

Section 4

SOURCES SYMBOLIQUES QUANTIQUES

La théorie statistique de l'information trouve un champ d'application particulièrement intéressant avec la mécanique quantique, qui étend le formalisme classique en spécifiant la signification et le contenu de la variable p dans la fonction $H(p)$, ce qui permet de préciser le concept de source symbolique. Dans la théorie classique les probabilités sont déduites a posteriori d'un décompte statistique des fréquences des symboles ou, si une telle opération est impossible, celles-ci sont estimées a priori selon un principe simple mais "passif" – et souvent discutable – comme l'équiprobabilité. En microphysique il en va tout autrement : "l'estimation" des probabilités est la conséquence d'une démarche active de l'observateur. Cette étape préalable est rendue nécessaire parce que, avant tout processus cognitif de la part de l'observateur, ce dernier ne peut, dès le départ et par principe même du calcul quantique, invoquer aucune connaissance quelle qu'elle soit concernant cette source. Comme le résume Mugur-Schächter en étendant l'épistémologie de la mesure quantique à l'étape initiale par laquelle un observateur extrait des chaînes de connaissance du réel où il est plongé et auquel il appartient, *"la toute première phase de tout processus de conceptualisation est foncièrement active, générative, relativisante, et la création de l'objet d'étude s'y accomplit en général indépendamment de la création des qualifications de cet objet"* [Mugur-Schächter, 1994, 1997]. La théorie quantique est un modèle mathématique du monde physique dont l'application met bien en évidence les différentes étapes du processus de conceptualisation du réel :

(i) Étape opérationnelle : *préparation* P d'un *micro-état* étiqueté ψ . Cette étape initiale est fondamentale, et radicalement nouvelle par rapport aux concepts classiques : la source $P\psi$ n'a pas d'existence *per se*, elle est factuellement la conséquence d'une opération active d'individualisation d'un objet par rapport au continuum contextuel (dénommé "le réel" dans la théorie microphysique).

(ii) Définition des observables : un *observable* A est une propriété d'un système physique (position, moment,...) qui peut être mesurée par principe, par le biais d'interactions entre un micro-état et un appareillage macroscopique. Comme en (i), dans cette étape l'observateur est actif : il décide en termes opératoires des qualifications qui caractérisent son objet d'étude.

(iii) Postulat de mesure : qualifiant l'entité $P\psi$, un observable est perceptible par ses

valeurs propres dont l'ensemble constitue son *spectre*. L'opération de mesure **MA** produit un résultat (une valeur propre A_i). Ce processus a deux caractéristiques fondamentales : a) en général, l'entité ψ n'existe plus (est détruite) après l'opération de mesure ; b) le résultat A_i est foncièrement, de façon inhérente, relatif à l'opération de mesure **MA** qui a permis de l'obtenir. Vue sous l'angle informationnel, la théorie quantique ne traite que de *couples* de sources symboliques ($P\psi$, **MA**), le système mesuré et le système mesurant, indissolublement associés.

(iv) Evolution dynamique et caractérisation pratique : l'évolution spatio-temporelle d'un micro-état ψ produit par $P\psi$ est unitaire. Elle est décrite par une certaine *fonction d'état* $\psi(\mathbf{r}, t)$ qui dès lors se substitue à la simple étiquette " ψ " dans la chaîne de conceptualisation pour désigner ce micro-état. C'est cette fonction qui génère les algorithmes permettant des prédictions (statistiques) précises et reproductibles.

Reprenons chacun de ces points en les détaillant, en suivant un plan adopté par un certain nombre d'auteurs qui ont publié récemment leurs réflexions sur le thème de la théorie quantique de l'information [notamment : Preskill, 1998 ; Smolin, 1996 ; Svozil, 1995 ; Vazirani, Chuang, 1997].

4.1. Introduction

4.1.1. Rappel succinct des axiomes de la mécanique quantique

4.1.1.1. États

L'état d'un système physique isolé \mathcal{S} est représenté par un vecteur $|\psi\rangle$ dans un espace de Hilbert \mathcal{H} .

