Introduction à l'information quantique

Bellac Le

Prétirage INLN 2003/08

INTRODUCTION A L'INFORMATION QUANTIQUE

Michel Le Bellac

Cours donné à l'Ecole Supérieure de Sciences Informatiques (ESSI)

Octobre 2003

Résumé. Ce cours a pour objectif d'exposer à un public de non physiciens les notions de physique quantique nécessaires pour comprendre l'information quantique et d'illustrer le calcul quantique en prenant comme exemple de l'algorithme de factorisation de Shor.

INSTITUT NON LINEAIRE DE NICE UMR 6638
1361 routes des Lucioles 06560 Valbonne
e-mail : michel.le_bellac@inln.cnrs.fr

Table des matières

Contents

Chapitre 1

Qu'est-ce qu'un qu-bit ?

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

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

La grande nouveauté, depuis le début des année 1980, est la possibilité pour les physiciens de *manipuler et d'observer des objets quantiques élémentaires individuels* : photons, atomes, ions etc., et pas seulement d'agir sur le comportement quantique collectif d'un grand nombre de tels objets. C'est cette possibilité de manipuler et d'observer des objets quantiques élémentaires qui est à l'origine de l'information quantique, où ces objets quantiques élémentaires permettront de construire physiquement les qu-bits. Cela dit, aucun concept fondamentalement nouveau n'a été introduit depuis les années 1930, et les pères fondateurs de la mécanique quantique (Heisenberg, Schrödinger, Dirac ...), s'ils ressuscitaient aujourd'hui, ne seraient pas surpris par l'informatique quantique, même s'ils seraient sûrement éblouis les prouesses des expérimentateurs, qui réalisent aujourd'hui des expériences qualifiées à l'époque de "gedanken experiment".

Il vaut aussi la peine de signaler que la miniaturisation croissante de l'électronique va trouver ses limites en raison des effets quantiques, qui vont devenir incontournables en dessous du nanomètre. Ainsi on estime que le loi de Moore pourrait ne plus être valable d'ici dix à quinze ans.

Quelques références

- Le livre de base est celui de Michael Nielsen et Isaac Chuang *Quantum Computation and Quantum Information*, Cambridge University Press, Cambridge (2000).
- Également intéressant est le cours de John Preskill que l'on trouve (avec des exercices) sur le site http ://www.theory.caltech.edu/˜ preskill/

Ce cours se place à un niveau plus avancé que celui de Nielsen et Chuang.
- Voir aussi le cours de David Mermin
http ://www.ccmr.cornell.edu/˜ mermin/qccomp/CS483.html
- Comme introduction sans équations à la physique quantique, je recommande le livre de Valerio Scarani, *Introduction à la physique quantique*, Vuibert (2003).
- Pour un texte plus complet, voir par exemple mon livre *Physique quantique*, EDPSciences/Éditions du CNRS, (2003).

1.1 Polarisation de la lumière

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

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

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

où ω est la fréquence de la vibration, $\omega = ck$, c étant la vitesse de propagation. Dans le plan $z = 0$

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

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

$$\begin{aligned} E_x &= E_0 \cos\theta \, \cos\omega t \\ E_y &= E_0 \sin\theta \, \cos\omega t \end{aligned} \tag{1.1}$$

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

$$\hat{p} = (\cos\theta, \sin\theta) \qquad \vec{E} = E_0 \hat{p} \, \cos\omega t \tag{1.2}$$

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

Pour étudier de façon quantitative la polarisation, nous allons nous servir d'un *ensemble polariseur/ analyseur*. Nous faisons d'abord passer la lumière dans un polariseur dont l'axe fait un angle θ avec l'axe Ox, puis dans un second polariseur, appelé analyseur, dont l'axe fait un angle α avec l'axe Ox (figure 1.1), avec

$$\hat{n} = (\cos\alpha, \sin\alpha) \tag{1.3}$$

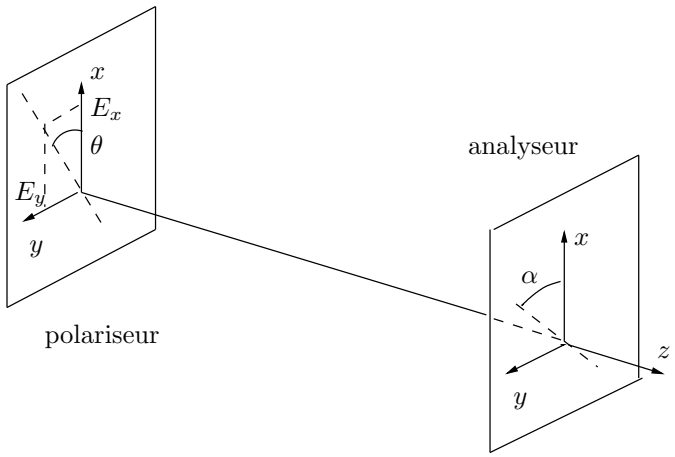

FIG. 1.1 – Ensemble polariseur-analyseur.

À la sortie de l'analyseur, le champ électrique $\vec{E}'$ s'obtient en projetant le champ (1.1) sur $\hat{n}$

$$\begin{aligned} \vec{E}' &= (\vec{E}\cdot\hat{n})\hat{n} = E_0\cos\omega t\,(\hat{p}\cdot\hat{n})\hat{n} \\ &= E_0\cos\omega t\,(\cos\theta\cos\alpha+\sin\theta\sin\alpha)\,\hat{n} \\ &= E_0\cos\omega t\,\cos(\theta-\alpha)\hat{n} \end{aligned} \tag{1.4}$$

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

$$I' = I\cos^2(\theta-\alpha) \tag{1.5}$$

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

$$\begin{aligned} E_x &= \frac{E_0}{\sqrt{2}}\,\cos\omega t \\ E_y &= \frac{E_0}{\sqrt{2}}\,\cos\left(\omega t-\frac{\pi}{2}\right) = \frac{E_0}{\sqrt{2}}\,\sin\omega t \end{aligned} \tag{1.6}$$

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

$$\begin{aligned} E_x &= E_0\cos\theta\cos(\omega t-\delta_x) = E_0\,\mathrm{Re}\left[\cos\theta\,\mathrm{e}^{-\mathrm{i}(\omega t-\delta_x)}\right] \\ E_x &= E_0\sin\theta\cos(\omega t-\delta_x) = E_0\,\mathrm{Re}\left[\sin\theta\,\mathrm{e}^{-\mathrm{i}(\omega t-\delta_y)}\right] \end{aligned} \tag{1.7}$$

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

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

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

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

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

Remarques

- Une lame biréfringente (figure 1.2) permet de séparer deux états de polarisation orthogonaux, tandis qu'un polaroïd absorbe une des deux polarisations en laissant passer la polarisation orthogonale.
- Considérons un ensemble analyseur/polariseur croisés, par exemple le polariseur suivant Ox et l'analyseur suivant Oy. Aucune lumière n'est transmise. Mais si on introduit un polariseur intermédiaire dont l'axe fait un angle θ avec Ox, alors une partie de la lumière est rétablie : une première projection donne un facteur $\cos\theta$ et une seconde un facteur $\sin\theta$, d'où l'intensité à la sortie de l'analyseur

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

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

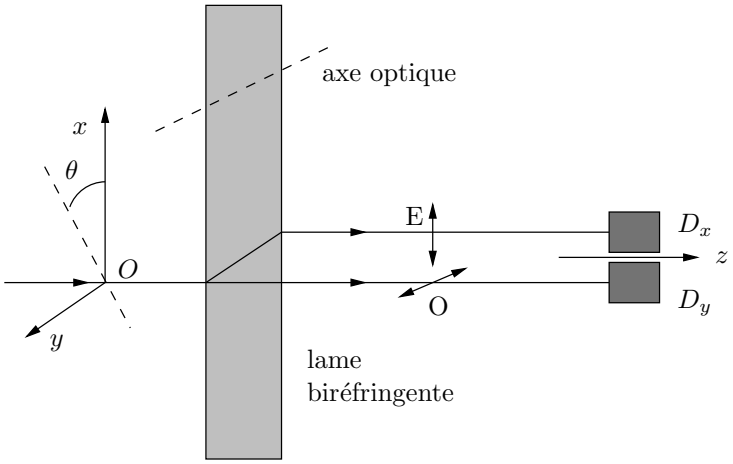

FIG. 1.2 – Décomposition de la polarisation par une lame biréfringente. Le rayon ordinaire O est polarisé horizontalement, le rayon extraordinaire E est polarisé verticalement.

1.2 Polarisation d'un photon

Depuis Einstein (1905), on sait que la lumière est composée de photons, ou particules de lumière. Si l'on réduit suffisamment l'intensité lumineuse, on devrait pouvoir étudier la polarisation des photons individuels, que l'on sait parfaitement détecter à l'aide de photomultiplicateurs. Supposons que l'expérience détecte N photons. Lorsque $N \to \infty$, on doit retrouver les résultats de l'optique ondulatoire que nous venons d'énoncer. Effectuons par exemple l'expérience suivante (figure 1.2) : une lame biréfringente sépare un faisceau lumineux dont la polarisation fait un angle θ avec Ox en un faisceau polarisé suivant Ox et un faisceau polarisé suivant Oy, les intensités étant respectivement $I\cos^2\theta$ et $I\sin^2\theta$. Réduisons l'intensité de telle sorte que les photons arrivent un à un, et plaçons deux photodétecteurs D_x et D_y derrière la lame. L'expérience montre D_x et D_y ne cliquent jamais simultanément[1] : un photon arrive entier soit sur D_x, soit sur D_y, un photon ne se divise pas. D'autre part l'expérience montre que la probabilité p_x (p_y) de détection d'un photon par D_x (D_y) est de $\cos^2\theta$ ($\sin^2\theta$). Si l'expérience détecte N photons, on aura donc N_x (N_y) photons détectés par D_x (D_y)

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

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

- **Premier problème.** Peut-on prévoir, pour un photon donné, s'il va déclencher D_x ou D_y ? La réponse de la théorie quantique est NON, énoncé qui a profondément choqué Einstein (Dieu ne

[1] Sauf cas de "dark count", où un compteur se déclenche spontanément.

joue pas aux dés !). Certains physiciens (dont Einstein) ont été tentés de supposer que la théorie quantique était incomplète, et qu'il y avait des "variables cachées" dont la connaissance permettrait de prévoir le sort individuel de chaque photon. Moyennant des hypothèses très raisonnables sur lesquelles je reviendrai au chapitre 3, on sait aujourd'hui que de telles variables cachées sont exclues. Les probabilités de la théorie quantique sont *intrinsèques*, elles ne sont pas liées à une connaissance imparfaite de la situation physique, comme c'est le cas par exemple dans le jeu de pile ou face.

- **Deuxième problème.** Recombinons[2] les deux faisceaux de la première lame biréfringente, en utilisant une seconde lame symétrique de la première (figure 1.3). Cherchons la probabilité qu'un photon traverse l'analyseur. Un photon peut choisir le trajet x avec une probabilité $\cos^2\theta$; il a ensuite une probabilité $\cos^2\alpha$ de traverser l'analyseur, soit une probabilité totale $\cos^2\theta\cos^2\alpha$. S'il choisit le trajet y, il aura une probabilité $\sin^2\theta\sin^2\alpha$ de traverser l'analyseur. La probabilité totale s'obtient en additionnant les probabilités des deux trajets possibles

$$\mathsf{p}'_{\rm tot} = \cos^2\theta\cos^2\alpha + \sin^2\theta\sin^2\alpha \tag{1.8}$$

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

$$\mathsf{p}_{\rm tot} = \cos^2(\theta-\alpha) \tag{1.9}$$

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

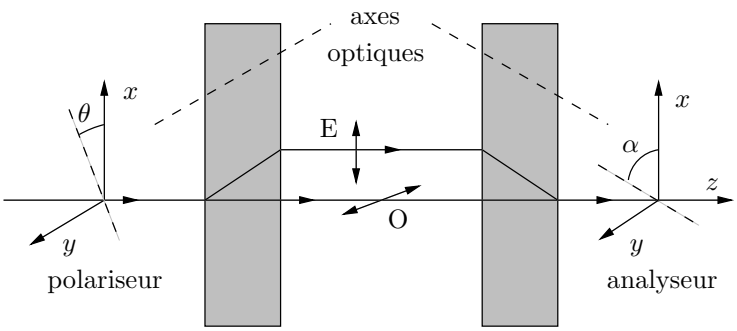

FIG. 1.3 – Décomposition et recombinaison de polarisations à l'aide de lames biréfringentes.

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

$$\begin{aligned} a(\theta\to x) &= \cos\theta \qquad & a(x\to\alpha) &= \cos\alpha \\ a(\theta\to y) &= \sin\theta \qquad & a(y\to\alpha) &= \sin\alpha \end{aligned}$$

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

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

ce qui redonne bien (1.9)

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

Supposons que l'on ait un moyen de savoir si le photon emprunte le trajet x ou le trajet y (impossible dans notre cas, mais des expériences analogues répondant à la question "Quel trajet ?" ont été réalisées avec des atomes). On pourrait alors diviser les photons en deux classes, ceux qui ont choisi le trajet x et ceux qui ont choisi le trajet y. Pour les photons ayant choisi le trajet x, on pourrait bloquer le trajet y par un cache sans rien changer, et inversement pour les photons ayant choisi le trajet y on pourrait bloquer le trajet x. Bien évidemment, le résultat ne peut être alors que (1.8). Si on arrive à discriminer entre les trajets, le résultat n'est plus le même, les trajets ne sont plus indiscernables. Dans les conditions expérimentales où il est impossible en principe de distinguer entre les trajets, on peut dire, au choix :

[2] Toutefois il faut prendre quelques précautions : voir *Physique quantique*, exercice 3.1.

- soit que le photon emprunte les deux trajets à la fois ;
- soit (ce qui a ma préférence) que cela n'a pas de sens de poser la question "Quel trajet ?", puisque les conditions expérimentales ne permettent pas d'y répondre, et, suivant Asher Peres "Unperformed experiments have no results !".

Il faut noter que si l'expérience permet de décider entre les deux trajets, le résultat est (1.8), même si l'on décide de ne pas les observer. Il suffit que les conditions expérimentales permettent, *en principe*, de distinguer entre les deux trajets.

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

1.3 Formulation mathématique : le qu-bit

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

$$yyxyxyyyx\cdots$$

Bob analyse la polarisation de ces photons à l'aide d'une lame biréfringente comme dans la figure 1.2 et en déduit le message d'Alice

$$001010001\cdots$$

Ce n'est évidemment pas une façon très efficace d'échanger des messages, mais c'est à la base de la cryptographie quantique. Cependant la question intéressante est maintenant : quelle est la valeur du bit que l'on peut attribuer par exemple à un photon polarisé à 45^o ? Suivant les résultats de la section précédente, un photon polarisé à 45^o est une *superposition linéaire* d'un photon polarisé suivant Ox et d'un photon polarisé suivant Oy. Un qu-bit est donc une entité beaucoup plus riche qu'un bit ordinaire, qui ne peut prendre que les valeurs 0 et 1. En un certain sens, un qu-bit peut prendre toutes les valeurs intermédiaires entre 0 et 1 et contiendrait donc une quantité infinie d'information ! Cependant cet énoncé optimiste est immédiatement démenti lorsque l'on se rend compte que la mesure du qu-bit ne peut donner que le résultat 0 ou 1, quelle que soit la base choisie. Malgré tout on peut se poser la question de cette "information cachée" dans la superposition linéaire, et nous verrons au chapitre 4 qu'on peut l'exploiter sous certaines conditions.

Afin de rendre compte de la possibilité des superpositions linéaires, il est naturel d'introduire pour la description mathématique de la polarisation un espace vectoriel à deux dimensions $\mathcal{H}$. À tout état de polarisation on va faire correspondre un vecteur de cet espace vectoriel. On peut par exemple choisir pour vecteurs de base de $\mathcal{H}$ les vecteurs $|x\rangle$ et $|y\rangle$ correspondant aux polarisations linéaires suivant Ox et Oy. Tout état de polarisation pourra se décomposer suivant cette base[3]

$$|\Phi\rangle = \lambda|x\rangle + \mu|y\rangle \tag{1.10}$$

Une polarisation linéaire sera décrite par des coefficients λ et μ réels, mais la description d'une polarisation circulaire (1.6) ou elliptique (1.7) exige de faire appel à des coefficients λ et μ complexes : l'espace $\mathcal{H}$ est donc un *espace vectoriel complexe*.

