De la sphère de Poincaré au chiffrement d'information quantique

Martin Courchesne
École d'informatique
Université McGill, Montréal
Octobre 2005

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfilment of the requirements of the degree of Master

Remerciements

Je désire remercier Claude Crépeau pour son soutien indéfectible, Valérie Poulain et Thomas Pedersen pour les discussions fructueuses que nous avons eues, Geneviève Arboit pour la révision du texte et Patrick Hayden pour ses précieux conseils. Je tiens également à remercier Simon-Pierre Desrosiers pour l'aide incroyable qu'il m'a apportée.

Résumé

Nous présentons dans ce mémoire plusieurs caractéristiques de la sphère de Poincaré, aussi appelée sphère de Bloch. Nous décrivons comment calculer la matrice de rotation de $\mathbb{R}^3$ qui correspond à une transformation unitaire donnée, et inversement, la matrice unitaire correspondant à chaque rotation. Nous montrons que tout ensemble d'états quantiques contenu dans un plan à travers la sphère de Poincaré peut être chiffré parfaitement à l'aide d'un bit de clé, et que les autres ensembles requièrent deux bits. Nous investiguons également le cas des espaces de dimension supérieure à deux. Nous étudions enfin la généralité de notre modèle de chiffre, le canal privé quantique, ou CQP. Nous introduisons la notion de CQP unitaire à ancille variable, et nous montrons que tout modèle de chiffre parfait se réduit à ce nouveau modèle. Nous montrons aussi que, peu importe le modèle utilisé, il faut $2n$ bits de clé pour chiffrer n qubits.

Abstract

In this thesis, we present many facts pertaining to the Poincaré sphere, also called the Bloch sphere. We demonstrate how to compute the $\mathbb{R}^3$ rotation matrix that corresponds to a given unitary matrix and, inversely, the unitary matrix corresponding to each rotation matrix. We show that any set of quantum states that is contained in a plane through the Bloch sphere can be encrypted perfectly using only one bit of key, while two bits are required to encrypt any other set. We also try to find equivalent statements for spaces of dimension higher than two. Last, we study the generality of our encryption model, the private quantum channel, or PQC. We introduce the notion of unitary PQC with variable ancilla, and show that any model for a perfect quantum cypher reduces to this new model. We also show that, using any model, $2n$ bits of key are always required to encrypt n qubits.

Introduction

La recherche qui a mené à l'écriture de ce mémoire provient d'une idée lancée en 2003 à l'Institut de recherche Bellairs de McGill, à la Barbade, lors de l'atelier de cryptographie quantique qu'y organise annuellement Claude Crépeau. On y présentait des résultats sur le chiffrement d'information quantique. Un des présentateurs, Alain Tapp, décrivait la notion de canal quantique privé (CQP), l'équivalent quantique du chiffre de Vernam. Dans ce modèle, il apparaît que deux bits (classiques) de clé doivent être partagés au préalable par Alice et Bob pour chaque qubit qu'Alice veut transmettre de manière confidentielle.

Lors d'une discussion informelle, à laquelle participaient notamment Claude Crépeau, Daniel Gottesman, Patrick Hayden, Louis Salvail, Alain Tapp et Andreas Winter, on remarqua qu'en restreignant l'ensemble des messages en clair qu'Alice pouvait chiffrer — traditionnellement, on lui permettait tout état quantique ρ d'une dimension donnée —, on observait un phénomène intrigant : on trouvait des ensembles requérant un, $\log 3$ ou deux bits de clé pour être chiffrés parfaitement, mais aucun ensemble ne semblait nécessiter une quantité de clé différente de $\log n$ bits, pour un nombre entier n quelconque. On émit donc l'hypothèse que n'importe quel ensemble pourrait être chiffré avec $\log n$ bits de clé, où n dépendrait de la structure de l'ensemble des messages en clair.

L'auteur de ce mémoire, présent lors de cette discussion, suivit le conseil de Claude

Crépeau, son directeur de recherche, et fit des questions qui venaient d'être soulevées son principal sujet de recherche. L'objectif premier était de découvrir une règle qui permettrait, étant donné un quelconque ensemble d'états Ω, de connaître la quantité de clé requise pour chiffrer Ω. Si cette valeur permet en quelque sorte de quantifier l'information que peut contenir l'ensemble Ω, alors on s'attendait à observer de véritables quanta d'information — car, pensait-on, seulement les quantités discrètes $\log n$ seraient permises.

Les réponses à ces questions demeurent très incomplètes, mais les investigations entreprises pour les trouver ont mené à d'autres résultats. Le premier exemple qui avait été donné d'ensembles pouvant être chiffrés à l'aide d'un seul bit de clé avait été illustré géométriquement : il s'agissait de considérer des cercles sur la sphère de Poincaré, qui est une représentation dans $\mathbb{R}^3$ des états quantiques dans un espace de dimension deux. À mesure que sa recherche avançait, l'auteur a vite compris qu'une compréhension approfondie de ce qu'est la sphère de Poincaré était certainement nécessaire à l'avancement de sa recherche.

Le deuxième chapitre, qui devait au départ être très court et ne servir qu'à introduire les concepts nécessaires à la compréhension du chapitre 3, est donc devenu un exposé détaillé de tout ce que l'auteur a pu découvrir au sujet de cette sphère au cours de sa recherche. On y montre entre autre — c'était déjà connu — qu'il existe une correspondance entre, d'une part, les transformations unitaires sur les états quantiques et, d'autre part, les rotations dans $\mathbb{R}^3$. L'originalité de ce chapitre vient du fait que cette correspondance sera donnée explicitement : étant donné une matrice unitaire, on montre comment calculer la rotation associée, et vice versa.

L'auteur s'est beaucoup interrogé sur l'existence d'une sphère de Poincaré pour les espaces de dimensions supérieures à deux. Il fut enthousiasmé de découvrir l'existence d'un article de Paolo Zanardi, [Zan98], qui confirmait l'état de sa recherche à

ce sujet et la poussait plus loin.

Les emprunts à Zanardi et à tout ouvrage de référence seront dûment notés. Plusieurs propositions présentés au cours du texte font partie du folklore de l'informatique quantique. Peu de résultats seront vraiment nouveaux. Par contre, toutes les démonstrations, à moins d'indication contraire, sont l'œuvre de l'auteur.

Le troisième chapitre tente de répondre à la question qui fut le moteur de cette recherche, à savoir quelle quantité de clé il est nécessaire d'utiliser pour chiffrer l'ensemble arbitraire d'états Ω. La question sera complètement résolue pour les Ω qui sont sous-ensembles d'un espace de dimension deux (un qubit) : un bit suffit si la représentation de Ω dans $\mathbb{R}^3$ est contenue dans un plan à travers la sphère de Poincaré ; autrement, deux bits sont nécessaires. On montre aussi qu'un bit est toujours nécessaire et que deux bits suffisent toujours. Pour les espaces à plus haute dimension, quelques cas particuliers seront résolus. Pour le cas général, une condition sera énoncée, à laquelle devra répondre tout chiffre $\mathcal{E}$ de l'ensemble Ω.

Les résultats du chapitre trois sont tous des résultats originaux, développés tantôt par l'auteur seul, tantôt en collaboration avec des collègues. On retrouvera, au fil du texte, des notes indiquant chaque collaboration.

Le quatrième et dernier chapitre s'interroge sur le modèle utilisé pour décrire les canaux privés quantiques. Suivant l'exemple de [AMTdW00], il était présumé, dans les chapitres précédents, que l'opérateur de chiffrement devait avoir une forme particulière. Le chapitre quatre devait être une démonstration du fait que ce modèle est aussi général que possible. Devant l'impossibilité de démontrer ce fait, l'auteur a développé la notion de CQP unitaire à ancille variable, un modèle légèrement plus général que le précédent. La notion de CQP général sera aussi introduite, pour inclure tous les opérateurs de chiffrement possibles. Il sera enfin démontré qu'à tout CQP général correspond un CQP unitaire à ancille variable équivalent.

Cet ultime chapitre présente aussi une preuve que $2n$ bits de clé sont nécessaires pour chiffrer n qubits, même dans le modèle de CQP général. Il avait été montré dans [DHT03] que $2n$ bits classiques étaient requis pour dissimuler[1] n qubits. On y laissait aussi entendre que la même preuve pourrait s'appliquer au chiffrement d'information quantique. C'est ce qui a été fait dans ce chapitre. L'auteur s'est aussi aidé des exercices donnés par Patrick Hayden dans le cadre de son cours d'informatique quantique, donné à l'hiver 2005 à l'université McGill.

[1]Nous utilisons le terme «dissimulation» pour traduire l'expression anglaise «data hiding».

Chapitre 1

Introduction à l'informatique quantique

Dans ce chapitre seront formulées les notions de base de l'informatique quantique. Beaucoup d'éléments sont empruntés à [NC00], mais nous verrons uniquement les notions nécessaires à la compréhension du présent document. Le traitement condensé qui en est fait sera idéal pour le lecteur qui désire rafraîchir sa mémoire et se familiariser avec la notation. Par contre, le lecteur pour qui une partie significative des notions présentées ici serait complètement nouvelle est invité à consulter des ouvrages de référence plus détaillés.

1.1 États quantiques et notation «braket»

Définition 1.1.1. *Un **espace de Hilbert** est un espace vectoriel complexe muni du produit interne canonique* $< \sum_i z_i \vec{e_i} \ , \ \sum_j w_j \vec{f_j} >= \sum_{i,j} \bar{z_i} w_j < \vec{e_i} \ , \ \vec{f_j} >$*, où* $\bar{z}$ *dénote le conjugué complexe de* z*. On dénote l'espace de Hilbert de dimension* d *par* H^d*.*

L'importance des espaces de Hilbert est la suivante : la mécanique quantique postule qu'à tout système physique isolé est associé un espace de Hilbert H^d. L'état d'un tel système peut être décrit complètement par un vecteur de norme un dans H^d.

L'espace des états quantiques dans H^d sera dénoté par H_1^d, où l'indice 1 rappelle la norme des vecteurs contenus dans ce sous-ensemble. Nous appellerons ces états des «états purs», par opposition aux états mélangés, un concept que nous définirons bientôt. De la même façon qu'on trace une flèche au-dessus d'une lettre qui dénote un vecteur, par exemple $\vec{v}$, nous inscrirons dans un **ket** un symbole qui représente un vecteur de norme un, comme ceci : $|\psi\rangle$. Les d vecteurs formant la base canonique de H^d seront représentés ainsi : $|0\rangle, |1\rangle, \ldots, |d-1\rangle$.

L'unité de base de l'information quantique est le **qubit**. Tandis qu'un bit classique peut prendre les valeurs 0 ou 1, un qubit peut prendre pour valeur n'importe quel vecteur de H_1^2, en particulier $|0\rangle$ ou $|1\rangle$.

On peut toujours représenter la valeur $|\psi\rangle \in H_1^2$ d'un qubit de la façon suivante :

$$|\psi\rangle = \alpha\,|0\rangle + \beta\,|1\rangle = \begin{bmatrix} \alpha \\ \beta \end{bmatrix},$$

où $\alpha, \beta \in \mathbb{C}$ et $|\alpha|^2 + |\beta|^2 = 1$.

Un **bra** est un ket conjugué et transposé :

$$\langle\psi| \stackrel{\text{def}}{=} |\psi\rangle^\dagger = \begin{bmatrix} \bar{\alpha} & \bar{\beta} \end{bmatrix}.$$

De façon générale, $A^\dagger \stackrel{\text{def}}{=} (\bar{A})^T$ est donnée par la matrice A conjuguée, puis transposée.

Puisque le ket $|\psi\rangle$ est un vecteur colonne et le bra $\langle\varphi|$ est un vecteur ligne, la juxtaposition d'un bra et d'un ket, $\langle\varphi|\psi\rangle$, pourra être interprétée comme un produit

matriciel, ou encore comme le produit interne des vecteurs $|\varphi\rangle$ et $|\psi\rangle$, puisque les deux interprétations concordent.

Voici quelques états dont il sera souvent question dans les pages à venir :

$$|0\rangle \stackrel{\text{def}}{=} \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \qquad |1\rangle \stackrel{\text{def}}{=} \begin{bmatrix} 0 \\ 1 \end{bmatrix},$$

$$|\nearrow\rangle \stackrel{\text{def}}{=} \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \qquad |\searrow\rangle \stackrel{\text{def}}{=} \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix},$$

$$|\circlearrowleft\rangle \stackrel{\text{def}}{=} \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ i \end{bmatrix}, \qquad |\circlearrowright\rangle \stackrel{\text{def}}{=} \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -i \end{bmatrix}.$$

1.2 Les fonctions trace et Vec

La **trace** d'une matrice est la somme de ses éléments diagonaux. On la définit comme suit :

$$\text{tr}\,(A) \stackrel{\text{def}}{=} \sum_{i=0}^{d-1} \langle i| A |i\rangle,$$

où A est une matrice $d \times d$ et $\{|i\rangle\}$ est une base orthonormale de H^d. Voici quelques propriétés utiles de la trace :

$$\begin{aligned} &\text{tr}\,(A+B) = \text{tr}\,(A) + \text{tr}\,(B), \\ &\text{tr}\,(\lambda A) = \lambda \text{tr}\,(A), \\ &\text{tr}\,(AB) = \text{tr}\,(BA). \end{aligned}$$

$\text{Vec}\,(A)$ est un vecteur colonne composé de toutes les colonnes de A placées les unes sous les autres. Si A est de taille $m \times n$ et qu'on dénote ses éléments par $a_{i,j}$, alors

$$\text{Vec}\,(A) \stackrel{\text{def}}{=} \Big[a_{1,1}, \ldots, a_{m,1}, a_{1,2}, \ldots, a_{m,2}, \ldots, a_{1,n}, \ldots, a_{m,n}\Big]^T.$$

Tout comme la trace, Vec est une transformation linéaire :

$$\begin{aligned}\mathrm{Vec}\,(A+B) &= \mathrm{Vec}\,(A) + \mathrm{Vec}\,(B)\,,\\ \mathrm{Vec}\,(\lambda A) &= \lambda \mathrm{Vec}\,(A)\,.\end{aligned}$$

Observation intéressante : le **produit interne de Hilbert-Schmidt** entre deux matrices A et B de taille $n \times n$, donné par $< A, B > \stackrel{\text{def}}{=} \mathrm{tr}\left(A^\dagger B\right)$, peut être récrit en ces termes :

$$< A, B > = < \mathrm{Vec}\,(A)\,, \mathrm{Vec}\,(B) >,$$

le terme de droite étant le produit interne canonique de H^{n^2}.

L'identité suivante est aussi digne de mention. La démonstration n'a pu en être trouvée dans aucun ouvrage de référence — c'est pourquoi elle est faite ici. On peut en sauter la lecture sans compromettre sa compréhension du reste du texte.

Proposition 1.2.1. *Si les dimensions des matrices A, B et C sont telles que le produit ABC existe, alors* $\mathrm{Vec}(ABC) = (C^T \otimes A)\mathrm{Vec}(B)$.[1]

Démonstration. Supposons que A soit une matrice $m \times n$, B une matrice $n \times p$ et C une matrice $p \times q$. Trouvons d'abord une expression décrivant l'élément à l'intersection de la i^{e} ligne et de la j^{e} colonne de la matrice ABC :

$$\begin{aligned}(ABC)_{ij} &= \mathrm{lgn}_i(AB)\mathrm{col}_j(C)\\ &= \mathrm{lgn}_i(A)B\mathrm{col}_j(C)\\ &= \sum_{k=1}^{p} \mathrm{lgn}_i(A)\mathrm{col}_k(B)C_{jk}.\end{aligned}$$

Pour $0 \leq j \leq q-1, 1 \leq i \leq m$, on peut maintenant écrire le $(mj+i)^{\mathrm{e}}$ élément de $\mathrm{Vec}(ABC)$ comme

$$\mathrm{Vec}(ABC)_{mj+i} = \sum_{k=1}^{p} \mathrm{lgn}_i(A)\mathrm{col}_k(B)C_{jk}.$$

[1] L'opérateur $\otimes$ sera défini à la section 1.7.

À présent, calculons le $(mj+i)^{\text{e}}$ élément de $(C^{\mathrm{T}} \otimes A)\mathrm{Vec}(B)$ pour $0 \leq j \leq q-1$ et $1 \leq i \leq m$:

$$\begin{aligned}\left((C^{\mathrm{T}} \otimes A)\mathrm{Vec}(B)\right)_{mj+i} &= \mathrm{lgn}_{mj+i}\left((C^{\mathrm{T}} \otimes A)\mathrm{Vec}(B)\right) \\ &= \sum_{k=1}^{p} C_{jk}\mathrm{lgn}_i(A)\mathrm{col}_k(B) \\ &= \mathrm{Vec}(ABC)_{mj+i}.\end{aligned}$$

□

1.3 Quelques matrices à connaître

Voici les principales familles de matrices avec lesquelles le lecteur devra se familiariser :

1. N est **normale** si $N^\dagger N = NN^\dagger$.
2. H est **hermitienne** si $H = H^\dagger$. On dénote l'ensemble des matrices hermitiennes $d \times d$ par $\mathcal{H}(d)$.
3. Q est **positive** si pour tout état $|\psi\rangle \in H_1^2$, $\langle\psi| Q |\psi\rangle \geq 0$ ou, autrement dit, si $\langle\psi| Q |\psi\rangle$ est toujours réel et non négatif. On dénote l'ensemble des matrices positives $d \times d$ par $\mathcal{P}(d)$.
4. ρ est une **matrice de densité** si ρ est positive et $\mathrm{tr}(\rho) = 1$. On dénote l'ensemble des matrices de densité $d \times d$ par $\mathcal{D}(d)$.
5. U est **unitaire** si $U^\dagger U = I$. On dénote l'ensemble des matrices unitaires $d \times d$ par $\mathcal{U}(d)$.
6. P est un **projecteur** si $P^\dagger = P$ et $P^2 = P$.

1.3.1 Matrices normales et décomposition spectrale

Soit N une matrice $d \times d$. Le théorème de la **décomposition spectrale** affirme que N est normale si et seulement si on peut trouver une base orthonormale $\{|\psi_i\rangle\}_{i=1}^{d}$ de H^d et des nombres complexes λ_i tels que

$$N = \sum_{i=1}^{d} \lambda_i |\psi_i\rangle \langle\psi_i| . \tag{1.1}$$

Si tel est le cas, les $|\psi_i\rangle$ sont les vecteurs propres de N et les λ_i sont les valeurs propres correspondantes.

La décomposition spectrale permet de catégoriser certaines matrices en fonction de leurs valeurs propres. En effet, l'équation 1.1 et un peu d'effort permettent de voir que si N est une matrice normale, alors :

1. N est hermitienne si et seulement si ses valeurs propres sont réelles.

2. N est positive si et seulement si ses valeurs propres sont réelles et non négatives.

3. N est une matrice de densité si et seulement si ses valeurs propres sont réelles, non négatives et ont pour somme un.

4. N est unitaire si et seulement si ses valeurs propres sont des nombres complexes de norme un.

5. N est un projecteur si et seulement si chacune de ses valeurs propres est soit zéro, soit un.

1.3.2 Matrices positives

Il existe plusieurs façons de décrire les matrices positives, chacune ayant ses avantages. On peut trouver la preuve du théorème qui suit dans [Ber92], au théorème 12.7.12.

Théorème 1.3.1. *Soit Q une matrice $d \times d$. Les conditions suivantes sont équivalentes :*

1. *Q est positive.*
2. *$Q^\dagger = Q$ et les valeurs propres de Q sont réelles et non négatives.*
3. *Il existe une matrice positive R telle que $Q = R^2$.*
4. *Il existe une matrice $d \times d$ quelconque S telle que $Q = S^\dagger S$.*

1.3.3 Matrices unitaires

Proposition 1.3.2. *Les conditions suivantes sont équivalentes :*

1. *U est une matrice unitaire.*
2. *U préserve le produit scalaire : $\forall \vec{v}, \vec{w}\,.\, < U\vec{v}\,, U\vec{w} >=< \vec{v}\,, \vec{w} >$.*
3. *U préserve la norme : $\forall \vec{v}\,.\, \|U\vec{v}\| = \|\vec{v}\|$.*
4. *U effectue un changement de base. Autrement dit, de l'application de U à une base orthonormale résulte une autre base orthonormale.*

Si $\{|\psi_i\rangle\}_{i=0}^{d-1}$ et $\{|\varphi_i\rangle\}_{i=0}^{d-1}$ sont deux bases orthonormales de H^d, on peut toujours trouver une matrice unitaire U telle que $|\psi_i\rangle = U\,|\varphi_i\rangle$.

Les matrices unitaires permettent de ramener les matrices normales à leur forme diagonale. En effet, si N est normale, la décomposition spectrale permet d'écrire $N = \sum_{i=0}^{d-1} \lambda_i |\psi_i\rangle \langle\psi_i|$, où $\{|\psi_i\rangle\}_{i=0}^{d-1}$ est une base orthonormale de H^d. Il existe alors une matrice unitaire U telle que $U|i\rangle = |\psi_i\rangle$ pour tous les i, de sorte que

$$N = \sum_{i=0}^{d-1} \lambda_i U |i\rangle \langle i| U^\dagger.$$

Il s'ensuit que

$$U^\dagger N U = \sum_{i=0}^{d-1} \lambda_i U^\dagger U |i\rangle \langle i| U^\dagger U = \sum_{i=0}^{d-1} \lambda_i |i\rangle \langle i|$$

est une matrice diagonale.

1.3.4 Matrices de Pauli et de Hadamard

Les quatre matrices suivantes, appelées matrices de Pauli, joueront un rôle central dans ce mémoire :

$$I \stackrel{\text{def}}{=} \sigma_0 \stackrel{\text{def}}{=} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},$$

$$X \stackrel{\text{def}}{=} \sigma_1 \stackrel{\text{def}}{=} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix},$$

$$Y \stackrel{\text{def}}{=} \sigma_2 \stackrel{\text{def}}{=} \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix},$$

$$Z \stackrel{\text{def}}{=} \sigma_3 \stackrel{\text{def}}{=} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}.$$

Le lecteur pourra vérifier que chacune de ces matrices est à la fois hermitienne et unitaire, et donc, satisfait les égalités suivantes : $\sigma_i^2 = \sigma_i \sigma_i^\dagger = I$.

Une autre matrice jouant un rôle important est celle de Hadamard. Elle aussi est hermitienne et unitaire :

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

Les relations suivantes faciliteront plusieurs calculs :

$$\begin{aligned} Y &= iXZ, \\ X &= HZH, \\ \sigma_i\sigma_j &= \begin{cases} -\sigma_j\sigma_i & \text{si } i,j > 0 \text{ et } i \neq j \\ \sigma_j\sigma_i & \text{autrement.} \end{cases} \end{aligned}$$

1.4 Matrices de densité

Un **mélange** d'états purs est une distribution de probabilité sur un ensemble fini d'états purs :

$$\Psi = \begin{cases} |\psi_1\rangle & \text{avec probabilité } p_1 \\ \vdots & \\ |\psi_n\rangle & \text{avec probabilité } p_n. \end{cases} \quad . \tag{1.2}$$

Nous voudrions pouvoir considérer un mélange d'états purs comme étant lui-même un état. Malheureusement, le formalisme des «vecteurs-états» ne le permet pas. C'est pourquoi nous présentons un nouveau formalisme : celui des matrices de densité, ou **opérateurs de densité**. Il est équivalent à celui des vecteurs-états en ce qui a trait aux états purs, mais il permet également de travailler avec des états mélangés.

Dans les sections qui suivent, chaque loi de la mécanique quantique sera présentée sous deux formes différentes : une fois dans le langage des vecteurs-états, une fois dans celui des matrices de densité.

Si on représente l'état d'un système quantique par le vecteur-état $|\psi\rangle$, alors la matrice de densité représentant le même système sera $|\psi\rangle\langle\psi|$. Par exemple, un qubit dans l'état $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$ sera représenté par la matrice densité

$$|\psi\rangle\langle\psi| = \begin{bmatrix}\alpha\\ \beta\end{bmatrix}\begin{bmatrix}\bar{\alpha} & \bar{\beta}\end{bmatrix} = \begin{bmatrix}\alpha\bar{\alpha} & \alpha\bar{\beta}\\ \beta\bar{\alpha} & \beta\bar{\beta}\end{bmatrix} = \begin{bmatrix}|\alpha|^2 & \alpha\bar{\beta}\\ \bar{\alpha}\beta & |\beta|^2\end{bmatrix}.$$

L'opérateur de densité représentant le mélange 1.2 sera :

$$\rho = \sum_{i=1}^{n} p_i |\psi_i\rangle\langle\psi_i|.$$

Dorénavant, on traitera ρ comme un état à part entière. On dira qu'il s'agit d'un **état mélangé**, par opposition aux états purs. L'ensemble de tous les états (mélangés et purs) dans un espace de dimension d, puisqu'il correspond à l'ensemble des matrices de densité $d \times d$, pourra être dénoté par $\mathcal{D}(d)$.

Proposition 1.4.1. *Un état $\rho \in \mathcal{D}(d)$ est pur si et seulement si* $\mathrm{tr}(\rho^2) = 1$.

Pour dénoter la matrice de densité associée à l'état pur $|\psi\rangle$, on écrira parfois ψ au lieu de $|\psi\rangle\langle\psi|$.

1.5 Évolution unitaire

Les matrices unitaires servent à décrire l'évolution des systèmes quantiques. En effet, si $|\psi\rangle_{t_0}$ est l'état d'un système quantique isolé au temps t_0, alors l'état $|\psi\rangle_{t_1}$ du système au temps t_1 est relié à l'état initial par une transformation unitaire U, comme ceci : $|\psi\rangle_{t_1} = U|\psi\rangle_{t_0}$.

Si l'état d'un système est décrit à l'aide d'une matrice de densité ρ, voici comment on peut en décrire l'évolution :

$$\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i| \overset{U}{\mapsto} \sum_i p_i U|\psi_i\rangle\langle\psi_i|U^\dagger = U\rho U^\dagger.$$

L'état du système au temps t_1 est donc obtenu de l'état initial ρ_{t_0} en effectuant une conjugaison par une matrice unitaire : $\rho_{t_1} = U\rho_{t_0}U^\dagger$.

1.6 Mesures

1.6.1 Mesures projectives

Si $P_0, P_1, \ldots, P_{n-1}$ sont des projecteurs tels que $\sum_{i=0}^{n-1} P_i = I$,[2] on représente la **mesure projective** correspondante par $\mathbb{M} = [P_0, P_1, \ldots, P_{n-1}]$. Si on effectue la mesure $\mathbb{M}$ sur un système qui est dans l'état $|\psi\rangle$, alors on obtient le résultat m avec probabilité $||P_m|\psi\rangle||^2 = \langle\psi|P_m|\psi\rangle$. Sachant que le résultat m a été obtenu, l'état du système après la mesure est $\frac{P_m|\psi\rangle}{||P_m|\psi\rangle||}$. On peut donc voir $\mathbb{M}$ comme un opérateur décrit par :

$$|\psi\rangle \overset{\mathbb{M}}{\mapsto} \frac{P_m|\psi\rangle}{||P_m|\psi\rangle||}, \text{ avec probabilité } ||P_m|\psi\rangle||^2.$$

Par exemple, si un qubit dans l'état $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$ est mesuré dans la base de calcul, c'est-à-dire à l'aide de la mesure $\mathbb{M} = [|0\rangle\langle 0|, |1\rangle\langle 1|]$, alors avec probabilité $|\alpha|^2$, on obtiendra le résultat 0 et l'état du qubit deviendra $|0\rangle$. Avec probabilité $|\beta|^2$, on obtiendra le résultat 1 et l'état $|1\rangle$.

