 6$ et $\text{Prob}(y) = 0$ pour $y = 1, 3, 5, 7$. On observe que $M/r = 2$. Puisque $M = 2^3 = 8$ on en déduit $r = 4$ ce qui est le bon résultat.

Ces dernières remarques indiquent que lors de la lecture de l'état sortant, par les techniques illustrées dans le paragraphe précédent, nous devrions observer des raies spectrales correspondantes aux états $|0\rangle = |000\rangle$, $|2\rangle = |010\rangle$, $|4\rangle = |100\rangle$, $|6\rangle = |110\rangle$. Discutons maintenant la réalisation expérimentale proprement dite.

Celle-ci utilise une solution contenant des molécules de synthèse contenant, entre autres, 7 noyaux actifs de spin $1/2$. Plus précisément, il y a 2 C^{13} et 5 F^{19}. Une représentation schématique de la molécule est donnée sur la Figure **17.6**, avec la numérotation utilisée pour les qubits. Les fréquences de Larmor déplacées ainsi que les couplages des 7 noyaux actifs peuvent être déterminés expérimentalement et sont donnés à titre d'indication dans la table. On peut penser à chaque molécule comme à un "circuit quantique" ou du moins un substrat pour le circuit.

La Figure **17.7** montre la séquence de pulses utilisée pour réaliser les portes à un qubit, les portes à deux qubits étant réalisées de façon naturelle via les couplages spin-spin.

Figure 17.7 Séquence temporelle des pulses de radio-fréquence agissant sur les qubits 1 à 7 de la molécule. [source: [VSBYSC] Nature vol 414 pp.883-887 (2001)].

Sur cette figure, les lignes de 1 a 7 représentent les 7 qubits (noyaux), l'axe horizontal représente le temps et les barres verticales indiquent les instants auxquels agissent les pulses. Comme expliqué plus haut, la durée de chaque pulse est négligeable (de l'ordre de 10^{-6} sec) par rapport aux intervalles de temps séparant les pulses (de l'ordre de 10^{-3} sec) correspondant à l'évolution naturelle. On distingue 4 phases séparées par les lignes verticales pointillées. La phase d'initialisation (0) nécessaire à la préparation de l'état initial, la phase (1) de préparation de l'état de superposition cohérente, la phase (2) qui est la plus longue (d'une durée totale de $\sim$ 400 ms) calcule l'exponentielle modulaire, et enfin la phase (3) qui effectue la QFT (d'une durée de $\sim$ 120 ms). La durée totale

Figure 17.8 Spectre expérimental des trois premiers qubits après la préparation de l'état initial. Il s'agit d'un état de mélange approximant bien l'état pur $|\uparrow\uparrow\uparrow\rangle$. [source: [VSBYSC] Nature vol 414 pp.883-887 (2001)].

de l'expérience est de 720 ms. Notons que la technique utilisée pour la préparation de l'état initial n'est pas triviale et fait partie des innovations de cette expérience. En effet, à priori, l'état des 7 qubits, n'est pas l'état pur $|0000001\rangle$, mais un état de mélange thermique qui est ramené à un état "pseudo-pur" par des opérations appropriées. Cette étape constitue en fait une limitation importante pour réaliser le calcul quantique quand le nombre de qubits augmente. Nous n'en disons pas plus ici.

Donnons quelques informations supplémentaires sur les différents types de pulses utilisés. Les barres hautes rouges sont des $\pi/2$ pulses correspondant à des rotations autour des axes x positif (pas de croix), x négatif (croix inférieure), et y positif (croix supérieure). Quand ces barres sont isolées elles représentent des portes de Hadamard alors que quand elles surviennent par paires, elles sont séparées par une évolution correspondant à une porte à deux qubits. Les barres bleues représentent les pulses utilisés pour la refocalisation. Ce sont des rotations d'angle π autour de l'axe x (positif $\rightarrow$ bleu foncé; négatif $\rightarrow$ bleu clair). Les petites barres vertes représentent des rotations autour de z.

On remarquera que la grande majorité des pulses utilisée ont trait à la refocalisation, et non pas aux portes du circuit quantique de Shor. La séquence de pulses nécessaire à la refocalisation a été calculée par simulations numériques, avant d'être implémentée dans l'expérience de RMN. De plus, un modèle de décohérence a été utilisé pour simuler la dynamique. Tous ces aspects dépassent de loin le cadre de cette introduction.

Passons finalement à l'étape finale qui consiste à lire *l'état sortant* des qubits. On utilise des $\pi/2$ pulses de lecture réglées sur les fréquences de résonance des 3 premiers qubits. Leurs spins basculent dans le plan xy et précessent autour de z, ce qui induit, en vertu de la loi de Faraday, un signal dans la bobine. Le spectre de Fourier de ce signal est donné sur la Figure **17.9**.

Figure 17.9 Spectres théorique idéal (première ligne), expérimental (deuxième ligne) et simulé avec un modèle de décohérence (troisième ligne) des trois qubits à la sortie du circuit. [source: [VSBYSC] Nature vol 414 pp.883-887 (2001)].

Ce spectre contient plusieurs fréquences à cause des nombreux effets de chemical shift. Néanmoins, on peut immédiatement en déduire que, dans l'état sortant, le qubit 0 était dans l'état $|\uparrow\rangle$ (c.-à-d. $|0\rangle$). D'autre part, les qubits 1 et 2 sont dans des états de superposition de $|\uparrow\rangle$ et $|\downarrow\rangle$ (c.-à-d. $|0\rangle$ et $|1\rangle$). On peut donc en déduire que l'état sortant était une superposition de

$$|000\rangle,\ |010\rangle,\ |100\rangle,\ |110\rangle.$$

Les entiers intervenant dans cette superposition sont $|0\rangle$, $|2\rangle$, $|4\rangle$, $|6\rangle$, comme prédit par la théorie. Cela permet de conclure que $M/r = 2$ et donc $r = 4$, puisque $M = 2^3 = 8$.