Les amplitudes de probabilité vont correspondre à un produit scalaire sur cet espace. Soit deux vecteurs $|\Phi\rangle$ (1.10) et $|\Psi\rangle$

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

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

$$\langle\Psi|\Phi\rangle = \nu^*\lambda + \sigma^*\mu = \langle\Phi|\Psi\rangle^* \tag{1.11}$$

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

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

$$||\Phi||^2 = \langle\Phi|\Phi\rangle = |\lambda|^2 + |\mu|^2 \tag{1.12}$$

Notez que les vecteurs $|x\rangle$ et $|y\rangle$ sont orthogonaux par rapport au produit scalaire (1.11) et qu'ils sont de norme unité

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

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

$$||\Phi||^2 = |\lambda|^2 + |\mu|^2 = 1 \tag{1.13}$$

Les états de polarisation seront donc représentés mathématiquement par des vecteurs unitaires (de norme unité) de l'espace $\mathcal{H}$. Un espace vectoriel muni d'un produit scalaire défini positif est appelé un *espace de Hilbert*, et $\mathcal{H}$ est *l'espace de Hilbert des états de polarisation.*

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

$$|\theta\rangle = \cos\theta\,|x\rangle + \sin\theta\,|y\rangle \tag{1.14}$$

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

$$a(\theta \to \alpha) = \cos(\theta - \alpha) = \langle\alpha|\theta\rangle \tag{1.15}$$

Elle est donc donnée par le produit scalaire des vecteurs $|\alpha\rangle$ et $|\theta\rangle$, et la probabilité de traverser l'analyseur est donnée par le module carré de cette amplitude (voir (1.9))

$$\mathsf{p}(\theta \to \alpha) = \cos^2(\theta - \alpha) = |\langle\alpha|\theta\rangle|^2 \tag{1.16}$$

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

$$a(\Phi \to \Psi) = \langle\Psi|\Phi\rangle \tag{1.17}$$

et la probabilité correspondante sera

$$\mathsf{p}(\Phi \to \Psi) = |a(\Phi \to \Psi)|^2 = |\langle\Psi|\Phi\rangle|^2 \tag{1.18}$$

N.B. En fait un vecteur d'état n'est défini qu'à une phase multiplicative près

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

car remplacer $|\Phi\rangle$ par

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

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

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

$$\begin{aligned} |\theta\rangle &= \cos\theta\,|x\rangle + \sin\theta\,|y\rangle \\ |\theta_\perp\rangle &= -\sin\theta\,|x\rangle + \cos\theta\,|y\rangle \end{aligned} \tag{1.19}$$

Si le polariseur prépare par exemple le photon dans l'état $|\theta\rangle$ et que l'analyseur est orienté suivant Ox, la probabilité de réussite du test est $\cos^2\theta$. Deux remarques sont essentielles

- Après le passage dans l'analyseur, l'état de polarisation du photon n'est plus $|\theta\rangle$, mais $|x\rangle$. La mesure modifie l'état de polarisation.
- Si le photon est polarisé elliptiquement, et non linéairement

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

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

On constate donc une différence de principe entre la mesure en physique classique et la mesure en physique quantique. En physique classique, *la quantité physique à mesurer préexiste à la mesure* : si un radar mesure la vitesse de votre voiture à 180 km/h sur l'autoroute, cette vitesse préexistait à sa mesure par le gendarme (ce qui lui donne la légitimité pour verbaliser). Au contraire, dans la mesure de la polarisation d'un photon $|\theta\rangle$ par un analyseur orienté suivant Ox, le fait que le test donne une polarisation suivant Ox ne permet pas de conclure que le photon testé avait au préalable sa polarisation suivant Ox. Si l'on reprend l'analogie de la voiture, on pourrait imaginer que comme dans (1.19) la voiture soit dans un état de superposition linéaire de deux vitesses[4], par exemple

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

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

1.4 Principes de la mécanique quantique

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

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

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

et la probabilité pour Φ de réussir le test Ψ est

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

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

Après le test, le système quantique est dans l'état $|\Psi\rangle$, ce qui veut dire du point de vue mathématique que l'on réalise une projection orthogonale sur $|\Psi\rangle$. Soit $\mathcal{P}_\Psi$ ce projecteur. Comme[5]

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

on peut écrire ce projecteur sous la forme très commode

$$\mathcal{P}_\Psi = |\Psi\rangle\langle\Psi| \tag{1.20}$$

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

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

La projection du vecteur d'état est appelée dans l'interprétation de Copenhague de la mécanique quantique "réduction du vecteur d'état", ou, pour des raisons historiques, "réduction du paquet d'ondes". Cette réduction du vecteur d'état est une fiction commode de l'interprétation de Copenhague, qui évite d'avoir à se poser des questions sur le processus de mesure, et elle est souvent ajoutée comme principe de base supplémentaire. Cependant on peut parfaitement se passer de ce principe si on prend en compte le processus de mesure. Un exemple en sera donné dans la section 4.4.

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

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

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

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

relation dite *relation de fermeture*, qui se généralise à une base orthonormée d'un espace de Hilbert $\mathcal{H}$ de dimension N

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

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

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

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

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

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

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

Les tests $|x\rangle$ et $|\theta\rangle$ sont dits *incompatibles*.

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

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

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

$$\mathsf{p}(\Psi) = |\langle\Psi|\Phi\rangle|^2 = \langle\Phi|\Psi\rangle\langle\Psi|\Phi\rangle = \langle\Phi(|\Psi\rangle\langle\Psi|\Phi\rangle) = \langle\Phi|\mathcal{P}_\Psi\Phi\rangle \tag{1.21}$$

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

$$\langle\Phi|M\Phi\rangle \equiv \langle M\rangle_\Phi$$

Au test $\mathcal{T} = |\Psi\rangle$ on peut donc associer le projecteur $\mathcal{P}_\Psi$ dont la valeur moyenne dans l'état $|\Phi\rangle$ donne suivant (1.21) la probabilité de réussite du test.

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

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

qui vérifie bien

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

La valeur moyenne de M est par définition

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

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

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

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

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

$$M = \sum_n a_n \mathcal{P}_n \tag{1.22}$$

$$\mathcal{P}_n\mathcal{P}_m = \mathcal{P}_n\delta_{mn} \qquad \sum_n \mathcal{P}_n = I \tag{1.23}$$

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

1.5 Générateur quantique de nombres aléatoires

L'utilisation des propriétés quantiques permet de réaliser expérimentalement des générateurs de nombres aléatoires, et non pseudo-aléatoires, ce qui est essentiel pour la cryptographie quantique, comme on le verra dans la section suivante. Un des dispositifs les plus simples utilise une lame semi-transparente, ou séparateur de faisceau. Si un rayon lumineux tombe sur une lame semi-transparente, une partie de la lumière est transmise et une partie est réfléchie. On peut s'arranger que ceci se fasse dans des proportions de 50%/50%. Si maintenant on diminue l'intensité de sorte que les photons arrivent un à un sur la lame, on constate que ces photons peuvent être, soit réfléchis et détectés par D_1, soit transmis et détectés par D_2 (figure 1.4). Il n'y a aucune corrrélation entre les détections, et on a un véritable jeu de pile ou face non biaisé.

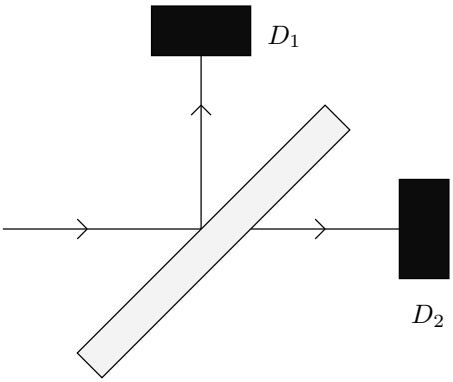

FIG. 1.4 – Lame semi-transparente et détection de photons.

Un prototype a été réalisé suivant ce principe par le groupe d'optique quantique de Genève. Il fournit des nombres aléatoires au taux de 10^5 nombres par seconde et l'absence de biais (ou de corrélations entre nombres supposés aléatoires) a été testée par des programmes standard.

1.6 Cryptographie quantique

La cryptographie quantique est une invention récente fondée sur l'incompatibilité de deux bases différentes d'états de polarisation linéaire. La cryptographie usuelle repose sur une clé de chiffrage connue seulement de l'expéditeur et du destinataire. Ce système est appelé à *clé secrète*. Il est en principe très sûr[6], mais il faut que l'expéditeur et le destinataire aient le moyen de se transmettre la clé sans que celle-ci soit interceptée par un espion. Or la clé doit être changée fréquemment, car une suite de messages codés avec la même clé est susceptible de révéler des régularités permettant le déchiffrage du message par une tierce personne. Le processus de transmission d'une clé secrète est un processus à risque, et c'est pour cette raison que l'on préfère maintenant les systèmes fondés sur un principe différent, dits systèmes à *clé publique*, où la clé est diffusée publiquement, par exemple sur Internet. Un système à clé publique courant[7] est fondé sur la difficulté de décomposer un nombre très grand N en facteurs premiers, alors que l'opération inverse est immédiate : sans calculette on obtiendra en quelques secondes $137 \times 53 = 7261$, mais étant donné 7261, cela prendra un certain temps à le décomposer en facteurs premiers. Avec les meilleurs algorithmes actuels, le temps de calcul sur ordinateur nécessaire pour décomposer un nombre N en facteurs premiers croît avec N comme $\exp[(\ln N)^{1/3}(\ln \ln N)^{2/3}]$. Il faut aujourd'hui quelques mois à une grappe de PC pour factoriser un nombre de 150 chiffres. Dans le système de chiffrage à clé publique, le destinataire, appelé conventionnellement Bob, diffuse publiquement à l'expéditeur, appelé conventionnellement Alice, un nombre très grand $N = pq$ produit de deux nombres premiers p et q, ainsi qu'un autre nombre c (voir l'annexe 1.6.1). Ces deux nombres N et c suffisent à Alice pour chiffrer le message, mais il faut disposer des nombres p et q pour le déchiffrer. Bien sûr un espion (appelé par convention Ève) disposant d'un ordinateur suffisamment puissant finira par casser le code, mais on peut en général se contenter de conserver secret le contenu du message pendant un temps limité. Cependant, on ne peut pas exclure que l'on dispose un jour d'algorithmes très performants pour décomposer un nombre en facteurs premiers, et de plus, si des ordinateurs quantiques voient le jour, aucun nombre ne pourra leur résister. Heureusement la mécanique quantique vient à point nommé pour contrecarrer les efforts des espions !

"Cryptographie quantique" est une expression médiatique, mais quelque peu trompeuse : en effet, il ne s'agit pas de chiffrer un message à l'aide de la physique quantique, mais d'utiliser celle-ci pour s'assurer que la transmission de la clé n'a pas été espionnée. Comme nous l'avons déjà expliqué, la transmission d'un message, chiffré ou non, peut se faire en utilisant les deux états de polarisation linéaire orthogonaux d'un photon, par exemple $|x\rangle$ et $|y\rangle$. On peut décider d'attribuer par convention la valeur 1 à la polarisation $|x\rangle$ et la valeur 0 à la polarisation $|y\rangle$: chaque photon transporte donc un bit d'information. Tout message, chiffré ou non, peut être écrit en langage binaire, comme une suite de 0 et de 1, et le message 1001110 sera codé par Alice grâce à la séquence de photons $xyyxxxy$, qu'elle expédiera à Bob par exemple par une fibre optique. À l'aide d'une lame biréfringente, Bob sépare les photons de polarisation verticale et horizontale comme dans la figure 1.2, et deux détecteurs placés derrière la lame lui permettent de décider si le photon était polarisé horizontalement ou verticalement : il peut donc reconstituer le message. S'il s'agissait d'un message ordinaire, il y aurait bien sûr des façons bien plus simples et efficaces de le transmettre ! Remarquons simplement que si Ève s'installe sur la fibre, détecte les photons et renvoie à Bob des photons de polarisation identique à ceux expédiés par Alice, Bob ne peut pas savoir que la ligne a été espionnée. Il en serait de même pour tout dispositif fonctionnant de façon classique (c'est-à-dire sans utiliser le principe de superposition) : si l'espion prend suffisamment de précautions, il est indétectable.

C'est ici que la mécanique quantique et le principe de superposition viennent au secours d'Alice et de Bob, en leur permettant de s'assurer que leur message n'a pas été intercepté. Ce message n'a pas besoin d'être long (le système de transmission par la polarisation est très peu performant). Il s'agira en général de transmettre une clé permettant de chiffrer un message ultérieur, clé qui pourra être remplacée

[6] Un chiffrage absolument sûr a été découvert par Vernam en 1935. Cependant la sécurité absolue suppose que la clé soit aussi longue que le message et ne soit utilisée qu'une seule fois !

[7] Appelé chiffrage RSA, découvert par Rivest, Shamir et Adleman en 1977.

à la demande. Alice envoie vers Bob quatre types de photons : polarisés suivant Ox : ↑ et Oy : → comme précédemment, et polarisés suivant des axes inclinés à $\pm 45^o$ Ox' : ↖ et Oy' : ↗, correspondant respectivement aux valeurs 1 et 0 des bits. De même Bob analyse les photons envoyés par Alice à l'aide d'analyseurs pouvant prendre quatre directions, verticale/horizontale, et $\pm 45^o$. Une possibilité serait d'utiliser un cristal biréfringent orienté aléatoirement soit verticalement, soit à 45^o de la verticale et de détecter les photons sortant de ce cristal comme dans la figure 1.3. Cependant, au lieu de faire tourner l'ensemble cristal+détecteurs, on utilise plutôt une cellule de Pockels, qui permet de transformer une polarisation donnée en une polarisation orientée de façon arbitraire et de maintenir fixe l'ensemble cristal+détecteur. La figure 1.5 donne un exemple : Bob enregistre 1 si le photon est polarisé ↑ ou ↖, 0 s'il est polarisé → ou ↗. Après enregistrement d'un nombre suffisant de photons, Bob annonce publiquement la suite des analyseurs qu'il a utilisés, mais non ses résultats. Alice compare sa séquence de polariseurs à celle de Bob et lui donne toujours publiquement la liste des polariseurs compatibles avec ses analyseurs. Les bits qui correspondent à des analyseurs et des polariseurs incompatibles sont rejetés (−), et, pour les bits restants, Alice et Bob sont certains que leurs valeurs sont les mêmes : ce sont les bits qui serviront à composer la clé, et ils sont connus seulement de Bob et Alice, car l'extérieur ne connaît que la liste des orientations, pas les résultats ! Le protocole décrit ci-dessus est appelé BB84, du nom de ses inventeurs Bennett et Brassard.

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

FIG. 1.5 – Cryptographie quantique : transmission de photons polarisés entre Bob et Alice.

Il reste à s'assurer que le message n'a pas été intercepté et que la clé qu'il contenait peut être utilisée sans risque. Alice et Bob choisissent au hasard un sous-ensemble de leur clé et le comparent publiquement. La conséquence de l'interception de photons par Ève serait une réduction de la corrélation entre les valeurs de leurs bits : supposons par exemple qu'Alice envoie un photon polarisé suivant Ox. Si Ève l'intercepte avec un polariseur orienté suivant Ox', et que le photon est transmis par son analyseur, elle ne sait pas que ce photon était initialement polarisé suivant Ox ; elle renvoie donc à Bob un photon polarisé dans la direction Ox', et dans 50% des cas Bob ne va pas trouver le bon résultat. Comme Ève a une chance sur deux d'orienter son analyseur dans la bonne direction, Alice et Bob vont enregistrer une différence dans 25% des cas et en conclure que le message a été intercepté. Cette discussion est bien sûr simplifiée : elle ne tient pas compte des possibilités d'erreurs qu'il faut corriger, et d'autre part il faut réaliser des impulsions à un seul photon et non des paquets d'états cohérents qui ne seraient pas inviolables.[8]. Néanmoins la méthode est correcte dans son principe et un prototype a été réalisé récemment pour des transmissions dans l'air sur plusieurs kilomètres. Il est difficile avec une fibre optique de contrôler la direction de la polarisation sur de longues distances, et c'est pourqui on utilise un support physique différent pour mettre en oeuvre le protocole BB84 avec des fibres. Dans ces conditions la transmission a pu être effectuée sur une centaine de kilomètres.

Annexe 1.6.1 : le codage RSA. Bob choisit deux nombres premiers p et q, $N = pq$, et un nombre c n'ayant pas de diviseur commun avec le produit $(p-1)(q-1)$. Il calcule d qui est l'inverse de c pour la

[8] Dans le cas de transmission de photons isolés, le théorème de non clonage quantique (§ 6.3.2) garantit qu'il est impossible à Ève de tromper Bob, même s'il lui est possible de faire moins de 50% d'erreurs en utilisant une technique d'interception plus sophistiquée.

multiplication modulo $(p-1)(q-1)$

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

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

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

et envoie b à Bob. Quand Bob reçoit le message il calcule

$$b^d \bmod N = a\,(!)$$

Le fait que le résultat soit précisément a, c'est-à dire le message original d'Alice, est un résultat de théorie des nombres. En résumé, sont envoyés sur voie publique, non sécurisée, les nombres N, c et b.