Si l'état d'un système est décrit par une matrice de densité ρ et qu'on applique la mesure projective $\mathbb{M} = [P_0, P_1, \ldots, P_{n-1}]$ à ce système, alors on obtient le résultat m avec probabilité

$$p_m = \mathrm{tr}\,(P_m\rho)$$

[2]Le lecteur pourra s'amuser à vérifier l'équivalence suivante, où les P_i sont des projecteurs $d\times d$: $\sum_{i=0}^{n-1} P_i = I \Leftrightarrow \left(\sum_{i=0}^{m-1} \mathrm{tr}\,(P_i) = d \text{ et } P_iP_j = 0 \text{ chaque fois que } i \neq j\right)$.

et l'état du système après l'obtention du résultat m devient

$$\frac{P_m \rho P_m}{\mathrm{tr}\,(P_m \rho)}.$$

1.6.2 Mesures générales

Une mesure quelconque peut toujours être décrite par une collection $\{A_i\}_{i=1}^n$ de matrices $d \times d$ satisfaisant l'équation

$$\sum_{i=1}^{n} A_i^\dagger A_i = I.$$

Nous appellerons **mesure générale**, ou tout simplement **mesure**, une telle collection $\{A_i\}_{i=1}^n$. Si on mesure l'état ρ, il se transformera selon la règle

$$\rho \overset{\mathbb{M}}{\mapsto} \frac{A_m \rho A_m^\dagger}{\mathrm{tr}\left(A_m \rho A_m^\dagger\right)}, \text{ avec probabilité } \mathrm{tr}\left(A_m \rho A_m^\dagger\right).$$

Les mesures projectives sont un cas particulier des mesures générales. Bien que nous n'en ferons pas ici la démonstration, il est possible de simuler toute mesure sur un système A à l'aide de la séquence d'opérations suivante :

- Adjonction d'un système ancillaire B,
- transformation unitaire sur le système conjoint AB,
- mesure projective du système ancillaire B,
- trace partielle du système B.

1.7 Produit tensoriel et juxtaposition de systèmes

Si A est un matrice $m \times n$ dont on dénote les éléments par $a_{i,j}$ et que B est de dimension $p \times q$, alors le **produit tensoriel** entre A et B est la matrice $mp \times nq$ définie par

$$A \otimes B \stackrel{\text{def}}{=} \begin{bmatrix} a_{1,1}B & \cdots & a_{1,n}B \\ \vdots & \ddots & \vdots \\ a_{m,1}B & \cdots & a_{m,n}B \end{bmatrix}.$$

Les principales propriétés du produit tensoriel sont les suivantes :

1. $(A \otimes B) \otimes C = A \otimes (B \otimes C)$.
2. $(A + B) \otimes (C + D) = A \otimes C + A \otimes D + B \otimes C + B \otimes D$.
3. $\lambda A \otimes \mu B = \lambda\mu(A \otimes B)$.
4. $(A \otimes B)(C \otimes D) = AC \otimes BD$.
5. $(A \otimes B)^\dagger = A^\dagger \otimes B^\dagger$.

Si V et W sont des espaces vectoriels ayant pour base $\{|i\rangle\}_{i=1}^m$ et $\{|j\rangle\}_{j=1}^n$, respectivement, alors $V \otimes W$ est un espace de dimension mn ayant pour base les vecteurs $|i\rangle \otimes |j\rangle$.

En termes de systèmes physiques, le produit tensoriel sert à décrire l'état d'un système composé de différents sous-systèmes juxtaposés. Si on juxtapose n systèmes, numérotés de 1 à n, et que l'état du système i est décrit par $|\psi_i\rangle \in H^{d_i}$, alors l'état du système global est décrit par $|\psi_1\rangle \otimes \cdots \otimes |\psi_n\rangle \in H^{d_1} \otimes \cdots \otimes H^{d_n}$.

De la même façon, si on décrit l'état du système i à l'aide de la matrice de densité ρ_i, l'état du système global sera décrit par $\rho_1 \otimes \cdots \otimes \rho_n$.

La proposition suivante s'avérera utile à la section 2.6 :

Proposition 1.7.1.[3] *Si A est une matrice $m \times m$ et B, une matrice $n \times n$, alors* $\det A \otimes B = (\det A)^n (\det B)^m$.

1.8 Trace partielle

Prenons deux systèmes, A et B, de dimension m et n, respectivement. Si l'état du système A est donné par la matrice ρ^A et celui de B, par σ^B, nous avons vu que l'état du système conjoint pouvait être décrit par le produit tensoriel $\rho^A \otimes \sigma^B$.

À l'inverse, la **trace partielle**[4] est la fonction $\mathrm{tr}_B : \mathcal{D}(mn) \to \mathcal{D}(m)$ qui, prenant pour entrée un état ρ^{AB} du système AB, a pour valeur une matrice de densité $\rho^A = \mathrm{tr}_B\left(\rho^{AB}\right)$ décrivant l'état du système A, quand on le considère comme un système isolé. En termes mathématiques,

$$\mathrm{tr}_B\left(\rho^{AB}\right) \stackrel{\mathrm{def}}{=} \sum_{k=1}^{n} \left(I^A \otimes \langle f_k|^B\right) \rho^{AB} \left(I^A \otimes |f_k\rangle^B\right),$$

où I^A est la matrice identité $m \times m$ opérant sur le système A et $\{|f_k\rangle\}_{k=1}^n$ est une base orthonormale de l'espace H^n associé au système B. La valeur de $\mathrm{tr}_B\left(\rho^{AB}\right)$ ne dépend aucunement du choix de la base $\{|f_k\rangle\}_{k=1}^n$.

[3] [Ber92], page 324.

[4] Pour une définition différente, mais équivalente, de ce concept, le lecteur est invité à consulter [NC00], où la trace partielle est définie comme étant la fonction linéaire satisfaisant l'équation $\mathrm{tr}_B\left(|a_1\rangle\langle a_2| \otimes |b_1\rangle\langle b_2|\right) = |a_1\rangle\langle a_2| \,\mathrm{tr}\left(|b_1\rangle\langle b_2|\right)$, chaque fois que $|a_1\rangle$ et $|a_2\rangle$ sont deux états purs du système A et $|a_1\rangle$ et $|a_2\rangle$, deux états purs du système B.

1.9 Super-opérateurs

Un super-opérateur est une transformation $\mathcal{E} : \mathcal{D}(d) \to \mathcal{D}(d)$ décrite par une famille $\{A_j\}_{j=1}^k$ de matrices $d \times d$ telle que $\sum_{j=1}^k A_j^\dagger A_j = I$. L'effet du super-opérateur $\mathcal{E}$ sur la matrice de densité ρ sera décrit comme ceci :

$$\mathcal{E}(\rho) = \sum_{j=1}^{k} A_j \rho A_j^\dagger.$$

En général, toute opération quantique n'apportant pas d'information sur le système à transformer peut être décrite par un super-opérateur.

Aussi, tout super-opérateur est équivalent à la séquence d'opérations suivantes : ajout d'une ancille, opération unitaire, trace partielle. Plus formellement, pour tout super-opérateur $\mathcal{E}$, il existe une matrice unitaire U telle que pour tout $\rho^A \in \mathcal{D}(d)$,

$$\mathcal{E}(\rho) = \mathrm{tr}_B \left(U \left(\rho^A \otimes |0\rangle \langle 0|^B \right) U^\dagger \right).$$

1.10 Entropie d'états quantiques

1.10.1 Entropie de von Neumann

Soit $\vec{p} = (p_1, p_2, \ldots, p_n)$ un vecteur de probabilité.[5] L'entropie de Shannon de la distribution $\vec{p}$ est donnée par

$$H(\vec{p}) \stackrel{\text{def}}{=} -\sum_{i=1}^{n} p_i \log p_i.$$

On peut définir l'entropie d'un état quantique de façon similaire. Soit $\rho \in \mathcal{D}(d)$ un état quantique ayant pour décomposition spectrale $\rho = \sum_{i=1}^d \lambda_i |\psi_i\rangle \langle\psi_i|$. Les valeurs

[5] Le vecteur $(p_1, \ldots, p_n)$ est un vecteur de probabilité si $0 \leq p_i \leq 1$ pour tout i et si $\sum_{i=1}^n p_i = 1$.

propres de ρ forment une distribution de probabilité. On peut définir l'entropie de von Neumann de l'état ρ comme étant l'entropie de Shannon de ses valeurs propres :

$$H(\rho) \stackrel{\text{def}}{=} -\sum_{i=1}^{n} \lambda_i \log \lambda_i.$$

Le lecteur désireux d'apprendre certaines propriétés et applications de l'entropie de von Neumann est invité à consulter [NC00].

1.10.2 Entropie conditionnelle

Si ρ est un état du système AB, nous écrirons parfois $H(A,B)_\rho$ au lieu de $H(\rho)$. Nous pouvons aussi définir $H(A)_\rho \stackrel{\text{def}}{=} H(\mathrm{tr}_B(\rho))$. L'**entropie conditionnelle** du système A, connaissant B, sera alors définie comme étant

$$H(A|B)_\rho \stackrel{\text{def}}{=} H(A,B)_\rho - H(B)_\rho.$$

Maintenant, si ρ est un état du système ABC, on peut définir l'**information mutuelle** entre A et B, connaissant C, comme étant

$$I(A;B|C)_\rho \stackrel{\text{def}}{=} H(A|C)_\rho - H(A|B,C)_\rho.$$

Voici maintenant une collection de faits qui seront utilisés dans la section 4.3. Nous nous contenterons d'en voir l'énoncé, sans démonstration.

Proposition 1.10.1.[6] $I(A;B|C)_\rho \leq 2\log\min(\dim A, \dim B)$.

Proposition 1.10.2.[7] $H(A,B)_{\rho\otimes\sigma} = H(A)_\rho + H(B)_\sigma$.

Proposition 1.10.3.[8] $H(A|B,C)_\rho \leq H(A|B)_\rho$.

[6] [Hay05].
[7] [NC00], exercice 11.13, page 514
[8] [NC00], théorème 11.15, page 522.

Proposition 1.10.4.[9] *L'entropie conditionnelle est concave, c'est-à-dire que pour tout λ tel que $0 \leq \lambda \leq 1$, $H(A|B)_{\lambda\rho_0+(1-\lambda)\rho_1} \geq \lambda H(A|B)_{\rho_0} + (1-\lambda)H(A|B)_{\rho_1}$.*

Proposition 1.10.5.[10] *Si ρ^{AB} est séparable, alors $H(A|B)_\rho \geq 0$.*

1.11 Faits relatifs à la commutativité de matrices

Théorème 1.11.1. [11] *Soit A une matrice normale et B une matrice hermitienne. Alors $AB = BA$ si et seulement s'il existe une base orthonormale telle que A et B sont toutes deux diagonales dans la même base.*

Démonstration. La preuve est la même que dans [NC00]. Il n'y a dans cette preuve qu'une seule restriction sur A, c'est qu'elle doit posséder une décomposition spectrale, ce qui est le cas lorsqu'elle est normale. La preuve reste valide tant que cette condition est remplie.

□

Corollaire 1.11.2. *Si U est une matrice unitaire telle que $U\rho = \rho U$ pour toute matrice hermitienne $\rho \in \mathcal{H}(d)$, alors $U = \lambda I$, où $\lambda \in \mathbb{C}, |\lambda| = 1$.*

Démonstration. Prenons une matrice U telle que dans l'énoncé du corollaire. Nous allons commencer par montrer que tout vecteur $|\psi\rangle$ est un vecteur propre de U.

[9] [NC00], corollaire 11.13, page 520.
[10] [Hay05].
[11] Adaptation du théorème 2.2 de [NC00], page 77.

Soit $|\psi\rangle$ un vecteur arbitraire. Alors $\psi \stackrel{\text{def}}{=} |\psi\rangle \langle\psi|$ est une matrice hermitienne et, par hypothèse, $U\psi = \psi U$. Par le théorème 1.11.1, il existe une base $\{|\psi_1\rangle, \ldots |\psi_d\rangle\}$ de H^d telle que U et ψ sont toutes deux diagonales dans la même base. Aussi, puisque ψ divise l'espace en un espace propre de dimension 1 correspondant à la valeur propre 1, et en un espace propre de dimension $d-1$ correspondant à la valeur propre 0, il faut que $\psi = |\psi_i\rangle \langle\psi_i|$ pour un certain i. Sans perte de généralité, $i = 1$ et nous pouvons écrire $U = \lambda |\psi\rangle \langle\psi| + \sum_{i=2}^{d} \lambda_i |\psi_i\rangle \langle\psi_i|$, ce qui montre que $U|\psi\rangle = \lambda|\psi\rangle$. Autrement dit, $|\psi\rangle$ est un vecteur propre de U.

Montrons maintenant que λ est la seule valeur propre de U, c'est-à-dire que $\lambda_i = \lambda$ pour tout i. Puisque chaque vecteur est un vecteur propre de U, alors, en particulier, $U(|\psi\rangle - |\psi_i\rangle) = z(|\psi\rangle - |\psi_i\rangle)$ pour un certain z. Mais comme il est aussi vrai que $U(|\psi\rangle - |\psi_i\rangle) = \lambda|\psi\rangle - \lambda_i|\psi_i\rangle$, il faut nécessairement que $z = \lambda_i = \lambda$.

Donc, $U = \lambda \sum_{i=1}^{d} |\psi_i\rangle \langle\psi_i| = \lambda I$. Finalement, λ doit avoir norme un parce que U est unitaire :

$$I = UU^\dagger = \lambda\lambda^* I \Rightarrow \lambda\lambda^* = 1.$$

□

Chapitre 2

La sphère de Poincaré

2.1 Introduction

La sphère de Poincaré est une représentation géométrique, dans $\mathbb{R}^3$, de l'ensemble $\mathcal{D}(2)$ des états quantiques à deux dimensions. Bien que cette représentation soit simple et naturelle, il est assez difficile d'en trouver les détails explicites dans les ouvrages courants, les auteurs y voyant peut-être autant de trivialités auxquelles il ne vaut pas la peine de consacrer plus de quelques lignes. L'auteur de ce mémoire a donc dû prouver lui-même[1] les faits colligés dans ce chapitre. Même s'ils sont effectivement simples et sans doute déjà connus de plusieurs membres de la communauté scientifique, les jeunes chercheurs en retireront sans doute une connaissance approfondie de plusieurs notions importantes.

[1] À moins d'indication contraire, chaque démonstration de ce mémoire est l'œuvre de l'auteur.

2.2 Notions préliminaires

2.2.1 Les groupes SU(2) et SO(3)

L'intérêt réel de la sphère de Poincaré découle d'une correspondance inusitée entre deux groupes qui n'ont à priori pas de relation évidente entre eux. Il s'agit des groupes $\mathcal{SU}(2)$ et $\mathcal{SO}(3)$, qui sont des groupes de *transformations* sur les espaces $\mathcal{H}(2)$ et $\mathbb{R}^3$, respectivement. Nous ne ferons dans cette section qu'un bref rappel de ce que sont ces deux groupes.

Le groupe $\mathcal{SU}(2)$ est bien connu des informaticiens quantiques : il s'agit du groupe des transformations unitaires de déterminant 1 sur un espace vectoriel complexe à deux dimensions. De façon générale, le déterminant d'une matrice unitaire est un nombre complexe de norme un, comme le montrent les égalités suivantes :

$$\begin{aligned} |U||U|^* &= |U||U^\dagger| \\ &= |UU^\dagger| \\ &= |I| \\ &= 1. \end{aligned}$$

La raison pour laquelle nous nous intéressons à $\mathcal{SU}(2)$ plutôt qu'à $\mathcal{U}(2)$, c'est que les états quantiques sont **invariants sous un changement de phase global**. Plus précisément, si z est un nombre complexe de norme 1, alors, pour un état $|\psi\rangle$ quelconque, on considère que $|\psi\rangle$ et $z\,|\psi\rangle$ représentent le même état, au sens où on ne peut physiquement les distinguer. Il s'ensuit que pour toute transformation unitaire U, les transformations zU, où z est complexe et de norme 1, sont toutes équivalentes. On choisit donc un représentant unique pour la classe d'équivalence, soit la matrice $\frac{1}{\sqrt{|U|}}U$, dont le déterminant est 1.

Le groupe $\mathcal{SO}(3)$ est aussi très usité. C'est le groupe des transformations orthogonales de déterminant 1 sur un espace vectoriel réel à trois dimensions ou, autrement dit, des rotations dans $\mathbb{R}^3$. Une matrice O est dite **orthogonale** si $OO^T = I$. On dénote l'espace des matrices orthogonales $d \times d$ par $\mathcal{O}(3)$. Les matrices orthogonales sont l'équivalent réel des matrices unitaires : elles sont, elles aussi, définies comme étant des isométries[2], la seule différence étant qu'elles agissent sur un espace réel plutôt que complexe. Celles de déterminant +1 sont des rotations et celles de déterminant -1, des réflexions.

2.3 Matrices de densité et espaces vectoriels

Nous nous intéresserons ici aux matrices de densité en tant qu'éléments d'un espace vectoriel. Tâchons donc de découvrir dans quel espace résident les matrices de densité.

Soit $\mathcal{D}(d)$ l'ensemble des matrices de densité $d \times d$. Première constatation : $\mathcal{D}(d)$ n'est pas un espace vectoriel. Par exemple, si $\rho \in \mathcal{D}(d)$, alors $-\rho \notin \mathcal{D}(d)$, simplement parce que $\mathrm{tr}(-\rho) = -1$. Deuxièmement, le plus petit espace vectoriel *complexe* contenant $\mathcal{D}(d)$ est $M(d, \mathbb{C})$, l'ensemble de toutes les matrices complexes $d \times d$. Fait plus intéressant, le plus petit espace vectoriel *réel* contenant $\mathcal{D}(d)$ est $\mathcal{H}(d)$, l'ensemble des matrices hermitiennes $d \times d$. C'est ce que nous montrera la proposition 2.3.2.

2.3.1 Matrices hermitiennes

Lemme 2.3.1. *Soit H une matrice hermitienne. Alors on peut trouver des matrices positives P et Q telles que $H = P - Q$.*

[2]Une isométrie est une transformation qui préserve la norme de chacun des vecteurs de son domaine.

Démonstration. Comme H est hermitienne, les coefficients λ_i de sa décomposition spectrale $H = \sum_{i=1}^{n} \lambda_i |\psi_i\rangle \langle\psi_i|$ sont réels. On peut supposer, sans perte de généralité, que les k premiers λ_i sont positifs ou nuls, tandis que les autres sont négatifs, de façon à ce qu'on puisse écrire

$$H = \sum_{i=1}^{k} |\lambda_i| |\psi_i\rangle \langle\psi_i| - \sum_{i=k+1}^{n} |\lambda_i| |\psi_i\rangle \langle\psi_i| .$$

On n'a plus qu'à poser $P = \sum_{i=1}^{k} |\lambda_i| |\psi_i\rangle \langle\psi_i|$ et $Q = \sum_{i=k+1}^{n} |\lambda_i| |\psi_i\rangle \langle\psi_i|$. Les matrices P et Q sont toutes deux positives, car leurs valeurs propres, les $|\lambda_i|$, sont non négatives.

□

En particulier, les matrices de Pauli peuvent s'écrire comme suit :

$$\begin{aligned}
Z &= |0\rangle \langle 0| - |1\rangle \langle 1| , \\
X &= HZH \\
&= H |0\rangle \langle 0| H - H |1\rangle \langle 1| H \\
&= |\nearrow\rangle \langle \nearrow| - |\searrow\rangle \langle \searrow| , \\
Y &= i |1\rangle \langle 0| - i |0\rangle \langle 1| \\
&= \frac{1}{2} (|0\rangle \langle 0| - i |0\rangle \langle 1| + i |1\rangle \langle 0| + |1\rangle \langle 1| - |0\rangle \langle 0| - i |0\rangle \langle 1| + i |1\rangle \langle 0| - |1\rangle \langle 1|) \\
&= \frac{1}{2}(|0\rangle + i |1\rangle)(\langle 0| - i \langle 1|) - \frac{1}{2}(|0\rangle - i |1\rangle)(\langle 0| + i \langle 1|) \\
&= |\circlearrowleft\rangle \langle \circlearrowleft| - |\circlearrowright\rangle \langle \circlearrowright| .
\end{aligned}$$

Proposition 2.3.2. *L'espace vectoriel réel engendré par l'ensemble des matrices de densité $d \times d$ est $\mathcal{H}(d)$.*

Démonstration. Il suffit de montrer qu'il est possible d'écrire toute matrice hermitienne comme une combinaison linéaire réelle de matrices de densité. Soit H une

matrice hermitienne. Écrivons $H = P - Q$, où P et Q sont positives, comme nous le permet le lemme 2.3.1. Alors nous pouvons aussi écrire $H = \mathrm{tr}\,(P)\,\tilde{P} - \mathrm{tr}\,(Q)\,\tilde{Q}$, où

$$\tilde{P} \stackrel{\mathrm{def}}{=} \begin{cases} \frac{1}{\mathrm{tr}(P)} P & : \quad \mathrm{tr}\,(P) > 0 \\ 0 & : \quad P = 0 \end{cases}$$

et $\tilde{Q}$ est définie de façon similaire. Dans la définition de $\tilde{P}$, les deux cas mentionnés sont les seuls cas possibles, car la trace d'une matrice positive est toujours réelle et plus grande ou égale à zéro, l'égalité tenant si et seulement si la matrice est elle-même zéro.[3]

Lorsque $\tilde{P}$ et $\tilde{Q}$ ne sont pas zéro, ce sont des matrices de densité, car elles ont une trace de valeur un et sont positives. Il s'ensuit que H est une combinaison linéaire réelle (potentiellement triviale) de matrices de densité.

□

Corollaire 2.3.3. *L'espace vectoriel réel engendré par l'ensemble des matrices positives $d \times d$ est $\mathcal{H}(d)$.*

Démonstration. Soit $\mathcal{P}(d)$ l'ensemble des matrices positives $d \times d$. Comme les matrices positives sont aussi hermitiennes[4], nous obtenons la chaîne d'inclusions suivante :

$$\mathcal{D}(d) \subseteq \mathcal{P}(d) \subseteq \mathcal{H}(d).$$

[3]Ceci découle de la décomposition spectrale : posons $P = \sum_i \lambda_i \left|\psi_i\right\rangle \left\langle\psi_i\right|$, où $\lambda_i \geq 0$. Clairement, $\mathrm{tr}\,(P) = \sum_i \lambda_i \geq 0$ et $\mathrm{tr}\,(P) = 0 \Leftrightarrow \forall i.\lambda_i = 0$.

[4]Une matrice P est positive si et seulement s'il existe une matrice S telle que $P = S^\dagger S$. Cela implique automatiquement que P est hermitienne : $P^\dagger = (S^\dagger S)^\dagger = S^\dagger S = P$.

Les espaces engendrés par ces ensembles doivent suivre la même chaîne d'inclusions :

$$\begin{aligned} & \operatorname{span}(\mathcal{D}(d)) \subseteq \operatorname{span}(\mathcal{P}(d)) \subseteq \operatorname{span}(\mathcal{H}(d)) \\ \Rightarrow \quad & \mathcal{H}(d) \subseteq \operatorname{span}(\mathcal{P}(d)) \subseteq \mathcal{H}(d), \text{ puisque } \operatorname{span}(\mathcal{D}(d)) = \operatorname{span}(\mathcal{H}(d)) = \mathcal{H}(d) \\ \Rightarrow \quad & \operatorname{span}(\mathcal{P}(d)) = \mathcal{H}(d). \end{aligned}$$

□

2.3.2 Une base pour $\mathcal{H}(2)$

Jusqu'à présent, la discussion sur les matrices de densité et les matrices hermitiennes était assez simple pour traiter du cas général à d dimensions. Par contre, à partir de maintenant et à moins d'avis contraire, nous nous en tiendrons au cas à deux dimensions.

La proposition 2.3.2 nous a appris que pour étudier les matrices de densité en tant qu'éléments d'un espace vectoriel réel, il faut se tourner vers l'espace des matrices hermitiennes. La prochaine étape est évidemment de trouver une «bonne» base pour $\mathcal{H}(2)$, ce qui fait l'objet du prochain lemme :

Proposition 2.3.4. *Les matrices de Pauli forment une base de* $\mathcal{H}(2)$.

Démonstration. Nous nous contenterons de démontrer que les matrices de Pauli engendrent $\mathcal{H}(2)$. Le lecteur pourra lui-même vérifier qu'elles sont linéairement indépendantes.

Soit $A \in \mathcal{H}(2)$. Par la propriété $A^\dagger = A$, les éléments diagonaux de A sont réels et les deux autres sont des conjugués complexes l'un de l'autre. On peut donc écrire

$$A = \begin{bmatrix} a & b - ci \\ b + ci & d \end{bmatrix},$$

où $a, b, c, d \in \mathbb{R}$. Maintenant, nous pouvons exprimer A comme une combinaison linéaire réelle des matrices de Pauli :

$$\begin{aligned} A &= \begin{bmatrix} \frac{a+d}{2} & 0 \\ 0 & \frac{a+d}{2} \end{bmatrix} + \begin{bmatrix} \frac{a-d}{2} & 0 \\ 0 & \frac{-a+d}{2} \end{bmatrix} + \begin{bmatrix} 0 & b \\ b & 0 \end{bmatrix} + \begin{bmatrix} 0 & -ci \\ ci & 0 \end{bmatrix} \\ &= \left(\frac{a+d}{2}\right) I + \left(\frac{a-d}{2}\right) Z + bX + cY. \end{aligned}$$

□

2.4 La sphère de Poincaré

Il est temps d'entrer dans le vif du sujet : nous allons voir comment représenter les matrices de densité par des vecteurs de $\mathbb{R}^3$. Définissons les ensembles B_3, la boule de rayon un dans $\mathbb{R}^3$ et S_2, la surface de B_3 :

$$\begin{aligned} S_2 &\stackrel{\text{def}}{=} \{\vec{r} \in \mathbb{R}^3 \,|\, \|\vec{r}\| = 1\}, \\ B_3 &\stackrel{\text{def}}{=} \{\vec{r} \in \mathbb{R}^3 \,|\, \|\vec{r}\| \leq 1\}. \end{aligned}$$

Nous verrons qu'il est possible de représenter chaque état pur $|\psi\rangle \in H_1^2$ par un vecteur de S_2, puis, qu'en général, il existe une correspondance entre les états quantiques $\rho \in \mathcal{D}(2)$ et les vecteurs de B_3. L'ensemble S_2 se nomme aussi **sphère de Poincaré**.