(i) L'évolution d'un système quantique isolé \mathcal{S} est toujours unitaire. Mais la difficulté est qu'il n'existe pas de tels systèmes réellement isolés, sans interaction avec tout autre système (excepté peut-être l'univers entier), aussi doit-on faire certaines approximations dans l'emploi des algorithmes quantiques. Une solution consiste à tenir compte de l'environnement \mathcal{T} du système \mathcal{S} , ce qui revient à considérer des couples de systèmes physiques. Anticipant sur ce qui sera exposé dans la suite de ce paragraphe, l'effet de l'interaction entre \mathcal{S} et \mathcal{T} est équivalent à un processus de mesure. L'évolution de \mathcal{S} n'est alors plus unitaire (car les projections ne le sont plus) : ce phénomène, appelé *décohérence*, a des conséquences importantes, notamment dans le domaine des calculateurs quantiques.

(ii) Un espace de Hilbert \mathcal{H} est un espace vectoriel sur le corps \mathbb{C} des nombres

complexes. Cet espace peut comprendre un nombre de dimensions fini dénombrable, infini dénombrable ou infini non dénombrable. Les vecteurs sont notés $|\cdot\rangle$, selon la notation de Dirac ("ket"). Pour distinguer les vecteurs entre eux, une lettre ou un chiffre ou un symbole quelconque est inséré : le vecteur $\psi \in \mathcal{H}$ est ainsi représenté par " $|\psi\rangle$ ". Une classe d'équivalence de vecteurs regroupe l'ensemble des vecteurs qui ne diffèrent que par la multiplication par un scalaire complexe non nul.

Principe de superposition : un espace de Hilbert étant linéaire par définition, toute combinaison linéaire de vecteurs est un vecteur. Donc s'il "existe" deux états représentés par les vecteurs $|\psi_1\rangle$ et $|\psi_2\rangle$, alors tout état représenté par le vecteur $|\phi\rangle$ "existe" également :

$$|\phi\rangle = \lambda_1 |\psi_1\rangle + \lambda_2 |\psi_2\rangle \in \mathcal{H} \quad \lambda_1, \lambda_2 \in \mathbb{C}$$

Ce principe s'applique aussi au cas d'un espace dont le nombre de dimensions est infini :

$$|\phi\rangle = \int_{t_1}^{t_2} \lambda(t) |t\rangle \in \mathcal{H}$$

Une telle combinaison linéaire de vecteurs est appelée *superposition cohérente des états* parce que la phase relative entre vecteurs de la superposition est physiquement significative :

$\lambda_1 |\psi_1\rangle + \lambda_2 |\psi_2\rangle$ et $e^{i\alpha} (\lambda_1 |\psi_1\rangle + \lambda_2 |\psi_2\rangle)$ appartiennent à la même classe, mais pas $\lambda_1 |\psi_1\rangle + e^{i\alpha} \lambda_2 |\psi_2\rangle$.

(iii) Un espace de Hilbert est muni :

- d'un produit interne $\langle\phi|\psi\rangle$ ("produit scalaire") qui lie une paire ordonnée de vecteurs à un nombre complexe, avec les propriétés suivantes :

- positivité : $\langle\psi|\psi\rangle > 0$
- linéarité : $\langle\phi|\psi_1 + \psi_2\rangle = \langle\phi|\psi_1\rangle + \langle\phi|\psi_2\rangle$ et $\langle\phi|\lambda\psi\rangle = \lambda \langle\phi|\psi\rangle$
- symétrie hermitienne : $\langle\phi|\psi\rangle = \langle\psi|\phi\rangle^*$ (* = complexe conjugué)

- d'une norme : $\|\psi\| = \langle\psi|\psi\rangle^{1/2}$, dont les propriétés sont :

- inégalité du triangle (ou de Minkovski) : $\|\phi + \psi\| \leq \|\phi\| + \|\psi\|$
- inégalité de Schwarz : $|\langle\phi|\psi\rangle| \leq \|\phi\| \cdot \|\psi\|$

- d'une distance $\text{dist}(\phi, \psi) = \|\phi - \psi\|$, qui lui confère une métrique. Propriété :

- $\text{dist}(\phi, \psi) \leq \text{dist}(\phi, \gamma) + \text{dist}(\gamma, \psi) \quad \forall \phi, \psi, \gamma \in \mathcal{H}$

Un espace de Hilbert est *complet* pour cette norme, ce qui, dans la pratique, revient presque toujours à supposer l'espace séparable, propriété que l'on retrouvera en section 6 avec

l'analyse de Fourier multidimensionnelle.

Exemples :

• \mathbb{R} ou \mathbb{C} muni du produit scalaire $x \cdot y$, de la norme $|x|$, et de la distance $|x-y|$.