Exemple.

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

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

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

Alice envoie donc à Bob le message 16. Bob calcule

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

et Bob récupère donc le message original $a = 4$.

Chapitre 2

Manipulations d'un qu-bit

Dans le chapitre précédent, j'ai examiné un qu-bit à un instant déterminé. Dans un espace de Hilbert $\mathcal{H}$, ce qu-bit est décrit par un vecteur unitaire $|\varphi\rangle$

$$|\varphi\rangle = \lambda|0\rangle + \mu|1\rangle \qquad |\lambda|^2 + |\mu|^2 = 1 \tag{2.1}$$

J'ai choisi une base orthonormée $\{|0\rangle, |1\rangle\}$ de $\mathcal{H}$ et décomposé le vecteur $|\varphi\rangle$ suivant cette base. Je me propose maintenant d'examiner l'évolution temporelle de ce qu-bit, ce qui expliquera comment nous pourrons le manipuler.

2.1 Sphère de Bloch, spin 1/2

Avant de passer à cette évolution temporelle, je voudrais donner une description un peu plus générale du qu-bit et de ses réalisations physiques. J'ai choisi en écrivant (2.1) une base orthonormée $\{|0\rangle, |1\rangle\}$ de $\mathcal{H}$, et les coefficients λ et μ peuvent être paramétrés, compte tenu de l'arbitraire de phase, par

$$\lambda = \mathrm{e}^{-\mathrm{i}\phi/2} \cos\frac{\theta}{2} \qquad \mu = \mathrm{e}^{\mathrm{i}\phi/2} \sin\frac{\theta}{2} \tag{2.2}$$

Les deux angles θ et ϕ peuvent être considérés comme des angles polaires et azimutal, et (θ, ϕ) paramètrent la position d'un point sur la surface d'une sphère de rayon unité, appelée *sphère de Bloch* (ou sphère de Poincaré pour le photon) (figure 2.1).

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

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

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

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

Une autre réalisation physique importante du qu-bit est le spin 1/2. La RMN (Résonance Magnétique Nucléaire) et l'IRM (Imagerie par Résonance Magnétique...nucléaire[1]) reposent sur le fait que le proton possède un spin 1/2, ce que l'on met en évidence de la façon suivante : on fait passer un faisceau de particules[2] de spin 1/2 dans un champ magnétique orienté suivant une direction $\hat{n}$ perpendiculaire à la direction du faisceau. On constate que le faisceau se scinde en deux sous-faisceaux, l'un est dévié dans la direction $\hat{n}$, l'autre dans la direction opposée $-\hat{n}$. C'est l'expérience de Stern-Gerlach (figure 2.2, avec $\hat{n} \parallel Oz$), qui est très analogue dans son principe à la séparation d'un rayon de lumière naturelle en deux

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

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

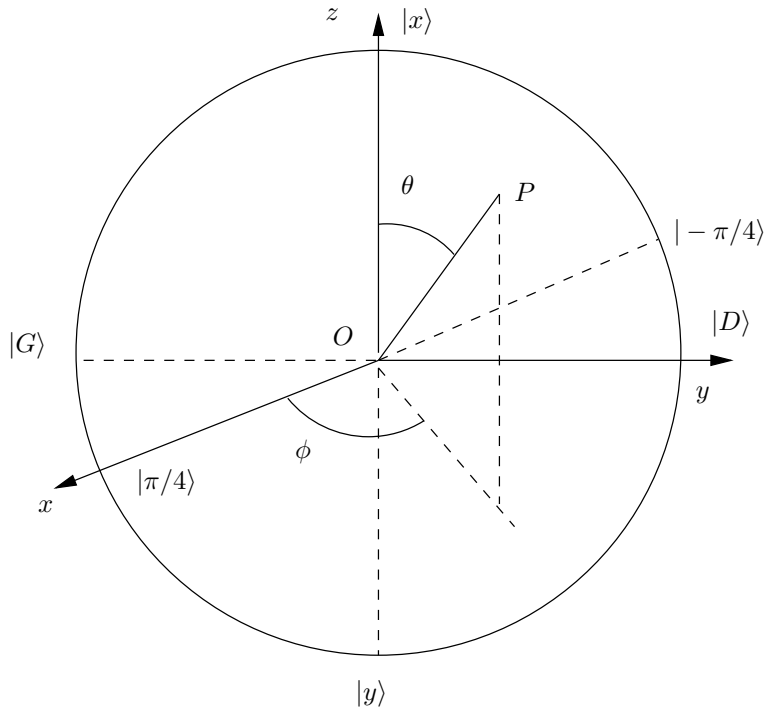

FIG. 2.1 – Sphère de Bloch.

rayons par un cristal biréfringent. On peut imaginer l'analogue d'une expérience analyseur/polariseur avec un spin 1/2 (figure 2.3). Toutefois on remarque que la situation polariseur/analyseur croisés correspond à $\theta = \pi$ et non à $\theta = \pi/2$ comme dans le cas des photons[3]. On construit une base de $\mathcal{H}$ en prenant pour vecteur de base les vecteurs $|+\rangle$ et $|-\rangle$, qui correspondent aux états préparés par un champ magnétique parallèle à Oz. Suivant (2.1) et (2.2), l'état de spin 1/2 le plus général est

$$|\varphi\rangle = \mathrm{e}^{-\mathrm{i}\phi/2} \cos\frac{\theta}{2} |+\rangle + \mathrm{e}^{\mathrm{i}\phi/2} \sin\frac{\theta}{2} |-\rangle \tag{2.3}$$

et on montre[4] que cet état est celui sélectionné par un champ magétique parallèle à $\hat{n}$, avec

$$\hat{n} = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta) \tag{2.4}$$

La sphère de Bloch possède dans ce cas une interprétation géométrique évidente : le spin 1/2 décrit par le vecteur (2.3) est orienté suivant la direction $\hat{n}$.

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

$$\sigma_1\,(\text{ou}\,\sigma_x) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad \sigma_2\,(\text{ou}\,\sigma_y) = \begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix} \qquad \sigma_3\,(\text{ou}\,\sigma_z) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{2.5}$$

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

$$M = \lambda_0 I + \sum_{i=1}^{3} \lambda_i \sigma_i \tag{2.6}$$

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

$$\sigma_i^2 = I \qquad \sigma_1\sigma_2 = \mathrm{i}\sigma_3 + \text{perm. circ.} \tag{2.7}$$

[3] Le photon a un spin 1, et non 1/2 !

[4] Ceci est une conséquence de l'invariance par rotation : voir *Physique quantique*, chapitre 3.

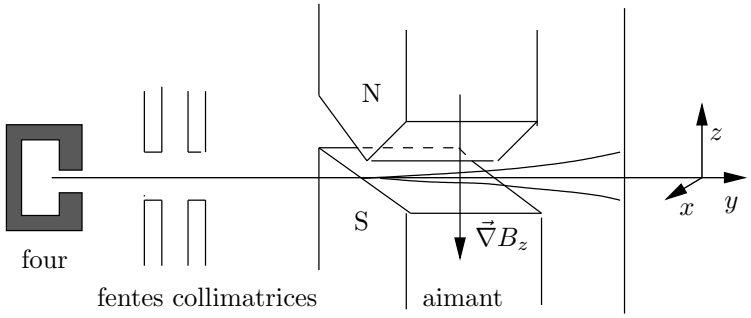

FIG. 2.2 – Expérience de Stern-Gerlach.

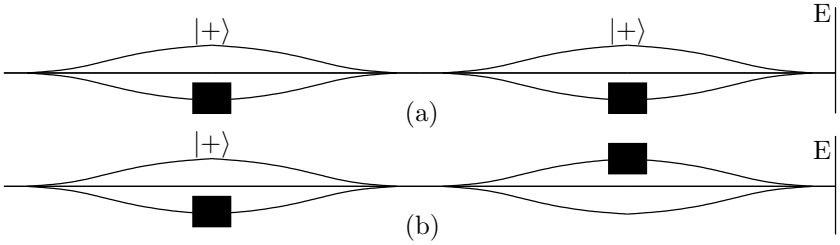

FIG. 2.3 – Polariseurs croisés pour le spin 1/2.

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

$$|+\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad |-\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \qquad \sigma_z|\pm\rangle = \pm|\pm\rangle \tag{2.8}$$

et on vérifie immédiatement que le vecteur $|\varphi\rangle$ (2.3) est vecteur propre de

$$\vec{\sigma}\cdot\hat{n} = \sigma_x n_x + \sigma_y n_y + \sigma_z n_z$$

avec la valeur propre $+1$

$$\vec{\sigma}\cdot\hat{n} = \begin{pmatrix} \cos\theta & \mathrm{e}^{-\mathrm{i}\phi}\sin\theta \\ \mathrm{e}^{\mathrm{i}\phi}\sin\theta & -\cos\theta \end{pmatrix} \tag{2.9}$$

Nous venons de voir la réalisation physique d'un qu-bit par un spin 1/2, mais il en existe bien d'autres, comme par exemple un atome à deux niveaux. Dans tous les cas on aura un espace de Hilbert de dimension 2, et l'état d'un qu-bit pourra toujours être représenté par un point sur la sphère de Bloch. Revenant à la notation $\{|0\rangle, |1\rangle\}$, on pourra (par exemple) faire l'identification $|+\rangle \to |0\rangle$ et $|-\rangle \to |1\rangle$.

2.2 Évolution dynamique

Nous introduisons explicitement le temps, en supposant que (2.1) est valable à $t = 0$

$$|\varphi(t=0)\rangle = \lambda(t=0)|0\rangle + \mu(t=0)|1\rangle \tag{2.10}$$

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

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

est linéaire et que la norme de $|\varphi\rangle$ reste égale à l'unité[5]

$$|\varphi(t)\rangle = \lambda(t)|0\rangle + \mu(t)|1\rangle \tag{2.11}$$
$$|\lambda(t)|^2 + |\mu(t)|^2 = 1 \tag{2.12}$$

La transformation $|\varphi(0)\rangle \to |\varphi(t)\rangle$ est donc une *transformation unitaire* $U(t,0)$

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

En général

$$|\varphi(t_2)\rangle = U(t_2,t_1)|\varphi(t_1)\rangle \qquad U^\dagger(t_2,t_1) = U^{-1}(t_2,t_1) \tag{2.13}$$

De plus U doit obéir à la propriété de groupe

$$U(t_2,t_1) = U(t_2,t')U(t',t_1) \tag{2.14}$$

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

$$\begin{aligned} U(t+\mathrm{d}t,t_0) &= U(t+\mathrm{d}t,t)U(t,t_0) \\ U(t+\mathrm{d}t,t_0) &\simeq U(t,t_0) + \mathrm{d}t\frac{\mathrm{d}}{\mathrm{d}t}U(t,t_0) \\ U(t+\mathrm{d}t,t)U(t,t_0) &\simeq \big[I - \mathrm{i}\,\mathrm{d}t\hat{H}(t)\big]U(t,t_0) \end{aligned}$$

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

$$\hat{H}(t) = \mathrm{i}\,\frac{\mathrm{d}U(t',t)}{\mathrm{d}t'}\Big|_{t'=t} \tag{2.15}$$

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

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

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

$$\boxed{\mathrm{i}\,\frac{\mathrm{d}U(t,t_0)}{\mathrm{d}t} = \hat{H}(t)U(t,t_0)} \tag{2.16}$$

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

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

$$\hat{H} = \begin{pmatrix} \omega_A & 0 \\ 0 & \omega_B \end{pmatrix} \tag{2.17}$$

ω_A et ω_B sont les *niveaux d'énergie* du spin 1/2. Si le champ magnétique est parallèle à Oz, les vecteurs propres de $\hat{H}$ ne sont autres que les vecteurs de base $|+\rangle \equiv |0\rangle$ et $|-\rangle \equiv |1\rangle$. Comme $\hat{H}$ est indépendant du temps, l'équation d'évolution (2.16)

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

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

s'intègre immédiatement

$$U(t,t_0) = \exp[-\mathrm{i}\hat{H}(t-t_0)] \tag{2.18}$$

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

$$U(t,t_0) = \begin{pmatrix} \mathrm{e}^{-\mathrm{i}\omega_A(t-t_0)} & 0 \\ 0 & \mathrm{e}^{-\mathrm{i}\omega_B(t-t_0)} \end{pmatrix} \tag{2.19}$$

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

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

alors

$$|\varphi(t)\rangle = \mathrm{e}^{-\mathrm{i}\omega_A t}\,\lambda|0\rangle + \mathrm{e}^{-\mathrm{i}\omega_B t}\,|1\rangle \tag{2.20}$$

soit

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

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

$$\omega_0 = \omega_B - \omega_A \tag{2.21}$$

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

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

La quantité ω_0 joue un rôle capital et elle est appelée *énergie (ou fréquence*[6]*) de résonance.*

J'en profite pour dire un mot sur une autre réalisation physique d'un qu-bit, *l'atome à deux niveaux.* L'atome possède deux niveaux d'énergie, un niveau fondamental ω_A et un niveau excité ω_B, $\omega_B > \omega_A$; s'il est porté dans son état excité, il revient spontanément dans son état fondamental en émettant un photon de fréquence $\omega_0 = \omega_B - \omega_A$. Si l'on envoie sur l'atome dans son état fondamental un faisceau laser de fréquence $\simeq \omega_0$, on observera un phénomène de résonance : l'absorption de la lumière laser sera d'autant plus importante que l'on sera proche de ω_0.

2.3 Manipulations de qu-bits : oscillations de Rabi

Comme nous le verrons au chapitre 4, pour les besoins du calcul quantique, il est nécessaire de pouvoir transformer par exemple un état $|0\rangle$ du qu-bit en une superposition linéaire de $|0\rangle$ et de $|1\rangle$. Pour ce faire, en prenant comme exemple le spin 1/2, la solution est d'appliquer au spin un champ magnétique tournant dans le plan xOy à une vitesse angulaire

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

C'est ce qui est fait dans la RMN. La forme de $\hat{H}(t)$ est alors

$$\hat{H}(t) = -\frac{1}{2}\begin{pmatrix} \omega_0 & \omega_1\mathrm{e}^{\mathrm{i}\omega t} \\ \omega_1\mathrm{e}^{-\mathrm{i}\omega t} & \omega_0 \end{pmatrix} \tag{2.22}$$

où ω_1 est proportionnelle à B_1, et donc ajustable[7]. La fréquence ω_1 est appelée *fréquence de Rabi.* Il reste à résoudre l'équation d'évolution (2.16). Celle-ci se transforme aisément en un système de deux équations différentielles du premier ordre couplées pour $\lambda(t)$ et $\mu(t)$, et la résolution de ce système ne pose aucune

[6] En toute rigueur,j'aurais dû préciser qu'énergie et fréquence sont reliées par la relation de Planck-Einstein $E = \hbar\omega$, où $\hbar$ est la constante de Planck. Afin de simplifier, je me suis placé dans un système d'unités où $\hbar = 1$.

[7] On peut comprendre l'origine de (2.22) en admettant que le hamiltonien est de la forme $\vec{\sigma}\cdot\vec{B}$

$$H \propto \vec{\sigma}\cdot\vec{B} = B_z\sigma_z + B_1(\sigma_x\cos\omega t - \sigma_y\sin\omega t)$$

Cette forme du hamiltonien provient du couplage entre le moment magnétique associé au spin 1/2 et le champ magnétique.

difficulté (voir l'annexe 2.3.1). On peut exprimer le résultat sous la forme suivante : si le qu-bit est au temps $t = 0$ dans l'état $|0\rangle$, il aura au temps t une probabilité $\mathsf{p}_{0\to1}(t)$ de se trouver dans l'état $|1\rangle$ donnée par

$$\mathsf{p}_{0\to1}(t) = \left(\frac{\omega_1}{\Omega}\right)^2 \sin^2\frac{\Omega t}{2} \qquad \Omega = \sqrt{(\omega-\omega_0)^2+\omega_1^2} \tag{2.23}$$

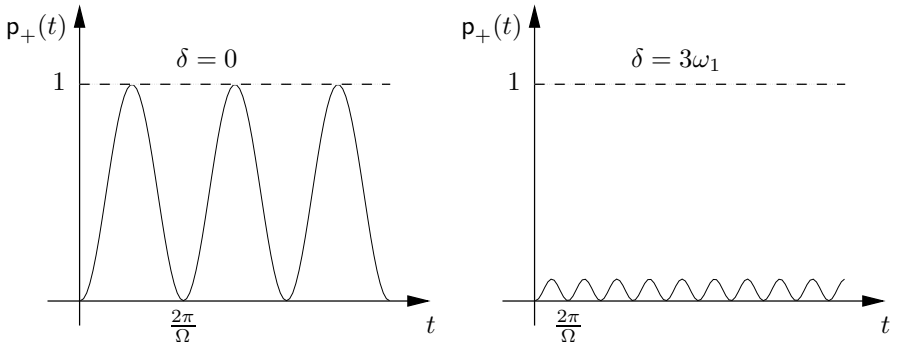

FIG. 2.4 – Oscillations de Rabi. Le *désaccord* $\delta = \omega - \omega_0$ et $\mathsf{p}_+(t) \equiv \mathsf{p}_{0\to1}(t)$.