2.4.1 Définition paramétrique de la sphère de Poincaré

La sphère de Poincaré est la représentation géométrique dans un espace réel, soit $\mathbb{R}^3$, de l'ensemble H_1^2 des états purs quantiques dans un espace complexe de dimension 2. Nous avons vu, à la section 1.1, que chaque état pur $|\psi\rangle \in H_1^2$ peut s'écrire sous la forme $|\psi\rangle = \alpha\,|0\rangle + \beta\,|1\rangle$, où $\|\alpha\|^2 + \|\beta\|^2 = 1$. Cette contrainte

sur la norme de α et de β peut aussi être prise en compte directement en écrivant $|\psi\rangle = e^{i\varphi_0} cos(\theta) |0\rangle + e^{i\varphi_1} sin(\theta) |1\rangle$, où $0 \leq \theta \leq \frac{\pi}{2}$. Maintenant, par l'invariance des états sous un changement de phase global,[5] on peut laisser tomber le coefficient $e^{i\varphi_0}$ pour écrire

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

où $0 \leq \varphi < 2\pi$ et $0 \leq \theta \leq \pi$. L'ajout du facteur $\frac{1}{2}$ devant θ permet d'obtenir la paramétrisation d'une sphère dans $\mathbb{R}^3$ — soit S_2 —, ce qui met en évidence une bijection entre S_2 et les états purs de H_1^2.

2.4.2 Une bijection entre états quantiques et points de la sphère

La proposition 2.3.4 nous apprenait que les matrices de Pauli forment une base de $\mathcal{H}(2)$. Une matrice de densité quelconque ρ, étant hermitienne, peut donc s'écrire sous la forme $\rho = aI + bX + cY + dZ$, avec $a, b, c, d \in \mathbb{R}$. Aussi, comme $\mathrm{tr}(\rho) = 1$, il faut que

$$\begin{aligned} 1 &= \mathrm{tr}(\rho) \\ &= \mathrm{tr}(aI + bX + cY + dZ) \\ &= a\mathrm{tr}(I) + b\mathrm{tr}(X) + c\mathrm{tr}(Y) + d\mathrm{tr}(Z) \\ &= 2a, \end{aligned}$$

puisque $\mathrm{tr}(I) = 2$, tandis que $\mathrm{tr}(X) = \mathrm{tr}(Y) = \mathrm{tr}(Z) = 0$.

[5] Pour tout $|\psi\rangle \in H_1^2$ et tout $\varphi \in \mathbb{R}$, $e^{i\varphi} |\psi\rangle$ et $|\psi\rangle$ représentent le même état quantique.

Pour reprendre la notation de [NC00], on peut poser $r_1 = 2b$, $r_2 = 2c$, $r_3 = 2d$, pour ensuite écrire

$$\rho = \frac{1}{2}(I + \vec{r} \cdot \vec{\sigma}), \tag{2.1}$$

où $\vec{r} = (r_1, r_2, r_3)$ et $\vec{\sigma} = (X, Y, Z)$.

Nous avons aussi vu, à la proposition 1.4.1 que $\mathrm{tr}(\rho^2) \leq 1$ et que l'égalité est atteinte si et seulement si ρ est un état pur. Nous utiliserons ce fait pour établir une contrainte sur $\vec{r}$. Mais, tout d'abord, il est nécessaire de présenter le lemme suivant, dont la signification ne sera pas apparente immédiatement, mais qui prendra tout son sens à la section 2.7 :

Lemme 2.4.1. *Soit $\vec{r} \in \mathbb{R}^3$. Alors $(\vec{r} \cdot \vec{\sigma})^2 = \|\vec{r}\|^2 I$.*

Démonstration.

$$\begin{aligned}
(\vec{r} \cdot \vec{\sigma})^2 &= \left(\sum_{i=1}^{3} r_i \sigma_i \right)^2 \\
&= \sum_{\substack{i,j=1 \\ i=j}}^{3} r_i r_j \sigma_i \sigma_j + \sum_{\substack{i,j=1 \\ i<j}}^{3} r_i r_j \sigma_i \sigma_j + \sum_{\substack{i,j=1 \\ i>j}}^{3} r_i r_j \sigma_i \sigma_j \\
&= \sum_{i=1}^{3} r_i^2 I + \sum_{\substack{i,j=1 \\ i<j}}^{3} r_i r_j \sigma_i \sigma_j + \sum_{\substack{i,j=1 \\ i<j}}^{3} r_j r_i \sigma_j \sigma_i \\
&= \|\vec{r}\|^2 I + \sum_{\substack{i,j=1 \\ i<j}}^{3} r_i r_j (\sigma_i \sigma_j + \sigma_j \sigma_i) \\
&= \|\vec{r}\|^2 I,
\end{aligned}$$

puisque pour $1 \leq i < j \leq 3$, $\sigma_i \sigma_j = -\sigma_j \sigma_i$.

□

Proposition 2.4.2. *Soit $\rho = \frac{1}{2}(I + \vec{r} \cdot \vec{\sigma})$ une matrice de densité quelconque. Alors $\|\vec{r}\| \leq 1$, et l'égalité est atteinte si et seulement si ρ est un état pur.*

Démonstration.

$$\begin{aligned} 1 &\geq \mathrm{tr}(\rho^2) \\ &= \mathrm{tr}(\frac{1}{4}(I + \vec{r} \cdot \vec{\sigma})^2) \\ &= \frac{1}{4}\mathrm{tr}(I + 2\vec{r} \cdot \vec{\sigma} + (\vec{r} \cdot \vec{\sigma})^2) \\ &= \frac{1}{4}(\mathrm{tr}\,(I) + 2\mathrm{tr}\,(\vec{r} \cdot \vec{\sigma}) + \mathrm{tr}(\|\vec{r}\|^2 I)), \text{ par le lemme 2.4.1} \\ &= \frac{1}{4}(2 + 0 + 2\|\vec{r}\|^2) \\ &= \frac{1}{2} + \frac{1}{2}\|\vec{r}\|^2. \end{aligned}$$

Il s'ensuit que $\|\vec{r}\| \leq 1$.

□

Nous voyons maintenant que toute matrice de densité peut être représentée par un vecteur de norme égale ou inférieure à un dans $\mathbb{R}^3$, ou autrement dit, par un point dans B_3.

Nous pouvons donc définir la fonction $m : \mathcal{D}(2) \to B_3$ de la façon suivante :

$$m\left(\frac{1}{2}(I + \vec{r} \cdot \vec{\sigma})\right) = \vec{r}. \tag{2.2}$$

Cette fonction est bien définie car la représentation $\rho = \frac{1}{2}(I + \vec{r} \cdot \vec{\sigma})$ est unique :

$$\begin{aligned} \frac{1}{2}(I + \vec{r} \cdot \vec{\sigma}) = \frac{1}{2}(I + \vec{s} \cdot \vec{\sigma}) &\Rightarrow \vec{r} \cdot \vec{\sigma} = \vec{s} \cdot \vec{\sigma} \\ &\Rightarrow (r_1 - s_1)X + (r_2 - s_2)Y + (r_3 - s_3)Z = 0 \\ &\Rightarrow r_1 = s_1\,, r_2 = s_2 \text{ et } r_3 = s_3, \end{aligned}$$

par indépendance linéaire des matrices de Pauli.

Si, au premier abord, on avait trouvé surprenant qu'il soit possible de représenter un ensemble de matrices complexes 2×2 par un sous-ensemble de $\mathbb{R}^3$, on pourra maintenant réaliser à quel point cette correspondance est naturelle. En effet, considérons $\mathcal{H}_0(2)$, l'ensemble des matrices hermitiennes 2×2 de trace zéro. Il s'agit d'un espace vectoriel réel à trois dimensions ayant pour base $\{X, Y, Z\}$. Les éléments de $\mathcal{H}_0(2)$ ont donc la forme $\vec{r} \cdot \vec{\sigma}$, où $\vec{r} \in \mathbb{R}^3$. $\mathcal{H}_1(2)$, l'ensemble des matrices hermitiennes de trace un, est une translation de $\mathcal{H}_0(2)$: on peut donc écrire $\mathcal{H}_1(2) = \{\frac{1}{2}(I + \vec{r} \cdot \vec{\sigma}) \,|\, \vec{r} \in \mathbb{R}^3\}$.

De la même façon, l'ensemble des matrices de densité $\mathcal{D}(2)$ n'est que la translation d'une boule de $\mathcal{H}_0(2)$ — nous l'avons vu, $\mathcal{D}(2) = \{\frac{1}{2}(I + \vec{r} \cdot \vec{\sigma}) \,|\, \vec{r} \in \mathbb{R}^3, \|\vec{r}\| \leq 1\}$. Aussi, il existe un isomorphisme évident — nommons-le h — entre $\mathcal{H}_0(2)$ et $\mathbb{R}^3$, soit $X \overset{h}{\mapsto} \vec{i}$, $Y \overset{h}{\mapsto} \vec{j}$, $Z \overset{h}{\mapsto} \vec{k}$. C'est ce qui explique l'existence de la bijection m, qu'on pourrait aussi écrire sous la forme $m(\rho) = h(2\rho - I)$.

2.4.3 Propriétés de la fonction m

Voici quelques autres propriétés de la fonction m qui pourront nous être utiles. On peut en trouver la preuve dans [Zan98].

Proposition 2.4.3.

1. *m est une bijection affine, au sens où $m(\sum_i \lambda_i \rho_i) = \sum_i \lambda_i m(\rho_i)$ chaque fois que $\sum_i \lambda_i = 1$.*
2. *$\langle \rho \,, \sigma \rangle = \frac{1}{2} + \frac{1}{2} < m(\rho), m(\sigma) >$.*
3. *$d(\rho, \sigma) = \frac{1}{2}\|m(\rho) - m(\sigma)\|$.*
4. *$m(\mathcal{D}_{pur}(2)) = S_2$ et $m(\mathcal{D}(2)) = B_3$.*

Pour conclure cette section, voyons la représentation dans B_3 de quelques uns des états les plus communs. Les symboles $\vec{\imath}$, $\vec{\jmath}$ et $\vec{k}$ désigneront les trois vecteurs de la base canonique de $\mathbb{R}^3$.

$$
\begin{aligned}
m(|0\rangle\langle 0|) &= m(\frac{1}{2}I + \frac{1}{2}Z) = (0,0,1) = \vec{k},\\
m(|1\rangle\langle 1|) &= m(\frac{1}{2}I - \frac{1}{2}Z) = (0,0,-1) = -\vec{k},\\
m(|\nearrow\rangle\langle\nearrow|) &= m(H|0\rangle\langle 0|H) = m(H(\frac{1}{2}I + \frac{1}{2}Z)H)\\
&= m(\frac{1}{2}I + \frac{1}{2}X) = (1,0,0) = \vec{\imath},\\
m(|\searrow\rangle\langle\searrow|) &= m(H|1\rangle\langle 1|H) = m(H(\frac{1}{2}I - \frac{1}{2}Z)H)\\
&= m(\frac{1}{2}I - \frac{1}{2}X) = (-1,0,0) = -\vec{\imath},\\
m(|\circlearrowleft\rangle\langle\circlearrowleft|) &= m\left(\left(\frac{1}{\sqrt{2}}(|0\rangle + i|1\rangle)\right)\left(\frac{1}{\sqrt{2}}(\langle 0| - i\langle 1|)\right)\right)\\
&= m\left(\frac{1}{2}(|0\rangle\langle 0| + |1\rangle\langle 1|) + i(|1\rangle\langle 0| - |0\rangle\langle 1|)\right)\\
&= m\left(\frac{1}{2}I + \frac{1}{2}Y\right) = (0,1,0) = \vec{\jmath},\\
m(|\circlearrowright\rangle\langle\circlearrowright|) &= m(Z|\circlearrowleft\rangle\langle\circlearrowleft|Z) = m\left(\frac{1}{2}I - \frac{1}{2}Y\right) = (0,-1,0) = -\vec{\jmath},\\
m\left(\frac{1}{2}I\right) &= (0,0,0).
\end{aligned}
$$

Cette dernière équation nous apprend que l'état $\frac{1}{2}I$, qu'on appelle parfois **état complètement mélangé**, correspond au centre de la sphère. Les autres états dont il est question plus haut correspondent aux points d'intersection entre la surface S_2 et les axes des x, des y et des z.

2.5 La conjugaison par une matrice unitaire

La section précédente met en lumière une correspondance entre les états quantiques et les points de B_3. Notre prochain objectif sera de découvrir une correspondance

analogue entre les transformations unitaires sur les états quantiques et certaines transformations sur B_3. Cet objectif ne sera atteint qu'à la section suivante. Avant d'y arriver, nous devrons prendre le temps d'étudier la fonction Ad, définie ci-dessous. Le symbole $\mathcal{L}(\mathcal{H}(2))$ désigne l'ensemble des transformations linéaires sur $\mathcal{H}(2)$.

Définition 2.5.1. $\mathrm{Ad} : \mathcal{SU}(2) \to \mathcal{L}(\mathcal{H}(2))$ *est la fonction définie par*

$$\mathrm{Ad}U(\tau) \stackrel{\mathrm{def}}{=} U\tau U^\dagger$$

pour toutes matrices $U \in \mathcal{SU}(2), \tau \in \mathcal{H}(2)$.

La première chose à vérifier, c'est que AdU est bel et bien une transformation linéaire sur $\mathcal{H}(2)$. D'abord, si $\tau \in \mathcal{H}(2)$, alors $(U\tau U^\dagger)^\dagger = U\tau^\dagger U^\dagger = U\tau U^\dagger$, ce qui montre que $\mathrm{Ad}U(\tau) \in \mathcal{H}(2)$. Ensuite, pour $\alpha, \beta \in \mathbb{C}$ et $\tau, \sigma \in \mathcal{H}(2)$, nous avons $U(\alpha\sigma + \beta\tau)U^\dagger = \alpha U\sigma U^\dagger + \beta U\tau U^\dagger$, donc Ad$U$ est une transformation linéaire.

Proposition 2.5.2.

1. Ad *est un homomorphisme du groupe multiplicatif* $\mathcal{SU}(2)$*, c'est-à-dire que pour* $U, V \in \mathcal{SU}(2)$*,* $\mathrm{Ad}(UV) = \mathrm{Ad}U \circ \mathrm{Ad}V$.
2. $\ker \mathrm{Ad} = \{I, -I\}$.

Démonstration.

1. Soit $U, V \in \mathcal{SU}(2)$. Pour tout $\tau \in \mathcal{H}(2)$, nous avons

$$\begin{aligned} \mathrm{Ad}(UV)(\tau) &= (UV)\tau(UV)^\dagger \\ &= U(V\tau V^\dagger)U^\dagger \\ &= U(\mathrm{Ad}V(\tau))U^\dagger \\ &= \mathrm{Ad}U(\mathrm{Ad}V(\tau)) \\ &= \mathrm{Ad}U \circ \mathrm{Ad}V(\tau). \end{aligned}$$

Donc, $\mathrm{Ad}(UV) = \mathrm{Ad}U \circ \mathrm{Ad}V$.

2. Montrons d'abord que $\operatorname{Ker}\mathrm{Ad} \subseteq \{I, -I\}$. Soit $U \in \operatorname{Ker}\mathrm{Ad}$, c'est-à-dire que U est l'identité sur $\mathcal{H}(2)$, ou encore, que $U\tau U^\dagger = \tau$ pour tout $\tau \in \mathcal{H}(2)$. En multipliant chaque côté de cette égalité par U par la droite, on trouve que $U\tau = \tau U$ pour tout τ. Or, toute matrice répondant à cette condition doit être, par le corollaire 1.11.2, un multiple de l'identité — c'est-à-dire que $U = zI$, où $z \in \mathbb{C}$. Comme $|zI| = z^2$ et, par hypothèse, $|U| = 1$, il faut que $U = \pm I$.

   L'inclusion inverse, $\{I, -I\} \subseteq \operatorname{Ker}\mathrm{Ad}$, suit du simple fait que pour tout $\tau \in \mathcal{H}(2)$, $(-I)\tau(-I) = I\tau I = \tau$.

   Donc $\operatorname{Ker}\mathrm{Ad} = \{I, -I\}$.

□

À présent, voyons quelles propriétés peut avoir $\mathrm{Ad}U$, en tant que transformation sur $\mathcal{H}(2)$. Il s'avérera que la fonction Ad est une surjection qui envoie les matrices unitaires à des transformations orthogonales sur $\mathcal{H}(2)$, au sens de la définition suivante :

Définition 2.5.3. *Une transformation $T : \mathcal{H}(2) \to \mathcal{H}(2)$ est dite* ***orthogonale*** *si le produit interne $\langle A, B \rangle = \operatorname{tr}\left(A^\dagger B\right)$ est invariant sous T, ou, autrement dit, si pour tout $A, B \in \mathcal{H}(2)$, $\langle TA, TB \rangle = \langle A, B \rangle$.*

Proposition 2.5.4. *Si U est une matrice unitaire, alors $\mathrm{Ad}U$ est une transformation orthogonale de $\mathcal{H}(2)$ qui préserve la trace de toute matrice $A \in \mathcal{H}(2)$.*

Démonstration. Vérifions d'abord que $\mathrm{Ad}U$ préserve le produit interne entre toutes

matrices $A, B \in \mathcal{H}(2)$:

$$\begin{aligned}
\langle \operatorname{Ad}U(A), \operatorname{Ad}U(B) \rangle &= \langle UAU^\dagger, UBU^\dagger \rangle \\
&= \operatorname{tr}\left((UAU^\dagger)^\dagger (UBU^\dagger)\right) \\
&= \operatorname{tr}\left(UA^\dagger U^\dagger UBU^\dagger\right) \\
&= \operatorname{tr}\left(A^\dagger B\right) \\
&= \langle A, B \rangle .
\end{aligned}$$

Aussi, $\operatorname{tr}(UAU^\dagger) = \operatorname{tr}(A)$, ce qui montre que $\operatorname{Ad}U$ préserve la trace de A.

□

Puisque, par la proposition 2.5.4, $\operatorname{Ad}U$ est une transformation orthogonale de $\mathcal{H}(2)$ qui préserve la trace, alors, en particulier, $\operatorname{Ad}U$ est une transformation orthogonale de $\mathcal{H}_0(2)$. Comme $\mathcal{H}_0(2)$ et $\mathbb{R}^3$ sont des espaces isomorphes, on s'attend à ce que $\operatorname{Ad}U$ corresponde, dans $\mathbb{R}^3$, à une transformation orthogonale. Nous nommerons cette transformation $\varphi(U)$ et nous verrons à la section suivante qu'en fait, $\varphi(U) \in \mathcal{SO}(3)$.

2.6 Représentation des transformations unitaires par des rotations de la sphère de Poincaré

Nous avons vu, à la section 1.5, que l'évolution, sur un intervalle de temps donné, d'un système quantique isolé dont l'état est décrit par la matrice de densité ρ, peut être représenté par la transformation $\rho \mapsto \operatorname{Ad}U(\rho)$. Quand ρ subit une telle transformation, qu'advient-il de $m(\rho)$? Autrement dit, quelle transformation de B_3 correspond à $\operatorname{Ad}U$? Il s'agira tout naturellement de la fonction φ définie ci-dessous :

Définition 2.6.1. $\varphi : \mathcal{SU}(2) \to \mathcal{SO}(3)$ *est la fonction définie par la règle suivante :* $\varphi(U) = m \circ \operatorname{Ad}U \circ m^{-1}$.

Le diagramme suivant illustre cette définition :

$$\begin{array}{ccc} \mathcal{D}(2) & \xrightarrow{\mathrm{Ad}U} & \mathcal{D}(2) \\ \downarrow m & & \downarrow m \\ B_3 & \xrightarrow{\varphi(U)} & B_3 \end{array}$$

Proposition 2.6.2. *φ est un homomorphisme.*

Démonstration.

$$\begin{aligned} \varphi(UV) &= m \circ \mathrm{Ad}(UV) \circ m^{-1} \\ &= m \circ \mathrm{Ad}U \circ \mathrm{Ad}V \circ m^{-1}, \text{ par la proposition 2.5.2} \\ &= m \circ \mathrm{Ad}U \circ m^{-1} \circ m \circ \mathrm{Ad}V \circ m^{-1} \\ &= \varphi(U) \circ \varphi(V). \end{aligned}$$

□

Rien n'indique à priori que le codomaine de φ doive être $\mathcal{SO}(3)$. Commençons par montrer que $\varphi(U)$ est nécessairement orthogonale :

Proposition 2.6.3. *Pour tout $U \in \mathcal{SU}(2)$, $\varphi(U) \in \mathcal{O}(3)$.*

Démonstration. Soit $U \in \mathcal{SU}(2)$. Il suffit de montrer que $\varphi(U)$ préserve le produit interne entre n'importe quels deux vecteurs $\vec{r}$ et $\vec{s}$ de norme un dans $\mathbb{R}^3$. Posons $\vec{r} = m(\rho)$ et $\vec{s} = m(\sigma)$. Alors,

$$\begin{aligned} < \varphi(U)(m(\rho)), \varphi(U)(m(\sigma)) > &= < m(U\rho U^\dagger), m(U\sigma U^\dagger) > \\ &= -1 + 2 < U\rho U^\dagger, U\sigma U^\dagger > , \text{ par la proposition 2.4.3} \\ &= -1 + 2 < \rho, \sigma > , \text{ par la proposition 2.5.4} \\ &= < m(\rho), m(\sigma) > . \end{aligned}$$

□

La proposition 2.6.5 montrera que les $\varphi(U)$ sont en fait des rotations. Tâchons pour l'instant de rendre ces concepts un peu plus concrets. Prenons une matrice unitaire, disons la matrice de Pauli X. À cette matrice doit correspondre une rotation $\varphi(X)$ sur $\mathbb{R}^3$. Pour découvrir laquelle, il suffit de connaître son effet sur les vecteurs $\vec{\imath}$, $\vec{\jmath}$ et $\vec{k}$:

$$\begin{aligned}
\varphi(X)(\vec{\imath}) &= m \circ (\mathrm{Ad}X) \circ m^{-1}(\vec{\imath}) \\
&= m \circ (\mathrm{Ad}X)(\frac{1}{2}I + \frac{1}{2}X) \\
&= m(X(\frac{1}{2}I + \frac{1}{2}X)X^\dagger) \\
&= m(\frac{1}{2}I + \frac{1}{2}X) \\
&= \vec{\imath}.
\end{aligned}$$

Cela démontre que $\varphi(X)$ fixe l'axe des x. Aussi,

$$\begin{aligned}
\varphi(X)(\vec{\jmath}) &= m \circ (\mathrm{Ad}X) \circ m^{-1}(\vec{\jmath}) \\
&= m(X(\frac{1}{2}I + \frac{1}{2}Y)X^\dagger) \\
&= m(\frac{1}{2}I - \frac{1}{2}Y) \\
&= -\vec{\jmath}
\end{aligned}$$

et, suivant le même raisonnement,

$$\begin{aligned}
\varphi(X)(\vec{k}) &= m(X(\frac{1}{2}I + \frac{1}{2}Z)X^\dagger) \\
&= m(\frac{1}{2}I - \frac{1}{2}Z) \\
&= -\vec{k},
\end{aligned}$$

ce qui démontre que $\varphi(X)$ est en fait une rotation d'un demi-tour autour de l'axe des x.

De la même façon, nous pourrions montrer que $\varphi(Y)$ est une rotation d'un demi-tour autour de l'axe des y et $\varphi(Z)$, une rotation d'un demi-tour autour de l'axe des z.

La méthode que nous venons d'utiliser pour déterminer explicitement $\varphi(X)$ peut être généralisée. Pour une matrice unitaire quelconque U, il suffit de déterminer l'effet de $\mathrm{Ad}U$ sur les matrices de Pauli de trace zéro pour construire explicitement la matrice de rotation associée à $\varphi(U)$:

Proposition 2.6.4. *Soit $U \in \mathcal{SU}(2)$ et soit A la matrice de la transformation $\mathrm{Ad}U$, c'est-à-dire la matrice telle que pour $1 \leq j \leq 3$, $U\sigma_j U^\dagger = \sum_{i=1}^3 A_{ij}\sigma_i$. Alors $\varphi(U) = A$.*[6]

Démonstration. Soit $\vec{r} \in S_2$ un vecteur arbitraire de norme un. Alors,

$$\begin{aligned}
\varphi(U)(\vec{r}) &= m \circ \mathrm{Ad}U \circ m^{-1}(\vec{r}) \\
&= m \circ \mathrm{Ad}U \left(\frac{1}{2}I + \frac{1}{2}\sum_{j=1}^{3} r_j\sigma_j \right) \\
&= m \left(U \left(\frac{1}{2}I + \frac{1}{2}\sum_{j=1}^{3} r_j\sigma_j \right) U^\dagger \right) \\
&= m \left(\frac{1}{2}I + \frac{1}{2}\sum_{j=1}^{3} r_j U\sigma_j U^\dagger \right) \\
&= m \left(\frac{1}{2}I + \frac{1}{2}\sum_{j=1}^{3} r_j \sum_{i=1}^{3} A_{ij}\sigma_i \right) \\
&= m \left(\frac{1}{2}I + \frac{1}{2}\sum_{i=1}^{3} \left(\sum_{j=1}^{3} A_{ij} r_j \right) \sigma_i \right) \\
&= m \left(\frac{1}{2}I + \frac{1}{2}\sum_{i=1}^{3} s_i\sigma_i \right), \text{ où } s_i = \sum_{j=1}^{3} A_{ij} r_j \\
&= \vec{s} \\
&= A\vec{r}.
\end{aligned}$$

[6]Plus formellement, la matrice de la transformation $\varphi(U)$ est A. L'auteur de ce mémoire se permet d'utiliser les concepts de transformation linéaire et de matrice associée à une transformation linéaire de manière interchangeable.

Puisque $\varphi(U)$ est une transformation linéaire, $\varphi(U)(\vec{r}) = A\vec{r}$ pour tout $\vec{r} \in \mathbb{R}^3$, ce qui termine la preuve.

□

Proposition 2.6.5. *Pour tout $U \in \mathcal{SU}(2)$, $\varphi(U) \in \mathcal{SO}(3)$.*

Démonstration. Soit $U \in \mathcal{SU}(2)$. Nous savons déjà (par la proposition 2.6.3) que $\varphi(U) \in \mathcal{O}(3)$. La proposition 2.6.4 nous apprend comment construire la matrice associée à $\varphi(U)$: il s'agit de la matrice A telle que $U\sigma_j U^\dagger = \sum_{i=1}^3 A_{ij}\sigma_i$, pour $1 \leq j \leq 3$. Il reste simplement à montrer que le déterminant de A est 1.

Construisons la matrice $\tilde{A}$ telle que $U\sigma_j U^\dagger = \sum_{i=0}^3 \tilde{A}_{ij}\sigma_i$, pour $0 \leq j \leq 3$. Nous avons simplement ajouté l'équation triviale $U\sigma_0 U^\dagger = \sigma_0$ à notre système. La matrice $\tilde{A}$ est donc une extension triviale de la matrice A :

$$\tilde{A} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & & & \\ 0 & & A & \\ 0 & & & \end{bmatrix}.$$

On peut voir aisément que $|\tilde{A}| = |A|$. Remarquons aussi que $\tilde{A}$ est la matrice de la transformation $\mathrm{Ad}U : \mathcal{H}(2) \to \mathcal{H}(2)$, par rapport à la base $\{\sigma_i\}_{i=0}^3$.