Part IV

Projets

18 Projets

18.1 Projet 2019

Inégalités de Bell (version CHSH): implémentation sur une machine quantique d'IBMQ

Le but de ce projet est d'implémenter un test de des inégalités de CHSH (vues en cours) sur un NISQ-device d'IBMQ. Dans un premier temps, vous allez préparer des états de Bell. Dans un deuxième temps, vous implémenterez le protocole des mesures faites par "Alice" et "Bob" puis calculerez le coefficient de corrélation de CHSH.

i) Vous pouvez créer vos circuits sur le "Composer" dans un premier temps et les tester d'abord sur le simulateur, puis faire l'expérience réelle.

ii) Ensuite vous pourrez créer des notebooks avec le langage Quiskit, simuler vos circuits avec le quasm-simulator, faire l'expérience sur les machines.

Tout d'abord: Allez sur le site https://www.ibm.com/quantum-computing/ et familiarisez-vous avec les informations. Ensuite, inscrivez-vous avec votre adresse email.

Composer: Vous pouvez utiliser directement le "Composer" pour créer des circuits, les simuler, puis faire les expériences.

Qiskit: familiarisez-vous grâce aux nombreux tutoriels sur le site. *Pour utiliser les vraies machines avec Qiskit*, créez un "account" et générez un API token. Une fois le compte sauvé ("load account" et "save account" dans Qiskit) il n'y a plus besoin de re-générer un nouveau token à chaque fois. Vous pouvez utiliser le simulateur, choisir une vraie machine avec une faible queue ou sélectionner la machine "least busy" (voir commandes sur Qiskit).

Étapes principales du projet: (à faire sur le Composer d'abord, puis avec le langage Qiskit).

1) Créez un circuit à deux qubits avec input $|00\rangle$ et output $|B_{00}\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$. Faites des mesures des deux qubits (dans la base computationnelle) et obtenez des his-

togrammes. Familiarisez-vous avec le résultat quand vous faites "1 shot", "1024 shots", etc. Observez les résultats sur le simulateur et sur les vraies machines.

2) Faites de même pour les 3 autres états de Bell. Observez que l'entrée du circuit est toujours $|00\rangle$. Vous devez donc ajouter des portes appropriées.

3) Maintenant reprenez le circuit qui prépare l'état de Bell. Le premier qubit est celui "d'Alice" et le second celui de "Bob." On rappelle le protocole expérimental pour tester l'inégalité de Bell:

- Mesures de type 1: A et B font N mesures dans les bases $\{|\alpha\rangle, |\alpha_\perp\rangle\}$ et $\{|\beta\rangle, |\beta_\perp\rangle\}$.
- Mesures de type 2: A et B font N mesures dans les bases $\{|\alpha\rangle, |\alpha_\perp\rangle\}$ et $\{|\beta'\rangle, |\beta'_\perp\rangle\}$.
- Mesures de type 3: A et B font N mesures dans les bases $\{|\alpha'\rangle, |\alpha'_\perp\rangle\}$ et $\{|\beta\rangle, |\beta_\perp\rangle\}$.
- Mesures de type 4: A et B font N mesures dans les bases $\{|\alpha'\rangle, |\alpha'_\perp\rangle\}$ et $\{|\beta'\rangle, |\beta'_\perp\rangle\}$.

Ils mettent en commun leurs résultats puis calculent le coefficient de corrélation:

$$X_{CHSH} = \text{Moy}_1(ab) + \text{Moy}_2(ab') - \text{Moy}_3(a'b) + \text{Moy}_4(a'b')$$

où Moy est une *moyenne empirique*. Pour les angles optimaux suivants $\alpha = 0, \alpha' = -\frac{\pi}{4}$, $\beta = \frac{\pi}{8}, \beta' = -\frac{\pi}{8}$ la théorie prévoit que $X_{CHSH}^{\text{theorie}} = 2\sqrt{2}$.

On veut tester ce résultat:

4) Créez 4 circuits qui implémentent les 4 types de mesures. Indication: Comme on ne peut faire des mesures que dans la base computationnelle, $\{|0\rangle, |1\rangle\}$ vous devez ajouter des portes quantiques, qui opèrent le changement de base, juste avant l'opération de mesure.

5) Collectez les histogrammes pour $N = 1024$ (ou $N = 8192$) pour chaque circuit.

6) En déduire $X_{CHSH}^{\text{simulateur}}$ (sur le simulateur) et $X_{CHSH}^{\text{experience}}$ (sur une vraie machine).

18.2 Projet 2020

Le but de ce projet est d'étudier le problème d'échantillonnage d'une distribution de probabilité classique grâce à l'utilisation d'un NISQ-device. À la fin de ce rapport, nous donnons des raisons qui poussent à s'intéresser à ce problème et des liens sur de la littérature récente. Il n'est cependant pas nécessaire de lire cette littérature (qui va au-delà de la matière enseignée en classe) pour pouvoir faire le projet.

Logez-vous sur https://quantum-computing.ibm.com/ et créez un compte avec e-mail epfl. Acceptez le contrat d'utilisateur. Une fois sur la page principale, générez un "token" que vous pouvez sauvegarder. Ce "token" n'est à générer qu'une fois et restera le même par la suite. Familiarisez-vous avec le site web IBMQ en essayant d'implémenter les circuits de base vus en cours.

Quelques ressources utiles (mais pas obligatoires) sont:
https://youtu.be/a1NZC5rqQD8 (une série de petites vidéos pour apprendre à travailler avec Qiskit)
https://qiskit.org/textbook/ch-algorithms/teleportation.html (ce lien montre aussi comment implémenter la téléportation quantique et est très bien pour apprendre comment construire des circuits et lancer des expériences.)