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

$$\mathsf{p}_{0\to1}(t) = \sin^2\frac{\omega_1 t}{2} \qquad \omega = \omega_0 \tag{2.24}$$

Pour passer de l'état $|0\rangle$ à l'état $|1\rangle$, il suffit d'ajuster le temps t d'application du champ tournant

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

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

$$|0\rangle \to \frac{1}{\sqrt{2}}\left(|0\rangle + |1\rangle\right) \tag{2.25}$$

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

Annexe 2.3.1 : solution de l'équation d'évolution. Suivant (2.22), $\lambda(t)$ et $\mu(t)$ obéissent au sytème d'équations différentielles couplées

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

Il est commode de définir

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

Le système d'équations différentielles de simplifie en

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

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

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

La solution du système est alors

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

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

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

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

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

Les facteurs de phase peuvent être réabsorbés dans une redéfinition des états $|0\rangle$ et $|1\rangle$ de façon à obtenir (2.25).

Chapitre 3

Corrélations quantiques

3.1 États à deux qu-bits

On pourrait s'attendre à ce que le passage d'un qu-bit à deux qu-bits n'apporte que peu de nouveauté. En fait nous allons voir que la structure à deux qu-bits est extraordinairement riche, car elle introduit des corrélations quantiques entre les deux qu-bits, et on ne peut pas en rendre compte par des considérations de probabilités classiques. Comme nous le verrons à la section 4, l'intrication est à la base des spécificités du calcul quantique.

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

$$|X_{00}\rangle = |0_A \otimes 0_B\rangle \quad |X_{01}\rangle = |0_A \otimes 1_B\rangle \quad |X_{10}\rangle = |1_A \otimes 0_B\rangle \quad |X_{11}\rangle = |1_A \otimes 1_B\rangle \tag{3.1}$$

Il n'est pas difficile de construire l'état où le qu-bit A est dans

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

et le qu-bit B dans

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

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

$$\begin{aligned} |\varphi_A \otimes \varphi_B\rangle &= \lambda_A\lambda_B|0_A \otimes 0_B\rangle + \lambda_A\mu_B|0_A \otimes 1_B\rangle + \mu_A\lambda_B|1_A \otimes 0_B\rangle + \mu_A\mu_B|1_A \otimes 1_B\rangle \\ &= \lambda_A\lambda_B|X_{00}\rangle + \lambda_A\mu_B|X_{01}\rangle + \mu_A\lambda_B|X_{10}\rangle + \mu_A\mu_B|X_{11}\rangle \end{aligned} \tag{3.2}$$

Nous avons construit l'espace $\mathcal{H}_A \otimes \mathcal{H}_B$ *produit tensoriel* des espaces $\mathcal{H}_A$ et $\mathcal{H}_B$. On note que le vecteur $|\varphi_A \otimes \varphi_B\rangle$ est bien de norme unité[1]. Les physiciens sont assez laxistes sur les notations, et on trouvera au lieu de $|\varphi_A \otimes \varphi_B\rangle$, soit $|\varphi_A\rangle \otimes |\varphi_b\rangle$, soit même $|\varphi_A\varphi_B\rangle$, en omettant le symbole du produit tensoriel.

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

$$\begin{aligned} |\Psi\rangle &= \alpha_{00}|0_A \otimes 0_B\rangle + \alpha_{01}|0_A \otimes 1_B\rangle + \alpha_{10}|1_A \otimes 0_B\rangle + \alpha_{11}|1_A \otimes 1_B\rangle \\ &= \alpha_{00}|X_{00}\rangle + \alpha_{01}|X_{01}\rangle + \alpha_{10}|X_{10}\rangle + \alpha_{11}|X_{11}\rangle \end{aligned} \tag{3.3}$$

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

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

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

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

$$|\Phi\rangle = \frac{1}{\sqrt{2}}\Big(|0_A \otimes 1_B\rangle - |1_A \otimes 0_B\rangle\Big) \tag{3.4}$$

En effet

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

On définit de même le produit tensoriel $M_A \otimes M_B$ de deux opérateurs M_A et M_B

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

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

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

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

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

Exercices 3.1.1. Montrer que l'opérateur

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

permute les valeurs des deux bits A et B

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

La notation $\vec{\sigma}_A \cdot \vec{\sigma}_B$ désigne à la fois un produit scalaire et un produit tensoriel. Montrer également que l'opérateur cNOT, représenté par la matrice 4×4

$$\text{cNOT} = \begin{pmatrix} I & 0 \\ 0 & \sigma_x \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

n'est pas un produit tensoriel. Montrer que l'action de cNOT sur le vecteur produit tensoriel

$$\frac{1}{\sqrt{2}}\big(|0\rangle + |1\rangle\big) \otimes |0\rangle$$

donne un état intriqué.

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

Montrons le d'abord sur un exemple, celui de l'état $|\Phi\rangle$ (3.4). Soit M une propriété physique du qu-bit A, et calculons sa valeur moyenne $\langle\Phi|M\Phi\rangle$

$$\begin{aligned} \langle M\rangle_\Phi = \langle\Phi|M\Phi\rangle &= \frac{1}{2}\big[\langle 0_A \otimes 1_B| - \langle 1_A \otimes 0_B|\big]\big[|(M\,0_A) \otimes 1_B\rangle - |(M\,1_A) \otimes 0_B\rangle\big] \\ &= \frac{1}{2}\big(\langle 0_A|M\,0_A\rangle + \langle 1_A|M\,1_A\rangle\big) \end{aligned} \tag{3.5}$$

où on a utilisé

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

Il n'existe pas d'état

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

tel que

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

En effet on aurait alors

$$\langle\varphi_A|M\varphi_A\rangle = |\lambda|^2\langle 0_A|M0_A\rangle + (\lambda^*\mu\langle 0_A|M1_A\rangle + \text{ c.c.}) + |\mu|^2\langle 1_A|M0_A\rangle$$

Pour reproduire (3.5), une condition nécessaire serait que $|\lambda| = |\mu| = 1/\sqrt{2}$, et les termes en $\lambda^*\mu$ ne seraient pas nuls. Le résultat (3.5) a une interprétation physique simple : l'état du qu-bit A est un mélange *incohérent* de 50% de l'état $|0_A\rangle$ et de 50% de l'état $|1_A\rangle$, et non une superposition linéaire. En résumé, on ne peut pas en général décrire une *partie* d'un système quantique par un vecteur d'état.

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

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

de même qu'une lumière polarisée circulairement à droite

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

3.2 Opérateur densité et entropies

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

$$|\Phi\rangle = \sum_{i,\mu} \alpha_{i\mu}|i\otimes\mu\rangle \tag{3.6}$$

Soit M une propriété physique du sous-système A

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

Calculons la valeur moyenne de M

$$\begin{aligned}\langle\Phi|M\Phi\rangle &= \sum_{j,\nu}\sum_{i,\mu}\alpha^*_{j\nu}\alpha_{i\mu}\langle j\otimes\nu|Mi\otimes\mu\rangle \\ &= \sum_{i,j}\sum_{\mu}\alpha^*_{j\mu}\alpha_{i\mu}\,\langle j|Mi\rangle = \sum_{i,j}\rho_{ij}\,\langle j|Mi\rangle = \sum_{i,j}\rho_{ij}\,M_{ji} = \mathrm{Tr}\,(\rho M)\end{aligned} \tag{3.7}$$

Pour obtenir (3.7) on a utilisé

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

et on a défini un objet qui joue un rôle crucial, la *matrice* (ou plus correctement *l'opérateur*) *densité* ρ du sous-système A

$$\rho_{ij} = \sum_{\mu}\alpha_{i\mu}\alpha^*_{j\mu} \tag{3.8}$$

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

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

Un état physique tel que ceux examinés dans le chapitre 1 sont appelés des *cas purs*. Il est facile de vérifier que l'opérateur densité d'un cas pur obéit à $\rho^2 = \rho$ et inversement tout opérateur densité tel que $\rho^2 = \rho$ décrit un cas pur. Mais la description la plus générale d'un système quantique se fait au moyen de l'opérateur densité.

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

$$\rho = \sum_i \mathsf{p}_i |i\rangle\langle i|$$

En raison de la positivité de ρ, $\mathsf{p}_i \geq 0$ et la condition $\mathrm{Tr}\,\rho = 1$ donne $\sum_i \mathsf{p}_i = 1$, ce qui fait que les p_i peuvent être interprétés comme des probabilités. On peut dire que ρ représente un *mélange statistique* d'états $|i\rangle$, chaque état $|i\rangle$ ayant une probabilité p_i. Ceci nous amène à définir une généralisation de l'entropie de Shannon, *l'entropie de von Neumann*

$$\boxed{H_{\mathrm{vN}} = -\sum_i \mathsf{p}_i \log \mathsf{p}_i = -\mathrm{Tr}\,\rho \log \rho} \tag{3.9}$$

où log est un logarithme de base 2. On note que l'entropie d'un cas pur est nulle, car tous les p_i sont nuls à l'exception d'un seul qui vaut un. Il est important de noter que ρ peut être interprété de plusieurs façons si l'on admet de choisir des états $|\alpha\rangle$ de norme un mais non orthogonaux ($\langle\alpha|\beta\rangle \neq \delta_{\alpha\beta}$), et qui par conséquent ne peuvent pas être parfaitement distingués[2]

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

Il y a en général une infinité de mélanges statistiques différents qui donnent le même opérateur densité, et on montre que

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

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

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

ou sous forme matricielle

$$\rho_{A,ij} = \sum_\mu [\rho_{AB}]_{i\mu,j\mu} \qquad \rho_{B,\mu\nu} = \sum_i [\rho_{AB}]_{i\mu,i\nu}$$

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

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

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

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

[2] On peut par exemple former un mélange de 70% de photons linéairement polarisés $|x\rangle$ et de 30% de photons $|\pi/4\rangle$.

où

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

L'inégalité de droite est la même pour les deux entropies, mais celle de gauche (inégalité de Araki-Lieb) est différente. Par exemple si ρ_{AB} est l'opérateur densité décrivant l'état pur de deux qu-bits (3.4), $H_{\text{vN}}(\rho_{AB}) = 0$, alors que

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

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

L'importance de la notion d'opérateur densité est confirmée par le théorème de Gleason, qui dit en gros que la description la plus générale d'un système quantique est donnée par un opérateur densité.

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

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

ainsi que

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

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

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

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

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

$$|\Phi\rangle = \sum_i \sqrt{\mathsf{p}_i}\; |i_A \otimes i'_B\rangle \tag{3.10}$$

avec

$$\langle i_A | j_A\rangle = \langle i'_B | j'_B\rangle = \delta_{ij}$$

Les états $|i_A\rangle$ et $|i'_B\rangle$ dépendent bien évidemment de $|\Phi\rangle$. Cette expression donne immédiatement[3] les opérateurs densité réduits ρ_A et ρ_B

$$\rho_A = \sum_i \mathsf{p}_i |i_A\rangle\langle i_A| \qquad \rho_B = \sum_{i'} \mathsf{p}_i |i'_B\rangle\langle i'_B| \tag{3.11}$$

Le nombre des p_i différents de zéro est le *nombre de Schmidt*. Si l'on applique sur un état $|\Phi\rangle$ quelconque une transformation unitaire qui est un produit tensoriel de transformations agissant sur A et B, $U_A \otimes U_B$, on ne peut pas changer le nombre de Schmidt en manipulant *séparément* les qu-bits A et B. On retrouve le résultat énoncé ci-dessus pour un produit tensoriel en remarquant que le nombre de Schmidt d'un produit tensoriel est un.

[3] Soit $|i\rangle$ une base orthonormée de $\mathcal{H}$ et M l'opérateur $|\varphi\rangle\langle\psi|$. Alors

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

car $\sum_i |i\rangle\langle i| = I$. De plus

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

Annexe 3.3.1. Exemple de construction d'un état intriqué. Prenons l'exemple de deux spins 1/2. Une interaction possible[4]entre ces deux spins est

$$\hat{H} = \frac{\omega}{2}\ \vec{\sigma}_1 \cdot \vec{\sigma}_2$$

Utilisons le résultat de l'exercice 3.1.1

$$\frac{1}{2}(I + \vec{\sigma}_1 \cdot \vec{\sigma}_2)|ij\rangle = |ji\rangle$$

pour montrer que

$$\begin{aligned}(\vec{\sigma}_1 \cdot \vec{\sigma}_2)\frac{1}{\sqrt{2}}\big(|10\rangle + |01\rangle\big) &= (\vec{\sigma}_1 \cdot \vec{\sigma}_2)|\Phi_+\rangle = |\Phi_+\rangle \\ (\vec{\sigma}_1 \cdot \vec{\sigma}_2)\frac{1}{\sqrt{2}}\big(|10\rangle - |01\rangle\big) &= (\vec{\sigma}_1 \cdot \vec{\sigma}_2)|\Phi_-\rangle = -3|\Phi_-\rangle\end{aligned}$$

Les vecteurs $|\Phi_+\rangle$ et $|\Phi_-\rangle$ sont vecteurs propres de $\vec{\sigma}_1 \cdot \vec{\sigma}_2$ avec les valeurs propres respectives $+1$ et -3[5]. Partons au temps $t = 0$ d'un état non intriqué $|\Phi(0)\rangle = |10\rangle$. Pour obtenir son évolution temporelle, il suffit de décomposer cet état sur $|\Phi_+\rangle$ et $|\Phi_-\rangle$

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

Écrire l'évolution temporelle est alors immédiat

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

Il suffit de choisir $\omega t = \pi/2$ pour obtenir l'état intriqué $|\Psi\rangle$

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

La difficulté vient de ce que $\hat{H}$ est en général une interaction *interne* au système, qui, contrairement aux interactions de type externe utilisées pour les qu-bits individuels, ne peut pas être branchée et débranchée facilement pour ajuster t. Si l'interaction est à courte distance, il est possible de rapprocher puis d'éloigner les deux qu-bits. On peut aussi appliquer aux deux qu-bits des interactions externes différentes, ce qui est la technique utilisée dans le cas de la RMN, où l'interaction interne est plus simple, en $\sigma_{1z}\sigma_{2z}$; il est commode de se servir de l'identité

$$\mathrm{cNOT} = H(\mathrm{c}\sigma_z)H$$

Il est aussi possible d'obtenir un état intriqué de deux objets en faisant intervenir un troisième objet auxiliaire, par exemple intriquer deux atomes en les faisant interagir avec un photon d'une cavité résonante.

3.3 Théorème de non clonage quantique

La condition indispensable pour que la méthode de cryptographie quantique décrite au § 1.6 soit parfaitement sûre est que l'espionne Ève ne puisse pas reproduire (cloner) l'état de la particule envoyée par Bob à Alice tout en conservant pour elle le résultat de sa mesure, ce qui rendrait l'interception du message indétectable. Que ceci ne soit pas possible est garanti par le théorème de non clonage quantique.