Maintenant, grâce à l'identité $\mathrm{Vec}(ABC) = \mathrm{Vec}(C^T \otimes A)\mathrm{Vec}(B)$[7], on peut écrire

$$\begin{aligned} (U^* \otimes U)\mathrm{Vec}(\sigma_j) &= \mathrm{Vec}(U\sigma_j U^\dagger) \\ &= \mathrm{Vec}(\sum_{i=0}^3 \tilde{A}_{i,j}\sigma_i) \\ &= \sum_{i=0}^3 \tilde{A}_{i,j}\mathrm{Vec}(\sigma_i). \end{aligned}$$

[7]Cette identité est l'objet de la proposition 1.2.1.

Cela montre que $\tilde{A}$ est la matrice de la transformation $U^* \otimes U$ sur l'espace $\mathrm{Vec}(\mathcal{H}(2))$, par rapport à la base $\{\mathrm{Vec}(\sigma_i)\}_{i=0}^3$. Il faut donc que $\tilde{A}$ et $U^* \otimes U$ aient le même déterminant, et nous obtenons que

$$\begin{aligned} |\varphi(U)| &= |A| \\ &= |\tilde{A}| \\ &= |(U^* \otimes U)| \\ &= |U^*|^2 |U|^2, \text{ par la proposition 1.7.1} \\ &= 1. \end{aligned}$$

□

Étudions un autre exemple qui nous permettra de concrétiser ces notions. La transformée de Hadamard a la représentation matricielle suivante :

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

Calculons $\varphi(H)$ à l'aide de la proposition 2.6.4. Tout d'abord, nous avons vu, à la section 1.3.4, que

$$\begin{aligned} HXH &= Z, \\ HYH &= -Y, \\ HZH &= X. \end{aligned}$$

La proposition 2.6.4 nous apprend donc que

$$\varphi(H) = \begin{bmatrix} 0 & 0 & 1 \\ 0 & -1 & 0 \\ 1 & 0 & 0 \end{bmatrix}.$$

Pour trouver l'axe de rotation ℓ de $\varphi(H)$, il suffit d'en trouver un vecteur propre :

$$\begin{aligned}
\varphi(H)(\vec{r}) = \vec{r} \quad &\Leftrightarrow \quad \begin{bmatrix} 0 & 0 & 1 \\ 0 & -1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} r_1 \\ r_2 \\ r_3 \end{bmatrix} \\
&\Leftrightarrow \quad \begin{bmatrix} r_3 \\ -r_2 \\ r_1 \end{bmatrix} = \begin{bmatrix} r_1 \\ r_2 \\ r_3 \end{bmatrix} \\
&\Leftrightarrow \quad \vec{r} = \begin{bmatrix} r_1 \\ 0 \\ r_1 \end{bmatrix}.
\end{aligned}$$

Conséquemment, $\ell = \{(r, 0, r) \,|\, r \in \mathbb{R}\}$ est dans le plan xz et forme un angle de $\frac{\pi}{4}$ avec ces deux axes. De plus, puisque $\varphi(H)^2 = I$, il s'agit d'une rotation d'un demi-tour.

Voyons maintenant de quelle façon tout cela peut nous simplifier la vie. Pour calculer, par exemple, $\mathrm{Ad}H(|\circlearrowleft\rangle\langle\circlearrowleft|)$, on peut utiliser l'identité $\mathrm{Ad}H = m^{-1} \circ \varphi(H) \circ m$ et travailler dans $\mathbb{R}^3$, en gardant en tête[8] que $m(|\circlearrowleft\rangle\langle\circlearrowleft|) = \vec{\jmath}$:

$$\begin{aligned}
\mathrm{Ad}H(|\circlearrowleft\rangle\langle\circlearrowleft|) &= m^{-1} \circ \varphi(H) \circ m(|\circlearrowleft\rangle\langle\circlearrowleft|) \\
&= m^{-1} \circ \varphi(H)(\vec{\jmath}) \\
&= m^{-1}(-\vec{\jmath}) \\
&= |\circlearrowright\rangle\langle\circlearrowright|.
\end{aligned}$$

[8] Voir l'équation correspondante en page 42.

2.7 Rotations d'un demi-tour

Dans ce chapitre, nous porterons notre attention sur un ensemble de matrices particulier, celui des matrices à la fois unitaires et hermitiennes. De façon équivalente, on peut dire qu'il s'agit de l'ensemble des matrices unitaires qui sont également des racines carrées de l'identité. En effet, pour $U \in \mathcal{U}(2)$,

$$\begin{aligned} U \in \mathcal{H}(2) &\Leftrightarrow U = U^\dagger \\ &\Leftrightarrow U^2 = UU^\dagger \\ &\Leftrightarrow U^2 = I. \end{aligned}$$

Continuons notre réflexion. Soit U une matrice unitaire et hermitienne. Puisque qu'elle est hermitienne, on peut écrire $U = \sum_{i=0}^{3} r_i\sigma_i$, pour $r_i \in \mathbb{R}$. Aussi,

$$\begin{aligned} I &= U^2 \\ &= \left(\sum_{i=0}^{3} r_i\sigma_i\right)^2 \\ &= (r_0 I + \vec{r}\cdot\vec{\sigma})^2 \text{, où } \vec{r} = (r_1, r_2, r_3) \\ &= r_0^2 I + 2r_0(\vec{r}\cdot\vec{\sigma}) + (\vec{r}\cdot\vec{\sigma})^2 \\ &= r_0^2 I + 2r_0(\vec{r}\cdot\vec{\sigma}) + \|\vec{r}\|^2 I, \text{ par le lemme 2.4.1} \\ &= \sum_{i=0}^{3} r_i^2 I + 2r_0(\vec{r}\cdot\vec{\sigma}). \end{aligned}$$

Nous pouvons récrire ce résultat de la manière suivante :

$$\left(-1 + \sum_{i=0}^{3} r_i^2\right) I + 2r_0 r_1 X + 2r_0 r_2 Y + 2r_0 r_3 Z = 0.$$

Par l'indépendance linéaire des matrices de Pauli, il faut que le coefficient de chacun des termes de cette équation soit zéro.

Il faut donc que les conditions suivantes soient remplies :

1. $\sum_{i=0}^{3} r_i^2 = 1.$
2. $r_0 r_1 = r_0 r_2 = r_0 r_3 = 0.$

La condition (2) implique $r_0 = 0$ ou bien $r_1 = r_2 = r_3 = 0$.

Le cas $r_1 = r_2 = r_3 = 0$ demande, par la condition (1), que $r_0^2 = 1$. En d'autres mots, nous sommes en présence du cas trivial $U = \pm I$.

Autrement, le cas $r_0 = 0$ nous permet de récrire la condition (1) de la façon suivante :

$$r_1^2 + r_2^2 + r_3^2 = 1. \tag{2.3}$$

Nous retrouvons donc une paramétrisation identique à celle d'une sphère !

En résumé, si U est à la fois hermitienne et unitaire, il n'y a que deux situations possibles : soit U est triviale — c'est-à-dire que $U = \pm I$ —, soit $U = \vec{r} \cdot \vec{\sigma}$, pour un certain $\vec{r} \in S_2$. Le cas trivial est sans intérêt, mais l'autre cas suggère une correspondance entre S_2 et $\mathcal{U}_{\mathcal{H}}(2)$, l'ensemble défini comme suit :

Définition 2.7.1. $\mathcal{U}_{\mathcal{H}}(2) \stackrel{\text{def}}{=} (\mathcal{H}(2) \cap \mathcal{U}(2)) \setminus \{I, -I\}$

Définition 2.7.2. *Nous nommerons ω la fonction $\omega : S_2 \to \mathcal{U}_{\mathcal{H}}(2)$ définie par $\omega(\vec{r}) \stackrel{\text{def}}{=} \vec{r} \cdot \vec{\sigma}$.*

Par la discussion précédente, cette fonction est surjective, car nous avons vu que toute matrice hermitienne et unitaire qui n'est pas $\pm I$ doit répondre à la condition 2.3. Elle est aussi clairement injective, car $\vec{r} \cdot \vec{\sigma} = \vec{s} \cdot \vec{\sigma}$ implique nécessairement que $\vec{r} = \vec{s}$. Nous venons de démontrer la proposition suivante :

Proposition 2.7.3. *La fonction ω est bijective.*

Les éléments de $\mathcal{U}_{\mathcal{H}}(2)$ étant des matrices U non triviales telles que $U^2 = I$, les rotations $\varphi(U)$ qui leur correspondent doivent être des rotations d'un demi-tour non triviales, puisque le noyau de la fonction φ contient seulement les matrices I et $-I$, et que

$$\begin{aligned}
\varphi(U)^2 &= m \circ \mathrm{Ad}U \circ m^{-1} \circ m \circ \mathrm{Ad}U \circ m^{-1} \\
&= m \circ \mathrm{Ad}(U^2) \circ m^{-1} \\
&= m \circ \mathrm{Ad}I \circ m^{-1} \\
&= I.
\end{aligned}$$

Soit $U \in \mathcal{U}_{\mathcal{H}}(2)$. Quel peut être l'axe de rotation de $\varphi(U)$? Pourrait-t-il y avoir un lien avec le vecteur associé à U par la fonction ω ? Miraculeusement, le vecteur $\omega^{-1}(U)$ *est* l'axe de rotation de $\varphi(U)$:

Proposition 2.7.4. *Soit $\vec{r} \in S_2$. Alors $\vec{r}$ est un point fixe de $\varphi(\omega(\vec{r}))$.*

Démonstration. Premièrement, posons $U = \omega(\vec{r}) = \vec{r} \cdot \vec{\sigma}$. Rappelons-nous aussi que $m^{-1}(\vec{r}) = \frac{1}{2}(I + \vec{r} \cdot \vec{\sigma})$, pour constater que $m^{-1}(\vec{r})$ est un point fixe de $\mathrm{Ad}U$.

$$\begin{aligned}
Um^{-1}(\vec{r})U^\dagger &= U\frac{1}{2}(I + \vec{r} \cdot \vec{\sigma})U \\
&= \frac{1}{2}U(I + U)U \\
&= \frac{1}{2}(U^2 + U^3) \\
&= \frac{1}{2}(I + U) \\
&= m^{-1}(\vec{r}).
\end{aligned}$$

On peut maintenant voir que $\vec{r}$ est un point fixe de $\varphi(U)$:

$$\begin{aligned}\varphi(U)(\vec{r}) &= m \circ (\mathrm{Ad}U) \circ m^{-1}(\vec{r}) \\ &= m(Um^{-1}(\vec{r})U^\dagger) \\ &= m(m^{-1}(\vec{r})) \\ &= \vec{r}.\end{aligned}$$

□

2.8 Trouver la matrice unitaire associée à une rotation arbitraire

La proposition 2.6.4 permet, étant donné une matrice unitaire U, d'écrire explicitement la matrice de $\varphi(U)$. Dans cette section, nous solutionnons le problème inverse : étant donné une rotation arbitraire $R : \mathbb{R}^3 \to \mathbb{R}^3$, comment trouver une matrice U telle que $\varphi(U) = R$?

La discussion de la section précédente nous apporte un début de solution : étant donné un axe $\vec{r}$, la matrice $U = \omega(\vec{r})$ est telle que $\varphi(U)$ est une rotation d'un demi-tour autour de $\vec{r}$. Pour développer encore les concepts de la section précédente, prenons $U \in \mathcal{U}_{\mathcal{H}}(2)$ et considérons le sous-groupe à un paramètre engendré par l'exponentielle de cette matrice :

$$\begin{aligned}e^{i\theta U} &\stackrel{\text{def}}{=} \sum_{j=0}^{\infty} \frac{(i\theta U)^j}{j!} \\ &= \sum_{j=0}^{\infty} \frac{(i\theta U)^{2j}}{(2j)!} + \sum_{j=0}^{\infty} \frac{(i\theta U)^{2j+1}}{(2j+1)!}\end{aligned}$$

$$
\begin{aligned}
&= I\sum_{j=0}^{\infty}\frac{(-1)^j\theta^{2j}}{(2j)!} + iU\sum_{j=0}^{\infty}\frac{(-1)^j\theta^{2j+1}}{(2j+1)!}\\
&= \cos(\theta)I + i\sin(\theta)U.
\end{aligned}
$$

Remarquons que $e^{i\theta U}$ est unitaire :

$$
\begin{aligned}
(\cos(\theta)I + i\sin(\theta)U)(\cos(\theta)I + i\sin(\theta)U)^\dagger &= (\cos(\theta)I + i\sin(\theta)U)(\cos(\theta)I - i\sin(\theta)U^\dagger)\\
&= \cos^2(\theta)I + \sin^2(\theta)UU^\dagger\\
&= I.
\end{aligned}
$$

Le théorème qui suit constitue l'objectif principal de la présente section. Étant donné une rotation quelconque $R \in \mathcal{SO}(3)$, il permet de construire une matrice unitaire V telle que $\varphi(V) = R$.

Théorème 2.8.1. *Soit $\vec{r} \in \mathbb{R}^3$ tel que $\|\vec{r}\| = 1$ et soit $\theta \in \mathbb{R}$. Posons $U = \vec{r}\cdot\vec{\sigma}$ et définissons $U(\theta) \stackrel{\text{def}}{=} e^{i\frac{\theta}{2}U} = \cos(\frac{\theta}{2})I + i\sin(\frac{\theta}{2})U$. Alors $\varphi(U(\theta))$ est une rotation de θ radians autour de $\vec{r}$.*[9]

Prenons soin de noter que $\varphi(U(\pi)) = \varphi(U)$. En effet,

$$
\begin{aligned}
\varphi(U(\pi)) &= \varphi(iU)\\
&= m\circ \mathrm{Ad}(iU)\circ m^{-1}\\
&= m\circ \mathrm{Ad}(U)\circ m^{-1}\\
&= \varphi(U).
\end{aligned}
$$

Avant de passer à la démonstration du théorème 2.8.1, nous aurons besoin des lemmes suivants :

[9] La rotation autour d'un vecteur se fait dans la direction donnée par la règle de la main droite : quand le pouce droit pointe vers le haut, on tourne dans la direction pointée par les doigts.

Lemme 2.8.2. *Soit $\theta \in \mathbb{R}$ et $Z(\theta) = \cos(\frac{\theta}{2})I + i\sin(\frac{\theta}{2})Z$. Alors $\varphi(Z(\theta))$ est une rotation de θ radians autour de $\vec{k}$.*

Démonstration. Comme $\varphi(Z)$ est une rotation d'un demi-tour autour de $\omega^{-1}(Z) = \vec{k}$, on s'attend à ce que $\varphi(Z(\theta))$ soit une rotation de θ radians autour du même axe. Vérifions cette assertion à l'aide de la proposition 2.6.4, en commençant par déterminer l'effet de $\mathrm{Ad}(Z(\theta))$ sur les matrices X, Y et Z.

$$\begin{aligned}
Z(\theta)XZ(\theta)^\dagger &= \left(\cos\left(\frac{\theta}{2}\right)I + i\sin\left(\frac{\theta}{2}\right)Z\right)X\left(\cos\left(\frac{\theta}{2}\right)I - i\sin\left(\frac{\theta}{2}\right)Z\right) \\
&= \cos^2\left(\frac{\theta}{2}\right)X - i\cos\left(\frac{\theta}{2}\right)\sin\left(\frac{\theta}{2}\right)XZ \\
&\quad + i\cos\left(\frac{\theta}{2}\right)\sin\left(\frac{\theta}{2}\right)ZX + \sin^2\left(\frac{\theta}{2}\right)ZXZ \\
&= \left(\cos^2\left(\frac{\theta}{2}\right) - \sin^2\left(\frac{\theta}{2}\right)\right)X + 2\cos\left(\frac{\theta}{2}\right)\sin\left(\frac{\theta}{2}\right)Y \\
&= \cos(\theta)X + \sin(\theta)Y.
\end{aligned}$$

Nous pouvons substituer Y à X dans les équations apparaissant ci-haut pour obtenir

$$\begin{aligned}
Z(\theta)YZ(\theta)^\dagger &= \cos(\theta)Y + \sin(\theta)iYZ \qquad (2.4) \\
&= \cos(\theta)Y - \sin(\theta)X.
\end{aligned}$$

Notons que dans la ligne 2.4, nous avons utilisé l'identité $Y = iXZ$ avant de substituer Y à X. Finalement,

$$\begin{aligned}
Z(\theta)ZZ(\theta)^\dagger &= \left(\cos\left(\frac{\theta}{2}\right)I + i\sin\left(\frac{\theta}{2}\right)Z\right)Z\left(\cos\left(\frac{\theta}{2}\right)I - i\sin\left(\frac{\theta}{2}\right)Z\right) \\
&= \cos^2\left(\frac{\theta}{2}\right)Z - i\cos\left(\frac{\theta}{2}\right)\sin\left(\frac{\theta}{2}\right)I \\
&\quad + i\cos\left(\frac{\theta}{2}\right)\sin\left(\frac{\theta}{2}\right)I + \sin^2\left(\frac{\theta}{2}\right)Z \\
&= \left(\cos^2\left(\frac{\theta}{2}\right) + \sin^2\left(\frac{\theta}{2}\right)\right)Z \\
&= Z.
\end{aligned}$$

La proposition 2.6.4 nous permet à présent de construire explicitement la matrice de $\varphi(Z(\theta))$:

$$\varphi(Z(\theta)) = \begin{bmatrix} \cos(\theta) & -\sin(\theta) & 0 \\ \sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Cette matrice représente bel et bien une rotation d'un angle θ autour de $\vec{k}$.

□

Lemme 2.8.3. *Soit $U \in \mathcal{U}_{\mathcal{H}}(2)$. Alors il existe une matrice unitaire $C \in \mathcal{SU}(2)$ telle que $U = CZC^{\dagger}$.*

Démonstration. Puisque U est hermitienne, il existe une matrice $D \in \mathcal{SU}(2)$ telle que $DUD^{\dagger} = \text{diag}(\lambda_1, \lambda_2)$, où $\lambda_1, \lambda_2 \in \mathbb{R}$. Comme U est également unitaire, on peut aussi affirmer que $|\lambda_1| = |\lambda_2| = 1$, ce qui implique que $\lambda_1, \lambda_2 \in \{1, -1\}$. De plus,

$$\begin{aligned} \lambda_1 + \lambda_2 &= \text{tr}\left(DUD^{\dagger}\right) \\ &= \text{tr}\left(U\right) \\ &= \text{tr}\left(\vec{r} \cdot \vec{\sigma}\right), \text{ où } \vec{r} \in S_2 \\ &= 0, \end{aligned}$$

puisque $\text{tr}\left(\sigma_i\right) = 0$ quand $1 \leq i \leq 3$. Il faut donc nécessairement que $\lambda_1 = -\lambda_2$. En d'autres mots, $DUD^{\dagger} = \pm Z$.

Si $DUD^{\dagger} = Z$, la preuve est terminée. Si $DUD^{\dagger} = -Z$, On pose $C = XD$, pour obtenir

$$\begin{aligned} CUC^{\dagger} &= XDUD^{\dagger}X \\ &= -XZX \\ &= Z. \end{aligned}$$

□

Démonstration du théorème 2.8.1. Soit $\vec{r}, U$ et $U(\theta)$ tels que dans l'énoncé du théorème. Nous verrons qu'un simple changement de base permet de passer du lemme 2.8.3 au résultat du présent théorème.

Par le lemme 2.8.3, on peut trouver $C \in \mathcal{SU}(2)$ telle que $CZC^\dagger = U$. Il s'ensuit immédiatement que

$$\begin{aligned} CZ(\theta)C^\dagger &= C\left(\cos\left(\frac{\theta}{2}\right) I + i\sin\left(\frac{\theta}{2}\right) Z\right) C^\dagger \\ &= \cos\left(\frac{\theta}{2}\right) CIC^\dagger + i\sin\left(\frac{\theta}{2}\right) CZC^\dagger \\ &= \cos\left(\frac{\theta}{2}\right) I + i\sin\left(\frac{\theta}{2}\right) U \\ &= U(\theta). \end{aligned}$$

Ce changement de base se transporte automatiquement dans $\mathbb{R}^3$:

$$\begin{aligned} \varphi(U(\theta)) &= \varphi(CZ(\theta)C^\dagger) \\ &= m \circ \mathrm{Ad}(CZ(\theta)C^\dagger) \circ m^{-1} \\ &= m \circ \mathrm{Ad}C \circ \mathrm{Ad}Z(\theta) \circ \mathrm{Ad}C^\dagger \circ m^{-1}, \text{ par la propostition 2.5.2} \\ &= m \circ \mathrm{Ad}C \circ m^{-1} \circ m \circ \mathrm{Ad}Z(\theta) \circ m^{-1} \circ m \circ \mathrm{Ad}C^\dagger \circ m^{-1} \\ &= \varphi(C) \circ \varphi(Z(\theta)) \circ \varphi(C^\dagger) \\ &= \varphi(C) \circ \varphi(Z(\theta)) \circ \varphi(C)^{-1}. \end{aligned}$$

L'égalité $\varphi(C)^{-1} = \varphi(C^\dagger)$, utilisée ci-dessus, découle des égalités suivante :

$$\begin{aligned} \varphi(C) \circ \varphi(C^\dagger) &= m \circ \mathrm{Ad}C \circ m^{-1} \circ m \circ \mathrm{Ad}C^\dagger \circ m^{-1} \\ &= m \circ \mathrm{Ad}C \circ \mathrm{Ad}C^\dagger \circ m^{-1} \\ &= m \circ \mathrm{Ad}(CC^\dagger) \circ m^{-1} \\ &= m \circ \mathrm{Ad}I \circ m^{-1} \\ &= m \circ I \circ m^{-1} \\ &= I. \end{aligned}$$

Puisque $\varphi(C) \in \mathcal{SO}(3)$ est une rotation, la conjugaison par $\varphi(C)$ est un changement de base qui préserve l'orientation. C'est donc dire que $\varphi(U(\theta))$ est simplement la rotation $\varphi(Z(\theta))$ autour d'un autre axe.

Pour comprendre que l'axe de rotation de $\varphi(U(\theta))$ est le même que celui de $\varphi(U)$, c'est-à-dire $\vec{r}$, il suffit de jeter un coup d'œil aux équations suivantes :

$$\begin{aligned}
\varphi(C)(\vec{k}) &= m \circ \mathrm{Ad}C \circ m^{-1}(\vec{k}) \\
&= m(C(\frac{1}{2}I + \frac{1}{2}Z)C^{\dagger}) \\
&= m(\frac{1}{2}I + \frac{1}{2}U) \\
&= m(\frac{1}{2}I + \frac{1}{2}\vec{r}\cdot\vec{\sigma}) \\
&= \vec{r}.
\end{aligned}$$

Ceci nous montre que $\vec{r}$ est un point fixe de $U(\theta)$:

$$\begin{aligned}
\varphi(U(\theta))(\vec{r}) &= \varphi(C) \circ \varphi(Z(\theta)) \circ \varphi(C)^{-1}(\vec{r}) \\
&= \varphi(C) \circ \varphi(Z(\theta))(\vec{k}) \\
&= \varphi(C)(\vec{k}) \\
&= \vec{r}.
\end{aligned}$$

□

Corollaire 2.8.4. *La fonction φ est surjective.*

Corollaire 2.8.5. *Les groupes $\mathcal{SU}(2)/\{I, -I\}$ et $\mathcal{SO}(3)$ sont isomorphes.*

Démonstration. Nous savons déjà que $\varphi : \mathcal{SU}(2) \to \mathcal{SO}(3)$ est un homomorphisme surjectif (grâce aux propositions 2.6.2 et 2.8.4). Aussi, le noyau de φ est le même

que celui de Ad, c'est-à-dire $\{I, -I\}$, car

$$\begin{aligned} U \in \operatorname{Ker}\varphi &\Leftrightarrow m \circ \mathrm{Ad}U \circ m^{-1} = I \\ &\Leftrightarrow \mathrm{Ad}U = I \\ &\Leftrightarrow U \in \operatorname{Ker}\mathrm{Ad}. \end{aligned}$$

On peut donc conclure que la fonction $\bar{\varphi} : \mathcal{SU}(2)/\{I, -I\} \to \mathcal{SO}(3)$ donnée par la règle $\bar{\varphi}(\bar{U}) = \varphi(U)$[10], est un isomorphisme, par le Premier théorème des isomorphismes[11].

□

2.9 Entropie de von Neumann

La notion d'entropie d'un état quantique a été définie à la section 1.10.1. La présente section vise à présenter une vision géométrique de cette quantité. Plus précisément, il apparaîtra que l'entropie de von Neumann d'un état ρ peut être calculé à partir de la quantité $\|m(\rho)\|$.

Soit $\rho \in \mathcal{D}(2)$. Traçons, à travers l'origine, la ligne $\ell \subset \mathbb{R}^3$ passant aussi par $m(\rho)$. Les deux points où ℓ intersecte la sphère S_2 définissent deux vecteurs de sens opposé, disons $\vec{r}$ et $-\vec{r}$. Notons tout d'abord que les états purs $m^{-1}(\vec{r})$ et $m^{-1}(-\vec{r})$ sont nécessairement orthogonaux :

Proposition 2.9.1. *Soit $\vec{r} \in \mathbb{R}^3$ tel que $\|\vec{r}\| = 1$ et posons $|\psi_0\rangle\langle\psi_0| = m^{-1}(\vec{r})$, $|\psi_1\rangle\langle\psi_1| = m^{-1}(-\vec{r})$. Alors $\langle\psi_0|\psi_1\rangle = 0$.*

[10]Ici, $\bar{U}$ dénote la classe d'équivalence de U, c'est-à-dire $\{U, -U\}$, dans le groupe quotient $\mathcal{SU}(2)/\{I, -I\}$.

[11]À ce sujet, le lecteur pourra consulter [Art91], page 68.

Résumé de la démonstration. Puisque $\vec{r}$ peut être obtenu de $\vec{k}$ par une simple rotation R, on s'attend à ce que $|\psi_0\rangle$ et $|\psi_1\rangle$ puissent être obtenu de $|0\rangle$ et $|1\rangle$ par un simple transformation unitaire, plus précisément qu'une matrice U telle que $R = \varphi(U)$ donne $U|\psi_0\rangle = |0\rangle$ et $U|\psi_1\rangle = |1\rangle$. Puisque qu'une matrice unitaire préserve le produit interne, on aura $\langle\psi_0|\psi_1\rangle = \langle 0|1\rangle = 0$.

□

Démonstration. Soit $\vec{r}$, $|\psi_0\rangle$ et $|\psi_1\rangle$ tels que dans l'énoncé de la proposition, soit $R \in SO(3)$ une rotation telle que $R(\vec{k}) = \vec{r}$ et $U \in SU(2)$ telle que $\varphi(U) = R$. Alors

$$\begin{aligned}
|\psi_0\rangle\langle\psi_0| &= m^{-1}(\vec{r}) \\
&= m^{-1}(R(\vec{k})) \\
&= m^{-1}(\varphi(U)(\vec{k})) \\
&= \mathrm{Ad}U \circ m^{-1}(\vec{k}), \text{ par définition de } \varphi \ (2.6.1) \\
&= U|0\rangle\langle 0|U^\dagger.
\end{aligned}$$

De façon similaire,

$$\begin{aligned}
|\psi_1\rangle\langle\psi_1| &= m^{-1}(-\vec{r}) \\
&= \mathrm{Ad}U \circ m^{-1}(-\vec{k}) \\
&= U|1\rangle\langle 1|U^\dagger.
\end{aligned}$$

Il s'ensuit que

$$\begin{aligned}
|\langle\psi_0|\psi_1\rangle|^2 &= \mathrm{tr}\left(|\psi_0\rangle\langle\psi_0|\psi_1\rangle\langle\psi_1|\right) \\
&= \mathrm{tr}\left(U|0\rangle\langle 0|U^\dagger U|1\rangle\langle 1|U^\dagger\right) \\
&= \mathrm{tr}\left(|0\rangle\langle 0|1\rangle\langle 1|\right) \\
&= 0.
\end{aligned}$$

Donc $\langle\psi_0|\psi_1\rangle = 0$.