• \mathbb{R}^n ou \mathbb{C}^n muni du produit scalaire $\sum_{i=1}^n x_i^* y_i$ et de la norme $\left(\sum_{i=1}^n |x_i|^2 \right)^{1/2}$

• Espace L_2 des séquences infinies dénombrables :

$$L_2 = \{ \varphi : \varphi = (x_1, x_2, \dots, x_i, \dots), x_i \in \mathbb{C}, \sum_{i=1}^{\infty} |x_i|^2 < \infty \}$$

• Espace L_2 des fonctions continues sur \mathbb{C} , de carré sommable, muni du produit scalaire

$$\int \varphi^* \psi dx \text{ et de la norme } \|\varphi\|^2 = \int |\varphi|^2 dx .$$

L'espace vectoriel dual de l'espace de Hilbert est noté \mathcal{H}^\dagger (exemple : espace des fonctions $L_1 \subset L_2$ versus espace de leurs transformées de Fourier). Ses vecteurs sont notés $\langle . |$ ("bra"), avec les règles suivantes :

$$\lambda^\dagger = \lambda^*$$

$$\langle \psi |^\dagger = |\psi\rangle \text{ et } |\psi\rangle^\dagger = \langle \psi |$$

$$\langle \psi | \varphi \rangle^\dagger = \langle \varphi | \psi \rangle = \langle \psi | \varphi \rangle^*$$

L'espace de Hilbert est autodual : $\mathcal{H}^\dagger = \mathcal{H}$

Les éléments de l'ensemble des vecteurs de la base orthonormée $\{ |i\rangle : i \in \mathbf{I} \}$ où \mathbf{I} est un ensemble d'index de même cardinalité que \mathcal{H} satisfont :

$$\langle i | j \rangle = \delta_{ij} = \begin{cases} 0 & \text{si } i \neq j \\ 1 & \text{si } i = j \end{cases} \quad (\delta_{ij} \text{ est le symbole de Kronecker})$$

si \mathbf{I} est dénombrable, et :

$$\langle i | j \rangle = \delta(x-y) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{i(x-y)t} dt$$

si \mathbf{I} est continu (la fonction de Dirac, introduite à cette occasion, se substituant au symbole de Kronecker).

Tout vecteur $|\psi\rangle$ s'écrit selon une combinaison linéaire de vecteurs de la base orthonormée $\{ |i\rangle : i \in \mathbf{I} \}$:

$$|\psi\rangle = \sum_{i \in \mathbf{I}} a_i \cdot |i\rangle \quad \text{avec} \quad a_i = \langle i|\psi\rangle \in \mathbf{C}$$

L'opérateur identité $\mathbf{1}$ (avec $a_i = 1$) s'écrit (les sommes deviennent des intégrales dans le cas continu, celui de l'analyse de Fourier par exemple) :

$$\mathbf{1} = \sum_{i \in \mathbf{I}} |i\rangle \langle i|$$

Par exemple, si l'index i est identifié à l'opérateur de position spatiale \mathbf{r} et si l'état $|\psi(t)\rangle$ dépend du temps, alors $\langle \mathbf{r}|\psi(t)\rangle = \psi(\mathbf{r}, t)$ n'est autre que la fonction d'onde de Schrödinger.

4.1.1.2. Observables

Un observable est représenté par un opérateur auto-adjoint. Un opérateur \mathbf{A} est une application linéaire qui lie un vecteur à un autre :

$$\mathbf{A} : |\psi\rangle \rightarrow \mathbf{A}|\psi\rangle$$

$$\mathbf{A} (\alpha|\psi\rangle + \beta|\psi'\rangle) = \alpha\mathbf{A}|\psi\rangle + \beta\mathbf{A}|\psi'\rangle$$

L'adjoint d'un opérateur est défini par :

$$\langle \varphi|\mathbf{A}|\psi\rangle = \langle \mathbf{A}^\dagger|\varphi|\psi\rangle \quad \forall |\varphi\rangle, |\psi\rangle$$

L'opérateur est auto-adjoint si $\mathbf{A} = \mathbf{A}^\dagger$.

Si \mathbf{A} et \mathbf{B} sont auto-adjoints, alors $\mathbf{A} + \mathbf{B}$ l'est aussi, car $(\mathbf{A} + \mathbf{B})^\dagger = \mathbf{A}^\dagger + \mathbf{B}^\dagger$, mais $(\mathbf{AB})^\dagger = \mathbf{B}^\dagger \mathbf{A}^\dagger$, aussi \mathbf{AB} est auto-adjoint seulement si \mathbf{A} et \mathbf{B} commutent.