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

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

Pour montrer ce théorème, supposons que l'on souhaite dupliquer un état quantique *inconnu* $|\chi_1\rangle$. Le système sur lequel on veut imprimer la copie est noté $|\varphi\rangle$: c'est l'équivalent de la feuille blanche. Par exemple si l'on veut cloner un état de spin 1/2 $|\chi_1\rangle$, $|\varphi\rangle$ est aussi un état de spin 1/2. L'évolution du vecteur d'état dans le processus de clonage doit être de la forme

$$|\chi_1 \otimes \varphi\rangle \rightarrow |\chi_1 \otimes \chi_1\rangle \tag{3.12}$$

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

$$|U(\chi_1 \otimes \varphi)\rangle = |\chi_1 \otimes \chi_1\rangle \tag{3.13}$$

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

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

Évaluons maintenant le produit scalaire

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

de deux façons différentes

$$\begin{array}{ll} (1) & X = \langle \chi_1 \otimes \varphi | \chi_2 \otimes \varphi\rangle = \langle \chi_1 | \chi_2\rangle \\ (2) & X = \langle \chi_1 \otimes \chi_1 | \chi_2 \otimes \chi_2\rangle = (\langle \chi_1 | \chi_2\rangle)^2 \end{array} \tag{3.14}$$

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

3.4 Inégalités de Bell

La peuve du caractère non classique des corrélations d'un état intriqué est donnée par les inégalités de Bell, que je vais expliquer sur un exemple. Supposons que nous ayons fabriqué des paires de photons A et B partant en sens inverse et dont les polarisations sont intriquées (figure 3.1)

$$|\Phi\rangle = \frac{1}{\sqrt{2}} \Big(|x_A x_B\rangle + |y_A y_B\rangle \Big) \tag{3.15}$$

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

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

Afin de donner une forme commode à ce résultat, on convient de décrire la corrélation des polarisations de la façon suivante (A_x et B_x ne sont pas autre chose que l'opérateur $M = \mathcal{P}_x - \mathcal{P}_y$ introduit dans la section 1.5)

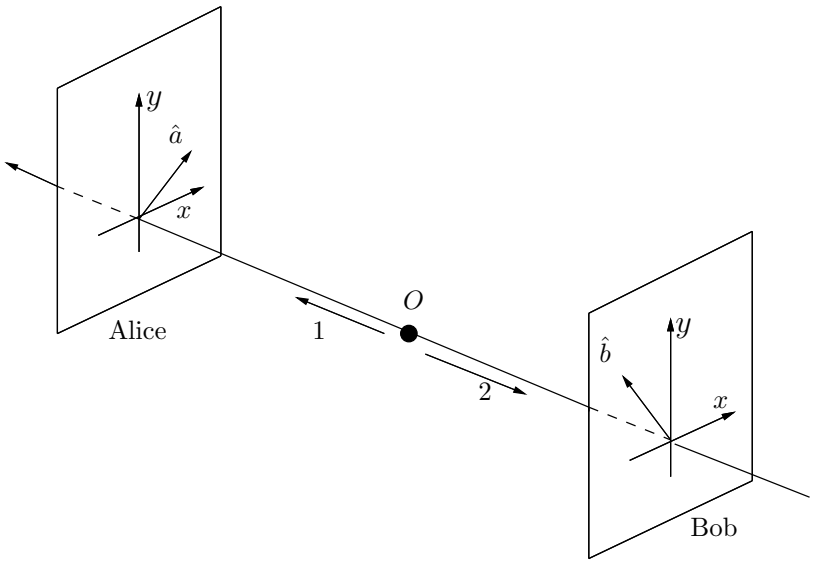

FIG. 3.1 – Configuration d'une expérience de type EPR.

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

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

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

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

$$\langle A_x B_x \rangle = 1 \tag{3.16}$$

À la réflexion ce résultat, pour l'instant, n'est pas trop surprenant. C'est une variante du "jeu[6] des deux douaniers" : deux voyageurs A et B partent en sens inverse depuis l'origine, chacun emportant une valise, et sont contrôlés ultérieurement par deux douaniers, Alice et Bob. L'une des valises contient une boule rouge, l'autre une boule verte, mais les voyageurs ont pris au hasard leur valise fermée et ils ne connaissent pas la couleur de la boule enfermée dans leur valise. Si Alice contrôle la valise du voyageur A, elle a 50% de chances de trouver une boule verte. Mais si elle trouve effectivement une boule verte, il est clair que Bob, avec une probabilité de 100%, va trouver une boule rouge ! Des corrélations ont été introduites au départ entre les valises, qui se retrouvent dans une corrélation des résultats d'Alice et Bob.

Cependant, comme l'ont remarqué pour la première fois Einstein, Podolsky et Rosen (EPR) dans un article célèbre[7] − sur un exemple différent, la version exposée ici est due à Bohm −, la situation devient nettement moins banale si Alice et Bob décident d'utiliser dans une autre série de mesures des orientations $\hat{\theta}$ et $\hat{\theta}_\perp$ au lieu des orientations Ox et Oy. En effet $|\Phi\rangle$ est invariant par rotation autour de Oz, car en utilisant (1.19), on montre immédiatement que $|\Phi\rangle$ s'écrit aussi

$$|\Phi\rangle = \frac{1}{\sqrt{2}} \Big(|\theta_A \theta_B\rangle + |\theta_{\perp A} \theta_{\perp B}\rangle \Big) \tag{3.17}$$

Si on remplace A_x par A_θ

[6] Inventé pour la circonstance !

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

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

Alors on a comme en (3.16)

$$\langle A_\theta B_\theta \rangle = 1 \tag{3.18}$$

Connaissant la polarisation du photon A suivant $\hat{\theta}$, on peut prédire avec certitude celle du photon B suivant $\hat{\theta}$ (ou $\hat{\theta}_\perp$), quel que soit le choix de θ. On a l'impression que Alice et Bob vont pouvoir s'envoyer des messages instantanément, même s'ils sont distants de plusieurs années lumière, et donc violer la relativité. Bien sûr ce n'est qu'une illusion, car pour pouvoir confronter leurs résultats et vérifier (3.18), Alice et Bob doivent pouvoir échanger des messages par une voie classique et donc à une vitesse inférieure à celle de la lumière. De plus on peut repoduire ces corrélations avec un modèle classique (figure 3.2), où les corrélations sont fixées à l'avance.

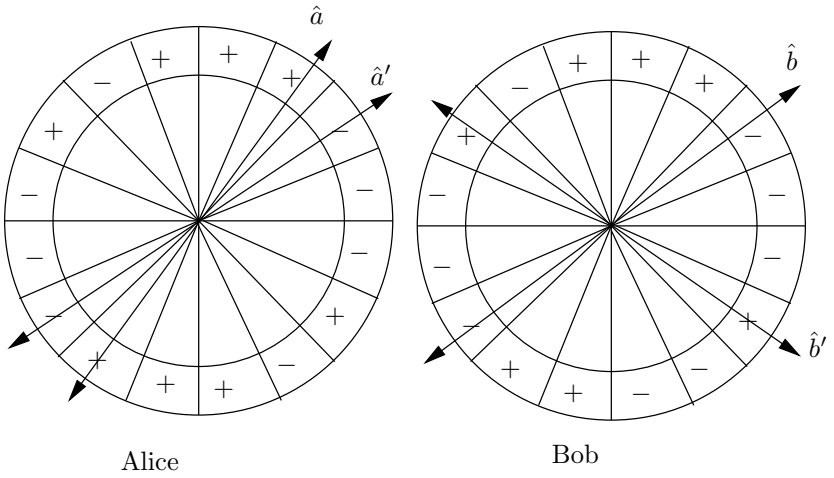

FIG. 3.2 – Modèle classique pour les corrélations EPR. Les valises des voyageurs A et B sont maintenant des cercles divisés en petits secteurs angulaires définissant des orientations $\hat{a}, \ldots, \hat{b}'$ dans le plan xOy, et qui peuvent être étiquetés + : polarisation suivant cette direction, ou − : polarisation orthogonale à cette direction. Les deux cercles sont identiques et deux points diamétralement opposés sont tous deux étiquetés + ou −. La figure correspond à $A_{\hat{a}} = +1,\ A_{\hat{a}'} = -1$, $B_{\hat{b}} = -1$ et $B_{\hat{b}'} = +1$.

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

$$E(\hat{a}, \hat{b}) = \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} A_n(\hat{a}) B_n(\hat{b}) \tag{3.19}$$

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

$$X_n = A_n B_n + A_n B'_n + A'_n B'_n - A'_n B_n = A_n(B_n + B'_n) + A'_n(B'_n - B_n) \tag{3.20}$$

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

$$|\langle X \rangle| = \Big| \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} X_n \Big| \leq 2 \tag{3.21}$$

La quantité X_n est "contrefactuelle", car elle ne peut pas être mesurée sur une seule paire : on a quatre choix possibles pour l'orientation des axes de mesure, mais une seule orientation peut être choisie pour une paire déterminée. Le point de vue EPR est que chaque photon transporte avec lui toute l'information sur sa propre polarisation et que les quatre combinaisons $A_nB_n\cdots A'_nB'_n$ existent pour toute paire n, même si on peut en mesurer une seule dans une expérience donnée. Cependant cela ne veut pas dire nécessairement que le point de vue EPR est erroné, car comme le dit très bien Feynman "It is not true that we can pursue science completely by using only those concepts which are directly subject to experiment". La falsification du point de vue EPR viendra de l'expérience.

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

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

en écrivant $|x_B\rangle$ et $|y_B\rangle$ en fonction de $|\theta_B$ et de $|\theta_{\perp B}$ (voir (1.19)). Il est alors immédiat de calculer les produits scalaires

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

soit

$$E(\hat{x},\hat{\theta}) = \frac{1}{2}\Big[2\cos^2\theta - 2\sin^2\theta\Big] = \cos(2\theta) \tag{3.22}$$

ou sous une forme manifestement invariante par rotation

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

Avec le choix d'angles de la figure 3.3 on trouve

$$|\langle X\rangle| = 2\sqrt{2} \simeq 2.82 \tag{3.23}$$

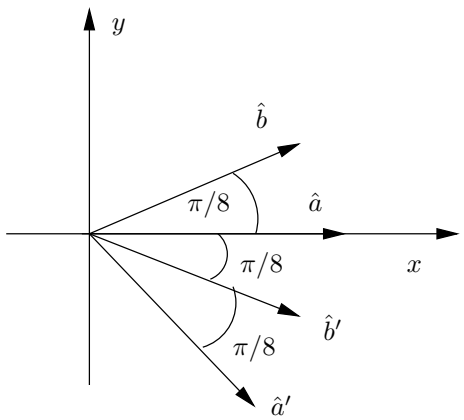

FIG. 3.3 – Configuration optimale des angles.

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

Remarquons aussi l'importance du théorème de non clonage si l'on veut éviter une propagation d'information à une vitesse supérieure à celle de la lumière. En effet Alice pourrait choisir d'utiliser la base $\{|x\rangle, |y\rangle\}$ ou la base $\{|\pi/4\rangle, -|\pi/4\rangle\}$ pour mesurer la polarisation de son photon, en attribuant la valeur 0 (1) du bit à la première (seconde) base. Si Bob était capable de cloner son photon, il pourrait mesurer sa polarisation, et en déduire instantanément la base choisie par Alice.

3.5 Téléportation

La *téléportation* est une application amusante des états intriqués, qui pourrrait servir à transférer l'information quantique (figure 3.4). Supposons qu'Alice souhaite transférer à Bob l'information sur l'état de spin $|\varphi\rangle_1$ d'une particule 1 de spin 1/2

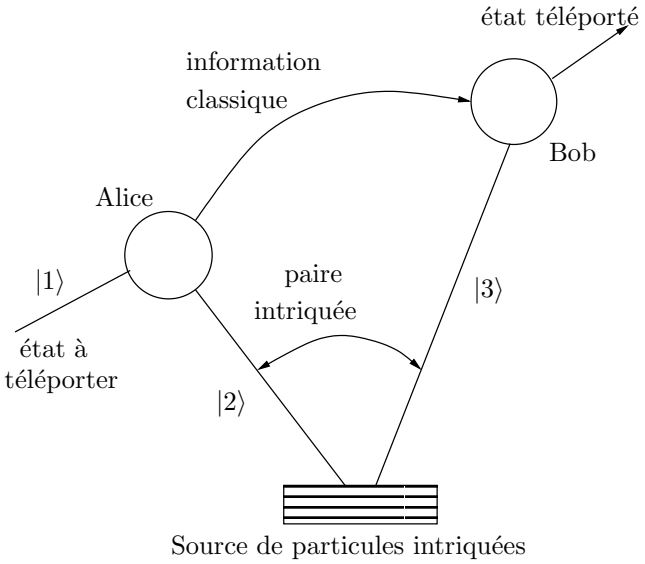

FIG. 3.4 – Téléportation : Alice effectue une mesure de Bell sur les particules 1 et 2 et informe Bob du résultat par une voie classique.

$$|\varphi\rangle_1 = \lambda|+\rangle_1 + \mu|-\rangle_1 \tag{3.24}$$

qui lui est *a priori* inconnu, sans lui transmettre directement cette particule. Elle ne peut pas faire une mesure du spin, car elle ne connaît pas l'orientation du spin de la particule 1, et toute mesure projetterait en général $|\varphi\rangle_1$ sur un autre état. Le principe du transfert de l'information consiste à utiliser une paire auxiliaire de particules intriquées 2 et 3 de spin 1/2 : la particule 2 est utilisée par Alice et la particule 3 est envoyée vers Bob. Ces particules 2 et 3 se trouvent par exemple dans l'état intriqué de spin

$$|\Phi^-\rangle_{23} = \frac{1}{\sqrt{2}}\Big(|+-\rangle_{23} - |-+\rangle_{23}\Big) \tag{3.25}$$

Alice va mesurer l'état de spin de la paire de particules 1 et 2 en utilisant une base particulière, la *base des états de Bell* : l'état de spin de la paire de particules 1 et 2 (non intriquées) peut se décomposer sur

cette base formée d'états intriqués

$$\begin{aligned}
|\Phi^+\rangle_{12} &= \frac{1}{\sqrt{2}}\Big(|+-\rangle_{12} + |-+\rangle_{12}\Big) \\
|\Phi^-\rangle_{12} &= \frac{1}{\sqrt{2}}\Big(|+-\rangle_{12} - |-+\rangle_{12}\Big) \\
|\Psi^+\rangle_{12} &= \frac{1}{\sqrt{2}}\Big(|++\rangle_{12} + |--\rangle_{12}\Big) \\
|\Psi^-\rangle_{12} &= \frac{1}{\sqrt{2}}\Big(|++\rangle_{12} - |--\rangle_{12}\Big)
\end{aligned} \tag{3.26}$$

L'état de spin des trois particules est

$$|\Psi\rangle_{123} = |\varphi\rangle_1 \otimes |\Phi^-\rangle_{23} = \frac{1}{\sqrt{2}}\Big(\lambda|++-\rangle_{123} + \mu|-+-\rangle_{123} - \lambda|+-+\rangle_{123} - \mu|--+\rangle_{123}\Big) \tag{3.27}$$

On utilise maintenant

$$|++\rangle_{12} = \frac{1}{\sqrt{2}}\Big(|\Psi^+\rangle_{12} + |\Psi^-\rangle_{12}\Big)$$

et trois relations analogues pour récrire

$$\begin{aligned}
|\Psi\rangle_{123} = \frac{1}{2}\Big[&|\Psi^+\rangle_{12}\Big(-\mu|+\rangle_3 + \lambda|-\rangle_3\Big) + |\Psi^-\rangle_{12}\Big(\mu|+\rangle_3 + \lambda|-\rangle_3\Big) \\
&+ |\Phi^+\rangle_{12}\Big(-\lambda|+\rangle_3 + \mu|-\rangle_3\Big) - |\Phi^-\rangle_{12}\Big(\lambda|+\rangle_3 + \mu|-\rangle_3\Big)\Big]
\end{aligned} \tag{3.28}$$

La mesure par Alice de l'état de spin de la paire de particules 1 et 2 projette cet état sur l'un des quatre états de base (3.28), ce qui projette l'état de la particule 3 sur l'état correspondant dans (3.26). Par exemple si la mesure projette sur $|\Phi^-\rangle_{12}$, l'état de spin de la particule 3 est $|\varphi\rangle_3 = \lambda|+\rangle_3 + \mu|-\rangle_3$. Alice transmet alors à Bob par une voie classique le résultat de sa mesure, et Bob sait que la particule 3 lui arrive précisément dans l'état inconnu de départ (3.24), mais qui reste tout aussi inconnu ! L'état de la particule 1 a été téléporté, mais il n'y a jamais eu de mesure de cet état. Si le résultat de la mesure d'Alice n'est pas $|\Phi^-\rangle_{12}$, Bob en sait assez pour faire la correction et appliquer un champ magnétique convenable pour réorienter son spin dans l'état (3.24).

Il est utile d'ajouter les remarques finales

- À aucun moment les coefficients λ et μ ne sont mesurés, et l'état $|\varphi\rangle_1$ est détruit au cours de la mesure faite par Alice. Il n'y a donc pas de contradiction avec le théorème de non clonage.
- Bob ne "connaît" l'état de la particule 3 que lorsqu'il a reçu le résultat de la mesure d'Alice. La transmission de cette information doit se faire par une voie classique, à une vitesse au plus égale à celle de la lumière. Il n'y a donc pas transmission instantanée de l'information à distance.
- Il n'y a jamais transport de matière dans la téléportation.