Le principe général

Considérez une distribution de probabilité $p(c_1, c_2, \ldots, c_n)$ de n variables binaires $c_i \in \{0, +1\}$. En d'autres mots, il y a 2^n entrées possibles et, pour chacune d'elles, $0 \leq p(c_1, c_2, \ldots, c_n) \leq 1$ et $\sum_{c_1,\ldots,c_n \in \{0,+1\}^n} p(c_1, c_2, \ldots, c_n) = 1$. Supposons que nous avons à disposition n bits quantiques, qui vivent dans l'espace de Hilbert $\mathbb{C}^{\otimes n}$. On dénote par $|c_1 c_2 \ldots c_n\rangle$ les états de la base computationnelle, où $c_i \in \{0, 1\}$. Étant donné la distribution p, supposons que nous puissions trouver une matrice unitaire $U : \mathbb{C}^{\otimes n} \to \mathbb{C}^{\otimes n}$ telle que

$$|\langle c_1 c_2 \ldots c_n | U | 0, 0, \ldots 0 \rangle|^2 = p(c_1, c_2, \ldots, c_n). \tag{18.1}$$

Si nous pouvons implémenter U avec un circuit quantique avec un nombre raisonnable (disons polynomial en n) de portes logiques élémentaires, alors nous pouvons échantillonner la distribution p en un temps polynomial (pour chaque échantillon).

Autrement dit, nous cherchons un circuit quantique tel que: 1) l'entrée du circuit est $|0\rangle^{\otimes n}$; 2) la sortie est $U|0\rangle^{\otimes n}$; 3) la mesure de la sortie dans la base computationnelle donne $|c_1 c_2 \ldots c_n\rangle$ avec probabilité donnée par l'équation (18.1).

Le circuit quantique implémente un échantillonneur. Pour n plus petit que 20 (approximativement), nous serons capable de l'implémenter physiquement sur un NISQ-device.

Une distribution de probabilité

Considérons la distribution de probabilité suivante pour $n \geq 3$:

$$p(c_1, \ldots, c_n) = \frac{1}{2^{2n}} \left| \sum_{b_1,\ldots,b_n \in \{0,1\}^n} (-1)^{\sum_{i=1}^n c_i b_i} \exp\left\{ i\theta \Big(\sum_{i=1}^{n-1} (-1)^{b_i}(-1)^{b_{i+1}} + (-1)^{b_n}(-1)^{b_1} \Big) \right\} \right|^2 \tag{18.2}$$

où $\theta \in [0, \pi]$. En particulier

$$p(0, \ldots, 0) = \frac{1}{2^{2n}} \left| \sum_{b_1,\ldots,b_n \in \{0,1\}^n} \exp\left\{ i\theta \Big(\sum_{i=1}^{n-1} (-1)^{b_i}(-1)^{b_{i+1}} + (-1)^{b_n}(-1)^{b_1} \Big) \right\} \right|^2 \equiv \frac{|Z(i\theta)|^2}{2^{2n}} \tag{18.3}$$

La quantité $Z(i\theta)$ est le facteur de normalisation (ou "fonction de partition") dudit modèle d'Ising, où $s_i \in \{-1, +1\}$

$$\pi(s_1, s_2, \ldots, s_n) = \frac{1}{Z(J)} \exp\left\{ J \Big(\sum_{i=1}^{n-1} s_i s_{i+1} + s_n s_1 \Big) \right\} \tag{18.4}$$

calculée pour $J = i\theta$. Ici, la fonction de partition peut être calculée exactement et nous admettons le résultat suivant (plus de détails dans l'Appendix)

$$Z(J) = 2^n \Big((\cosh J)^n + (\sinh J)^n \Big) \quad \text{et} \quad Z(i\theta) = 2^n \Big((\cos\theta)^n + i^n (\sin\theta)^n \Big) \tag{18.5}$$

Ceci donne la formule explicite

$$p(0, \ldots, 0) = |(\cos\theta)^n + i^n (\sin\theta)^n|^2. \tag{18.6}$$

Le circuit quantique pour l'échantillonnage de $p(c_1, \ldots, c_n)$

Considérez le circuit de la Figure **18.1**. L'état d'entrée est $|0\rangle^{\otimes n}$. On applique une porte de Hadamard à chaque qubit, ensuite on applique une matrice unitaire D suivie d'une deuxième série de portes de Hadamard. Finalement, nous mesurons dans la base computationnelle. Le résultat est une suite aléatoire de bits classiques $c_1, \ldots, c_n$.

Dans la Section **18.2**, vous serez guidés à travers une série de questions dont les objectifs sont:

- Découvrir quelle matrice unitaire D est telle que les résultats de la mesure sont des échantillonnages de la distribution (18.2). Autrement dit, pour quel D nous avons:

$$|\langle c_1, \ldots, c_n | H^{\otimes n} D H^{\otimes n} | 0, \ldots, 0 \rangle|^2 = p(c_1, \ldots, c_n) \tag{18.7}$$

- Trouver un circuit simple pour D ?
- Implémenter le circuit sur les machines NISQ d'IBMQ, disponibles au public et produire des histogrammes pour des petites valeurs de n et quelques valeurs de θ.
- Comparer les courbes théoriques pour $p(0, \ldots, 0)$ en fonction de θ avec les courbes (les points) expérimentales.