□

De la même façon que tout point de B_3 est situé sur une ligne ℓ passant par deux points opposés à la surface, tout état ρ est situé sur une ligne entre deux états purs orthogonaux, au sens où ρ est une combinaison convexe de ces deux états :

Proposition 2.9.2. *Soit $\rho \in \mathcal{D}(2)$ et supposons que $m(\rho) = k\vec{r}$, où $k \in [0,1]$ et $\|\vec{r}\| = 1$. Posons à nouveau $|\psi_0\rangle\langle\psi_0| = m^{-1}(\vec{r})$, $|\psi_1\rangle\langle\psi_1| = m^{-1}(-\vec{r})$. Alors $\rho = p\,|\psi_0\rangle\langle\psi_0| + (1-p)\,|\psi_1\rangle\langle\psi_1|$, où $p = \frac{k+1}{2}$.*

Démonstration. Posons $p = \frac{k+1}{2}$. Puisque, par la proposition 2.4.3, m est une transformation affine,

$$\begin{aligned} m\left(p\,|\psi_0\rangle\langle\psi_0| + (1-p)\,|\psi_1\rangle\langle\psi_1|\right) &= pm\left(|\psi_0\rangle\langle\psi_0|\right) + (1-p)m\left(|\psi_1\rangle\langle\psi_1|\right) \\ &= p\vec{r} + (1-p)(-\vec{r}) \\ &= (2p-1)\vec{r} \\ &= \left(2\left(\frac{k+1}{2}\right) - 1\right)\vec{r} \\ &= k\vec{r}. \end{aligned}$$

Donc $\rho = m^{-1}(k\vec{r}) = p\,|\psi_0\rangle\langle\psi_0| + (1-p)\,|\psi_1\rangle\langle\psi_1|$.

□

Corollaire 2.9.3. *Soit $\rho \in \mathcal{D}(2)$ et supposons que $\|m(\rho)\| = k \in [0,1]$. Alors $H(\rho) = -p\log(p) - (1-p)\log(1-p)$, où $p = \frac{k+1}{2}$.*

Démonstration. La proposition 2.9.2 nous donne deux états purs $|\psi_0\rangle$ et $|\psi_1\rangle$ tels que $\rho = p\,|\psi_0\rangle\langle\psi_0| + (1-p)\,|\psi_1\rangle\langle\psi_1|$, où $p = \frac{k+1}{2}$. De plus, la proposition 2.9.1 précise que

$|\psi_0\rangle$ et $|\psi_1\rangle$ sont orthogonaux. Nous avons donc devant les yeux la décomposition spectrale de ρ, dont les valeurs propres sont p et $(1-p)$.

□

2.10 La sphère de Poincaré pour les espaces de dimension d

On entend souvent qu'il n'y a pas d'équivalent à la sphère de Poincaré pour les espaces de dimension supérieure à deux. Il est vrai qu'on n'y retrouve pas d'isomorphisme semblable à celui du corollaire 2.8.5. Cela ne veut toutefois pas dire que la généralisation de la sphère de Poincaré aux espaces de dimensions supérieures à deux est dénuée d'intérêt.

2.10.1 Une sphère dans $\mathbb{R}^{d^2-1}$

Toute matrice de densité $\rho \in \mathcal{D}(d)$, puisque sa trace vaut un, répond à la condition $\rho - \frac{1}{d}I \in \mathcal{H}_0(d)$. Puisque les matrices hermitiennes $d \times d$ de trace zéro forment un espace vectoriel réel de dimension $D \stackrel{\text{def}}{=} d^2 - 1$, on peut toujours trouver une base $\{\tau_i\}_{i=1}^{D}$ de $\mathcal{H}_0(d)$ et des coefficients réels r_i tels que $\rho - \frac{1}{d}I = \frac{1}{2}\sum_{i=1}^{D} r_i\tau_i$, ou encore,

$$\rho = \frac{1}{d}I + \frac{1}{2}\sum_{i=1}^{D} r_i\tau_i.$$

Par convention, nous choisissons les τ_i tels que $< \tau_i, \tau_j >= \mathrm{tr}\,(\tau_i\tau_j) = 2\delta_{i,j}$. Par exemple, quand $d = 2^n$, on peut prendre pour base $\{\tau_i\}_{i=1}^{D}$ les $4^n - 1 = D$ produits tensoriels (renormalisés) de n matrices de Pauli qui ne sont pas toutes l'identité :

$$\tau_i = \frac{1}{2^{n-1}} \bigotimes_{k=1}^{n} \sigma_{i_k},$$

où $i_k \in \{0, 1, 2, 3\}$ et au moins un des i_k est différent de zéro.

Définissons B_D comme étant la boule de $\mathbb{R}^D$ qui contient tous les vecteurs dont la norme vaut au plus $R_d = \sqrt{2(1 - \frac{1}{d})}$:

$$B_D \stackrel{\text{def}}{=} \left\{ \vec{r} \in \mathbb{R}^D \;\middle|\; \|\vec{r}\| \leq \sqrt{2\left(1 - \frac{1}{d}\right)} \right\}.$$

Le fait que les états purs ψ correspondent à des vecteurs de norme R_d peut être vérifié grâce au deuxième point de la proposition 2.10.1, énoncée plus bas.

$$\begin{aligned} 1 &= \operatorname{tr}\left(\psi^2\right) \\ &= <\psi, \psi> \\ &= \frac{1}{d} + \frac{1}{2} < m(\psi), m(\psi) > \\ &= \frac{1}{d} + \frac{1}{2}\|m(\psi)\|^2, \end{aligned}$$

ce qui montre que $\|m(\psi)\| = \sqrt{2(1 - \frac{1}{d})}$.

La fonction m définie par l'équation 2.2 se généralise aisément en une fonction $m : \mathcal{H}_0(d) \to \mathbb{R}^D$ définie comme ceci :

$$m(\frac{1}{d}I + \frac{1}{2}\vec{r} \cdot \vec{\tau}) = \vec{r}.$$

Proposition 2.10.1.[12]

1. *m est une bijection affine.*
2. *$\langle \rho, \sigma \rangle = \frac{1}{d} + \frac{1}{2} < m(\rho), m(\sigma) >$.*
3. *$d(\rho, \sigma) = \frac{1}{2}\|m(\rho) - m(\sigma)\|$.*
4. *$m(\mathcal{D}_{pur}(d)) \subseteq S_{D-1}$ et $m(\mathcal{D}(d)) \subseteq B_D$.*

[12]On retrouve la preuve de cette proposition dans [Zan98].

Cette fonction généralisée conserve essentiellement les caractéristiques que nous lui connaissons, à une exception près. Comme dans le cas $d = 2$, les $m(\rho)$ sont contenus dans B_D, étant situés sur la surface S_{D-1} quand ρ est pur. Seulement, quand $d > 2$, certains points de S_{D-1} et de B_D ne correspondent pas à un état quantique.

Par exemple, si $\vec{r} \in S_{D-1}$, alors $m^{-1}(\vec{r})$ et $m^{-1}(-\vec{r})$ ne peuvent être tous deux des états quantiques. En effet, si tel était le cas, on pourrait poser $|\psi_0\rangle\langle\psi_0| = m^{-1}(\vec{r})$ et $|\psi_1\rangle\langle\psi_1| = m^{-1}(-\vec{r})$ et on obtiendrait

$$\begin{aligned} < |\psi_0\rangle\langle\psi_0|, |\psi_1\rangle\langle\psi_1| > &= \operatorname{tr}\left(|\psi_0\rangle\langle\psi_0|\psi_1\rangle\langle\psi_1|\right) \\ &= |\langle\psi_0|\psi_1\rangle|^2 \\ &\geq 0. \end{aligned}$$

Mais la proposition 2.10.1 affirme que

$$\begin{aligned} < |\psi_0\rangle\langle\psi_0|, |\psi_1\rangle\langle\psi_1| > &= \frac{1}{d} + \frac{1}{2} < m(|\psi_0\rangle\langle\psi_0|), m(|\psi_1\rangle\langle\psi_1|) > \\ &= \frac{1}{d} + \frac{1}{2} < \vec{r}, -\vec{r} > \\ &= \frac{1}{d} - \frac{1}{2}\|\vec{r}\|^2 \\ &= \frac{1}{d} - \frac{1}{2}\left(2\left(1 - \frac{1}{d}\right)\right) \\ &= \frac{2}{d} - 1 \\ &< 0, \text{ lorsque } d > 2, \end{aligned}$$

ce qui contredit l'hypothèse voulant que $m^{-1}(\vec{r})$ et $m^{-1}(-\vec{r})$ soient tous deux des états quantiques.

Si on définit m sur le domaine $\mathcal{H}_0(d)$ plutôt que sur $\mathcal{D}(d)$, c'est justement parce que $m(\mathcal{D}(d)) \subsetneq B_D$ quand $d > 2$. En prenant $\mathcal{H}_0(d)$ pour domaine, on s'assure que m reste bijective — car $m(\mathcal{H}_0(d)) = \mathbb{R}^D$ — et donc, que m^{-1} existe.

Cela permet de définir $\varphi : \mathcal{SU}(d) \to \mathcal{SO}(D)$ de la manière habituelle, c'est-à-dire $\varphi(U) = m \circ \mathrm{Ad}U \circ m^{-1}$. Plusieurs résultats des sections 2.5 et 2.6 s'appliquent

toujours quand $d > 2$. C'est le cas des propositions 2.5.2, 2.5.4, 2.6.2, 2.6.3, 2.6.4 et 2.6.5.

La différence majeure avec les résultats précédents est que dans le cas où $d > 2$, φ n'est pas surjective. Autrement dit, $\varphi(\mathcal{SU}(d)) \subsetneq \mathcal{SO}(D)$.

Pour illustrer ce fait, reprenons l'exemple précédent. Nous avons vu qu'étant donné $\vec{r} \in S_{D-1}$, les matrices hermitiennes $m^{-1}(\vec{r})$ et $m^{-1}(-\vec{r})$ ne peuvent pas être toutes deux des états quantiques. Prenons un vecteur $\vec{r}$ tel que $m(|\psi_0\rangle \langle\psi_0|) = \vec{r}$. Il existe nécessairement une transformation $R \in \mathcal{SO}(D)$ telle que $R(\vec{r}) = -\vec{r}$. S'il existait aussi une transformation $U \in \mathcal{SU}(d)$ telle que $\varphi(U) = R$, alors on pourrait conclure que

$$\begin{aligned} -\vec{r} &= R(\vec{r}) \\ &= \varphi(U)(\vec{r}) \\ &= m \circ \mathrm{Ad}U \circ m^{-1}(\vec{r}) \\ &= m(U |\psi_0\rangle \langle\psi_0| U^\dagger) \\ &= m(|\psi_1\rangle \langle\psi_1|), \text{ où } |\psi_1\rangle = U |\psi_0\rangle , \end{aligned}$$

ce qui nous amènerait à une contradiction. Il n'existe donc pas de matrice $U \in \mathcal{SU}(d)$ telle que $\varphi(U) = R$, ce qui montre que φ n'est pas surjective.

Chapitre 3

Canaux quantiques privés

Ce chapitre traite de la transmission confidentielle d'information quantique. Les résultats présentés ici se veulent une extension de ceux de [AMTdW00]. Le modèle utilisé est l'équivalent quantique du célèbre chiffre de Vernam, dont nous reverrons à l'instant les principales caractéristiques.

3.1 Le chiffre de Vernam

Considérons un scénario à deux participants, Alice et Bob, dans lequel Alice veut envoyer un message à Bob sur un canal public, par exemple une ligne téléphonique, de façon confidentielle. On veut qu'une espionne, Ève, écoutant la ligne, ne soit capable d'obtenir aucune information sur le message en transit.

Si Alice et Bob partagent une clé secrète, c'est-à-dire une séquence de bits connue d'eux seuls, il existe une solution parfaite au problème : supposons qu'Alice et Bob partagent une clé secrète de n bits, $k \in (\mathbb{Z}_2)^n$, choisie au hasard selon la distribution uniforme sur $(\mathbb{Z}_2)^n$. Alors Alice peut chiffrer le message $m \in (\mathbb{Z}_2)^n$ de la façon

suivante :

$$e_k(m) = m + k.$$

Autrement dit, le i^{e} bit du message chiffré est le ou-exclusif du i^{e} bit du message en clair avec le i^{e} bit de la clé. On déchiffre ensuite avec la même fonction, c'est-à-dire que $d_k = e_k$:

$$d_k(m + k) = m + k + k = m.$$

3.1.1 Confidentialité parfaite

Soit $\mathcal{M}$ un ensemble de messages en clair, $\mathcal{C}$ un ensemble de cryptogrammes et $\mathcal{K}$ un ensemble de clés. Un chiffre associe à chaque clé $k \in \mathcal{K}$ une fonction $e_k : \mathcal{M} \to \mathcal{C}$. Disons qu'Alice chiffre un message m et que son cryptogramme $e_k(m)$ soit intercepté par Ève. Puisque Ève ne connaît pas la clé k, mais connaît la distribution de probabilité selon laquelle a été choisie k, elle peut considérer la clé comme une variable aléatoire K et le cryptogramme comme une variable aléatoire $e_K(m)$.

Pour qu'un chiffre soit parfaitement confidentiel, il faut que l'obtention par Ève du cryptogramme $e_K(m)$ ne lui donne aucun avantage pour deviner lequel des messages de $\mathcal{M}$ a été chiffré par Alice. En termes mathématiques, nous dirons qu'un chiffre est **parfaitement confidentiel** si l'égalité $e_K(m_1) = e_K(m_2)$ vaut pour toute paire de messages $m_1, m_2 \in \mathcal{M}$.[1]

Le chiffre de Vernam est parfaitement confidentiel.[2] Dans ce chiffre, la clé K prend ses valeurs uniformément dans $(\mathbb{Z}_2)^n$. L'entropie de K vaut $H(K) = n$. En ce sens, nous dirons que n bits de clé sont suffisants pour chiffrer les messages $m \in (\mathbb{Z}_2)^n$.

[1] Nous considérons que deux variables aléatoires X et Y sont égales si $P(X = x) = P(Y = x)$ pour toute valeur x prise par X ou par Y.

[2] Voir [Sti95], page 49.

Le chiffre de Vernam est également optimal, au sens où n bits sont nécessaires pour réaliser cette même tâche. Plus précisément, on peut montrer[3] qu'il faut absolument que $H(K) \geq H(M)$[4] pour obtenir une confidentialité parfaite. Si tous les messages en clair sont équiprobables, et donc que $H(M) = n$, il faudra que $H(K) \geq n$.

3.2 Canaux quantiques

Portons maintenant notre attention sur la transmission d'information quantique. Nous voulons qu'Alice puisse envoyer un état quantique ρ à Bob de façon à ce que celui-ci puisse recouvrer ρ avec certitude, mais, dans l'éventualité où Ève intercepte le message chiffré $\mathcal{E}(\rho)$, qu'elle ne puisse obtenir aucune information sur ρ.

En permettant à Alice et Bob de partager des paires EPR, la tâche deviendrait triviale : ils n'auraient qu'à téléporter[5] l'état à transmettre. Le cas qui nous intéresse est plus restrictif : comme dans le cas classique, nos deux amis partageront une séquence de bits aléatoire, mais aucune intrication quantique.

3.2.1 Modèle du processus de chiffrement

Récapitulons brièvement les notions principales énoncées dans [AMTdW00]. Dans ce modèle, Alice et Bob partagent publiquement une liste $\{U_k\}_{k=1}^n$ de matrices unitaires et un état ρ_a. La clé K est une variable aléatoire qui prend la valeur $k \in \{1, \ldots, n\}$ avec probabilité p_k. Supposons qu'Alice et Bob en partagent la valeur $K = k$. Pour chiffrer l'état ρ, Alice lui adjoindra l'ancille ρ_a et appliquera la transformation $\mathrm{Ad}U_k$

[3] Voir [Zém00], page 42.

[4] Si chaque message en clair m a, a priori, une probabilité $p(m)$ d'être chiffré par Alice, on peut voir le message comme une variable aléatoire M telle que $P(M = m) = p(m)$.

[5] Tel que décrit dans [BBC$^+$93].

à l'état $\rho \otimes \rho_a$. Bob, pour recouvrer ρ, n'aura qu'à appliquer $\mathrm{Ad}U_k^\dagger$ au système reçu d'Alice et à jeter le système B contenant l'ancille.

Du point de vue de Bob, donc, celui-ci reçoit $\mathcal{E}_k(\rho) = U_k(\rho^A \otimes \rho_a^B)U_k^\dagger$ et applique la fonction de déchiffrement $\mathcal{D}_k(\tau) = \mathrm{tr}_B\left(U_k^\dagger \tau U_k\right)$. Pour Ève, qui ne connaît pas la valeur i de la clé, la matrice de densité décrivant l'état envoyé par Alice sera $\mathcal{E}(\rho) = \sum_{k=1}^n p_k U_k(\rho \otimes \rho_a)U_k^\dagger$.

Comme dans le cas classique, un chiffre parfait devra être tel que le cryptogramme $\mathcal{E}(\rho)$ ne permette aucunement de deviner lequel des messages possibles a été chiffré. C'est ce qui motive la définition ci-dessous.

Soit $\Omega \subseteq \mathcal{D}(d)$ un ensemble d'états, $U = (U_1, \ldots, U_n)$ un vecteur de matrices unitaires, $p = (p_1, \ldots, p_n)$ un vecteur de probabilité et ρ_a et ϕ, deux matrices de densité. On dira du quintuplet $[\Omega, U, p, \rho_a, \phi]$ qu'il est un **canal quantique privé**, ou CQP, si pour tout $\rho \in \Omega$, $\sum_{k=1}^n p_k U_k(\rho \otimes \rho_a)U_k^\dagger = \phi$.

Parfois, on décrira aussi un CQP à l'aide du triplet $[\Omega, \mathcal{E}, \phi]$, où $\mathcal{E}$ est le super-opérateur décrit par $\mathcal{E}(\rho) = \sum_{k=1}^n p_k U_k(\rho \otimes \rho_a)U_k^\dagger$. Ce triplet devra être tel que $\mathcal{E}(\rho) = \phi$ pour tout $\rho \in \Omega$.

On dira que n bits de clé sont **nécessaires pour chiffrer** Ω si $H(p) \geq n$ chaque fois que $[\Omega, U, p, \rho_a, \phi]$ est un CQP.

On dira aussi que n bits de clé sont **suffisants pour chiffrer** Ω s'il existe un CQP $[\Omega, U, p, \rho_a, \phi]$ tel que $H(p) \leq n$.

Le résultat principal de [AMTdW00] peut être formulé comme suit :

Théorème 3.2.1. *$2n$ bits de clé sont nécessaires pour chiffrer $\mathcal{D}(2^n)$.*

La démonstration de ce théorème ne sera pas présentée ici. Le lecteur devra attendre le théorème 4.3.3, qui implique le théorème 3.2.1.

Théorème 3.2.2. *$2n$ bits de clé sont suffisants pour chiffrer $\mathcal{D}(2^n)$.*

Plutôt que de présenter une preuve formelle de ce théorème, nous nous contenterons de voir comment chiffrer un qubit avec deux bits de clé. Pour chiffrer n qubits, il suffit en fait de chiffrer chaque qubit individuellement.

Soit $\mathcal{E} = \{\frac{1}{2}\sigma_k\}_{k=0}^{3}$ le super-opérateur qui applique, à probabilités égales, les quatre matrices de Pauli. Rappelons que, quand $i > 0$, σ_i anti-commute avec les deux autres matrices de Pauli qui ne sont pas l'identité[6] :

$$\sigma_i\sigma_k = -\sigma_k\sigma_i \text{ lorsque } i > 0 \text{ et } k \in \{1,2,3\} \setminus \{i\}.$$

Pour tout $\rho \in \mathcal{D}(2)$, nous pouvons écrire $\rho = \frac{1}{2}I + \frac{1}{2}\vec{r}\cdot\vec{\sigma}$ et calculer

$$\begin{aligned}
\mathcal{E}(\rho) &= \mathcal{E}\left(\frac{1}{2}I + \frac{1}{2}\vec{r}\cdot\vec{\sigma}\right) \\
&= \frac{1}{4}\sum_{k=0}^{3}\sigma_k\left(\frac{1}{2}I + \frac{1}{2}\vec{r}\cdot\vec{\sigma}\right)\sigma_k \\
&= \frac{1}{2}I + \frac{1}{8}\sum_{k=0}^{3}\sum_{i=1}^{3} r_i\sigma_k\sigma_i\sigma_k \\
&= \frac{1}{2}I + \frac{1}{8}\sum_{i=1}^{3} r_i\left(\sigma_0\sigma_i\sigma_0 + \sigma_i\sigma_i\sigma_i + \sum_{k\in\{1,2,3\}\setminus\{i\}}\sigma_k\sigma_i\sigma_k\right) \\
&= \frac{1}{2}I + \frac{1}{8}\sum_{i=1}^{3} r_i(2\sigma_i - 2\sigma_i) \\
&= \frac{1}{2}I.
\end{aligned}$$

[6] Revoir, à ce sujet, la section 1.3.4.

3.3 Chiffrement d'états coplanaires

Maintenant que nous savons à quoi nous en tenir en ce qui a trait aux ensembles $\Omega = \mathcal{D}(2^n)$ de n qubits, à savoir que $2n$ bits de clé sont nécessaires et suffisants pour les chiffrer, nous tenterons, pour le reste du chapitre, de calculer la quantité de clé requise pour chiffrer des ensembles $\Omega \subsetneq \mathcal{D}(d)$ qui contiennent seulement une partie de tous les messages possibles. Dans le cas où $d = 2$, nous obtiendrons une solution complète. Quand $d > 2$, la situation se complique. Quelques résultats seront énoncés, mais beaucoup de travail reste à faire, ce qui constitue une piste de recherche intéressante.

Penchons-nous d'abord sur un fait en apparence anodin dont il est question dans [AMTdW00] :

Lemme 3.3.1. *Si* $\Omega = \{\cos(\theta)\,|0\rangle + \sin(\theta)\,|1\rangle \,|\, 0 \leq \theta < 2\pi\}$ *et* $\mathcal{E} = \{\frac{1}{\sqrt{2}}I, \frac{1}{\sqrt{2}}Y\}$, *alors* $[\Omega, \mathcal{E}, \frac{1}{2}I]$ *est un CQP .*

Démonstration. Soit $|\psi\rangle \in \Omega$. Écrivons $|\psi\rangle = \cos(\theta)\,|0\rangle + \sin(\theta)\,|1\rangle$. Alors

$$\begin{aligned}
\mathcal{E}(\psi) &= \frac{1}{2}\psi + \frac{1}{2}Y\psi Y \\
&= \frac{1}{2}\psi + \frac{1}{2}XZ\psi ZX \\
&= \frac{1}{2}\begin{bmatrix} \cos^2(\theta) & \cos(\theta)\sin(\theta) \\ \cos(\theta)\sin(\theta) & \sin^2(\theta) \end{bmatrix} + \frac{1}{2}XZ\begin{bmatrix} \cos^2(\theta) & \cos(\theta)\sin(\theta) \\ \cos(\theta)\sin(\theta) & \sin^2(\theta) \end{bmatrix}ZX \\
&= \frac{1}{2}\begin{bmatrix} \cos^2(\theta) & \cos(\theta)\sin(\theta) \\ \cos(\theta)\sin(\theta) & \sin^2(\theta) \end{bmatrix} + \frac{1}{2}X\begin{bmatrix} \cos^2(\theta) & -\cos(\theta)\sin(\theta) \\ -\cos(\theta)\sin(\theta) & \sin^2(\theta) \end{bmatrix}X \\
&= \frac{1}{2}\begin{bmatrix} \cos^2(\theta) & \cos(\theta)\sin(\theta) \\ \cos(\theta)\sin(\theta) & \sin^2(\theta) \end{bmatrix} + \frac{1}{2}\begin{bmatrix} \sin^2(\theta) & -\cos(\theta)\sin(\theta) \\ -\cos(\theta)\sin(\theta) & \cos^2(\theta) \end{bmatrix}X \\
&= \frac{1}{2}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.
\end{aligned}$$

□

Maintenant, revoyons l'énoncé du lemme 3.3.1 sous un aspect géométrique. L'ensemble $\Omega = \{\cos(\theta)\,|0\rangle + \sin(\theta)\,|1\rangle \,|\, 0 \leq \theta < 2\pi\}$ correspond à un méridien sur la sphère de Poincaré. Ce méridien, c'est-à-dire $m(\Omega)$, passe par les deux pôles, $m(|0\rangle) = \vec{k}$ et $m(|1\rangle) = -\vec{k}$, et passe à l'équateur par $m(\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)) = \vec{\imath}$ et $m(\frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)) = -\vec{\imath}$.

Le lemme 3.3.1 nous apprend qu'il est possible de chiffrer cet ensemble avec la transformation $\mathcal{E} = \{\frac{1}{\sqrt{2}}I, \frac{1}{\sqrt{2}}Y\}$, ce qui n'est pas surprenant, puisque la transformation de $\mathbb{R}^3$ correspondant à Y, soit $\varphi(Y)$, est une rotation d'un demi-tour autour de l'axe des y. Cela revient à dire que le méridien défini par Ω tourne sur lui-même d'un demi-tour. On ne s'étonnera donc pas que pour tout $|\psi\rangle \in \Omega$, $m(\mathcal{E}(|\psi\rangle\langle\psi|))$ soit le point milieu entre $m(|\psi\rangle)$ et $m(Y\,|\psi\rangle)$, c'est-à-dire le centre de la sphère, qui correspond à l'état complètement mélangé $\frac{1}{2}I$.

Ce que ce raisonnement met en lumière, c'est qu'il doit être possible de chiffrer, à l'aide d'un seul bit de clé, n'importe quel ensemble Ω tel que $m(\Omega)$ est contenu dans un cercle $C \subset S_2$. Il s'agira d'utiliser un super-opérateur qui, à probabilités égales, applique ou bien l'identité, ou bien une rotation d'un demi-tour de C sur lui-même. Si on considère aussi les états mélangés, un tel super-opérateur permet en fait de chiffrer tout ensemble Ω tel que $m(\Omega)$ est contenu dans le plan P dont l'intersection avec la sphère S_2 donne le cercle C — c'est-à-dire le plan P tel que $P \cap S_2 = C$.

Proposition 3.3.2. *Soit $l \in [0,1]$ et $\vec{r} \in \mathbb{R}^3$ tel que $\|\vec{r}\| = 1$. Définissons le plan $P = \{l\vec{r}+\vec{t}\,|\,\vec{r}\cdot\vec{t} = 0\}$. Si $\Omega \subseteq \mathcal{D}(2)$ est tel que $m(\Omega) \subseteq P$ et $\mathcal{E}$ est un super-opérateur défini par $\mathcal{E} = \{\frac{1}{\sqrt{2}}I, \frac{1}{\sqrt{2}}\vec{r}\cdot\vec{\sigma}\}$, alors $[\Omega, \mathcal{E}, m^{-1}(l\vec{r})]$ est un* CQP.