Chapitre 4

Introduction au calcul quantique

Le schéma d'un calcul sur un ordinateur quantique est dessiné sur la figure 4.1 : n qu-bits sont préparés dans l'état $|0\rangle$ au temps $t = t_0$. C'est la phase de préparation du système quantique, et le vecteur d'état initial appartient à un espace de Hilbert à 2^n dimensions, $\mathcal{H}^{\otimes n}$. Ces qu-bits subissent ensuite une évolution purement quantique décrite par un opérateur unitaire $U(t, t_0)$ agissant dans $\mathcal{H}^{\otimes n}$. La difficulté expérimentale consiste d'ailleurs à éviter toute interaction avec l'environnement, car le phénomène de *décohérence* rendrait l'évolution non unitaire. En effet, dans le cas d'une interaction avec l'environnement, l'évolution unitaire se fait dans un espace de Hilbert plus grand que $\mathcal{H}^{\otimes n}$, car il faut non seulement tenir compte des degrés de liberté des qu-bits, mais aussi de ceux de l'environnement. Une fois l'évolution quantique achevée, une mesure est effectuée au temps t sur les qu-bits afin d'obtenir le résultat du calcul. Un point important est que *l'on ne peut pas observer l'état du calcul* entre t_0 et t, car toute mesure modifierait l'évolution unitaire. Un autre point essentiel est que l'évolution unitaire est *réversible*[1] ; connaissant le vecteur d'état au temps t, on peut remonter à celui au temps t_0 par $U^{-1}(t, t_0) = U(t_0, t)$.

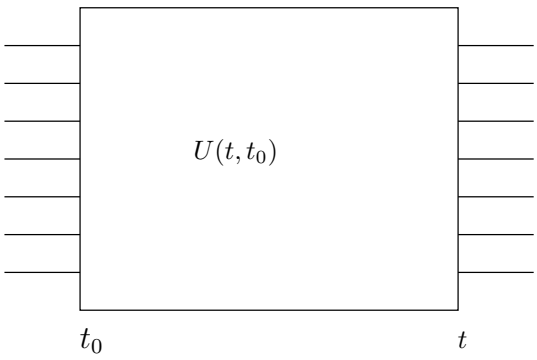

FIG. 4.1 – Schéma de principe d'un calcul quantique : n qu-bits sont préparés dans l'état $|0\rangle$. Ils subissent une évolution unitaire dans l'espace $\mathcal{H}^{\otimes n}$ de l'instant $t = t_0$ à l'instant t, décrite par un opérateur unitaire $U(t, t_0)$ agissant dans $\mathcal{H}^{\otimes n}$. Une mesure des qu-bits est effectuée au temps t.

4.1 Calcul réversible

La plupart des portes logiques usuelles sont irréversibles, car elles correspondent à un passage (2 bits→ 1 bit), et l'état final d'un bit ne permet pas de remonter à l'état initial de deux bits. Par exemple la porte NAND

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

[1] Note pour les physiciens : il ne faut surtout pas confondre évolution réversible et invariance par rapport au renversement du sens du temps, le renversement du temps étant représenté dans $\mathcal{H}$ par une opération antiunitaire.

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

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

On sait que la porte NAND et l'opération COPY suffisent à construire tous les circuits logiques. Une question intéressante est de savoir si toutes les opérations logiques habituelles pourraient être conduites de façon réversible.

La question a d'abord eu un intérêt théorique, mis en avant principalement par Landauer et Bennett, qui se sont demandé s'il était possible de calculer sans dissipation d'énergie. En effet, en dépit de son caractère abstrait, l'information est nécessairement portée par un support physique[2]. Comme bonus de cette étude, Bennett a pu donner une solution enfin satisfaisante (après plus d'un siècle !) du paradoxe du démon de Maxwell (voir l'annexe 4.1.1). Un calcul où entrent des opérations irréversibles dissipe obligatoirement de l'énergie, car effacer un bit coûte au minimum une entropie thermodynamique $k_B \ln 2$, où k_B est la constante de Boltzmann ($k_B = 1.38 \times 10^{-23}\,\mathrm{J/K}$), et donc une dissipation d'énergie dans l'environnement de $\Delta E = k_B T \ln 2$, où T est la température absolue de l'ordinateur. Le problème est pour le moment académique, car sur un PC actuel on a déjà $\Delta E \sim 500 k_B T$ par bit effacé, simplement en raison de la consommation électrique, et on n'en est donc pas à $k_B T$ près. Mais il est possible que la question devienne intéressante un jour d'un point de vue pratique.

Le principal intérêt du calcul réversible est la transposition au calcul quantique des algorithmes classiques. Une transposition directe est impossible, car le calcul quantique est réversible, et il est au préalable nécessaire de remplacer l'opération NAND par une opération réversible et de trouver l'équivalent de l'opération COPY sans entrer en conflit avec le théorème de non clonage. La solution fait intervenir deux portes logiques, la porte cNOT et la porte de Toffoli (figure 4.2). Si les bits d'entrée sont (x, y), la porte cNOT a l'effet suivant

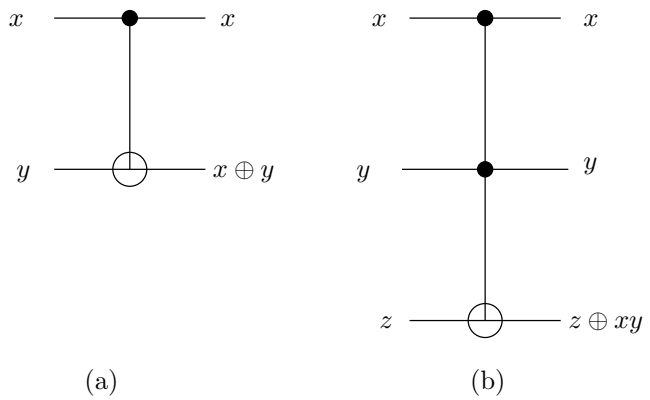

FIG. 4.2 – Portes cNOT (a) et de Toffoli (b). Les points noirs représentent les bits de contrôle et le cercles les bits cible.

$$\text{cNOT} \ : \qquad (x, y) \to (x, x \oplus y) \tag{4.1}$$

La porte cNOT copie le bit x si $y = 0$ et donne $\neg x$ si $y = 1$. Elle est réversible car il y a correspondance biunivoque entre état initial et état final : l'opération cNOT est une simple permutation des vecteurs de base (voir (4.3). Il est facile de montrer qu'avec les portes à un bit

$$x \to 1 \oplus x \quad \text{ou} \quad x \to \neg x$$

et la porte cNOT on ne peut construire que des fonctions linéaires. Il faut donc ajouter une porte supplémentaire, la porte de Toffoli, qui est une porte à trois bits d'entrée et de sortie

$$\text{Toffoli} \ : \qquad (x, y, z) \to (x, y, z \oplus xy) \tag{4.2}$$

[2] "Information is physical", disait Landauer, qui en déduisait (à mon sens abusivement), que les mathématiques et l'informatique étaient des branches de la physique !

Si $z = 1$, la porte de Toffoli effectue l'opération NAND. Le bit x dans la porte cNOT et les bits (x, y) dans la porte de Toffoli sont appelés *bits de contrôle* et le dernier bit le *bit cible*. Avec la porte de Toffoli, on peut reproduire de façon réversible tous les circuits logiques classiques : la porte de Toffoli est une porte universelle pour toutes les opérations réversibles de la logique booléenne.

Annexe 4.1.1 : Le démon de Maxwell. En 1871, Maxwell a imaginé le dispositif suivant : une enceinte contenant un gaz à la température absolue T est divisée en deux compartiments de volume identique, séparés par une cloison percée d'une petite ouverture. Un démon peut actionner sans dépense d'énergie une porte qui ouvre ou ferme l'ouverture, et il peut observer la vitesse des molécules. Les molécules dans l'enceinte ont une vitesse moyenne de quelques centaines de mètres par seconde à la température ambiante ($T \simeq 300\,\mathrm{K}$), mais certaines sont plus rapides et d'autres plus lentes. Le démon ouvre la porte quand il voit arriver vers l'ouverture une molécule rapide du compartiment de gauche et allant vers celui de droite, et aussi quand il voit une molécule lente du compartiment de droite se diriger vers celui de gauche. La vitesse moyenne des molécules du compartiment de droite va augmenter, celle des molécules du compartiment de gauche diminuer, l'énergie totale du gaz restant constante. Comme la vitesse moyenne est reliée à T et à la masse m des molécule par

$$v \simeq \sqrt{\frac{k_B T}{m}}$$

le compartiment de droite va devenir plus chaud que celui de gauche. On pourra alors se servir des ces deux compartiments comme de deux sources de chaleur à des températures différentes pour faire fonctionner une machine thermique, obtenant ainsi du travail en partant d'une seule source de chaleur, en contradiction avec le second principe de la thermodynamique (de façon équivalente, on peut aussi remarquer que l'on a fabriqué un réfrigérateur sans moteur, ce qui est aussi interdit par le second principe).

En 1929, le problème a été réduit à sa plus simple expression par Szilard, qui a considéré un gaz limité à une seule molécule. Cette molécule peut être localisée dans l'un ou l'autre des compartiments sans dépense d'énergie, et elle fournit du travail en repoussant un piston jusqu'à occuper l'ensemble de l'enceinte, en prenant de l'énergie à l'extérieur sous forme de chaleur. L'expansion se faisant à température constante, le travail fourni est donné par

$$W_0 = k_B T \int_{V/2}^{V} \frac{\mathrm{d}V'}{V'} = k_B T \ln 2$$

où V est le volume de l'enceinte. On peut recommencer N fois l'opération et obtenir ainsi un travail arbitrairement grand $W = N W_0 = N k_B T \ln 2$, à partir d'une seule source de chaleur.

Le paradoxe a été élucidé par Bennett en 1982 : Bennett a remarqué que le dispositif *ne fonctionne pas suivant un cycle*, ce qui est la condition de validité du second principe, car la localisation de la molécule dans l'un ou l'autre des compartiments au cours des N opérations suppose que cette information soit stockée dans une mémoire de N bits. Si l'on veut effacer le contenu de cette mémoire pour repartir à zéro et effectuer un cycle complet, cela va rejeter dans l'environnement une entropie au moins égale à $N k_B \ln 2$, et donc dissiper dans l'environnement une énergie d'au moins $N k_B T \ln 2$, ce qui convertit tout le travail obtenu sous forme de chaleur.

4.2 Portes logiques quantiques

L'opération quantique la plus générale est une transformation unitaire dans l'espace de Hilbert de dimension 2^n des n qu-bits, $\mathcal{H}^{\otimes n}$: la porte logique la plus générale est une matrice $2^n \times 2^n$ opérant dans $\mathcal{H}^{\otimes n}$. Un théorème d'algèbre linéaire permet de se ramener aux opérations sur un qu-bit et sur deux qu-bits.

Théorème Toute transformation unitaire sur $\mathcal{H}^{\otimes n}$ peut se décomposer en produit de transformations unitaires sur un qu-bit et de portes cNOT.

Il est clair qu'opérer individuellement sur les qu-bits ne peut pas donner une transformation unitaire générique de $\mathcal{H}^{\otimes n}$, car une telle opération est de la forme d'un produit tensoriel

$$U = U^{(1)} \otimes U^{(2)} \otimes \cdots \otimes U^{(n)}$$

et il faut au minimum se donner des opérations non triviales sur deux qu-bits. Le théorème ci-dessus garantit que cela suffit. Il est utile de donner la représentation comme matrice 4×4 de l'opération cNOT, qui en termes de qu-bits correspond à

$$|00\rangle \to |00\rangle \quad |01\rangle \to |01\rangle \quad |10\rangle \to |11\rangle \quad |11\rangle \to |10\rangle$$

Dans la base $\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$, cette représentation matricielle est donc

$$\text{cNOT} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} I & 0 \\ 0 & \sigma_x \end{pmatrix} \tag{4.3}$$

Sous cette forme il est clair que cNOT ne peut pas être un produit tensoriel. La généralisation de la porte cNOT est la porte CONTROL-U (cU), où la matrice σ_x est remplacée par une matrice 2×2 unitaire U

$$\text{cU} = \begin{pmatrix} I & 0 \\ 0 & U \end{pmatrix}$$

Il existe une construction de cU à partie de la porte cNOT (figure 4.3). Il faut trouver trois opérateurs unitaires A, B, C tels que

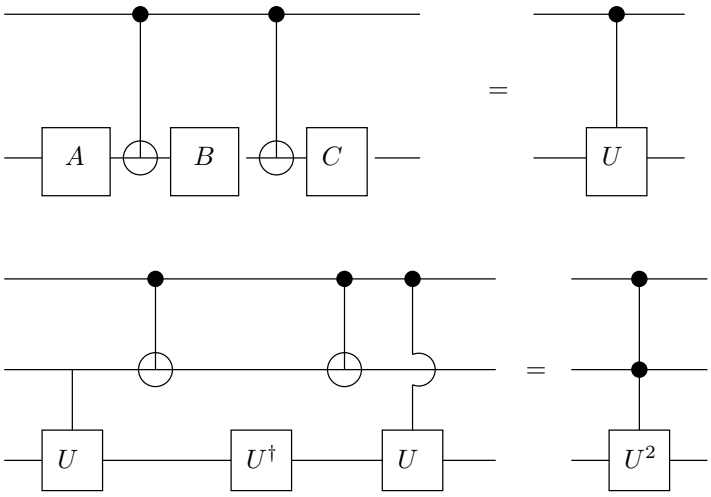

FIG. 4.3 – Construction de la porte cU et de la porte de Toffoli. Les diagrammes se lisent de gauche à droite, alors que les produits d'opérateurs s'effectuent de droite à gauche.

$$CBA = I \qquad C\sigma_x B\sigma_x A = U$$

La porte de Toffoli se construit à partir portes cU et de portes cNOT (figure 4.3) et de l'équation

$$\sqrt{\sigma_x} = \frac{1}{1+\mathrm{i}} \begin{pmatrix} 1 & \mathrm{i} \\ \mathrm{i} & 1 \end{pmatrix}$$

Compte tenu des résultats de la section 4.1, on voit que si l'on dispose d'un circuit logique classique permettant de calculer une fonction $f(x)$, alors on pourra construire un circuit quantique possédant essentiellement le même nombre de portes.

Exercices 4.2.1. (1) Justifier les circuits de la figure 4.3.

(2) Supposons que la mesure des qu-bits soit effectuée immmédiatement *après* une porte cU. Montrer que les probabilités de trouver le qu-bit cible dans les états $|0\rangle$ ou $|1\rangle$ et ses états finaux sont les mêmes que

si le bit de contrôle était mesuré *avant* la porte et que le bit cible était transformé ou non selon que le bit de contrôle a été trouvé dans l'état $|0\rangle$ ou dans l'état $|1\rangle$. Cette observation permet de remplacer la porte à deux qu-bits cU par une porte à un seul qu-bit sur le bit cible, ce qui est une grande simplification technologique. Mais ceci n'est valable qu'à la fin des calculs, pas sur une porte cU intermédiaire !

Sachant qu'il existe un circuit logique quantique capable d'évaluer une fonction $f(x)$, on peut maintenant avoir recours au parallélisme quantique. Commençons par le cas où le registre de données, celui de x, est un registre à un qu-bit, et de même pour $f(x)$. On construit une transformation U_f qui effectue les opérations

$$U_f \ : \ (x, y) \rightarrow (x, y \oplus f(x))$$

où $\oplus$ est l'addition modulo 2. Si la valeur initiale est $y = 0$, on a simplement

$$U_f \ : \ (x, 0) \rightarrow (x, f(x))$$

On peut se demander pourquoi on n'effectue pas tout simplement une transformation $x \rightarrow f(x)$. La réponse est que cette transformation ne peut pas être unitaire si la correspondance entre x et $f(x)$ n'est pas biunivoque, et elle ne convient pas pour un algorithme quantique. Au contraire il est facile de se convaincre que U_f est unitaire, car elle est de carré unité

$$U_f \ : \ (x, [y \oplus f(x)]) \rightarrow (x, [y \oplus f(x)] \oplus f(x)) = (x, y)$$

En effet $f(x) \oplus f(x) = 0$ quel que soit $f(x)$. La transformation U_f fait correspondre à un vecteur de base un autre vecteur de base, et comme $U_f^2 = I$, cette correspondance ne peut être qu'une simple permutation des vecteurs de base, et donc une transformation unitaire. En notation opératorielle

$$U_f|x \otimes 0\rangle = |x \otimes f(x)\rangle \qquad U_f|x \otimes y\rangle = |x \otimes [y \oplus f(x)]\rangle$$

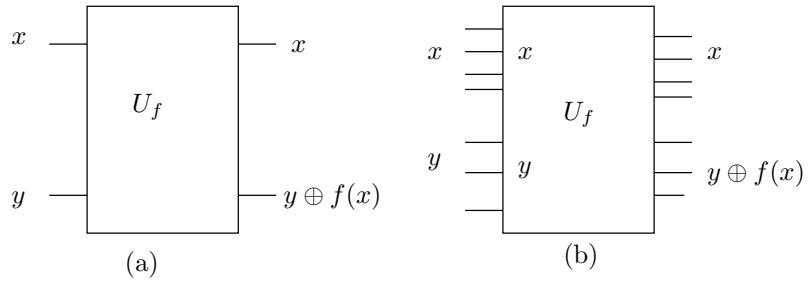

FIG. 4.4 – La construction U_f : (a) 2 qu-bits (b) $n + m$ qu-bits

Appliquons sur l'état $|0\rangle_x$ une transformation de Hadamard H (à ne pas confondre avec le hamiltonien $\hat{H}$!)

$$H = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \tag{4.4}$$

soit

$$H|0\rangle = \frac{1}{\sqrt{2}}\big(|0\rangle + |1\rangle\big)$$

Alors, si le second qu-bit est dans l'état initial $|0\rangle$, le vecteur d'état final des deux qu-bits est l'état intriqué

$$|\Psi\rangle = U_f \frac{1}{\sqrt{2}}(|0 \otimes 0\rangle + |1 \otimes 0\rangle) = \frac{1}{\sqrt{2}}\big(|0 \otimes f(0)\rangle + |1 \otimes f(1)\rangle\big) \tag{4.5}$$

Le vecteur d'état $|\Psi\rangle$ contient *à la fois* l'information sur $f(0)$ et sur $f(1)$.