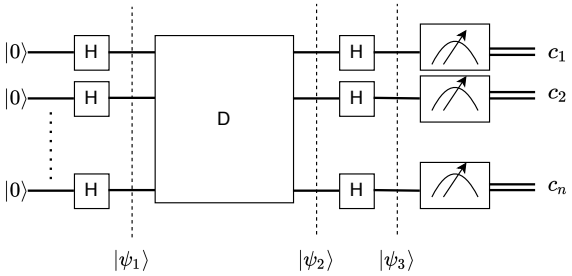


Figure 18.1 Circuit d'échantillonnage : le temps évolue de gauche à droite

Questions guidées

Sortie d'un circuit pour D diagonal

Considérons la classe des matrices D diagonales. Exprimé dans la base computationnelle, on a $D|b_1,\ldots,b_n\rangle = e^{i\varphi(b_1,\cdots,b_n)}|b_1,\ldots,b_n\rangle$ où $\varphi : \{0,1\}^n \to \mathbb{R}$ est une fonction réelle.

a) Calculez $|\psi_1\rangle$ puis $|\psi_2\rangle$. *Aide:* Commencez avec $n = 1$, $n = 2$ puis généraliser pour n'importe quel n.

b) Prouvez que la sortie du circuit juste avant la mesure est

$$|\psi_3\rangle = \frac{1}{2^n} \sum_{c_1,\ldots,c_n\in\{0,1\}} \left\{ \sum_{b_1\ldots b_n\in\{0,1\}^n} (-1)^{\sum_{i=1}^n b_i c_i} e^{i\varphi(b_1\ldots b_n)} \right\} |c_1\ldots c_n\rangle \tag{18.8}$$

Aide: Vérifiez que $H|b_i\rangle = \frac{1}{\sqrt{2}} \sum_{c_i=0,1} (-1)^{b_i c_i}|c_i\rangle$ et utilisez cette formule.

c) Quelle est la distribution de probabilité obtenue par une mesure dans la base computationnelle?

Trouver la matrice D

En comparant avec l'équation (18.2), nous devons trouver une matrice D de taille $2^n \times 2^n$ telle que $\varphi(b_1,\ldots,b_n) = \theta\Big(\sum_{i=1}^{n-1}(-1)^{b_i}(-1)^{b_{i+1}} + (-1)^{b_n}(-1)^{b_1}\Big)$.

d) Vérifiez que pour deux qubits, on a $e^{i\theta Z_1\otimes Z_2}|b_1b_2\rangle = e^{i\theta(-1)^{b_1}(-1)^{b_2}}|b_1b_2\rangle$ où $Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$. Déduisez-en D exprimée comme combinaisons de matrices de Pauli Z_i qui agissent sur les qubits $i = 1,\ldots,n$.

e) Le circuit pour D est représenté à la Figure **18.2**, où  représente la porte $e^{i\theta Z\otimes Z}$.

Montrez que cette porte est équivalente (à une phase globale près, qui n'a pas d'importance) à CNOT$(\mathbb{1} \otimes R(\theta))$CNOT où $R(\theta)$ est la porte de phase $\begin{pmatrix} 1 & 0 \\ 0 & e^{-2i\theta} \end{pmatrix}$.

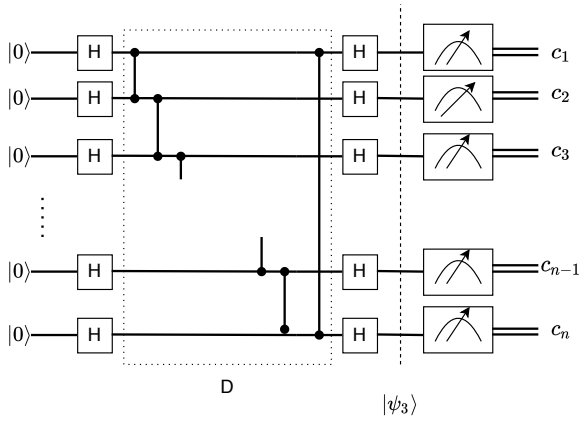


Figure 18.2 circuit d'échantillonnage avec l'implémentation de D

Implémentation sur IBMQ

Commencez par vous familiariser avec le site web de IBMQ:
https://quantum-computing.ibm.com/ Jouez avec le *Composer* pour vous habituer avec les circuits et utilisez le *simulateur* ainsi que les *vraies machines NISQ*. Découvrez également les bases du langage de Qiskit à l'aide des exemples donnés en classe et à travers les nombreux tutos disponibles sur le web.

f) D'abord, utilisez le *composer* (l'interface graphique) pour implémenter le circuit avec un petit nombre de qubits et choisissez le *simulateur* pour produire les histogrammes théoriques $p(c_1, \ldots, c_n)$. Ensuite, lancez l'expérience avec une vraie machine et comparez les histogrammes. Faites ceci pour $n = 3, \theta = \pi/5, \pi/3$ et$n = 4, \theta = \pi/5, \pi/3$ (la machine peut avoir plus que 3 ou 4 qubits; ignorez simplement les qubits en plus en ne faisant aucune opération dessus; notez cependant que vous devrez peut-être inclure l'opération de mesure à ces qubits de trop.

Remarque: la porte $e^{i\theta Z\otimes Z}$ est déjà implémentée sur Qiskit. Cependant, nous préférerons utiliser l'implémentation qui provient de la question **e)** au-dessus.

g) Écrivez un code Qiskit qui implémente le circuit. Vérifiez que

$$p(0, \cdots, 0) = |(\cos\theta)^n + i^n(\sin\theta)^n|^2$$

en lançant le *simulateur* et l'*expérience* sur des vraies machines NISQ pour $\theta \in [0, \pi]$ pour $n = 4, 5, 6, 7, 8, 9, 10$. Choisissez un pas de taille $\delta\theta = \pi/32$ pour les points expérimentaux (c.-à-d., pour chaque n, il y a 32 points expérimentaux). Vous avez le choix

d'utiliser les machines quantiques de votre préférence et comparez les niveaux de bruits (jusqu'à 5 qubits, il y a beaucoup de machines différentes qui sont disponibles, donc essayez avec deux ou trois différentes, mais pour $n \geq 6$, il n'y a actuellement qu'une seule machine disponible.