Démonstration. Posons d'abord $U = \omega(\vec{r}) = \vec{r}\cdot\vec{\sigma}$ (définition 2.7.2). Par la proposition 2.7.4, $\varphi(U)$ est une rotation d'un demi-tour autour de $\vec{r}$. Il suit immédiatement que $\varphi(U)(\vec{r}) = \vec{r}$ et que pour tout $\vec{t}$ orthogonal à $\vec{r}$, $\varphi(U)(\vec{t}) = -\vec{t}$. Alors pour tout

$l\vec{r}+\vec{t} \in P$, on obtient

$$\begin{aligned} \varphi(U)(l\vec{r}+\vec{t}) &= l\varphi(U)(\vec{r})+\varphi(U)(\vec{t}) \\ &= l\vec{r}-\vec{t}. \end{aligned}$$

Maintenant, soit $\rho \in \Omega$. Comme $m(\Omega) \subseteq P$, on peut écrire $m(\rho) = l\vec{r}+\vec{t}$, où $\vec{r}\cdot\vec{t} = 0$. Conséquemment,

$$\begin{aligned} m(\mathcal{E}(\rho)) &= m(\frac{1}{2}\rho + \frac{1}{2}U\rho U^{\dagger}) \\ &= \frac{1}{2}m(\rho) + \frac{1}{2}m(\mathrm{Ad}U(\rho)) \\ &= \frac{1}{2}m(\rho) + \frac{1}{2}\varphi(U)(m(\rho)) \\ &= \frac{1}{2}(l\vec{r}+\vec{t}) + \frac{1}{2}\varphi(U)(l\vec{r}+\vec{t}) \\ &= \frac{1}{2}(l\vec{r}+\vec{t}) + \frac{1}{2}(l\vec{r}-\vec{t}) \\ &= l\vec{r}. \end{aligned}$$

Nous avons montré que $\mathcal{E}(\rho) = m^{-1}(l\vec{r})$ pour tout $\rho \in \Omega$. Cela permet de conclure que $[\Omega, \mathcal{E}, m^{-1}(l\vec{r})]$ est un CQP.

□

3.4 Moins d'un bit de clé est inutile

Maintenant que nous avons vu deux possibilités distinctes quand au nombre de bits de clé requis pour chiffrer un ensemble d'états $\Omega \subseteq \mathcal{D}(2)$, soit un bit quand $m(\Omega)$ est contenu dans un plan et deux bits pour l'ensemble $\Omega = \mathcal{D}(2)$ de tous les états, il est légitime de se demander si ce sont là les seules possibilités, ou s'il existe des ensembles requérant un nombre fractionnaire de bits de clé. Nous verrons, en deux étapes, qu'il n'existe pas de tels ensembles intermédiaires.

Premièrement, il est impossible de chiffrer un ensemble non-trivial Ω — un ensemble contenant plus d'un état — avec moins d'un bit de clé et ce, peu importe la dimension de l'espace :

Théorème 3.4.1.[7] *Soit $\rho_1, \rho_2 \in \mathcal{D}(d)$ deux matrices de densité distinctes et $\mathcal{E}$ un super-opérateur donné par $\mathcal{E}(\rho) = \sum_{i=1}^{n} p_i U_i[\rho \otimes \rho_a]U_i^\dagger$ et tel que $\mathcal{E}(\rho_1) = \mathcal{E}(\rho_2)$. Alors un bit de clé est nécessaire pour chiffrer l'ensemble $\Omega = \{\rho_1, \rho_2\}$.*

Lemme 3.4.2. *Si U est une matrice unitaire et X une matrice quelconque, alors $\|\mathrm{Vec}(UXU^\dagger)\| = \|\mathrm{Vec}(X)\|$.*

Démonstration. Grâce à l'identité $\mathrm{Vec}(ABC) = (C^{\mathrm{T}} \otimes A)\mathrm{Vec}(B)$, on trouve que $\|\mathrm{Vec}(UXU^\dagger)\| = \|(U^* \otimes U)\mathrm{Vec}(X)\| = \|\mathrm{Vec}(X)\|$, puisque $U^* \otimes U$ est un opérateur unitaire et, donc, préserve la norme.

□

Démonstration du théorème. Soit ρ_1, ρ_2 et $\mathcal{E}$ tels que dans l'énoncé du théorème. Par hypothèse, $\mathcal{E}(\rho_1) = \mathcal{E}(\rho_2)$. Puisque $\mathcal{E}$ est une transformation linéaire, il faut que $\mathcal{E}(\rho_1 - \rho_2) = 0$. Posons $A = \rho_1 - \rho_2$. On sait donc que

$$\mathcal{E}(A) = \sum_{i=1}^{n} p_i U_i[A \otimes \rho_a]U_i^\dagger = \mathbf{0}. \tag{3.1}$$

En, considérant les termes de cette somme comme des vecteurs dans $\mathbb{C}^{d^2 d_a^2}$ — où d_a est la dimension du système ancillaire : $\rho_a \in \mathcal{D}(d_a)$ —, le lemme 3.4.2 nous apprend que chaque terme a une norme égale à

$$\begin{aligned} \|\mathrm{Vec}(p_i U_i[A \otimes \rho_a]U_i^\dagger)\| &= p_i \|\mathrm{Vec}(A \otimes \rho_a)\| \\ &= p_i c, \end{aligned}$$

[7] Merci à Simon-Pierre Desrosiers d'avoir collaboré à la preuve de ce théorème.

où $c = \|\mathrm{Vec}(A \otimes \rho_a)\|$. Notons que $c \neq 0$ puisque ni A ni ρ_a ne sont nulles. Quand une somme de vecteurs est nulle, l'un de ces vecteurs ne saurait être plus long que la somme de la longueur des autres, ce qui force $p_i \leq 1/2$ pour chaque i. Plus formellement, fixons i et montrons que $p_i \leq 1/2$. Tout d'abord, par l'équation (3.1),

$$p_i U_i [A \otimes \rho_a] U_i^\dagger = -\sum_{j \neq i} p_j U_j [A \otimes \rho_a] U_j^\dagger.$$

Donc,

$$\begin{aligned}
\|\mathrm{Vec}(p_i U_i [A \otimes \rho_a] U_i^\dagger)\| &= \|\mathrm{Vec}(-\sum_{j \neq i} p_j U_j [A \otimes \rho_a] U_j^\dagger)\| \\
&= \|\sum_{j \neq i} \mathrm{Vec}(p_j U_j [A \otimes \rho_a] U_j^\dagger)\| \\
&\leq \sum_{j \neq i} \|\mathrm{Vec}(p_j U_j [A \otimes \rho_a] U_j^\dagger)\|, \text{ par l'inégalité triangulaire.}
\end{aligned}$$

En d'autres mots,

$$\begin{aligned}
p_i c &\leq \sum_{j \neq i} p_j c \\
&= (1 - p_i) c \\
\Rightarrow p_i &\leq (1 - p_i) \\
\Rightarrow p_i &\leq \frac{1}{2}.
\end{aligned}$$

Comme le choix de i était arbitraire, $p_i \leq \frac{1}{2}$ pour tous les i.

Pour toute distribution de probabilité $\{p_i\}$ telle que $p_i \leq \frac{1}{2}$ pour tout i,

$$\begin{aligned}
H(\{p_i\}) &= -\sum_i p_i \log p_i \\
&\geq -\sum_i p_i \log \frac{1}{2} \\
&= \sum_i p_i \\
&= 1,
\end{aligned}$$

ce qui montre qu'il faut au moins un bit de clé pour chiffrer Ω.

□

3.5 S'il faut plus d'un bit, il en faut deux

Nous allons maintenant conclure notre enquête pour les ensembles $\Omega \subseteq \mathcal{D}(2)$ sur un qubit. Nous verrons que les seuls ensembles pouvant être chiffrés à l'aide de moins de deux bits de clé s'inscrivent dans un plan passant à travers la sphère de Poincaré. De tels ensembles, nous l'avons vu dans la proposition 3.3.2, peuvent être chiffrés par des super-opérateurs de la forme $\mathcal{E} = \{\frac{1}{\sqrt{2}}I, \frac{1}{\sqrt{2}}\vec{r}\cdot\vec{\sigma}\}$ — c'est-à-dire avec un bit de clé.

Nous dirons que les matrices de densité $\rho_1, \rho_2, \ldots, \rho_n$ sont **coplanaires** si les vecteurs réels associés, $m(\rho_1), m(\rho_2), \ldots, m(\rho_n)$, définissent des points de $\mathbb{R}^3$ qui font tous partie d'un même plan. Plus formellement, les ρ_i sont coplanaires si trois d'entre eux, disons ρ_1, ρ_2 et ρ_3, sont tels que pour tous les i entre 1 et n, on puisse représenter $m(\rho_i)$ de la façon suivante :

$$m(\rho_i) = \sum_{j=1}^{3} a_j^{(i)} m(\rho_j),$$

où $\sum_{j=1}^{3} a_j^{(i)} = 1$ pour chaque i.

Théorème 3.5.1.[8] *Soit $\Omega = \{\rho_i\}_{i=1}^4 \subseteq \mathcal{D}(2)$ un ensemble de quatre matrices de densité distinctes, soit ϕ un état donné et $\mathcal{E}(\rho) = \sum_{i=k}^{n} p_k U_k[\varphi \otimes \rho_0]U_k^\dagger$ un super-opérateur tel que $\mathcal{E}(\rho_i) = \phi$ quand $1 \leq i \leq 4$. Si les quatre matrices ρ_i ne sont pas coplanaires, alors pour toute matrice de densité $\rho \in \mathcal{D}(2)$, $\mathcal{E}(\rho) = \phi$.*

[8] Merci à Valérie Poulain d'avoir collaboré à la preuve de ce théorème.

Il apparaîtra que le théorème 3.5.1 est une conséquence directe des lemmes 3.5.3 et 3.5.4. Avant d'y arriver, nous aurons besoin du résultat suivant :

Lemme 3.5.2. *L'ensemble de matrices de densité* $R = \{\rho_1, \ldots, \rho_n\}$ *est linéairement indépendant sur* $\mathbb{R}$ *si et seulement si l'ensemble* $Q = \{\rho_1 - \rho_n, \ldots, \rho_{n-1} - \rho_n\}$ *est linéairement indépendant sur* $\mathbb{R}$.

Démonstration. Supposons que l'ensemble Q soit linéairement dépendant. On peut trouver des nombres $a_i \in \mathbb{R}$ qui ne sont pas tous nuls et tels que $\sum_{i=1}^{n-1} a_i(\rho_i - \rho_n) = 0$. Mais alors, on peut aussi écrire

$$\sum_{i=1}^{n-1} a_i\rho_i - \left(\sum_{i=1}^{n-1} a_i\right)\rho_n = 0,$$

ce qui démontre que l'ensemble R est linéairement dépendant.

Supposons maintenant que l'ensemble R soit linéairement dépendant. On peut trouver des a_i réels qui ne sont pas tous nuls et tels que $\sum_{i=1}^{n} a_i\rho_i = 0$. Puisque les ρ_i sont des matrices de densité, leur trace vaut un, donc

$$\begin{aligned}\sum_{i=1}^{n} a_i &= \sum_{i=1}^{n} a_i \mathrm{tr}\,(\rho_i) \\ &= \mathrm{tr}\left(\sum_{i=1}^{n} a_i\rho_i\right) \\ &= 0.\end{aligned}$$

On peut voir à présent que

$$\begin{aligned}\sum_{i=1}^{n-1} a_i(\rho_i - \rho_n) &= \sum_{i=1}^{n-1} a_i\rho_i - \sum_{i=1}^{n-1} a_i\rho_n \\ &= \sum_{i=1}^{n-1} a_i\rho_i + a_n\rho_n, \text{ puisque } a_n = -\sum_{i=1}^{n-1} a_i \\ &= \sum_{i=1}^{n} a_i\rho_i \\ &= 0,\end{aligned}$$

ce qui montre que l'ensemble Q est linéairement dépendant.

□

Lemme 3.5.3. *Si quatre matrices de densité $\rho_i \in \mathcal{D}(2), i = 1, \ldots, 4$, ne sont pas coplanaires, alors ces matrices sont linéairement indépendantes sur $\mathbb{R}$.*

Démonstration. Prouvons l'énoncé contraposé. Supposons que les quatre matrices ρ_i soient linéairement dépendantes sur $\mathbb{R}$. Alors on peut exprimer l'une d'entre elles, disons ρ_4, comme une combinaison linéaire réelle des trois autres :

$$\rho_4 = \sum_{i=1}^{3} a_i \rho_i,$$

où $a_i \in \mathbb{R}$ et $\sum_{i=1}^{3} a_i = 1$. Cette dernière condition découle de la linéarité de la trace :

$$\begin{aligned} 1 &= \mathrm{tr}\,(\rho_4) \\ &= \mathrm{tr}\left(\sum_{i=1}^{3} a_i \rho_i\right) \\ &= \sum_{i=1}^{3} a_i \mathrm{tr}\,(\rho_i) \\ &= \sum_{i=1}^{3} a_i. \end{aligned}$$

Mais alors,

$$\begin{aligned} m(\rho_4) &= m(\sum_{i=1}^{3} a_i \rho_i) \\ &= \sum_{i=1}^{3} a_i m(\rho_i), \end{aligned} \tag{3.2}$$

puisque, par la proposition 2.4.3, m est une transformation affine. L'équation 3.2 montre que $m(\rho_4)$ réside dans un plan défini par les points $m(\rho_1)$, $m(\rho_2)$ et $m(\rho_3)$.

□

Lemme 3.5.4. *Soit $\Omega \subseteq \mathcal{D}(d)$ un ensemble de d^2 matrices de densité qui sont linéairement indépendantes sur $\mathbb{R}$, soit $\phi \in \mathcal{D}(d)$ une matrice de densité quelconque et $\mathcal{E}$ un super-opérateur tel que $\mathcal{E}(\rho_i) = \phi$ pour tout $\rho_i \in \Omega$. Alors pour tout $\rho \in \mathcal{D}(d)$, $\mathcal{E}(\rho) = \phi$.*

Démonstration. Soit $n = d^2$. Supposons que $\Omega = \{\rho_1, \ldots, \rho_n\}$ est un ensemble de matrices de densité qui sont linéairement indépendantes sur $\mathbb{R}$. Alors l'ensemble $Q = \{\rho_1 - \rho_n, \ldots, \rho_{n-1} - \rho_n\}$ est aussi un ensemble linéairement indépendant sur $\mathbb{R}$, par le lemme 3.5.2. Il forme donc une base pour l'espace des matrices hermitiennes de trace zéro, puisque cet espace a une dimension de $n-1$. Il s'ensuit que pour toute matrice de densité ρ de dimension d, puisque $\rho - \rho_n$ est une matrice hermitienne de trace zéro, on peut écrire $\rho - \rho_n = \sum_{i=1}^{n-1} a_i(\rho_i - \rho_n)$, et donc

$$\begin{aligned} \rho &= \rho_n + \sum_{i=1}^{n-1} a_i(\rho_i - \rho_n) \\ &= \sum_{i=1}^{n-1} a_i\rho_i + (1 - \sum_{i=1}^{n-1} a_i)\rho_n \\ &= \sum_{i=1}^{n} a_i\rho_i, \end{aligned}$$

si on pose $a_n = 1 - \sum_{i=1}^{n-1} a_i$. Remarquons que $\sum_{i=1}^{n} a_i = 1$. Il s'ensuit qu'en imposant $\mathcal{E}(\rho_i) = \phi$ pour tout i, on force

$$\begin{aligned} \mathcal{E}(\rho) &= \mathcal{E}(\sum_{i=1}^{n} a_i\rho_i) \\ &= \sum_{i=1}^{n} a_i\mathcal{E}(\rho_i) \text{ par linéarité de } \mathcal{E} \\ &= \sum_{i=1}^{n} a_i\phi \\ &= \phi. \end{aligned}$$

□

Théorème 3.5.5. *Soit $\Omega \subseteq \mathcal{D}(2)$. Si les états contenus dans Ω sont coplanaires, alors un bit de clé est nécessaire et suffisant pour chiffrer Ω. Sinon, deux bits de clé sont nécessaires et suffisants pour chiffrer Ω.*

Démonstration. Tout d'abord, le théorème 3.4.1 montre qu'un bit est nécessaire pour chiffrer tout ensemble Ω. Ensuite, la proposition 3.3.2 nous apprend qu'un ensemble Ω d'états coplanaires — tel que $m(\Omega)$ est contenu dans le plan $P_{l\vec{r}} = \{l\vec{r}+\vec{t} \,|\, \vec{r}\cdot\vec{t} = 0\}$ — peut être chiffré à l'aide du super-opérateur $\mathcal{E} = \{\frac{1}{\sqrt{2}}I, \frac{1}{\sqrt{2}}\vec{r}\cdot\vec{\sigma}\}$, donc avec un seul bit de clé.

Le théorème 3.2.2 montre que deux bits de clé suffisent pour chiffrer tout ensemble $\Omega \subseteq \mathcal{D}(2)$. Le théorème 3.5.1, quand à lui, permet d'affirmer que lorsque $\Omega \subseteq \mathcal{D}(2)$ contient des états non coplanaires et que $[\Omega, \mathcal{E}, \phi]$ est un CQP, alors, nécessairement, $[\mathcal{D}(2), \mathcal{E}, \phi]$ est un CQP. Le théorème 3.2.1 permet de conclure que deux bits de clé sont nécessaires pour chiffrer de tels ensembles Ω.

□

3.6 États perpendiculaires

Maintenant que nous pouvons facilement déterminer s'il faut un ou deux bits de clé pour chiffrer un ensemble $\Omega \subseteq \mathcal{D}(2)$ — il suffit de savoir si les états qui le composent sont coplanaires —, nous tenterons de résoudre le même problème pour les dimensions supérieures à deux.

3.6.1 Chiffrement d'information classique

Considérons l'ensemble $\Omega = \{|0\rangle, |1\rangle, \ldots, |N-1\rangle\} \subseteq \mathcal{D}(d)$, où $N \leq d$. Puisque les états contenus dans Ω peuvent être distingués parfaitement, transmettre des états tirés de Ω équivaut à transmettre de l'information classique. On s'attend donc à ce que les résultats de la théorie de l'information classique s'appliquent toujours, ce qui est le cas : $\log N$ bits de clé sont nécessaires et suffisants pour chiffrer Ω.

Proposition 3.6.1. *Il suffit d'utiliser* $\log N$ *bits de clé pour chiffrer l'ensemble* $\Omega = \{|0\rangle, |1\rangle, \ldots, |N-1\rangle\} \subseteq \mathcal{D}(d)$.

Démonstration. Soit X_N la transformation unitaire donnée par $X_N |i\rangle = |i+1 \,(\mathrm{mod} N)\rangle$ quand $0 \leq i \leq N-1$, et $X_N |i\rangle = |i\rangle$ quand $i \geq N$. Définissons le super-opérateur $\mathcal{E}_{X_N} = \{\frac{1}{\sqrt{N}} X_N^k\}_{k=0}^{N-1}$. Alors, pour tout i entre 0 et $N-1$,

$$\begin{aligned}
\mathcal{E}_{X_N}(|i\rangle\langle i|) &= \frac{1}{N}\sum_{k=0}^{N-1} X_N^k |i\rangle\langle i| (X_N^k)^\dagger \\
&= \frac{1}{N}\sum_{k=0}^{N-1} |k\rangle\langle k| \\
&= \frac{1}{N}\mathrm{diag}(\underbrace{1,\ldots,1}_{N \text{ fois}},0,\ldots,0).
\end{aligned}$$

Puisque la quantité de clé qui est nécessaire pour appliquer le chiffre $\mathcal{E}_{X_N}$ est de $H(\underbrace{\frac{1}{N},\ldots,\frac{1}{N}}_{N fois}) = \log N$ bits, le résultat suit.

□

Proposition 3.6.2. *Il est nécessaire d'utiliser* $\log N$ *bits de clé pour chiffrer l'ensemble* $\Omega = \{|0\rangle, |1\rangle, \ldots, |N-1\rangle\} \subseteq \mathcal{D}(d)$.

Le développement du présent chapitre ne permet pas de présenter ici une preuve de la proposition 3.6.2. Le lecteur devra patienter jusqu'au chapitre suivant, où le théorème 4.3.3 démontre que, même en utilisant un chiffre quantique, $2n$ bits de clé sont nécessaires pour chiffrer des messages classiques de $2n$ bits. Cette preuve s'adapte aisément pour démontrer la proposition 3.6.2.

3.6.2 Hypercercles

Pour la discussion qui suit, nous nous en tiendrons uniquement aux états purs. En dimension deux, les ensembles Ω pouvant être chiffrés à l'aide d'un bit de clé forment des cercles sur la sphère de Poincaré, car nous avons vu que les états contenus dans de tels ensembles Ω doivent être coplanaires, et l'intersection d'un plan avec S_2 donne un cercle $C \subseteq S_2$.

En choisissant une base appropriée, un cercle C aura une symétrie de rotation autour de l'axe des z. On peut utiliser la forme paramétrique de la section 2.4.1 pour écrire

$$C = \left\{ cos(\frac{\theta}{2}) \left|0\right\rangle + e^{i\varphi} sin(\frac{\theta}{2}) \left|1\right\rangle \;\middle|\; 0 \leq \varphi < 2\pi \right\}.^{9}$$

Un tel ensemble pourra être chiffré avec un bit de clé, à l'aide du super-opérateur $\mathcal{E} = \{\frac{1}{\sqrt{2}} I, \frac{1}{\sqrt{2}} Z\}$.

Un phénomène semblable se produit en dimensions supérieures à deux. Soit $\{r_j\}_{j=0}^{N-1}$ une collection de nombres réels tels que $\sum_{j=0}^{N-1} r_j^2 = 1$. Considérons l'ensemble d'états

$$\Omega_{\vec{r}} = \left\{ r_0 \left|0\right\rangle + r_1 e^{i\varphi_1} \left|1\right\rangle + \ldots + e^{i\varphi_{N-1}} r_{N-1} \left|N-1\right\rangle \;\middle|\; 0 \leq \varphi_j < 2\pi \right\}. \qquad (3.3)$$

Nous montrerons bientôt qu'il est possible de chiffrer $\Omega_{\vec{r}}$ avec $\log N$ bits de clé, mais il faut d'abord prouver un lemme dont nous aurons besoin lors de cette démonstration.

[9] Cet ensemble est un parallèle sur la sphère de Poincaré. Sa latitude est donnée par l'angle θ, qui prend des valeurs entre 0 et π.

L'idée est simple : si on place N points sur le cercle de rayon un dans le plan complexe de façon à ce qu'il y ait un angle de $2\pi m/N$ radians entre chaque point et le suivant ou, en d'autres mots, si l'étoile qu'ils forment (en traçant une ligne du centre vers chacun des points) possède une symétrie de rotation de $2\pi m/N$ radians, alors la somme de ces N points est zéro. Plus formellement :

Lemme 3.6.3 (Lemme de l'étoile)**.** *Pour tout $m \in \mathbb{Z}$ tel que N ne divise pas m et pour tout $\theta \in \mathbb{R}$, l'égalité suivante est vraie :*

$$\sum_{k=0}^{N-1} e^{i\left(\theta+\frac{2\pi km}{N}\right)} = 0.$$

Démonstration. Posons

$$S = \sum_{k=0}^{N-1} e^{i\left(\theta+\frac{2\pi km}{N}\right)}.$$

Assurons-nous d'abord que la somme S n'est pas affectée par une rotation de $2\pi m/N$ radians :

$$\begin{aligned}
e^{i\frac{2\pi m}{N}} S &= e^{\frac{2\pi mi}{N}} \sum_{k=0}^{N-1} e^{i\left(\theta+\frac{2\pi km}{N}\right)} \\
&= \sum_{k=0}^{N-1} e^{i\left(\theta+\frac{2\pi(k+1)m}{N}\right)} \\
&= \sum_{k=1}^{N} e^{i\left(\theta+\frac{2\pi km}{N}\right)} \\
&= \sum_{k=0}^{N-1} e^{i\left(\theta+\frac{2\pi km}{N}\right)} \text{ , puisque } e^{i\frac{2\pi Nm}{N}} = e^{i\frac{2\pi 0m}{N}} \\
&= S.
\end{aligned}$$

Il s'ensuit que soit $S = 0$, soit on peut diviser chaque terme par S pour obtenir $e^{\frac{2\pi mi}{N}} = 1$. Le deuxième cas ne peut être vrai que si m/N est un nombre entier, ce qui est exclu par l'hypothèse que N ne divise pas m. Il faut donc que $S = 0$.

□

Théorème 3.6.4. *Pour chiffrer l'ensemble $\Omega_{\vec{r}}$ défini à l'équation 3.3,* $\log N$ *bits de clé suffisent.*

Démonstration. Soit Z_N la transformation unitaire donnée par $Z_N |j\rangle \stackrel{\text{def}}{=} e^{i\frac{2\pi j}{N}} |j\rangle$ quand $0 \leq j \leq N-1$, et $Z_N |j\rangle = |j\rangle$ quand $j \geq N$. Considérons aussi l'état $|\psi\rangle = \sum_{j=0}^{N-1} r_j e^{i\varphi_j} |j\rangle \in \Omega_{\vec{r}}$. Pour voir quel effet sur ψ auront k applications successives de la transformation Z_N, étudions la progression d'équations suivante :

$$\begin{aligned}
Z_N |j\rangle &= e^{i\frac{2\pi j}{N}} |j\rangle \\
Z_N^k |j\rangle &= e^{i\frac{2\pi jk}{N}} |j\rangle \\
Z_N^k |\psi\rangle &= \sum_{j=0}^{N-1} e^{i\left(\varphi_j + \frac{2\pi jk}{N}\right)} r_j |j\rangle \\
Z_N^k |\psi\rangle \langle\psi| (Z_N^k)^\dagger &= \sum_{j=0}^{N-1}\sum_{l=0}^{N-1} e^{i\left(\varphi_j - \varphi_l + \frac{2\pi k(j-l)}{N}\right)} r_j r_l |j\rangle \langle l| .
\end{aligned}$$

Définissons à présent le super-opérateur

$$\mathcal{E}_{Z_N} \stackrel{\text{def}}{=} \left\{ \frac{1}{\sqrt{N}} I, \frac{1}{\sqrt{N}} Z_N, \ldots, \frac{1}{\sqrt{N}} Z_N^{N-1} \right\}.$$

Dans le calcul de $\mathcal{E}_{Z_N}(\psi)$, nous briserons la triple somme en deux parties, soit une pour le cas $j = l$ et l'autre pour le cas $j \neq l$. Dans le premier cas, les coefficients complexes deviendront tous $e^0 = 1$, tandis que dans l'autre, la somme interne (sur les k) sera toujours nulle, par le lemme de l'étoile (3.6.3).

$$\begin{aligned}
\mathcal{E}_{Z_N}(\psi) &= \sum_{k=0}^{N-1} \frac{1}{N} Z_N^k |\psi\rangle \langle\psi| (Z_N^k)^\dagger \\
&= \frac{1}{N} \sum_{k=0}^{N-1}\sum_{j=0}^{N-1}\sum_{l=0}^{N-1} e^{i\left(\varphi_j - \varphi_l + \frac{2\pi k(j-l)}{N}\right)} r_j r_l |j\rangle \langle l| \\
&= \frac{1}{N} \sum_{j=0}^{N-1}\sum_{k=0}^{N-1} e^{i\left(0 + \frac{2\pi k \cdot 0}{N}\right)} r_j^2 |j\rangle \langle j| + \frac{1}{N} \sum_{j=0}^{N-1} \sum_{\substack{l=0 \\ l\neq j}}^{N-1} \left(\sum_{k=0}^{N-1} e^{i\left(\varphi_j - \varphi_l + \frac{2\pi k(j-l)}{N}\right)} \right) r_j r_l |j\rangle \langle l|
\end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{N}\sum_{j=0}^{N-1} N r_j^2 \left|j\right\rangle\left\langle j\right| + \frac{1}{N}\sum_{j=0}^{N-1}\sum_{\substack{l=0\\ l\neq j}}^{N-1} 0\cdot r_j r_l \left|j\right\rangle\left\langle l\right| \\
&= \sum_{j=0}^{N-1} r_j^2 \left|j\right\rangle\left\langle j\right|.
\end{aligned}
$$

Puisque les valeurs r_j sont les mêmes pour tout $|\psi\rangle \in \Omega_{\vec{r}}$, nous pouvons poser $\phi = \sum_{j=0}^{N-1} r_j^2 \left|j\right\rangle\left\langle j\right|$ et affirmer que $[\Omega_{\vec{r}}, \mathcal{E}_{Z_N}, \phi]$ est un CQP.