La généralisation de ce résultat consiste à prendre un registre de données à n qu-bits et un registre de résultats[3] à m qu-bits, où m est le nombre de bits nécessaire pour écrire $f(x)$. Prenons comme exemple le cas $n = 3$. Dans la notation $|x\rangle$, le nombre x est un des huit nombres (en écriture binaire)

$$|000\rangle,\ |001\rangle,\ \ |010\rangle,\ \ |011\rangle,\ \ |100\rangle,\ \ |101\rangle,\ \ |110\rangle,\ \ |111\rangle$$

Si l'on se restreignait à des états de cette base, appelée *base de calcul (computational basis)*, pour l'entrée et la sortie (figure 4.1) en choisissant des transformations unitaires

$$|y\rangle = U|x\rangle$$

alors le calcul quantique se limiterait à simuler un ordinateur classique, ce qui ne serait pas très passionnant. La richesse de l'ordinateur quantique consiste à faire des combinaisons linéaire de vecteurs de la base de calcul grâce à l'opération H qui donne dans le cas particulier $n = 3$

$$H|000\rangle = \frac{1}{\sqrt{8}} \sum_{x=0}^{7} |x\rangle$$

En général

$$H|0_n\rangle = \frac{1}{2^{n/2}} \sum_{x=0}^{2^n-1} |x\rangle$$

x est une notation condensée pour la représentation binaire[4] du nombre x et le vecteur d'état $|x\rangle = |x_0x_1x_2\rangle$, où x_0, x_1, x_2 prenent les valeurs 0 ou 1. L'opération U_f est définie en généralisant la définition précédente par

$$U_f|x \otimes y\rangle = |x \otimes [y \oplus f(x)]\rangle$$

où $\oplus$ est l'addition modulo 2 *sans retenue*

$$1101 \oplus 0111 = 1010$$

Rappelons que

$$|x\rangle = |x_0x_1 \ldots x_{n-1}\rangle \qquad |y\rangle = |y_0y_1 \ldots y_{m-1}\rangle$$

avec $x_i, y_j = 0$ ou 1. Ceci assure que $U_f^2 = I$ et U_f est unitaire. Si l'on prend $|0_m\rangle$ comme état initial du registre de résultats, alors

$$U_f|x \otimes 0_m\rangle = |x \otimes f(x)\rangle$$

Si enfin on applique H sur le registre de données dans l'état $|0_n\rangle$ avant U_f, le vecteur d'état de l'état final sera par linéarité

$$|\Psi_{\text{fin}}\rangle = U_f|(H\,0_n) \otimes 0_m\rangle = \frac{1}{2^{n/2}} \sum_{x=0}^{2^n-1} |x \otimes f(x)\rangle \tag{4.6}$$

Ce vecteur d'état contient en principe 2^n valeurs de la fonction $f(x)$. Par exemple si $n = 100$, il contient $\sim 10^{30}$ valeurs de $f(x)$: c'est le miracle du "parallélisme quantique". Mais bien sûr une mesure ne donnera qu'une seule de ces valeurs. On peut cependant extraire des informations utiles sur des *relations* entre valeurs de $f(x)$ pour un ensemble de valeurs de x différentes, mais bien sûr au prix de la perte de ces valeurs individuelles, alors qu'un ordinateur classique devrait évaluer $f(x)$ pour toutes ces valeurs de x de façon indépendante. Nous allons en voir un exemple sur la transformation de Fourier quantique.

[3] J'ai traduit "input register" par "registre de données" et "output register" par "registre de résultats", plutôt que par "registre d'entrée" et "registre de sortie", afin d'éviter toute confusion avec les qu-bits à l'entrée du calcul au temps t_0 et à la sortie au temps t (figure 4.1).

[4] Il est commode de numéroter les n qu-bits $0, 1, \ldots, n-1$.

4.3 Transformation de Fourier quantique

Soit un nombre entier x, $0 \leq x \leq 2^n - 1$, écrit avec n bits

$$x = 0, 1, \cdots, 2^n - 1$$

et $|x\rangle$ un vecteur de la base de calcul

$$|x\rangle = |x_0 x_1 \cdots x_{n-1}\rangle \qquad x_i = 0 \text{ ou } 1$$

On définit une transformation unitaire[5] $U_{\rm FT}$ dont les éléments de matrice dans la base de calcul sont

$$\langle y|U_{\rm FT}\, x\rangle = \frac{1}{2^{n/2}}\, \mathrm{e}^{2\mathrm{i}\pi xy/2^n} \tag{4.7}$$

La transformation $U_{\rm FT}$ est réalisée physiquement dans la boîte $U_{\rm FT}$ de la figure 4.6(a), et comme nous allons le voir, un circuit possible est donné par la figure 4.6(b). Si $|\Psi\rangle$ est une combinaison linéaire normalisée de vecteurs $|x\rangle$

$$|\Psi\rangle = \sum_{x=0}^{2^n-1} f(x)|x\rangle \qquad \sum_{x=0}^{2^n-1} |f(x)|^2 = 1 \tag{4.8}$$

où $f(x) = \langle x|\Psi\rangle$, alors l'amplitude pour trouver à la sortie de la boîte $U_{\rm FT}$ un état de la base de calcul $|y\rangle$ (noter que $|y\rangle$ relève du registre de données et non de celui des résultats !) est d'après (1.17)

$$\begin{aligned} a(\Psi \to y) &= \langle y|\Psi\rangle = \sum_{x=0}^{2^n-1} \langle y|U_{\rm FT}\, x\rangle\langle x|\Psi\rangle \\ &= \frac{1}{2^{n/2}} \sum_{x=0}^{2^n-1} \mathrm{e}^{2\mathrm{i}\pi xy/2^n} f(x) = \tilde{f}(y) \end{aligned} \tag{4.9}$$

L'amplitude de probabilité $a(\Psi \to y)$ n'est donc pas autre chose que la transformée de Fourier discrète (ou sur réseau) $\tilde{f}(y)$ de $f(x)$.

Pour construire la boîte $U_{\rm FT}$, il est commode d'écrire $U_{\rm FT}|x\rangle$ sous la forme

$$U_{\rm FT}|x\rangle = \sum_{y=0}^{2^n-1} |y\rangle\langle y|U_{\rm FT}\, x\rangle = \frac{1}{2^{n/2}} \sum_{y=0}^{2^n-1} \mathrm{e}^{2\mathrm{i}\pi xy/2^n}\, |y\rangle \tag{4.10}$$

Je vais transformer (4.10) afin de l'écrire sous la forme d'un état manifestement non intriqué, en utilisant une technique standard des transformées de Fourier rapides. Soit

$$\begin{aligned} x &= x_0 + 2x_1 + \cdots + 2^{n-1}x_{n-1} \\ y &= y_0 + 2y_1 + \cdots + 2^{n-1}y_{n-1} \end{aligned} \tag{4.11}$$

la décomposition binaire de x et de y. Pour fixer les idées, on peut prendre l'exemple $n = 3$, $N = 8$; compte tenu de ce que $\exp(2\mathrm{i}\pi p) = 1$ pour p entier, on peut remplacer dans l'exponentielle de (4.10) le produit $xy/8$ par

$$\begin{aligned} \frac{xy}{8} &\to y_0\left(\frac{x_2}{2} + \frac{x_1}{4} + \frac{x_0}{8}\right) + y_1\left(\frac{x_1}{2} + \frac{x_0}{4}\right) + y_2\,\frac{x_0}{2} \\ &= y_0.x_2x_1x_0 + y_1.x_1x_0 + y_2.x_0 \end{aligned}$$

[5] En effet

$$\sum_{y=0}^{2^n-1} (U_{\rm FT}^\dagger)_{x'y}(U_{\rm FT})_{yx} = \sum_{y=0}^{2^n-1} (U_{\rm FT})^*_{yx'}(U_{\rm FT})_{yx} = \frac{1}{2^{n/2}} \sum_{y=0}^{2^n-1} \mathrm{e}^{2\mathrm{i}\pi(x'-x)y/2^n} = \delta_{x'x}$$

Le résultat s'obtient en remarquant que la somme sur y est une série géométrique.

où l'on a introduit la notation (représentation binaire d'un nombre inférieur à un)

$$.x_p x_{p-1} \cdots x_1 x_0 = \frac{x_p}{2} + \frac{x_1}{2^{p-1}} + \cdots + \frac{x_0}{2^p} \tag{4.12}$$

On peut alors factoriser la somme sur y en sommes sur $y_0, \ldots, y_{n-1}$, $y_i = 0$ ou 1, $|y\rangle = |y_0 \cdots y_{n-1}\rangle$

$$U_{\rm FT}|x\rangle = \frac{1}{2^{n/2}} \left(|0\rangle_0 + \mathrm{e}^{2\mathrm{i}\pi.x_{n-1}\cdots x_0}|1\rangle_0\right) \cdots \left(|0\rangle_{n-1} + \mathrm{e}^{2\mathrm{i}\pi.x_0}|1\rangle_{n-1}\right) \tag{4.13}$$

ce qui met manifestement $|x\rangle$ sous la forme d'un produit tensoriel. Donnons un exemple pour $n = 2$

$$\begin{aligned} U_{\rm FT}|x\rangle \equiv U_{\rm FT}|x_0 x_1\rangle &= \frac{1}{4}\left(|00\rangle + \mathrm{e}^{2\mathrm{i}\pi.x_0}|01\rangle + \mathrm{e}^{2\mathrm{i}\pi.x_1 x_0}|10\rangle + \mathrm{e}^{2\mathrm{i}\pi(.x_1 x_0 + .x_0)}|11\rangle\right) \\ &= \frac{1}{4}\left(|0\rangle_0 + \mathrm{e}^{2\mathrm{i}\pi.x_1 x_0}|1\rangle_0\right)\left(|0\rangle_1 + \mathrm{e}^{2\mathrm{i}\pi.x_0}|1\rangle_1\right) \end{aligned}$$

Par exemple si $|x\rangle = |01\rangle$, $x_0 = 0$, $x_1 = 1$

$$\begin{aligned} U_{\rm FT}|01\rangle &= \frac{1}{4}\left(|00\rangle + |01\rangle + \mathrm{e}^{\mathrm{i}\pi}\,|10\rangle + \mathrm{e}^{\mathrm{i}\pi}\,|11\rangle\right) \\ &= \frac{1}{4}(|0\rangle_0 + \mathrm{e}^{\mathrm{i}\pi}\,|1\rangle_0)(|0\rangle_1 + |1\rangle_1) \end{aligned}$$

Un circuit logique possible pour effectuer cette transformée de Fourier est donné dans la figure 4.4. La porte R_d est définie par l'opérateur R_d

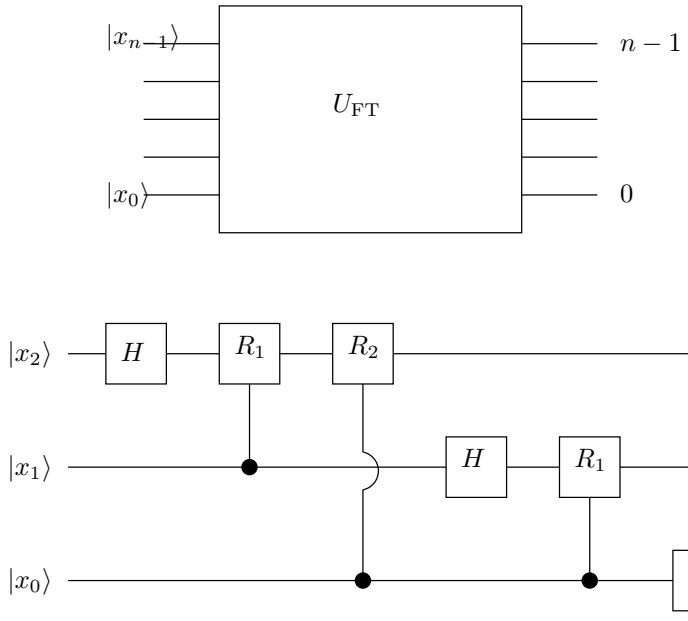

FIG. 4.5 – (a) Boîte $U_{\rm FT}$. (b) Circuit construisant $U_{\rm FT}$ dans le cas $n = 3$. La dernière opération (SWAP) consiste à permuter les bits 0 et 2.

$$R_d = \begin{pmatrix} 1 & 0 \\ 0 & \mathrm{e}^{\mathrm{i}\pi/2^d} \end{pmatrix} \tag{4.14}$$

En effet l'action de la porte H est

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

ce qui résume l'action sur le premier bit $|x_2\rangle$ par

$$H|x_2\rangle = \frac{1}{\sqrt{2}}\left(|0\rangle + \mathrm{e}^{2\mathrm{i}\pi.x_2}\,|1\rangle\right) \tag{4.15}$$

Notons $c_i R_d^j$ l'action sur le bit j de R_d contrôlé par le bit i ; alors

$$\begin{aligned} x_1 = 0 \qquad (\mathrm{c_1R_1^2})H|x_2\rangle &= \frac{1}{\sqrt{2}}\left(|0\rangle + \mathrm{e}^{2\mathrm{i}\pi.x_2}\,|1\rangle\right) \\ x_1 = 1 \qquad (\mathrm{c_1R_1^2})H|x_2\rangle &= \frac{1}{\sqrt{2}}\left(|0\rangle + \mathrm{e}^{2\mathrm{i}\pi.x_2}\,\mathrm{e}^{\mathrm{i}\pi/2}|1\rangle\right) \end{aligned}$$

ce qui se résume en

$$(\mathrm{c_1R_1^2})H|x_2\rangle = \frac{1}{\sqrt{2}}\left(|0\rangle + \mathrm{e}^{2\mathrm{i}\pi.x_2x_1}\,|1\rangle\right) \tag{4.16}$$

Il est clair que la procédure se poursuit par

$$(\mathrm{c_1R_2^2})(\mathrm{c_1R_1^2})H|x_2\rangle = \frac{1}{\sqrt{2}}\left(|0\rangle + \mathrm{e}^{2\mathrm{i}\pi.x_2x_1x_0}\,|1\rangle\right) \tag{4.17}$$

et on obtient l'état

$$|\Psi'\rangle = \frac{1}{\sqrt{8}}\left(|0\rangle_0 + \mathrm{e}^{2\mathrm{i}\pi.x_0}|1\rangle_0\right)\left(|0\rangle_1 + \mathrm{e}^{2\mathrm{i}\pi.x_1x_0}|1\rangle_1\right)\left(|0\rangle_2 + \mathrm{e}^{2\mathrm{i}\pi.x_2x_1x_0}|1\rangle_2\right)$$

Il faut donc permuter les bits 0 et 2 pour obtenir (4.13). Cette dernière opération n'est pas indispensable, il suffit de renuméroter les qu-bits à la sortie. Le nombre de portes nécessaire se décompose en n portes H et en

$$n + (n-1) + \cdots + 1 \simeq \frac{1}{2}n^2$$

portes conditionnelles, soit $\mathcal{O}(n^2)$ portes.