Motivations et informations complémentaires

Récemment, GOogle a annoncé avoir prouvé la *avantage quantique* (ils l'ont appelé *suprématie quantique*) dans le problème de l'échantillonnage d'une distribution faite à partir des *permanents* d'une classe spéciale de matrices. Ils ont utilisé des machines NISQ avec une cinquantaine de qubits et ont donc échantillonné à partir d'une distribution avec 2^{50} états (classiques). Puisqu'on ne sait pas comment calculer efficacement les permanents avec les méthodes classiques et que l'espace d'état est énorme, c'est un problème non trivial de produire des certificats qui certifient que le dispositif NISQ produit vraiment des échantillons corrects.

Le problème étudié dans ce mini-projet est trop simple pour aborder des questions liées à un avantage quantique. Néanmoins, il s'agit d'une illustration non triviale de la manière dont un dispositif quantique peut en principe être utilisé comme échantillonneur.

Pour plus d'informations, voir l'article de synthèse : "Quantum sampling problems, BosonSampling and quantum supremacy" par A. P. Lund, M. J. Bremner et T. C. Ralph ; npj Quantum Information (2017)3:15 ; www.nature.com/npjqi

Annexe : fonction de partition du modèle d'Ising unidimensionnel

La lecture de cette section n'est pas nécessaire pour le projet et est là pour être complet.

Le modèle d'Ising est un modèle mathématique en physique statistique pour le phénomène du magnétisme. On considère qu'un aimant est constitué par l'ensemble d'un nombre macroscopique de degrés de liberté magnétiques d'atomes. Ces degrés de liberté magnétiques sont grossièrement modélisés par des variables n s_i, $1 \leq i \leq n$, prenant des valeurs dans $\{\pm 1\}$. Ici, nous considérons juste une version simple du modèle où les atomes sont disposés dans un réseau cristallin cyclique "unidimensionnel" (voir Figure **18.3**) et chaque atome interagit avec ses deux voisins (ce modèle a été introduit par Lenz et étudié par Ising dans les années 1920).

Nous montrons ici comment calculer le facteur de normalisation de la distribution (18.4). Ce facteur de normalisation est appelé "fonction de partition" et joue un rôle

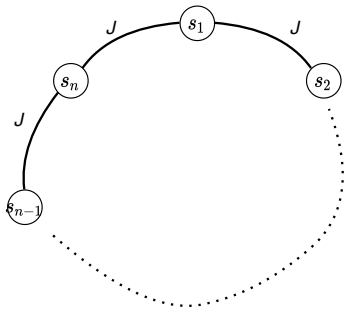


Figure 18.3 Modèle d'Ising avec condition au bord périodique

central dans la théorie. Il est donné par

$$Z(J) = \sum_{s_1,\ldots,s_n \in \{-1,+1\}^n} \exp\Big\{J\Big(\sum_{i=1}^{n-1} s_i s_{i+1} + s_n s_1\Big)\Big\}.$$

Nous utilisons une méthode appelée la méthode *matrice de transfert*. Considérons la matrice T,

$$T = \begin{pmatrix} e^J & e^{-J} \\ e^{-J} & e^J \end{pmatrix} \equiv \begin{pmatrix} T_{++} & T_{+-} \\ T_{+-} & T_{--} \end{pmatrix}$$

Avec un léger changement de notation $s_i = +, -$ et $s_{n+1} = s_1$ la fonction de partition peut être écrite comme

$$Z(J) = \sum_{\underline{s} \in \{\pm\}^n} \prod_{i=1}^{n} T_{s_i, s_{i+1}}$$

On remarque que cela équivaut à

$$Z = tr[T^n]$$

Pour calculer la trace de T^n on calcule d'abord les valeurs propres de T, et on les appelle λ_1 et λ_2. Notez que

$$T = e^J I + e^{-J} X$$

où $I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ est la matrice identité et $X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ une des matrices de Pauli. Les valeurs propres sont donc $\lambda_1 = e^J + e^{-J} = 2\cosh J$ et $\lambda_2 = e^J - e^{-J} = 2\sinh J$. Ainsi

$$Z(J) = trT^n = \lambda_1^n + \lambda_2^n = 2^n((\cosh J)^n + (\sinh J)^n)$$

18.3 Projet 2021

Un jeu quantique... le carré magique de Mermin-Peres

Les carrés magiques sont abondants en mathématiques et impliquent généralement des contraintes satisfaites par toutes les colonnes et lignes d'une grille de nombres. Parfois, leur conception peut devenir ardue, voire impossible.

Prenons l'exemple suivant. Supposons que nous voulions construire un carré magique 3×3 avec des nombres dans $\mathcal{S} = \{-1, 1\}$ tels que leur produit fasse 1 pour chaque ligne et -1 pour chaque colonne.