Dans un espace de Hilbert de cardinalité finie, des opérateurs auto-adjoints bornés sont hermitiens et peuvent s'écrire sous forme matricielle. La matrice adjointe d'une matrice donnée est la matrice transposée conjuguée.

Un opérateur auto-adjoint a une représentation spectrale :

$$\mathbf{A} = \sum_n a_n \cdot \mathbf{P}_n$$

où les a_n sont les valeurs propres de \mathbf{A} , et où les \mathbf{P}_n sont les projections orthogonales dans l'espace des vecteurs propres, les états propres de \mathbf{A} formant une base orthonormée dans \mathcal{H} . Si a_n est non-dégénérée, alors $\mathbf{P}_n = |a_n\rangle \langle a_n|$ est la projection sur le vecteur propre correspondant.

Deux observables \mathbf{A} et \mathbf{B} sont *compatibles*, c'est-à-dire "indépendants", définissables simultanément avec une précision arbitraire, si leur *commutateur* $[\mathbf{A}, \mathbf{B}] = \mathbf{AB} - \mathbf{BA}$ est nul.

Exemple :

Dans la base $\{ |x\rangle : x \in \mathbb{R} \}$, l'opérateur de position est $\mathbf{x} = x$, et l'opérateur de quantité de mouvement est $\mathbf{p}_x = p_x = \frac{\hbar}{i} \frac{\partial}{\partial x}$ (avec $\hbar = h/2\pi$). Ces deux opérateurs ne commutent pas :

$$[\mathbf{x}, \mathbf{p}_x] = \mathbf{x} \mathbf{p}_x - \mathbf{p}_x \mathbf{x} = x \frac{\hbar}{i} \frac{\partial}{\partial x} - \frac{\hbar}{i} \frac{\partial}{\partial x} x = i\hbar \neq 0$$

On montre que cette propriété conduit aux relations d'indétermination de Heisenberg :

$$\Delta x \Delta p_x \geq \frac{\hbar}{2\pi} \quad \text{avec} \quad \Delta x = \sqrt{x^2 - \langle x \rangle^2} ; \quad \Delta p_x = \sqrt{p_x^2 - \langle p_x \rangle^2}$$

4.1.1.3. Mesure

Une mesure de l'observable \mathbf{A} est une valeur propre de \mathbf{A} . Juste après une mesure, le système est dans un état propre de \mathbf{A} avec la valeur propre mesurée, et n'est plus en état de superposition cohérente. Ce principe est connu sous le nom de "collapse" de la fonction d'onde ψ . On compare souvent celle-ci – comme le faisait d'ailleurs Schrödinger lui-même [Wheeler, Zurek, 1983] – à un "catalogue des possibles" qui décrit non pas l'état du système en soi à un moment donné, mais l'état du système relativement au système de mesure, autrement dit l'état des "connaissances" que "l'observateur" a sur ce système mesuré à cet instant. En tant que représentation particulière de l'état du système dans une certaine base, la fonction d'onde est une représentation de l'information à laquelle l'observateur peut accéder. Par rapport à la théorie statistique classique de l'information, la théorie quantique fournit donc une étape supplémentaire dans le schéma de calcul de la théorie (voir schéma xx), qui consiste à préciser la loi ou plutôt l'ensemble des lois de probabilités possibles dans le cadre (et dans ce cadre seulement) de cette expérience à cet instant. Alors que la théorie classique ne connaît pas de loi de probabilité autre qu'une loi évidente d'équiprobabilité – sauf à faire des statistiques antérieurement à la mesure –, la théorie quantique propose une description du genre de probabilités avec lesquelles on peut mener les calculs, ces probabilités ne dépendant pas seulement du système physique, mais aussi de l'environnement de mesure qui l'englobe.