□

Portons une attention spéciale à la famille $\Omega = \{\frac{1}{\sqrt{N}}\sum_{k=0}^{N-1} Z_N^k \left|k\right\rangle\}_{j=0}^{N-1}$. Elle contient N états perpendiculaires, soit les N états de la base de Fourier.[10] Cela montre encore une fois que $\log N$ bits de clé suffisent pour chiffrer N états orthogonaux.

3.7 Morceaux de plusieurs qubits

Nous avons vu, au lemme 3.5.4, une condition suffisante pour déterminer qu'un chiffre pour Ω soit en fait un chiffre pour $\mathcal{D}(d)$ en entier : il suffit que Ω contienne d^2 états dont les matrices de densité sont linéairement indépendantes sur $\mathbb{R}$. Nous pouvons en déduire la proposition suivante :

Proposition 3.7.1. *Soit $\Omega \subseteq \mathcal{D}(2^n)$ un ensemble contenant 2^{2n} matrices de densité qui sont linéairement indépendantes. Alors, $2n$ bits de clé sont nécessaires et suffisants pour chiffrer Ω.*

Démonstration. Posons $d = 2^n$. Supposons que Ω contienne 2^{2n} matrices linéairement indépendantes et que $[\Omega, \mathcal{E}, \phi]$ soit un CQP. Alors $\mathcal{E}(\rho) = \phi$ pour tout $\rho \in \Omega$, ce qui

[10] Les états de la base de Fourier sont ceux de la base de calcul $\{|0\rangle, |1\rangle, \ldots, |N-1\rangle\}$ auxquels on applique la transformée de Fourier, c'est-à-dire $\mathcal{F}(|j\rangle) = \sum_{k=0}^{N-1} e^{i\frac{2\pi jk}{N}} |k\rangle$.

implique, par le lemme 3.5.4, que $\mathcal{E}(\rho) = \phi$ pour tout $\rho \in \mathcal{D}(2^n)$. Autrement dit, $[\mathcal{D}(d), \mathcal{E}, \phi]$ est un CQP.

Puisque tout chiffre pour Ω est aussi un chiffre pour $\mathcal{D}(2^n)$, le résultat du théorème 3.2.1 s'applique : $2n$ bits de clé sont nécessaires et suffisants pour chiffrer Ω.

□

Si Ω ne répond pas à cette condition, la quantité de clé requise pour chiffrer Ω demeure un mystère. Nous nous contenterons d'énoncer une condition à laquelle devra répondre tout opérateur $\mathcal{E}$ qui chiffre Ω.

Choisissons d'abord un état arbitraire $\mu \in \Omega$. Comme la trace est une fonction linéaire, la trace d'une différence $\rho - \sigma$ d'états quantiques vaut zéro. L'espace Ω_0 engendré par l'ensemble $\{\rho - \mu \,|\, \rho \in \Omega\}$ est donc un sous-espace de $\mathcal{H}_0(d)$. Prenons une base $\{\tau_i\}_{i=1}^m$ de Ω_0. On peut écrire tout $\rho \in \Omega$ sous la forme

$$\rho = \mu + \sum_{i=1}^{m} r_i \tau_i,$$

où $r_i \in \mathbb{R}$. Si le super-opérateur $\mathcal{E}$ est tel que $\mathcal{E}(\tau_i) = 0$ quand $1 \leq i \leq m$, il faut nécessairement que pour tout $\rho \in \Omega$, $\mathcal{E}(\rho) = \mathcal{E}(\mu)$.

L'inverse est également vrai :

Proposition 3.7.2. *Soit $\Omega \subseteq \mathcal{D}(d)$, $\mu \in \Omega$ et $\mathcal{E}$ un super-opérateur quelconque. Supposons que $\{\tau_i\}_{i=1}^m$ soit une base de Ω_0, le sous-espace de $\mathcal{H}_0(d)$ engendré par l'ensemble $\{\rho - \mu \,|\, \rho \in \Omega\}$. Alors $\mathcal{E}(\rho) = \mathcal{E}(\mu)$ pour tout $\rho \in \Omega$ si et seulement si $\mathcal{E}(\tau_i) = 0$ pour tout i entre 1 et m.*

Démonstration. Nous avons déjà fait la moitié de la preuve ; il reste à montrer la condition «seulement si». Supposons donc que $\mathcal{E}(\rho) = \mathcal{E}(\mu)$ pour tout $\rho \in \Omega$. Alors,

pour tout $\rho \in \Omega$,

$$\begin{aligned} 0 &= \mathcal{E}(\rho) - \mathcal{E}(\mu) \\ &= \mathcal{E}(\rho - \mu). \end{aligned}$$

Il faut donc, par linéarité, que $\mathcal{E}(\sigma) = 0$ pour tous les $\sigma \in \Omega_0$. En particulier, $\mathcal{E}(\tau_i) = 0$ pour tout i entre 1 et m.

□

Tâchons de concrétiser ces notions à l'aide d'un exemple familier. Considérons le disque $D_c = \{(r_1, r_2, c) \,|\, r_1^2 + r_2^2 \leq 1 - c^2\} \subseteq B_3$. Supposons que $\Omega \subseteq \mathcal{D}(2)$ soit l'ensemble tel que $m(\Omega) = D_c$. Alors les états contenus dans Ω on tous la forme $\rho = \frac{1}{2}(I + r_1 X + r_2 Y + cZ)$, où $r_1^2 + r_2^2 \leq 1 - c^2$.

Si μ est un état quelconque de Ω, alors pour tout autre état ρ de Ω, on peut trouver des nombres réels a et b tels que $\rho - \mu = aX + bY$, et l'espace vectoriel réel engendré par les $\rho - \mu$ est $\Omega_0 = \{aX + bY \,|\, a, b \in \mathbb{R}\}$. Pour chiffrer Ω il faut trouver un super-opérateur $\mathcal{E}$ tel que $\mathcal{E}(X) = \mathcal{E}(Y) = 0$. Il suffit bien sûr de prendre $\mathcal{E} = \{\frac{1}{\sqrt{2}} I, \frac{1}{\sqrt{2}} Z\}$, puisque

$$\begin{aligned} \frac{1}{2} IXI + \frac{1}{2} ZXZ &= \frac{1}{2} X - \frac{1}{2} X \\ &= 0, \\ \frac{1}{2} IYI + \frac{1}{2} ZYZ &= \frac{1}{2} Y - \frac{1}{2} Y \\ &= 0. \end{aligned}$$

Chapitre 4

Modèle

4.1 Introduction

Le modèle de CQP utilisé au chapitre précédent est-il assez général pour caractériser tous les chiffres où Alice et Bob partagent une clé secrète classique et où Bob peut recouvrer exactement et à coup sûr l'état qui lui est envoyé par Alice ? C'est ce que l'intuition porte à croire. Pourtant, au moment d'écrire ces lignes, l'auteur n'est toujours pas parvenu à le démontrer.

Dans ce chapitre, la notion de CQP général sera présentée, ainsi que des arguments qui indiquent que cette notion englobe vraiment toutes les opérations de chiffrement qu'on puisse imaginer. Dans un second temps, il sera démontré que l'utilisation du modèle de CQP général ne modifie pas les résultats de [AMTdW00]. Finalement, un modèle baptisé «CQP unitaire à ancille variable» sera présenté, qui diffère légèrement du modèle du chapitre précédent. Il sera démontré que tout CQP général peut se réduire à ce nouveau modèle.

4.2 Le modèle le plus général

Disons que la variable aléatoire K — la clé secrète — prend des valeurs entières entre 1 et n. Quand Alice et Bob partagent la valeur $K = k$ de cette variable, Alice chiffre son message ρ à l'aide de l'opération $\mathcal{E}_k$ et envoie $\mathcal{E}_k(\rho)$ à Bob, qui déchiffre grâce à l'opération $\mathcal{D}_k$. Nous exigeons que pour tout message ρ et pour tout k, $\mathcal{D}_k(\mathcal{E}_k(\rho)) = \rho$ et ce, avec une probabilité de un.

Considérons les opérations $\mathcal{E}_k$ et $\mathcal{D}_k$ comme étant des circuits quantiques. De tels circuits pourraient être, de façon générale, une succession de super-opérateurs et de mesures. Mais, grâce au principe de la mesure différée[1], nous pourrons supposer que toute mesure est en fait la dernière opération du circuit. Maintenant, il reste à se convaincre que toute mesure est inutile.

Supposons qu'il y ait une mesure à la fin du circuit $\mathcal{D}_k$. Par la condition $\mathcal{D}_k(\mathcal{E}_k(\rho)) = \rho$, il faut que $\mathcal{D}_k$, étant donné l'état chiffré $\mathcal{E}_k(\rho)$, permette de recouvrer ρ, *peu importe le résultat de la mesure*. Donc, ce résultat n'a pas besoin d'être connu de Bob. Mais, une mesure dont on ignore le résultat peut être décrite par un super-opérateur. Il s'ensuit qu'un super-opérateur suffit à décrire le circuit de chiffrement $\mathcal{D}$.

Reprenons le même argument de manière plus formelle pour montrer qu'une mesure dans le circuit $\mathcal{E}_k$ est inutile. Supposons qu'il y ait, à la fin du circuit $\mathcal{E}_k$, la mesure $\left\{A_m^{(k)}\right\}_{m=1}^{M}$. Ainsi, l'état envoyé à Bob par Alice est $\frac{1}{q_m} A_m^{(k)} \rho\prime \left(A_m^{(k)}\right)^\dagger$ avec probabilité q_m, où $q_m = \mathrm{tr}\left(A_m^{(k)} \rho\prime \left(A_m^{(k)}\right)^\dagger\right)$ et $\rho\prime$ est l'état du circuit avant la mesure.

Peu importe le résultat m de la mesure, il faut que $\mathcal{D}_k\left(\frac{1}{q_m} A_m^{(k)} \rho \left(A_m^{(k)}\right)^\dagger\right) = \rho$. Alice peut donc remplacer la mesure $\left\{A_m^{(k)}\right\}_{m=1}^{M}$ par le super-opérateur $\left\{A_m^{(k)}\right\}_{m=1}^{M}$

[1] Ce principe est décrit à la page 186 de [NC00].

— c'est-à-dire qu'elle enverra à Bob l'état $\sum_{m=1}^{M} A_m^{(k)} \rho \boldsymbol{\prime} \left(A_m^{(k)}\right)^\dagger$ — sans altérer la capacité de Bob à déchiffrer le message :

$$\begin{aligned}\mathcal{D}_k\left(\sum_{m=1}^{M} A_m^{(k)} \rho \boldsymbol{\prime} \left(A_m^{(k)}\right)^\dagger\right) &= \sum_{m=1}^{M} q_m \mathcal{D}_k\left(\frac{1}{q_m} A_m^{(k)} \rho \left(A_m^{(k)}\right)^\dagger\right) \\ &= \sum_{m=1}^{M} q_m \rho \\ &= \rho.\end{aligned}$$

Cela n'affectera pas non plus la confidentialité du chiffre : puisqu'il fallait que $\sum_{k=1}^{n} p_k A_m^{(k)} \rho \boldsymbol{\prime} \left(A_m^{(k)}\right)^\dagger = \phi$ pour chaque m — le message chiffré devait être ϕ peu importe le résultat de la mesure —, nous aurons maintenant

$$\begin{aligned}\sum_{k=1}^{n} p_k \left(\sum_{m=1}^{M} A_m^{(k)} \rho \boldsymbol{\prime} \left(A_m^{(k)}\right)^\dagger\right) &= \sum_{m=1}^{M} \left(\sum_{k=1}^{n} p_k A_m^{(k)} \rho \boldsymbol{\prime} \left(A_m^{(k)}\right)^\dagger\right) \\ &= \sum_{m=1}^{M} p_k \phi \\ &= \phi.\end{aligned}$$

Nous sommes maintenant prêts à définir le modèle le plus général possible. Soit $S \subseteq \mathcal{D}(d)$ un ensemble d'états, $\mathcal{E} = (\mathcal{E}_1, \ldots, \mathcal{E}_n)$ et $\mathcal{D} = (\mathcal{D}_1, \ldots, \mathcal{D}_n)$ deux vecteurs de super-opérateurs, $p = (p_1, \ldots, p_n)$ un vecteur de probabilité (tel que la somme des p_k vaille un) et ϕ une matrice de densité. On dira du quintuplet $[S, \mathcal{E}, \mathcal{D}, p, \phi]$ qu'il est un **CQP général** si, pour tout $\rho \in S$ et pour tout i entre 1 et n, $\mathcal{D}_k(\mathcal{E}_k(\rho)) = \rho$ et $\sum_{i=1}^{n} p_k \mathcal{E}_k(\rho) = \phi$.

Nous utiliserons aussi le symbole $\mathcal{E}$ pour désigner l'opérateur de chiffrement, du point de vue d'un adversaire qui ne connaît pas la clé :

$$\mathcal{E}(\rho) \stackrel{\text{def}}{=} \sum_{i=1}^{n} p_k \mathcal{E}_k(\rho).$$

Pour faire contraste avec la notion de CQP général, un CQP tel que défini à la section 3.2 sera nommé **CQP unitaire**.

4.3 Il faut deux bits de clé pour chiffrer un qubit

L'article [AMTdW00] fait la preuve que tout CQP unitaire $[S, U, p, \sigma, \phi]$ sur n qubits doit être tel que $H(p) \geq 2n$. Dans cette section, nous montrons que ce résultat tient toujours si on considère un CQP général plutôt qu'un CQP unitaire. Il s'agit essentiellement d'une version quantique de la preuve de Shannon : un CQP pouvant chiffrer un message quantique de n qubits permet de chiffrer un message classique de $2n$ bits. Nous montrons ensuite que même dans un contexte quantique, $2n$ bits de clé sont nécessaires pour chiffrer un message classique de $2n$ bits.

Cette version quantique de la preuve de Shannon est une adaptation d'une preuve tirée de [DHT03], qui montre que $2n$ bits classiques sont nécessaires pour dissimuler n qubits, ainsi que d'exercices donnés par Patrick Hayden à ses élèves à l'hiver 2005.

Considérons l'expérience suivante : Alice veut envoyer un message classique à Bob, choisi au hasard selon la distribution uniforme sur les chaînes de $2n$ bits. Alice et Bob partagent déjà un état maximalement intriqué $|\Phi\rangle^{AB} = 2^{-\frac{n}{2}} \sum_{i=0}^{2^n-1} |i\rangle^A |i\rangle^B$. Alice utilisera le codage dense pour coder son message classique m, c'est-à-dire qu'elle appliquera l'opérateur de Pauli σ_m à sa moitié de l'état $|\Phi\rangle^{AB}$:

$$|\Phi_m\rangle^{AB} \stackrel{\text{def}}{=} \left(\sigma_m^A \otimes I^B\right) |\Phi\rangle^{AB}, \tag{4.1}$$

où $0 \leq m \leq 2^{2n-1}$. Ensuite, elle chiffre sa moitié et l'envoie à Bob qui obtient, en combinant cette moitié avec la sienne, l'état $(\mathcal{E}_k \otimes I) |\Phi_m\rangle$. Bob, connaissant k, peut déchiffrer la moitié d'Alice et mesurer $|\Phi_m\rangle$ dans la base de Bell, $\{|\Phi_i\rangle\}_{i=0}^{2^{2n}-1}$, ce qui lui permet de retrouver m.

De cette façon, Alice transmet un message classique de $2n$ bits ; on s'attend donc à ce que l'entropie de la clé K (on traite maintenant la clé comme une variable aléatoire) soit d'au moins $2n$ bits. Nous consacrons le reste de cette section à démontrer la véracité de cette affirmation.

Lemme 4.3.1.[2] *Soit $\mathcal{E}$ un super-opérateur et ϕ un opérateur de densité tel que $\mathcal{E}(\varphi) = \phi$ pour tout état pur φ. Si $\{|j\rangle\}_{j=0}^{d-1}$ est une base de l'espace sur lequel agit $\mathcal{E}$, alors $\mathcal{E}(|j\rangle\langle k|) = \delta_{j,k}\phi$.*

Démonstration. Choisissons j et k tels que $0 \leq j, k \leq d-1$ et $j \neq k$. Considérons l'état pur $\varphi = \frac{1}{2}(|j\rangle + |k\rangle)(\langle j| + \langle k|)$. Par hypothèse, $\mathcal{E}(\varphi) = \phi$. Mais alors,

$$\begin{aligned}
2\phi &= 2\mathcal{E}(\varphi) \\
&= \mathcal{E}((|j\rangle + |k\rangle)(\langle j| + \langle k|)) \\
&= \mathcal{E}(|j\rangle\langle j|) + \mathcal{E}(|j\rangle\langle k|) + \mathcal{E}(|k\rangle\langle j|) + \mathcal{E}(|k\rangle\langle k|) \\
&= \phi + \mathcal{E}(|j\rangle\langle k|) + \mathcal{E}(|k\rangle\langle j|) + \phi \\
&\Rightarrow \mathcal{E}(|j\rangle\langle k|) = -\mathcal{E}(|k\rangle\langle j|).
\end{aligned}$$

De façon similaire, on peut poser $\tilde{\varphi} = \frac{1}{2}(|j\rangle + i|k\rangle)(\langle j| - i\langle k|)$ pour obtenir :

$$\begin{aligned}
2\phi &= 2\mathcal{E}(\tilde{\varphi}) \\
&= \mathcal{E}(|j\rangle\langle j|) + i\mathcal{E}(|j\rangle\langle k|) - i\mathcal{E}(|k\rangle\langle j|) + \mathcal{E}(|k\rangle\langle k|) \\
&= \phi + i\mathcal{E}(|j\rangle\langle k|) - i\mathcal{E}(|k\rangle\langle j|) + \phi \\
&\Rightarrow \mathcal{E}(|j\rangle\langle k|) = \mathcal{E}(|k\rangle\langle j|).
\end{aligned}$$

La seule conclusion possible est $\mathcal{E}(|j\rangle\langle k|) = \mathcal{E}(|k\rangle\langle j|) = 0$.

Par contre, quand $j = k$, $|j\rangle\langle k|$ est un état pur et $\mathcal{E}(|j\rangle\langle k|) = \phi$, par hypothèse.

□

[2] Ce lemme est pratiquement identique à celui qu'on retrouve en annexe dans [AMTdW00], la différence étant qu'on présente ici une preuve plus simple au coût d'une hypothèse plus forte : on demande que $\mathcal{E}(\varphi) = \phi$ pour tout état pur φ, tandis que [AMTdW00] impose cette condition aux seuls états φ qui sont des produits tensoriels de n qubits.

Nous avons défini plus haut l'état $|\Phi\rangle = 2^{-\frac{n}{2}} \sum_{i=0}^{2^n-1} |i\rangle\,|i\rangle$. Comme les transformations unitaires locales ne changent pas la quantité d'intrication entre deux systèmes, les états suivants sont eux aussi maximalement intriqués :

$$|\Phi\rangle_U = (U^A \otimes I^B)\,|\Phi\rangle\,.$$

Le lemme qui suit montre que le fait de chiffrer une moitié A d'un état $|\Phi\rangle_U^{AB}$ brise toute corrélation entre les systèmes A et B :

Lemme 4.3.2. *Soit $\mathcal{E}$ un super-opérateur et ϕ un opérateur de densité tel que $\mathcal{E}(\varphi) = \phi$ pour tout état pur φ. Alors $(\mathcal{E} \otimes I)(|\Phi\rangle_U\langle\Phi|_U) = \phi \otimes \frac{1}{d} I$.*

Démonstration. Posons $U\,|j\rangle = |j\rangle_U$. Alors $|\Phi\rangle_U = 2^{-\frac{n}{2}} \sum_{j=0}^{2^n-1} |j\rangle_U\,|j\rangle$. Donc,

$$\begin{aligned}
(\mathcal{E} \otimes I)(|\Phi\rangle_U\langle\Phi|_U) &= (\mathcal{E} \otimes I)\left(\frac{1}{2^n}\sum_{j,k=0}^{2^n-1} |j\rangle_U\langle k|_U \otimes |j\rangle\,\langle k|\right) \\
&= \frac{1}{2^n}\sum_{j,k=0}^{2^n-1} \mathcal{E}(|j\rangle_U\langle k|_U) \otimes |j\rangle\,\langle k| \\
&= \frac{1}{2^n}\sum_{j,k=0}^{2^n-1} \delta_{j,k}\phi \otimes |j\rangle\,\langle k|\,, \text{ par le lemme 4.3.1} \\
&= \frac{1}{2^n}\sum_{j=0}^{2^n-1} \phi \otimes |j\rangle\,\langle j| \\
&= \phi \otimes \frac{1}{2^n} I.
\end{aligned}$$

□

Théorème 4.3.3. *Si $[\mathcal{D}(2^n), \mathcal{E}, \mathcal{D}, p, \phi]$ est un CQP général chiffrant des messages de n qubits, alors $H(p) \geq 2n$.*

Démonstration. Supposons qu'Alice et Bob réalisent l'expérience décrite à la page 94, à l'aide du CQP général $[\mathcal{D}(2^n), \mathcal{E}, \mathcal{D}, p, \phi]$. En guise de clé secrète, ils partageront chacun une moitié de l'état $\sum_k p_k \left|k,k\right\rangle \left\langle k,k\right|^K$.

Ils partagent aussi chacun une moitié de l'état maximalement intriqué $\left|\Phi\right\rangle^{AB}$ sur $2n$ qubits. Le codage dense permettra donc à Alice de coder 2^{2n} messages distincts. Disons qu'Alice conserve son message classique $\left|m\right\rangle$ dans l'espace M de dimension 2^{2n}. Puisque les messages sont choisis selon la distribution uniforme, $H(M) = 2n$.

Si C est le canal quantique par lequel Alice transmet sa moitié (chiffrée) de $\left|\Phi_m\right\rangle$ et B contient l'autre moitié de $\left|\Phi_m\right\rangle$, alors le système global $KMCB$ est décrit par l'opérateur de densité suivant :

$$\tau^{KMCB} = \sum_{i,m} p_k \left|k,k\right\rangle \left\langle k,k\right|^K \otimes \frac{1}{\dim M} \left|m\right\rangle \left\langle m\right| \otimes (\mathcal{E}_k \otimes I)(\left|\Phi_m\right\rangle \left\langle \Phi_m\right|)^{CB}.$$

Nous aurons besoin, pour construire notre démonstration, des trois résultats suivants :

1. $H(M|C,B) = H(M)$,
2. $H(M|K,C,B) = 0$,
3. $H(K|M,C,B) \geq 0$.

Voici la preuve de ces trois énoncés, dans l'ordre :

1. $H(M|C,B) = H(M)$.

   Cet énoncé dit essentiellement que l'état intriqué $\left|\Phi_m\right\rangle$, dont une moitié a été chiffrée, n'apporte aucune information sur le message en clair. D'abord, si on

oublie la clé, l'état τ devient

$$\begin{aligned}
\tau^{MCB} &\stackrel{\text{def}}{=} \operatorname{tr}_K\left(\tau^{KMCB}\right) \\
&= \sum_i \left(\langle i|^K \otimes I^{MCB}\right) \tau^{KMCB} \left(|i\rangle^K \otimes I^{MCB}\right) \\
&= \frac{1}{\dim M} \sum_m |m\rangle\langle m|^M \otimes \left(\left(\sum_i p_i \mathcal{E}_i\right) \otimes I\right) \left(|\Phi_m\rangle\langle\Phi_m|\right)^{CB} \\
&= \frac{1}{\dim M} \sum_m |m\rangle\langle m|^M \otimes \left(\mathcal{E} \otimes I\right) \left(|\Phi_m\rangle\langle\Phi_m|\right)^{CB} \\
&= \frac{1}{\dim M} \sum_m |m\rangle\langle m|^M \otimes \phi^C \otimes \frac{1}{\dim B} I^B, \text{ par le lemme 4.3.2.}
\end{aligned}$$

Donc,

$$\begin{aligned}
H(M|C,B) &= H(M,C,B) - H(C,B) \\
&= H(M) + H(C,B) - H(C,B), \text{ par la proposition 1.10.2} \\
&= H(M).
\end{aligned}$$

2. $H(M|K,C,B) = 0$.

   Cela découle simplement du fait qu'on puisse déchiffrer sans erreur.

3. $H(K|M,C,B) \geq 0$.

   Comme τ est séparable, de part et d'autre des systèmes K et MCB, on n'a qu'à appliquer la proposition 1.10.5 et le résultat suit.

Les résultats 1 et 2 permettent de voir que

$$\begin{aligned}
I(K:M|C,B) &= I(M:K|C,B), \text{ par symétrie de l'information mutuelle} \\
&= H(M|C,B) - H(M|K,C,B), \text{ par définition} \\
&= H(M).
\end{aligned}$$

Alors,

$$\begin{aligned} H(K) - H(K|M,C,B) &\geq H(K|C,B) - H(K|M,C,B) \\ &= I(K:M|C,B) \\ &= H(M) \end{aligned}$$

et le résultat 3 nous amène à conclure que $H(K) \geq H(M) = 2n$.

□

4.4 Tout CQP général se réduit à un CQP unitaire à ancille variable

À la lumière des résultats de la section précédente, il appert que l'utilisation d'un CQP général plutôt que d'un CQP unitaire ne permet pas de diminuer la quantité de clé requise pour chiffrer de l'information quantique. Dans cette section, nous cheminons vers un résultat plus fort : peut-être existe-t-il, pour tout CQP général, un CQP unitaire équivalent ? Ignorant toujours si tel est le cas, l'auteur de ce mémoire se contentera de présenter une version édulcorée de cette affirmation.