Question 1: *Soit*

$$M = \left[\begin{array}{c|c|c} 1 & 1 & 1 \\ \hline -1 & 1 & -1 \\ \hline 1 & -1 & -1 \end{array}\right] \tag{18.9}$$

Est-ce que M satisfait les contraintes ci-dessus sur les lignes et les colonnes ? Montrer qu'en fait aucun carré magique 3×3 *n'existe avec les contraintes ci-dessus.*

Astuce : considérez $L_i = \prod_j M_{i,j}$ *et* $C_j = \prod_i M_{i,j}$*. Calculez* $\prod_i L_i$ *et* $\prod_j C_j$*. Que pouvez-vous en conclure ?*

Malgré ce problème, Alice et Bob sont mis au défi par Eve avec le **jeu du carré magique**, qui est joué de la manière suivante:

1. Au début, Alice et Bob peuvent discuter ensemble aussi longtemps qu'ils le souhaitent et planifier une stratégie.
2. Ils sont ensuite isolés dans deux pièces séparées sans possibilité de communication.
3. Eve tire un nombre aléatoire $i \in \{1, 2, 3\}$ uniformément et l'envoie à Alice uniquement. Alice doit remplir la ligne i avec trois nombres a_{i1}, a_{i2}, a_{i3} dont le produit vaut 1, et envoyer secrètement les nombres à Eve. Bob n'a pas accès à ces informations.
4. Eve tire un nombre aléatoire $j \in \{1, 2, 3\}$ uniformément et l'envoie uniquement à Bob. Il doit ensuite remplir la colonne j avec trois nombres b_{1j}, b_{2j}, b_{3j} dont le produit vaut -1, et envoyer secrètement les trois nombres à Eve. Alice n'a pas cette information.
5. La ligne i et la colonne j se coupent au niveau de "l'élément de matrice" ij. Eve déclare qu'Alice et Bob gagnent la partie si $a_{ij} = b_{ij}$, autrement dit si les choix d'Alice et de Bob sont compatibles. Sinon Eve déclare qu'ils perdent la partie.

Question 2 : *Expliquez pourquoi il n'est pas possible pour Alice et Bob de concevoir une stratégie qui gagne toujours la partie. Concevez une stratégie simple telle qu'Alice et Bob gagnent le jeu avec une probabilité maximale (aucune preuve formelle demandée).*

Heureusement, Alice et Bob savent que nous ne vivons pas dans un monde classique.

En effet, ils ont suivi des cours de physique quantique à l'université et ils sont même capables de construire de simples dispositifs quantiques ! Par conséquent, avant d'être isolés, ils préparent une stratégie quantique (cette stratégie satisfait à l'exigence de non-communication après leur séparation).

Ils préparent 2 qubits intriqués au maximum (paires EPR ou Bell) dans l'état :

$$|\psi\rangle_{AB} = \frac{1}{\sqrt{2}}|B_{00}\rangle_{A,B} \otimes \frac{1}{\sqrt{2}}|B_{00}\rangle_{A,B} \tag{18.10}$$

$$= \frac{1}{\sqrt{2}}\left(|00\rangle_{A_1,B_1} + |11\rangle_{A_1,B_1}\right) \otimes \frac{1}{\sqrt{2}}\left(|00\rangle_{A_2,B_2} + |11\rangle_{A_2,B_2}\right) \tag{18.11}$$

Notez qu'ils préparent un réservoir suffisamment grand de tels états pour pouvoir jouer plusieurs tours du jeu (mais pour fixer les idées, nous allons simuler un seul tour dans notre discussion).

Par conséquent, ils partagent quatre qubits au total. Lorsqu'ils sont séparés, Alice emporte les qubits A_1 et A_2, tandis que Bob reçoit B_1 et B_2. Avant d'aller plus loin dans le jeu, ils se sont également mis d'accord sur un ensemble mystérieux de 9 observables et ont rempli le carré magique avec ces observables :

$$Q = \left[\begin{array}{c|c|c} \sigma_x \otimes \sigma_x & \sigma_x \otimes \mathbb{1} & \mathbb{1} \otimes \sigma_x \\ \hline \sigma_y \otimes \sigma_y & -\sigma_x \otimes \sigma_z & -\sigma_z \otimes \sigma_x \\ \hline \sigma_z \otimes \sigma_z & \mathbb{1} \otimes \sigma_z & \sigma_z \otimes \mathbb{1} \end{array}\right] \tag{18.12}$$

Pour aller plus loin, il est utile de garder à l'esprit les aspects suivants du postulat de mesure :

1. Une observable est une grandeur mesurable décrite par une matrice hermitienne.
2. L'appareil de mesure projette l'état (ou fonction d'onde) sur l'un des vecteurs de la base propre $|v\rangle$ de l'observable.
3. La valeur mesurée par l'observable est donnée par la valeur propre associée au vecteur propre.
4. Les mesures simultanées de plusieurs observables ne sont possibles que pour les observables qui commutent puisque dans ce cas, elles ont une base propre commune.
5. La règle de Born indique que $\mathbb{P}\left[|\psi\rangle \to |v\rangle\right] = |\langle v|\psi\rangle|^2$. Si une valeur propre est dégénérée, la probabilité de mesurer cette valeur propre est la somme de ces probabilités sur les vecteurs propres correspondants.

Une remarque utile pour plus tard est la suivante: lorsque les valeurs propres sont dégénérées, les vecteurs propres correspondants (et donc la base propre) ne sont pas uniques.

Question 3 : *Vérifiez par exemple l'observable $Q_{1,2} = \sigma_x \otimes \mathbb{1}$. Quelles sont les valeurs propres possibles de cette observable ? Donnez deux ensembles de vecteurs qui forment une base propre possible.*

La stratégie quantique d'Alice et Bob est la suivante. D'abord ils préparent et

partagent l'état $|\psi\rangle_{AB}$. Après avoir reçu une ligne i, Alice fait la mesure décrite par les observables $Q_{i,1}, Q_{i,2}, Q_{i,3}$, stocke les résultats dans a_{i1}, a_{i2}, a_{i3}, et envoie ces trois nombres à Eve. Bob procède de la même manière, il stocke les résultats d'une mesure simultanée des observables $Q_{1,j}, Q_{2,j}, Q_{3,j}$ dans b_{1j}, b_{2j}, b_{3j}, et envoie les trois nombres à Eve. Il s'avère que c'est toujours une stratégie gagnante. Vous allez être guidé à travers la théorie, puis vous implémenterez le jeu sur les appareils NISQ de IBMQ !