Si l'état quantique juste avant la mesure est $|\psi\rangle$, alors le résultat a_n est obtenu avec la probabilité :

$$p(a_n) = \| \mathbf{P}_n |\psi\rangle \|^2 = \langle \psi | \mathbf{P}_n | \psi \rangle$$

Si le résultat a_n est réalisé, l'état quantique (normalisé) devient :

$$\frac{\mathbf{P}_n |\psi\rangle}{(\langle \psi | \mathbf{P}_n | \psi \rangle)^{1/2}}$$

La valeur moyenne de l'observable $\mathbf{A} = \sum_n a_n \cdot \mathbf{P}_n$, où $\mathbf{P}_n = |a_n\rangle \langle a_n|$, dans l'état $|\psi\rangle$, est donnée par :

$$\langle \mathbf{A} \rangle = \langle \psi | \mathbf{A} | \psi \rangle = \sum_n a_n \langle \psi | a_n \rangle \langle a_n | \psi \rangle = \sum_n a_n |\langle \psi | a_n \rangle|^2$$

L'état du système dépend du fait qu'il est ou non mesuré. C'est ce qui fait dire que *l'information est physique* [Landauer, 1991]. Cet aspect de l'information était seulement "soupçonné" à travers la théorie classique, avec la célèbre équation de Brillouin $\Delta(I + S) \geq 0$, mettant sur un pied d'égalité une grandeur purement "mathématique", l'information, mesurée par le biais d'une unité arbitraire (le shannon par exemple), et l'entropie, grandeur physique dotée elle d'une dimension physique bien déterminée (le Joule/Kelvin). Mais cette vue des choses était contestable, et en fait indémontrable. Avec la théorie quantique les choses se précisent : un processus informationnel est un phénomène physique, qui a pleinement sa place au sein de la théorie. Ce n'est pas une ombre, un fantôme, qui viendrait se surajouter au monde physique, ajout qui serait seulement du fait du théoricien. Au contraire, les deux systèmes, émetteur et récepteur, sont intrinsèquement liés : un contenu d'information (le système mesuré) est couplé à un contenu d'identification (le système de mesure et/ou l'observateur) au sein d'un contexte commun (le processus de mesure, qui est le processus informationnel). L'opération de mesure, modélisée dans la pratique par un opérateur particulier nommé *opérateur de densité*, introduit une dissymétrie et donc une distinction définitive entre le système mesuré et son environnement, donc entre la source émettrice et la source réceptrice.

4.1.1.4. Dynamique

L'évolution temporelle d'un état quantique est générée par un opérateur auto-adjoint nommé *l'hamiltonien* \mathbf{H} du système, qui apparaît dans *l'équation de Schrödinger* :

$$\frac{d}{dt} |\psi(t)\rangle = -i\mathbf{H} |\psi(t)\rangle$$

En dérivant au premier ordre par rapport au temps :

$$|\psi(t + dt)\rangle = (\mathbf{1} - i\mathbf{H}dt) |\psi(t)\rangle$$

où l'opérateur $\mathbf{U}(dt) = \mathbf{1} - i\mathbf{H}dt$ est unitaire et linéaire ($\mathbf{U}\mathbf{U}^\dagger = \mathbf{1}$) : cette propriété est capitale en interférométrie, mécanisme de base des calculateurs quantiques. Il vient :

$$|\psi(t)\rangle = \mathbf{U}(t)|\psi(0)\rangle$$

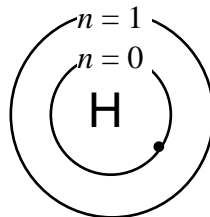
avec $\mathbf{U}(t) = e^{-it\mathbf{H}}$ si \mathbf{H} est indépendant du temps.

On voit donc que le modèle d'évolution d'un état quantique, représenté par $\psi(t)$, est dualiste : unitaire et déterministe si on spécifie $\psi(0)$, probabiliste si une mesure est effectuée. Le niveau microphysique met en évidence l'importance du fait informationnel, où le mot informer rejoint son sens étymologique de *former dedans, mettre en forme* : la mesure, c'est-à-dire le processus informationnel, est ici une mise en forme du réel. On verra par la suite que cette remarque dépasse largement le seul domaine microphysique.

4.1.2. Bit quantique

Nous allons nous intéresser au cas d'une source symbolique binaire quantique, dont les symboles élémentaires sont des bits quantiques, dénommés en abrégé par les auteurs anglo-saxons *qubits*.

Un électron aux niveaux d'énergie inférieurs de l'atome d'hydrogène en est un exemple.



• Figure 4.1 : bit quantique représenté par l'électron d'un atome d'hydrogène.