4.4 Période d'une fonction

L'algorithme de factorisation de Shor repose sur la possibilité de trouver "rapidement", c'est-à-dire en un temps polynômial en n, la période d'une fonction $f(x)$, dans le cas de Shor la fonction $a^x \bmod N$. Soit donc une fonction $f(x)$ de période r, $f(x) = f(x+r)$, avec

$$x = 0, 1, \cdots, 2^n - 1 \tag{4.18}$$

La réussite de l'algorithme suppose que $2^n > N^2$. Un algorithme classique utilise $\mathcal{O}(N)$ opérations élémentaires (la fonction $a^x \bmod N$ donne l'impression d'un bruit aléatoire sur une période), mais l'algorithme quantique décrit ci-dessous utilise seulement $\mathcal{O}(n^3)$ opérations élémentaires. La variable x est stockée dans un registre $|x\rangle$ et la fonction $f(x)$ dans un registre $|z\rangle$ correspondant à m qu-bits. On part de l'état initial de $n+m$ qu-bits

$$|\Phi\rangle = \frac{1}{2^{n/2}}\left(\sum_{x=0}^{2^n-1}|x\rangle\right) \otimes |0\cdots 0\rangle \tag{4.19}$$

On utilise ensuite la boîte U_f qui calcule la fonction $f(x)$

$$|\Psi_f\rangle = U_f|\Phi\rangle = \frac{1}{2^{n/2}}\sum_{x=0}^{2^n-1}|x \otimes f(x)\rangle \tag{4.20}$$

Ceci demande $\mathcal{O}(n)$ opérations. Si l'on mesure le registre de résultats et que l'on trouve le résultat f_0, le vecteur d'état du registre de données est après cette mesure

$$|\Psi\rangle = \frac{1}{\mathcal{N}}\sum_{x;f(x)=f_0}|x\rangle \tag{4.21}$$

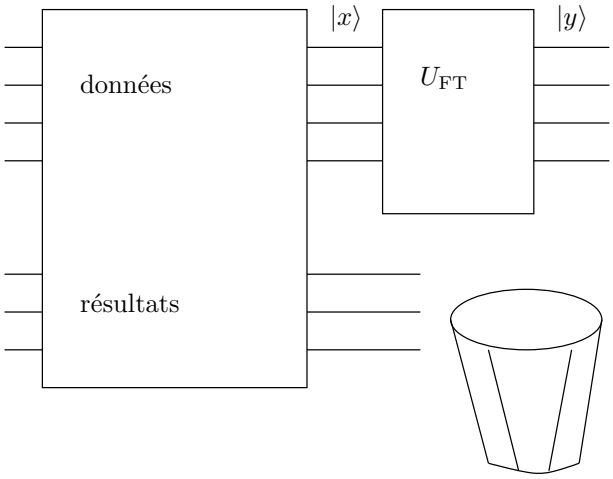

FIG. 4.6 – Schéma du calcul déterminant la période. Les qu-bits du registre de résultats sont mis à la poubelle.

où la somme porte sur les valeurs de x telles que $f(x) = f_0$ et $\mathcal{N}$ est un facteur de normalisation. Revenons au cas d'une fonction périodique : je supposerai que $f(x+s) = f(x)$ implique que $s = pr$, p entier, autrement dit que la fonction $f(x)$ ne prend jamais deux fois la même valeur sur une période, ce qui est le cas de la fonction $a^x \bmod N$. Le vecteur normalisé $|\Psi\rangle$ du registre de données est alors, avec $f(x_0) = f_0$ et x_0 la plus petite des valeurs de x telle que $f(x_0) = f_0$

$$|\Psi\rangle = \frac{1}{\sqrt{K}} \sum_{k=0}^{K-1} |x_0 + kr\rangle \tag{4.22}$$

où[6] $K \simeq 2^n/r$. *En réalité il n'est pas nécessaire de faire une mesure du registre de résultats.* À la sortie de la boîte U_f de la figure 4.6, les qu-bits du registre de données et ceux du registre de résultats sont intriqués (voir (4.20)), et si l'on observe seulement les qu-bits du registre de données, il faut prendre la trace sur le registre de résultats pour obtenir l'opérateur densité des qu-bits du registre de données : l'état physique des qu-bits du registre de données sera en général décrit par un opérateur densité, et non par un vecteur de $\mathcal{H}^{\otimes n}$. En d'autres termes, l'état physique du registre de données est une superposition

[6] Plus précisément $K = [2^n/r]$ ou bien $K = [2^n/r] + 1$, où $[z]$ désigne la partie entière de z.

incohérente [7] de vecteurs $|\Psi_i\rangle$

$$|\Psi_i\rangle = \frac{1}{\sqrt{K_i}} \sum_{k=0}^{K_i-1} |x_i + kr\rangle \tag{4.23}$$

où $f(x_i) = f_i$ et x_i la plus petite des valeurs de x telle que $f(x_i) = f_i$. Comme le raisonnement ci-dessous ne dépend pas de x_i, on peut parfaitement se passer de la mesure du registre de résultats : en d'autres termes il est tout à fait inutile de faire appel au postulat de réduction du paquet d'ondes.

Le vecteur d'état (4.22) correspond dans (4.8) au choix $f(x) = 1/\sqrt{K}$ si x est de la forme $x_0 + kr$ et $f(x) = 0$ dans le cas contraire. D'après (4.9), l'amplitude $a(\Psi \to y$ est donc

$$a(\Psi \to y) = \frac{1}{2^{n/2}} \frac{1}{\sqrt{K}} \sum_{k=0}^{K-1} \mathrm{e}^{2\mathrm{i}\pi y(x_0+kr)/2^n} \tag{4.24}$$

et la probabilité de mesurer la valeur y (c'est-à-dire de trouver l'état $|y_0 y_1 \dots y_{n-1}\rangle$ de la base de calcul à la sortie de la boîte U_{FT}) est donc

$$\mathsf{p}(y) = |a(\Psi \to y)|^2 = \frac{1}{2^n K} \Big| \sum_{k=0}^{K-1} \mathrm{e}^{2\mathrm{i}\pi kry/2^n} \Big|^2 \tag{4.25}$$

On constate que $\mathsf{p}(y)$ est indépendant de x_i, et on aurait pu partir de n'importe lequel des vecteurs $|\Psi_i\rangle$ de (4.23). On utilise ensuite la série géométrique[8]

$$\sum_{k=0}^{K-1} \mathrm{e}^{2\mathrm{i}\pi ykr/2^n} = \frac{1 - \mathrm{e}^{2\mathrm{i}\pi yKr/2^n}}{1 - \mathrm{e}^{2\mathrm{i}\pi yr/2^n}} = \mathrm{e}^{\mathrm{i}\pi(K-1)r/2^n} \ \frac{\sin(\pi yKr/2^n)}{\sin(\pi yr/2^n)}$$

et on écrit, avec j entier

$$y_j = j\frac{2^n}{r} + \delta_j \tag{4.26}$$

[7] Formellement, l'opérateur densité *total* (données + résultats) ρ_{tot} est, d'après (4.20)

$$\rho_{\mathrm{tot}} = \frac{1}{2^n} \sum_{x,z} |x \otimes f(x)\rangle\langle z \otimes f(z)|$$

L'opérateur densité du registre de données s'obtient en prenant la trace partielle sur le registre de résultats (voir la note 3 du chapitre 3 pour la technique de calcul)

$$\rho_{\mathrm{don}} = \mathrm{Tr}_{\mathrm{res}} \rho_{\mathrm{tot}} = \frac{1}{2^n} \sum_{x,z} |x\rangle\langle z| \ \langle f(z)|f(x)\rangle$$

Supposons que la fonction $f(x)$ prenne N_0 fois la valeur f_0 et N_1 fois la valeur f_1, $N_0 + N_1 = 2^n$. Alors

$$\rho_{\mathrm{don}} = \frac{1}{2^n} \left(\sum_{x,z;f(x)=f(z)=f_0} |x\rangle\langle z| + \sum_{x,z;f(x)=f(z)=f_1} |x\rangle\langle z| \right)$$

car $\langle f(x)|f(z)\rangle = 1$ si $f(x) = f(z)$ et $\langle f(x)|f(z)\rangle = 0$ si $f(x) \neq f(z)$. Ceci correspond à une superposition incohérente avec des probabilités $\mathsf{p}_0 = N_0/2^n$ et $\mathsf{p}_1 = N_1/2^n$ de vecteurs normalisés

$$|\Psi_0\rangle = \frac{1}{\sqrt{N_0}} \sum_{x;f(x)=f_0} |x\rangle \qquad |\Psi_1\rangle = \frac{1}{\sqrt{N_1}} \sum_{x;f(x)=f_1} |x\rangle$$

Dans le cas de la fonction périodique qui nous intéresse

$$\rho_{\mathrm{don}} = \frac{1}{2^n} \sum_{i=0}^{r-1} \sum_{k_i,k_j=0}^{K_i-1} |x_i + k_i r\rangle\langle x_i + k_j r|$$

où x_i est la plus petite des valeurs de x pour laquelle $f(x) = f_i$.

[8] Le problème est évidemment analogue à celui de la diffraction, par exemple la diffraction de neutrons par un cristal. Si a est la distance entre deux sites ($a = 1$ dans le texte), la maille du réseau est ra. Le (quasi-)vecteur d'onde q peut prendre les valeurs $q = 2\pi p/(2^n a)$, $p = 0, 1, \cdots, 2^n - 1$ (p et $p' = p + 2^n$ sont équivalents). Les pics de diffraction se produisent lorsque q est un multiple entier de la maille $2\pi/(ra)$ du réseau réciproque, soit $q = j2\pi/(ra)$, $j = 0, 1, \cdots, r - 1$.

ce qui donne la probabilité $\mathsf{p}(y_j)$

$$\mathsf{p}(y_j) = \frac{1}{2^n K} \frac{\sin^2(\pi\delta_j K r/2^n)}{\sin^2(\pi\delta_j r/2^n)} \tag{4.27}$$

Si y_j est multiple entier de $2^n/r$, $y_j = j2^n/r$, $\mathsf{p}(y_j)$ prend sa valeur maximale $K/2^n$; en général la fonction $\mathsf{p}(y)$ a des maxima aigus lorsque la valeur de y est proche de $j2^n/r$. En utilisant l'encadrement de $\sin x$

$$\frac{2}{\pi}\, x \leq \sin x \leq x \qquad 0 \leq x \leq \frac{\pi}{2}$$

on montre que si $|\delta_j| < 1/2$, la probabilité de trouver une des valeurs (4.26) est au moins de $4/\pi^2$

$$\mathsf{p}(y_j) \geq \frac{4}{\pi^2}\frac{K}{2^n} = \frac{4}{\pi^2}\frac{1}{r}$$

Comme $0 \leq j \leq r-1$ et que $r \gg 1$, il y a au moins 40% de chances ($4/\pi^2 \simeq 0.406$) de trouver une valeur de y_j proche de $j2^n/r$. De plus, comme $|\delta_j| \leq 1/2$

$$\Big|\frac{y_j}{2^n} - \frac{j}{r}\Big| \leq \frac{1}{2^{n+1}} \tag{4.28}$$

Comme n et y_j sont connus (y_j est un nombre entier $0 \leq y_j \leq 2^n - 1$ qui est le résultat de la mesure du registre de données), nous avons donc une estimation de la fraction j/r. Montrons maintenant que la mesure de y_j permet de déterminer j et r (toujours avec une probabilité d'au moins 40%). Supposons que nous fassions varier y_j d'une unité ; d'après (4.28) qui implique

$$|y_j r - j2^n| \leq \frac{r}{2}$$

nous avons

$$|(y_j \pm 1)r - j2^n| \geq \frac{r}{2}$$

et donc

$$\Big|\frac{y_j \pm 1}{2^n} - \frac{j}{r}\Big| \geq \frac{1}{2^{n+1}}$$

ce qui est en contradiction avec (4.28), et la valeur de y_j est bien déterminée. Grâce à notre choix $2^n > N^2$ qui implique $2^n > r^2$, nous avons obtenu une estimation de j/r qui diffère de la valeur exacte par moins de $1/(2r^2)$. La valeur de j/r comme fraction irréductible $j_0/r_0 = j/r$ peut alors être extraite d'un développement en fractions continues. Si l'on a de la chance et que j et r sont premiers entre eux, on en déduit $j = j_0$ et $r = r_0$, ce que l'on vérifie en calculant $f(x)$ et $f(x + r_0)$ avec un ordinateur classique. La probabilité que deux grands nombres soient premiers entre eux est plus grande que $1/2$ (en fait $6/\pi^2$), et le succès est au rendez-vous plus d'une fois sur deux. Si $f(x) \neq f(x + r_0)$, on peut essayer les premiers multiples de r_0, $2r_0, 3r_0, \cdots$, et si ces essais ne donnent rien, il faut recommencer toute l'opération, qui prend $\mathcal{O}(n^3)$ opérations élémentaires $\mathcal{O}(n^2)$ pour la transformation de Fourier et $\mathcal{O}(n)$ pour le calcul de a^x.

La détermination de la période r suffit à casser le code RSA. En effet (annexe 1.6.1), Ève dispose du message chiffré d'Alice, b, et des nombres N et c,qui sont diffusés publiquement. Elle calcule d' comme $cd' \equiv 1 \bmod r$ et ensuite $b^{d'} \bmod N$

$$b^{d'} = a^{cd'} = a^{1+mr} = a(a^r)^m \equiv a \bmod N$$

car $a^r \equiv 1 \bmod N$, et Ève récupère le message original a.

Si l'on veut en plus factoriser N il faut écrire

(i)

$$a^r - 1 = \left(a^{r/2} - 1\right)\left(a^{r/2} + 1\right) \equiv\ 0 \bmod N$$

(ii)

$$a^{r/2} \not\equiv\ \pm 1 \bmod N$$

Si on a de la chance, que (i) r est entier et que (ii) est vérifié, alors le produit de nombres entiers

$$\left(a^{r/2}-1\right)\left(a^{r/2}+1\right)$$

est divisible par $N = pq$. Il est donc nécessaire que p divise $\left(a^{r/2}-1\right)$ et q divise $\left(a^{r/2}+1\right)$. Les valeurs de p et q sont obtenues en cherchant les pgcd

$$p = \text{pgcd}\ \left(N, a^{r/2}-1\right) \qquad q = \text{pgcd}\ \left(N, a^{r/2}+1\right)$$

Si l'on n'a pas de chance, il faut recommencer, mais la probabilité de réussite est de plus de 50%.

4.5 Réalisations physiques

On en est encore aux premiers balbutiments de réalisations physiques d'ordinateurs quantiques, et les dispositifs dont la liste figure ci-dessous ont réussi au mieux à intriquer deux qu-bits (et encore !), à l'exception de la RMN qui est allée jusqu'à 7 qu-bits. Il est tout à fait prématuré d'essayer de prévoir aujourd'hui quel dispositif sera effectivement utilisé pour un ordinateur quantique pouvant traiter plusieurs centaines de qu-bits (s'il en existe un jour), vraisemblablement aucun de ceux listés ci-dessous. Cela dit, il serait aussi présomptueux d'affirmer qu'un tel ordinateur ne fonctionnera pas en 2050 que d'affirmer le contraire.

L'ennemi numéro un de l'ordinateur quantique est l'interaction avec l'environnement, qui conduit au phénomène de *décohérence*, la perte de la phase dans la superposition linéaire de qu-bits. Les calculs doivent être effectués en un temps inférieur au temps de décohérence τ_D. Si une opération élémentaire (porte logique) sur un qu-bit prend un temps τ_{op}, la figure de mérite d'un ordinateur quantique est le rapport

$$n_{\text{op}} = \frac{\tau_D}{\tau_{\text{op}}}$$

le nombre maximum d'opérations que l'ordinateur quantique peut effectuer. Les dispositifs imaginés jusquà présent sont (liste non exhaustive) :

- l'ordinateur quantique photonique exploitant l'effet Kerr non linéaire ;
- les cavités optiques résonantes ;
- les cavités micro-ondes résonantes ;
- les pièges à ions ;
- la RMN ;
- les jonctions Josephson ;
- les points quantiques ;
- les atomes provenant d'un condensat de Bose-Einstein piégés dans un réseau optique.

Le record du nombre de qu-bits appartient à la RMN, avec 7 qu-bits, ce qui est le nombre minimum de qu-bits nécessaire pour appliquer l'algorithme de Shor à la factorisation de 15. En effet a peut prendre les valeurs 2, 4, 7, 8, 11 ou 13, et la plus grande période de a^x mod N est de $r = 4$. Pour voir deux périodes, il faut donc prendre $x = 0, 1, \ldots, 7 = 2^3 - 1$, et bien sûr $f(x) = 0, 1, \ldots, 15 = 2^4 - 1$, soit un registre de données à 3 qu-bits et un registre de résultats à 4 qu-bits. L'avantage de la RMN est de pouvoir intriquer des qu-bits de façon relativement simple en les manipulant avec un champ magnétique externe en raison de la forme de l'interaction entre qu-bits. L'expérience de la factorisation de 15 a été menée avec succès en 2001. Malgré ce résultat spectaculaire (?), la RMN n'est pas une solution d'avenir, car il faut synthétiser une molécule possédant autant de sites discernables que de qu-bits, et surtout parce que le signal décroît exponentiellement avec le nombre de qu-bits. En effet la RMN n'utilise pas des objets quantiques individuels, mais un ensemble de $\sim 10^{18}$ molécules actives : le signal est un signal *collectif*, et c'est ce qui conduit à la diminution du rapport signal/bruit lorsque le nombre de qu-bits augmente.