Question 4 : *Vérifiez les propriétés suivantes du carré magique quantique Q :*

1. *Vérifiez que les observables dans les lignes et les colonnes commutent. Par conséquent, les mesures simultanées (par Alice) dans une ligne donnée et les mesures simultanées (par Bob) dans une colonne donnée sont autorisées.*
2. *Vérifiez également que si les observables n'appartiennent pas à la même ligne ou colonne, ils ne commutent pas nécessairement. Notez que la stratégie d'Alice et Bob ne nécessite pas de telles mesures simultanées !*
3. *Calculez les produits $\prod_j Q_{i,j}$ pour chaque j, et les produits $\prod_i Q_{i,j}$ pour chaque i. Qu'observez-vous ? Comparez avec le carré magique classique où nous n'utilisions que des nombres dans $\mathcal{S} = \{-1, +1\}$.*

Afin de mieux comprendre les observations précédentes, rappelons que si deux opérateurs commutent, alors ils peuvent être diagonalisés dans une base commune. Par exemple, dans la base $\mathcal{B}_1^A = \{|{+}{+}\rangle, |{+}{-}\rangle, |{-}{+}\rangle, |{-}{-}\rangle\}$, les opérateurs $Q_{1,1} = \sigma_x \otimes \sigma_x$, $Q_{1,2} = \sigma_x \otimes \mathbb{1}$, $Q_{1,3} = \mathbb{1} \otimes \sigma_x$ peuvent être diagonalisé car ils s'écrivent :

$$Q_{1,1} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \qquad Q_{1,2} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \qquad Q_{1,3} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \tag{18.13}$$

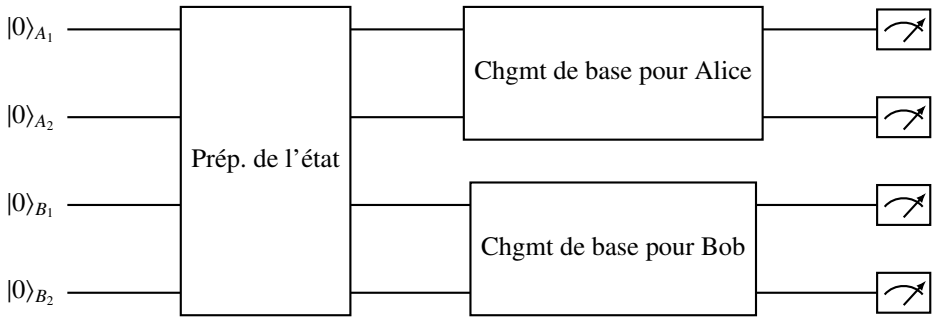
Question 5 : *Pour chacun des 4 résultats possibles dans $\mathcal{B}_1^A$ pour Alice lorsqu'elle obtient la ligne $i = 1$, spécifiez quel résultat est stocké dans $[a_{i1}, a_{i2}, a_{i3}]$, et vérifiez que le produit est égal à 1.*

Question 6 : *Voyons maintenant ce qui se passe du côté de Bob lorsqu'il obtient la colonne $j = 2$. Trouvez la base commune $\mathcal{B}_2^B$ pour les opérateurs $Q_{1,2}, Q_{2,2}, Q_{3,2}$, et pour chacun des 4 résultats possibles dans $\mathcal{B}_2^B$, expliquez quel résultat est stocké dans $[b_{1j}, b_{2j}, b_{3j}]$. Vérifiez également le produit $b_{1j}b_{2j}b_{3j}$.*

Question 7 : *Continuons à supposer qu'Alice a obtenu $i = 1$ et que Bob a obtenu $j = 2$. Pour gagner le jeu, ils doivent avoir mesuré $a_{12} = b_{12}$. Calculer la probabilité $P(a_{12} = b_{12}) = P(a_{12} = 1, b_{12} = 1) + P(a_{12} = -1, b_{12} = -1)$ en utilisant comme entrée la fonction d'onde $|\psi\rangle_{A,B}$ mentionnée précédemment. Que concluez-vous ?*

Expérience

Question 8 : *Afin de tester ces résultats expérimentalement, vous utiliserez un appareil NISQ de IBMQ et répliquerez les résultats possibles pour le cas spécifique $i = 1$ et $j = 2$. Concevez le circuit quantique approprié. Notez qu'IBMQ vous permet uniquement d'effectuer des mesures dans la base de calcul $\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$ vous devrez donc trouver un moyen de contourner ce problème en utilisant un changement de base. Votre circuit doit avoir la forme suivante :*



Maintenant, exécutez votre circuit sur la machine "simulateur" IBM. Notez les sorties et discutez les résultats : quelle sortie correspond à quels résultats a_{i1}, a_{i2}, a_{i3} et b_{1j}, b_{2j}, b_{3j} ? Alice et Bob respectent-ils les règles du jeu ? Gagnent-ils le jeu tout le temps ?