L'électron a les amplitudes de probabilité a et b pour être au repos (niveau $n = 0$) ou dans un état excité ($n = 1$) : en fait il existe dans les deux états à la fois, ce qui est noté sous la forme :

$$|\psi\rangle = a|0\rangle + b|1\rangle \quad a, b \in \mathbb{C}$$

La probabilité totale, qui représente la seule information certaine – l'électron existe – est égale à 1 :

$$|a|^2 + |b|^2 = 1$$

Le paradoxe est qu'il semble qu'un bit quantique contienne une quantité infinie d'information, puisque son état est représenté par deux degrés de liberté *continus*. Mais la théorie, comme cela a été vu plus haut, indique qu'une mesure (unique) de l'état de l'électron fournit soit $|0\rangle$ soit $|1\rangle$ avec une probabilité respectivement égale à $|a|^2$ ou $|b|^2$, c'est-à-dire un bit unique (classique) d'information. C'est seulement dans l'hypothèse où un grand nombre (en

toute rigueur une infinité) de bits quantiques sont préparés de façon identique que l'on peut connaître les quantités a et b . Cela pourrait suggérer que chaque bit quantique contienne une information "cachée" (les valeurs de a et b qui lui sont propres), mais rien dans la théorie n'indique que cela soit : tout ce que l'on est en droit de conclure, c'est que *la source symbolique elle-même*, dans sa globalité, et non tel ou tel message, contient les informations a et b que l'on ne peut, dans la pratique, qu'atteindre asymptotiquement.

Mathématiquement, un bit quantique est un état dans un espace de Hilbert bidimensionnel, dont les vecteurs $\{|0\rangle, |1\rangle\}$ forment une base orthonormée. Comme cela a été suggéré en introduction de cette section, point (ii), il nous faut, étant au niveau de la racine de la conceptualisation d'une source symbolique binaire quantique, construire une interprétation de celle-ci. Pour cela, nous avons toute latitude pour effectuer nos choix, et adopter par exemple une interprétation différente de celle qui vient d'être donnée (atome d'hydrogène). Nous choisirons préférentiellement l'interprétation suivante : $|\psi\rangle$ est l'état de spin d'un objet de spin $1/2$. La raison de ce choix est : la simplicité, l'efficacité et l'universalité du modèle mathématique qui sous-tend cette interprétation.

□ Description du modèle de bit quantique fondé sur le concept de spin

• *Rappel : rotations dans le plan*

Soit une rotation du vecteur $\overset{\mathbf{r}}{A} = \begin{pmatrix} a_x \\ a_y \end{pmatrix} \rightarrow \overset{\mathbf{r}}{B} = \begin{pmatrix} b_x \\ b_y \end{pmatrix}$, d'un angle θ par rapport à l'origine

O. Il existe un isomorphisme pour l'opérateur de rotation entre (\underline{A} et \underline{B} étant représentations complexes des vecteurs) :

$$\underline{B} = \mathcal{R}(O, \theta) \cdot \underline{A} \quad \text{avec} \quad \mathcal{R}(O, \theta) = e^{i\theta}, \quad \underline{A} = a_x + ia_y, \quad \underline{B} = b_x + ib_y$$

$$\text{et :} \quad \begin{pmatrix} b_x \\ b_y \end{pmatrix} = \mathbf{R} \begin{pmatrix} a_x \\ a_y \end{pmatrix} \quad \text{avec} \quad \mathbf{R} = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}$$

• *Spineur*

Soient deux vecteurs $\overset{\hat{}}{A}$ et $\overset{\hat{}}{B}$ ($a_r, b_r \in \mathbb{R}, r = x, y, z$), dans cet ordre, égaux et orthogonaux ($|\overset{\hat{}}{A}| = |\overset{\hat{}}{B}|$ et $\overset{\hat{}}{A} \cdot \overset{\hat{}}{B} = 0$). On construit un vecteur "complexe" $\overset{\hat{}}{U} = \overset{\hat{}}{A} + i \overset{\hat{}}{B}$, le coefficient i permettant de marquer l'ordre des vecteurs $\overset{\hat{}}{A}, \overset{\hat{}}{B}$. Soit $\overset{\hat{}}{U} = \{u_x, u_y, u_z \in \mathbb{C}\}$. $\overset{\hat{}}{U}$ est *isotrope* si :

$$u_x^2 + u_y^2 + u_z^2 = 0 \Leftrightarrow u_z = -(u_x^2 + u_y^2) = -(u_x + iu_y)(u_x - iu_y)$$

On pose : $u_x + iu_y \equiv -2\alpha^2$; $u_x - iu_y \equiv 2\beta^2$ avec $\alpha, \beta \in \mathbb{C}$

D'où $u_x = a_x + ib_x = \beta^2 - \alpha^2$

$$u_y = a_y + ib_y = i(\beta^2 + \alpha^2)$$

$$u_z = a_z + ib_z = \pm 2\alpha\beta \quad (\text{on choisit : } u_z = -2\alpha\beta)$$

Un $1/2$ -spineur (i.e. : spin $1/2$) dans \mathbb{R}^3 est un vecteur complexe isotrope écrit sous la forme d'un couple de deux nombres complexes appelé *spineur à deux composantes* $\begin{pmatrix} \beta \\ \alpha \end{pmatrix}$.