Soit $S \subseteq \mathcal{D}(d)$ un ensemble d'états, $U = (U_1, \ldots, U_n)$ un vecteur de matrices unitaires, $p = (p_1, \ldots, p_n)$ un vecteur de probabilité, $\sigma = (\sigma_1, \ldots, \sigma_n)$ un vecteur de matrices de densité et ϕ une matrice de densité. Si pour tout $\rho \in S$, $\sum_{i=1}^{n} U_i(\rho \otimes \sigma_i)U_i^\dagger = \phi$, alors nous dirons que $[S, \mathcal{U}, p, \sigma, \phi]$ est un **CQP unitaire à ancille variable**.

Nous dirons qu'un CQP général $[S, \mathcal{E}, \mathcal{D}, p, \phi]$ **se réduit à un CQP unitaire à ancille variable** s'il existe un CQP unitaire à ancille variable $[S, \mathcal{U}, p, \sigma, \phi \otimes \phi_1]$ tel que pour tout état ρ et valeur de clé i, $U_i(\rho \otimes \sigma_i)U_i^\dagger = \mathcal{E}_i(\rho) \otimes \phi_1$.

Pour le reste du chapitre, nous nous emploierons à démontrer le théorème suivant :

Théorème 4.4.1. *Tout CQP général se réduit à un CQP unitaire à ancille variable.*

Nous avons vu qu'un super-opérateur $\mathcal{E}$ peut toujours être décrit par une équation de la forme $\mathcal{E}(\rho) = \mathrm{tr}_B\left(U(\rho \otimes \sigma)U^\dagger\right)$. Intuitivement, la différence entre un CQP général et un CQP unitaire à ancille variable, c'est que dans le premier cas, Alice jette le contenu d'un registre B, tandis que dans le second cas, elle envoie tout à Bob. Nous montrerons que dans le premier cas, le contenu du système B est nécessairement constant et indépendant du message comme de la clé, ce qui impliquera qu'Alice peut le transmettre sans même le chiffrer. C'est ce qui permet de réduire le premier cas au second.

Avant de présenter la preuve du théorème 4.4.1, nous aurons besoin de présenter divers faits, rassemblés dans les pages qui suivent en une succession de lemmes. Le premier affirme que lorsqu'on se débarrasse d'un système B, on peut toujours présumer qu'il ait d'abord été mesuré, d'une mesure projective complète (sans, bien sûr, qu'on ait appris le résultat) :

Lemme 4.4.2.[3] *Soit τ un état d'un système bipartite AB et $\mathbb{M} = [I\otimes\Pi_1, \ldots, I\otimes\Pi_n]$ une mesure projective complète*[4] *du système B. Alors*

$$\mathrm{tr}_B(\tau) = \mathrm{tr}_B\left(\sum_i \mathbb{M}_i(\tau)\right),$$

où $\mathbb{M}_i(\tau) = (I \otimes \Pi_i)\tau(I \otimes \Pi_i)$.

Démonstration. Soit $\{|i\rangle\}_{i=1}^n$ une base orthonormale du système B qui est telle que

[3] Présenté sans preuve dans [NC00] sous le nom de «principle of implicit measurement».

[4] C'est-à-dire qu'il existe une base orthonormale $\{|i\rangle\}_{i=1}^n$ du système B telle que $|i\rangle\langle i| = \Pi_i$ pour chaque i.

$|i\rangle\langle i| = \Pi_i$. Alors,

$$\begin{aligned}
\mathrm{tr}_B\left(\sum_{i=1}^{n}\mathbb{M}_i(\tau)\right) &= \sum_{j=1}^{n}(I\otimes\langle j|)\left(\sum_{i=1}^{n}\mathbb{M}_i(\tau)\right)(I\otimes|j\rangle)\\
&= \sum_{j=1}^{n}(I\otimes\langle j|)\left(\sum_{i=1}^{n}(I\otimes|i\rangle\langle i|)\tau(I\otimes|i\rangle\langle i|)\right)(I\otimes|j\rangle)\\
&= \sum_{i,j=1}^{n}(I\otimes\langle j|i\rangle\langle i|)\tau(I\otimes|i\rangle\langle i|j\rangle)\\
&= \sum_{i=1}^{n}(I\otimes\langle i|)\tau(I\otimes|i\rangle)\\
&= \mathrm{tr}_B(\tau).
\end{aligned}$$

□

Le lemme qui suit caractérise les états purs : un état pur est un état qui ne peut être décrit comme une distribution de probabilité sur plusieurs états distincts. Il s'agit en fait d'une simple constatation géométrique.

Lemme 4.4.3. *Si $\psi \in \mathcal{D}(d)$ est un état pur et que $\psi = \sum_{i=1}^{n} p_i\rho_i$, où les ρ_i sont des états quelconques et $\sum_{i=1}^{n} p_i = 1$, alors $\rho_i = \psi$ pour tout i entre 1 et n.*

Démonstration. Si ψ est pur, alors $m(\psi)$ est à la surface de la boule B_{d^2-1} de rayon $R_d = \sqrt{2(1-\frac{1}{d})}$ dans $\mathbb{R}^{d^2-1}$. Il suffit de montrer qu'une combinaison convexe de points de la boule B_{d^2-1} doit être strictement à l'intérieur de celle-ci, à moins qu'il ne s'agisse d'une combinaison triviale d'un seul point.

Pour plus de simplicité, disons que $m(\psi) = R_d\vec{e_1}$, où $\vec{e_1} = (1, 0, \ldots, 0)$. Nous nommerons π le projecteur $\pi : \mathbb{R}^{d^2-1} \to \mathbb{R}$ qui donne la première composante d'un vecteur, de sorte que $\pi(m(\psi)) = R_d$. Puisque $\|m(\rho_i)\| \leq R_d$ pour tout i, il faut, en

particulier, que $\pi(m(\rho_i)) \leq R_d$. Donc, les équations

$$\begin{aligned} R_d &= \pi(m(\psi)) \\ &= \pi(m(\sum_{i=1}^{n} p_i \rho_i)) \\ &= \sum_{i=1}^{n} p_i \pi(m(\rho_i)) \end{aligned}$$

montrent qu'il faut que $\pi(m(\rho_i)) = R_d$ pour chaque i, car autrement, on aurait

$$\sum_{i=1}^{n} p_i \pi(m(\rho_i)) < \sum_{i=1}^{n} p_i R_d = R_d.$$

Il faut donc que pour chaque i, $m(\rho_i) = R_d \vec{e_1}$ — si une coordonnée de $m(\rho_i)$, autre que la première, était différente de zéro, la norme de $m(\rho_i)$ serait supérieure à R_d. En conclusion, $\rho_i = \psi$ pour chaque i.

□

Un état bipartite ρ^{AB} peut être écrit comme un produit tensoriel $\rho = \rho_A \otimes \rho_B$ si (et, trivialement, seulement si) toute mesure projective complète du système B laisse le système A inchangé :

Lemme 4.4.4.[5] *Soit ρ un état du système AB, $\rho_A = \mathrm{tr}_B(\rho)$ et $\rho_B = \mathrm{tr}_A(\rho)$. Supposons que pour toute mesure complète du système B, le système A après la mesure est décrit par la matrice de densité ρ_A et ce, peu importe le résultat obtenu. En termes mathématiques, nous supposons que*

$$\mathrm{tr}_B \left(\frac{(I \otimes \Pi)\rho(I \otimes \Pi)}{\mathrm{tr}\left((I \otimes \Pi)\rho\right)} \right) = \rho_A \tag{4.2}$$

pour tout projecteur Π de rang un agissant sur le système B. Alors $\rho = \rho_A \otimes \rho_B$.

[5] Merci à Patrick Hayden d'avoir fourni la preuve de cet énoncé.

Démonstration. Posons $\rho = \mathrm{tr}_B(\tau)$. Supposons que l'équation 4.2 soit vraie pour tout projecteur Π de rang un, conformément à l'hypothèse du lemme. Nous pouvons récrire cette équation ainsi :

$$\begin{aligned}
& \mathrm{tr}_B\left((I\otimes\Pi)\rho(I\otimes\Pi)\right) = \mathrm{tr}\left((I\otimes\Pi)\rho\right)\rho_A \\
\Rightarrow\ & \mathrm{tr}_B\left((I\otimes\Pi^2)\rho(I\otimes I)\right) = \mathrm{tr}\left(\mathrm{tr}_A\left((I\otimes\Pi)\rho\right)\right)\rho_A \\
\Rightarrow\ & \mathrm{tr}_B\left((I\otimes\Pi)\rho\right) = \mathrm{tr}\left(\Pi\mathrm{tr}_A(\rho)\right)\rho_A \\
\Rightarrow\ & \mathrm{tr}_B\left((I\otimes\Pi)\rho\right) = \mathrm{tr}\left(\Pi\rho_B\right)\rho_A. \qquad (4.3)
\end{aligned}$$

Soit $\{X_i \otimes Y_j\}$ une base orthonormale de l'espace vectoriel réel que forment les matrices hermitiennes du système AB. Nous pouvons écrire

$$\rho = \sum_{i,j} \alpha_{i,j} X_i \otimes Y_j.$$

Première remarque : bien que les Y_j ne soient pas nécessairement des projecteurs (ni même positifs), et qu'ils ne puissent donc pas être considérés comme des éléments d'une mesure projective (ni d'un POVM), l'équation 4.3 demeure vraie quand on remplace Π par Y_j. En effet, prenons Y_j, dont la décomposition spectrale est, disons, $Y_j = \sum_k \lambda_k \Pi_k$, et voyons ce que nous obtenons :

$$\begin{aligned}
\mathrm{tr}_B\left((I\otimes Y_j)\rho\right) &= \mathrm{tr}_B\left(\left(I\otimes\sum_k \lambda_k\Pi_k\right)\rho\right) \\
&= \sum_k \lambda_k \mathrm{tr}_B\left((I\otimes\Pi_k)\rho\right) \\
&= \sum_k \lambda_k \mathrm{tr}\left(\Pi_k\rho_B\right)\rho_A \text{ , par l'équation 4.3} \\
&= \mathrm{tr}\left(\left(\sum_k \lambda_k\Pi_k\right)\rho_B\right)\rho_A \\
&= \mathrm{tr}\left(Y_j\rho_B\right)\rho_A.
\end{aligned}$$

Nous pouvons maintenant calculer les coefficients $\alpha_{i,j}$. Observons tout d'abord que

$$\begin{aligned} < X_i \otimes Y_j, \rho > &= < X_i \otimes Y_j, \sum_{k,l} \alpha_{k,l} X_k \otimes Y_l > \\ &= \sum_{k,l} \alpha_{k,l} < X_i \otimes Y_j, X_k \otimes Y_l > \\ &= \sum_{k,l} \alpha_{k,l} \delta_{i,k} \delta_{j,l} \\ &= \alpha_{i,j}, \end{aligned}$$

puisque les $X_i \otimes Y_j$ sont orthonormaux. Continuons le calcul des $\alpha_{i,j}$:

$$\begin{aligned} \alpha_{i,j} &= < X_i \otimes Y_j, \rho > && (4.4) \\ &= \mathrm{tr}\left((X_i \otimes Y_j)\rho\right) && (4.5) \\ &= \mathrm{tr}\left(\mathrm{tr}_B\left((X_i \otimes Y_j)\rho\right)\right) && (4.6) \\ &= \mathrm{tr}\left(X_i \,\mathrm{tr}_B\left((I \otimes Y_j)\rho\right)\right) && (4.7) \\ &= \mathrm{tr}\left(X_i \,\mathrm{tr}\left(Y_j \rho_B\right) \rho_A\right) && (4.8) \\ &= \mathrm{tr}\left(Y_j \,\rho_B\right) \mathrm{tr}\left(X_i \rho_A\right). && (4.9) \end{aligned}$$

Maintenant, écrivons $\rho_A \otimes \rho_B = \sum_{i,j} \beta_{i,j} (X_i \otimes Y_j)$ et calculons les $\beta_{i,j}$:

$$\begin{aligned} \beta_{i,j} &= \mathrm{tr}\left((X_i \otimes Y_j)(\rho_A \otimes \rho_B)\right) && (4.10) \\ &= \mathrm{tr}\left(X_i \,\rho_A \otimes Y_j \rho_B\right) && (4.11) \\ &= \mathrm{tr}\left(X_i \,\rho_A\right) \mathrm{tr}\left(Y_j \rho_B\right). && (4.12) \end{aligned}$$

Comme $\alpha_{i,j} = \beta_{i,j}$ pour tout i, j, nous pouvons conclure que $\rho = \rho_A \otimes \rho_B$.

□

Le prochain résultat combine les trois lemmes précédents pour en tirer un fait crucial : supposons que Bob puisse toujours déchiffrer parfaitement, au sens où, pour

tout ψ, $\mathcal{D}(\mathcal{E}(\psi)) = \psi$. Voyons l'opération $\mathcal{D}$ comme étant la succession d'opérations suivantes : ajout d'une ancille, transformation unitaire, trace partielle. En d'autres mots,

$$\mathcal{D}(\mathcal{E}(\rho)) = \mathrm{tr}_{BC}\left(U(\mathcal{E}(\rho)^{AB} \otimes |0\rangle\langle 0|^{C})U^{\dagger}\right). \tag{4.13}$$

Alors, le théorème qui suit montre qu'avant de faire la trace partielle des systèmes B et C, Bob possède un état ayant la forme d'un produit tensoriel :

$$U(\mathcal{E}(\rho)^{AB} \otimes |0\rangle\langle 0|^{C})U^{\dagger} = \rho^{A} \otimes \sigma^{BC}.$$

Théorème 4.4.5. *Soit $[\mathcal{E}, \mathcal{D}]$ une paire de super-opérateur tels que pour tout état pur ψ, $\mathcal{D}(\mathcal{E}(\psi)) = \psi$. Alors il existe une matrice unitaire U et un état σ tel que pour tout ρ^{A}, $U(\mathcal{E}(\rho)^{AB} \otimes |0\rangle\langle 0|^{C})U^{\dagger} = \rho^{A} \otimes \sigma^{BC}$.*

Démonstration. Soit ψ un état pur. Choisissons une matrice unitaire U qui valide l'équation 4.13. Posons aussi $\tau = U(\mathcal{E}(\psi)^{AB} \otimes |0\rangle\langle 0|^{C})U^{\dagger}$. Si $\mathcal{E}$ et $\mathcal{D}$ sont tels que dans l'énoncé du théorème, alors, nécessairement, $\mathrm{tr}_{BC}(\tau) = \psi$. Prenons une mesure complète $\mathbb{M}$ du système BC. Pour alléger les équations, posons $p_i = \mathrm{tr}\left((I \otimes \Pi_i)\tau\right)$ et $\mathbb{M}_i(\tau) = (I \otimes \Pi_i)\tau(I \otimes \Pi_i)$. Nous pouvons utiliser le lemme 4.4.2 pour obtenir :

$$\begin{aligned}
\psi &= \mathrm{tr}_{BC}(\tau) \\
&= \mathrm{tr}_{BC}\left(\sum_i \mathbb{M}_i(\tau)\right) \\
&= \mathrm{tr}_{BC}\left(\sum_i p_i \frac{\mathbb{M}_i(\tau)}{p_i}\right) \\
&= \sum_i p_i \,\mathrm{tr}_{BC}\left(\frac{\mathbb{M}_i(\tau)}{p_i}\right).
\end{aligned}$$

Autrement dit, ψ peut être exprimé comme une combinaison convexe des états $\rho_i \stackrel{\mathrm{def}}{=} \mathrm{tr}_{BC}\left(\frac{\mathbb{M}_i(\tau)}{p_i}\right)$. Dans ce cas, le lemme 4.4.3 nous permet de conclure que $\rho_i = \psi$

pour tout i. En d'autres mots,

$$\mathrm{tr}_{BC}\left(\frac{(I\otimes\Pi_i)\tau(I\otimes\Pi_i)}{\mathrm{tr}\left((I\otimes\Pi_i)\tau\right)}\right)=\psi \text{ pour tout } i.$$

Ce que nous venons de montrer, c'est qu'en effectuant une mesure complète arbitraire du système BC, le système A reste dans l'état ψ et ce, peu importe le résultat i obtenu. Il s'ensuit par le lemme 4.4.4 que $\tau=\psi\otimes\sigma$.

Il ne nous reste qu'à remarquer une chose : l'état σ doit être le même, peu importe l'état ψ qui a été chiffré au départ. Autrement, on se trouverait en présence d'un circuit qui transforme un état ψ_1 en $\psi_1\otimes\sigma_1$ et un état ψ_2 en $\psi_2\otimes\sigma_2$, où $\sigma_1\neq\sigma_2$. Cela permettrait de distinguer (au moins partiellement) entre les états ψ_1 et ψ_2 sans les perturber, ce qui n'est pas permis par la mécanique quantique.[6]

Le résultat s'adapte aux états mélangés grâce à la linéarité des transformations $\mathcal{E}$, $\rho\mapsto\rho\otimes|0\rangle\langle 0|$ et $\mathrm{Ad}U$. En effet, si on pose $\tau=U(\mathcal{E}(\rho)^{AB}\otimes|0\rangle\langle 0|^C)U^\dagger$, et que ρ est un état mélangé ayant pour décomposition spectrale $\rho=\sum_i\lambda_i\psi_i$, alors

$$\begin{aligned}
\tau &= U\left(\mathcal{E}\left(\sum_i\lambda_i\psi_i\right)^{AB}\otimes|0\rangle\langle 0|^C\right)U^\dagger\\
&= \sum_i\lambda_i U\left(\mathcal{E}(\psi_i)^{AB}\otimes|0\rangle\langle 0|^C\right)U^\dagger\\
&= \sum_i\lambda_i(\psi_i\otimes\sigma)\\
&= \left(\sum_i\lambda_i\psi_i\right)\otimes\sigma\\
&= \rho\otimes\sigma.
\end{aligned}$$

□

[6] Voir [NC00], proposition 12.18, page 586.

Corollaire 4.4.6. *Soit $[\mathcal{E}, \mathcal{D}]$ une paire de super-opérateurs telle que $\mathcal{D}(\mathcal{E}(\psi)) = \psi$ pour tout état pur ψ. Alors il existe un état σ et une transformation unitaire U telle que pour tout état ρ,*

$$U(\rho \otimes \sigma)U^\dagger = \mathcal{E}(\rho) \otimes |0\rangle \langle 0| .$$

Nous sommes maintenant prêts à prouver le résultat principal de cette section :

Démonstration du théorème 4.4.1. Soit $[\mathcal{S}, \mathcal{E}, \mathcal{D}, p, \phi]$ un CQP général. Par le corollaire 4.4.6, il existe des états σ_i et des transformations unitaires U_i telles que $U_i(\rho \otimes \sigma_i)U_i^\dagger = \mathcal{E}_i(\rho) \otimes |0\rangle \langle 0|$.

Posons $U = (U_1, \ldots, U_n)$ et $\sigma = (\sigma_1, \ldots, \sigma_n)$. Alors $[S, U, p, \sigma, \phi \otimes |0\rangle \langle 0|]$ est un CQP car, du point de vue d'Ève qui ne connaît pas la clé, tout message ρ chiffré par Alice sera décrit par la matrice de densité $\phi \otimes |0\rangle \langle 0|$:

$$\begin{aligned}
\mathcal{E}(\rho) &= \sum_i p_i U_i(\rho \otimes \sigma_i)U_i^\dagger \\
&= \sum_i p_i \left(\mathcal{E}_i(\rho) \otimes |0\rangle \langle 0|\right) \\
&= \left(\sum_i p_i \mathcal{E}_i(\rho)\right) \otimes |0\rangle \langle 0| \\
&= \phi \otimes |0\rangle \langle 0| .
\end{aligned}$$

L'égalité $\sum_i p_i \mathcal{E}_i(\rho) = \rho$ découle de l'hypothèse voulant que $[S, \mathcal{E}, p, \phi]$ soit un CQP.

□

Conclusion

Dans ce mémoire, une grande attention a été portée à la sphère de Poincaré. On y décrit explicitement la bijection m entre les états quantiques et les points de la sphère, ainsi que la correspondance φ entre les transformations unitaires et les rotations dans $\mathbb{R}^3$. Il a été montré que des fonctions m et φ semblables existaient aussi pour les espaces de dimension supérieure à deux, à la différence qu'elles ne sont alors plus surjectives. Autrement dit, certains points de la boule $B_D \subseteq \mathbb{R}^D$ ne correspondent à aucun état quantique, de même que certaines «rotations» de $\mathcal{SO}(D)$ ne correspondent à aucune transformation unitaire.

Il serait intéressant de savoir exactement quelle partie de B_D correspond à des états quantiques, de même que de découvrir un critère qui permette de savoir si une transformation orthogonale de $\mathcal{SO}(D)$ correspond ou non à une transformation unitaire.

D'apporter une réponse à ces questions sera sans doute un premier pas vers la solution du problème que le chapitre trois laisse irrésolu : comment calculer la quantité de clé requise pour chiffrer un ensemble arbitraire d'états $\Omega \subseteq \mathcal{D}(d)$. En deux dimensions, une solution complète a été exposée : il faut un bit si Ω est contenu dans un plan, et deux bits autrement. En dimension supérieure à deux, toutefois, la question demeure entière et constitue sans doute une avenue de recherche intéressante.

Une autre avenue intéressante serait d'étudier la même question dans le modèle de [HLSW04]. Ce modèle, au lieu de considérer uniquement les chiffres répondant à la condition voulant que $\mathcal{E}(\rho) = \phi$ pour chaque état ρ, tolère une certaine marge d'erreur au niveau de la sécurité. On y demande que le chiffre réponde à la condition $\|\mathcal{E}(\rho) - \frac{1}{d}I\|_\infty \leq \frac{\epsilon}{d}$, où $\|\sigma\|_\infty = \max_{i=1}^{d} |\lambda_i|$ quand les λ_i sont les valeurs propres de σ. À ce sujet, on pourra aussi consulter [AS04], où on utilise essentiellement le même modèle (à la différence qu'on utilise la norme de trace plutôt que la norme ∞), mais où on montre comment réaliser de tels chiffres de façon plus pratique.

Au quatrième chapitre, il fut montré que tout CQP général se réduit à un CQP unitaire à ancille variable. Il reste tout de même un fossé à combler entre les modèles de CQP unitaire et de CQP unitaire à ancille variable. Bien sûr, puisque que la preuve a été faite que même dans le modèle de CQP général, $2n$ bits de clé sont nécessaires pour chiffrer n qubits, les deux modèles sont équivalents de ce point de vue. Mais, il reste qu'il serait bon d'avoir un modèle qui soit à la fois complètement général et le plus simple possible. Quand, à l'avenir, on voudra savoir s'il est possible qu'un chiffre ait une certaine propriété, on pourra supposer d'emblée que ce chiffre est décrit par notre modèle le plus simple.

Bibliographie

[AMTdW00] Andris Ambainis, Michele Mosca, Alain Tapp, and Ronald de Wolf. Private quantum channels. In *41th Annual Symposium on Foundations of Computer Science : Proceedings*, pages 347–553, Los Alamitos, CA, 2000.

[Art91] Michael Artin. *Algebra*. Prentice Hall, New Jersey, 1991.

[AS04] Andris Ambainis and Adam Smith. Small pseudo-random families of matrices : Derandomizing approximate quantum encryption, 2004.

[BBC+93] Charles Bennett, Gilles Brassard, Claude Crépeau, Richard Josza, Asher Peres, and William Wooters. Teleporting an unknown state via dual classical and epr channels. *Physical review letters*, 1993.

[Ber92] Sterling K. Berberian. *Linear Algebra*. Oxford University Press, 1992.

[DHT03] David P. DiVincenzo, Patrick Hayden, and Barbara M. Terhal. Hiding quantum data. *FOUND.PHYS.*, 33 :1629, 2003.

[Hay05] Patrick Hayden. Entretien privé, 2005.

[HLSW04] Patrick Hayden, Debbie Leung, Peter W. Shor, and Andreas Winter. Randomizing quantum states : Constructions and applications. *Communications in Mathematical Physics*, 250 :371, 2004.

[Kna02] Anthony W. Knapp. *Lie Groups Beyond an Introduction*. Birkhäuser, Boston/Basel/Berlin, second edition, 2002.

[NC00] Michael A. Nielsen and Isaac L. Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press, Cambridge, United Kingdom, 2000.

[Sti95] Douglas Robert Stinson. *Cryptography : theory and practice*. CRC Press LLC, 1995.

[Zan98] Paolo Zanardi. A note on quantum cloning in d dimensions. `http ://xxx.lanl.gov/abs/quant-ph/9804011`, 1998.

[Zém00] Gilles Zémor. *Cours de cryptographie*. Cassini, Paris, 2000.

Table des matières

Remerciements **3**

Résumé **5**

Abstract **7**

Introduction **9**

1 Introduction à l'informatique quantique **13**

1.1 États quantiques et notation «braket» 13

1.2 Les fonctions trace et Vec 15

1.3 Quelques matrices à connaître 17

1.3.1 Matrices normales et décomposition spectrale 18

1.3.2 Matrices positives 19

1.3.3 Matrices unitaires 19

1.3.4 Matrices de Pauli et de Hadamard 20

1.4 Matrices de densité 21

1.5 Évolution unitaire 22

1.6 Mesures 23

1.6.1 Mesures projectives 23

1.6.2 Mesures générales 24

1.7 Produit tensoriel et juxtaposition de systèmes 25

1.8 Trace partielle 26

1.9 Super-opérateurs 27

1.10 Entropie d'états quantiques 27

1.10.1 Entropie de von Neumann 27

1.10.2 Entropie conditionnelle 28

1.11 Faits relatifs à la commutativité de matrices 29

2 La sphère de Poincaré **31**

2.1 Introduction 31

2.2 Notions préliminaires 32

2.2.1 Les groupes SU(2) et SO(3) 32

2.3 Matrices de densité et espaces vectoriels 33

2.3.1 Matrices hermitiennes 33

2.3.2 Une base pour $\mathcal{H}(2)$ 36

2.4 La sphère de Poincaré 37

2.4.1 Définition paramétrique de la sphère de Poincaré 37

2.4.2 Une bijection entre états quantiques et points de la sphère . . 38

2.4.3 Propriétés de la fonction m 41

2.5 La conjugaison par une matrice unitaire 42

2.6 Représentation des transformations unitaires par des rotations de la sphère de Poincaré 45

2.7 Rotations d'un demi-tour 52

2.8 Trouver la matrice unitaire associée à une rotation arbitraire 55

2.9 Entropie de von Neumann 61

2.10 La sphère de Poincaré pour les espaces de dimension d 64

2.10.1 Une sphère dans $\mathbb{R}^{d^2-1}$ 64

3 Canaux quantiques privés **69**

3.1 Le chiffre de Vernam 69

3.1.1 Confidentialité parfaite 70

3.2 Canaux quantiques 71

3.2.1 Modèle du processus de chiffrement 71

3.3 Chiffrement d'états coplanaires 74

3.4 Moins d'un bit de clé est inutile 76

3.5 S'il faut plus d'un bit, il en faut deux 79

3.6 États perpendiculaires 83

3.6.1 Chiffrement d'information classique 84

3.6.2 Hypercercles 85

3.7 Morceaux de plusieurs qubits 88

4 Modèle **91**

4.1 Introduction 91

4.2 Le modèle le plus général 92

4.3 Il faut deux bits de clé pour chiffrer un qubit 94

4.4 Tout CQP général se réduit à un CQP unitaire à ancille variable . . . 99

Conclusion **109**

Bibliographie **111**

Cette table **113**