Question 9 : *Si vous êtes satisfait de votre circuit quantique, vous pouvez maintenant le faire fonctionner sur une vraie machine quantique ! (Selon la disponibilité des ressources, la file d'attente de lancement peut prendre jusqu'à quelques minutes). Observez-vous une différence avec les résultats de la question 8 ? Pourquoi?*

Question 10 : *Supposons maintenant qu'Alice reçoive $i = 3$ et Bob $j = 1$.*

1. *Quelle est la base commune $\mathcal{B}_3^A$ pour les observables d'Alice ? Quelle est la base commune $\mathcal{B}_1^B$ pour les observables de Bob ?*
2. *Proposer un circuit quantique comme à la question 8, exécutez-le sur "simulateur" et notez les sorties avec les résultats correspondants a_{i1}, a_{i2}, a_{i3} et b_{1j}, b_{2j}, b_{3j}. Exécutez le circuit sur une vraie machine quantique.*

Question 11 : *Supposons maintenant qu'Alice reçoive $i = 2$ et Bob $j = 3$.*

1. *Quelle est la base commune $\mathcal{B}_2^A$ pour les observables d'Alice ? Quelle est la base commune $\mathcal{B}_3^B$ pour les observables de Bob ?*
2. *Proposer un circuit quantique comme à la question 8, exécutez-le sur "simulateur" et notez les sorties avec les résultats correspondants a_{i1}, a_{i2}, a_{i3} et b_{1j}, b_{2j}, b_{3j}. Exécutez le circuit sur une vraie machine quantique.*

Question bonus. *Écrivez un code Qiskit qui traite toutes les possibilités de lignes et de colonnes pour Alice et Bob (il y a donc* 9 *circuits possibles). L'entrée doit être la ligne* i *et la colonne* j*. La sortie devrait être un histogramme avec les réponses d'Alice et de Bob. Exécutez-le sur le simulateur, puis sur une vraie machine quantique.*

18.4 Projet 2022

Grover et le problème 3-SAT

Le problème 3-SAT est une instance bien connue de la classe de problèmes NP-complets. Elle consiste à disposer d'une liste de bits classiques $b_1, \ldots, b_n$ et à trouver la solution d'un prédicat donné sous forme de conjonctions de trois-disjonctions. Par exemple, voici un prédicat 3-SAT :

$$f(x, y, z) = (x \vee y \vee \neg z) \wedge (\neg x \vee \neg y \vee z) \wedge (x \vee \neg y \vee z) \tag{18.14}$$

On cherche des solutions de $f(x, y, z) = 1$. On peut vérifier avec une recherche exhaustive pour cet exemple que les solutions sont 000, 100, 011, 101, 111. L'utilisation de l'algorithme de Grover peut aider à trouver des solutions d'un prédicat avec une accélération quadratique sur une recherche exhaustive naïve. Nous allons donc l'implémenter dans ce mini-projet.

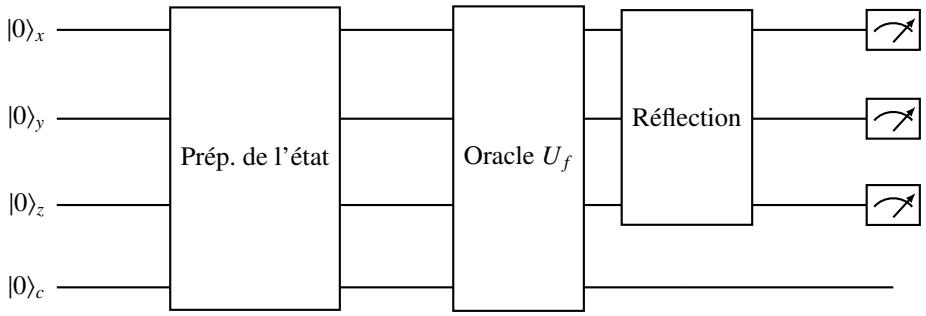
Question 1 *Nous considérons le prédicat 3-SAT f défini comme :*

$$f(xyz) = (\neg x \vee \neg y \vee \neg z) \wedge (\neg x \vee \neg y \vee z) \wedge (\neg x \vee y \vee z) \wedge (x \vee \neg y \vee z) \wedge (x \vee y \vee \neg z) \wedge (x \vee y \vee z) \tag{18.15}$$

Trouvez toutes les solutions possibles.

Dans les questions suivantes, vous implémenterez des circuits de style Grover pour trouver des solutions. Combien de fois devrez-vous appliquer l'opérateur Grover et pourquoi ?

Nous allons maintenant concevoir un circuit pour résoudre le problème ci-dessus. dont la forme générale est:



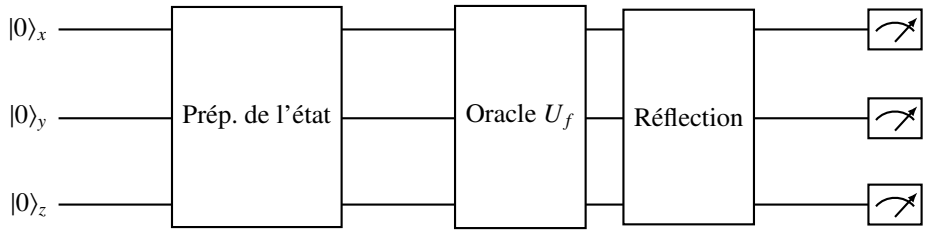
Question 2 *Concevez d'abord la porte oracle U_f qui doit être telle que $U_f|xyzc\rangle = |xyz\rangle \otimes |c \oplus f(xyz)\rangle$*

Question 3 *Concevez l'opérateur de réflexion et montrez la preuve analytique de votre conception. Vous devrez peut-être trouver une extension appropriée de l'opérateur*

de réflexion vu dans le cours.

Question 4 *Exécutez votre circuit dans IBMQ. Pouvez-vous récupérer les solutions? Y'a-t-il du bruit ?*

Supposons que nous ne pouvons utiliser que 3 qubits sur les machines quantiques. Le circuit peut en réalité être simplifié sous la forme suivante :



Question 5 *Trouvez un bloc U_f tel que $U_f|xyz\rangle = (-1)^{f(xyz)}|xyz\rangle$ et relancez votre circuit sur l'IBMQ Machines. Comparez vos résultats avec la question précédente.*

Astuce : dans les circuits ci-dessus, vous pouvez utiliser la porte de Toffoli.

Notes