Un spineur dans \mathbb{R}^3 est unitaire si : $\beta\beta^* + \alpha\alpha^* = 1$.

Pour passer de (x, y, z) à (β, α) , on peut définir une application linéaire par une matrice complexe d'ordre 2 associée à \hat{U} :

$$\begin{cases} u_z\beta + (u_x - iu_y)\alpha = 0 \\ (u_x + iu_y)\beta - u_z\alpha = 0 \end{cases} \Leftrightarrow \begin{pmatrix} u_z & u_x - iu_y \\ u_x + iu_y & -u_z \end{pmatrix} \begin{pmatrix} \beta \\ \alpha \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

• Matrices de Pauli

On introduit trois matrices 2x2, dites "matrices de Pauli", complexes, de déterminant -1 et de trace nulle, formant une base orthonormée de l'espace des $1/2$ -spineurs dans \mathbb{R}^3 :

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} ; \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} ; \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} ; \quad \text{plus } \sigma_0 = 1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

dont les vecteurs propres sont respectivement $\begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix}$; $\begin{pmatrix} 1/\sqrt{2} \\ i/\sqrt{2} \end{pmatrix}$; $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ pour une

valeur propre unité. Ces matrices vérifient les relations (caractérisant une algèbre de Lie) :

$$[\sigma_i, \sigma_j]_- = \sigma_i \cdot \sigma_j - \sigma_j \cdot \sigma_i = 2i\sigma_k \quad (\text{commutation})$$

$$[\sigma_i, \sigma_j]_+ = \sigma_i \cdot \sigma_j + \sigma_j \cdot \sigma_i = 2\delta_{ij} \sigma_0 \quad (\text{anticommutation})$$

$$(\sigma_i)^2 = \sigma_0$$

• Isomorphisme entre $SU(2)$ et $SO(3)$

Les matrices $\sigma_x, \sigma_y, \sigma_z$ sont les générateurs du groupe des transformations d'un bit quantique. C'est le groupe $SU(2)$ des matrices 2x2 hermitiennes, isomorphe au groupe $SO(3)$ des rotations d'un solide dans l'espace euclidien à trois dimensions, par l'application de Clifford ϕ qui à tout vecteur \hat{U} associe la matrice :

$$\begin{pmatrix} u_z & u_x - iu_y \\ u_x + iu_y & -u_z \end{pmatrix} = u_x \sigma_x + u_y \sigma_y + u_z \sigma_z = \vec{\sigma} \cdot \vec{U} = \phi(\hat{U})$$

où $\hat{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$ est le "vecteur de spin" à composantes matricielles.

Considérons deux spineurs unitaires représentés par leurs vecteurs isotropes complexes \hat{U} et \hat{V} de composantes respectives (u_x, u_y, u_z) et (v_x, v_y, v_z) . Soit $\theta/2$ l'angle (\hat{U}, \hat{V}) . Soit \hat{n} le vecteur unitaire de composantes (n_x, n_y, n_z) de l'axe de la rotation $\mathcal{R}(\hat{n}, \theta)$ transformant \hat{U} en \hat{V} . Cette rotation est caractérisée par :

$$\text{- un produit scalaire : } \hat{U} \cdot \hat{V} = \cos \frac{\theta}{2} = u_x v_x + u_y v_y + u_z v_z$$

$$\text{- un produit vectoriel : } \hat{U} \times \hat{V} = \hat{n} \sin \frac{\theta}{2} = \begin{pmatrix} n_x \\ n_y \\ n_z \end{pmatrix} \sin \frac{\theta}{2} = \begin{vmatrix} \hat{n}_x & \hat{n}_y & \hat{n}_z \\ u_x & u_y & u_z \\ v_x & v_y & v_z \end{vmatrix} = \begin{pmatrix} u_y v_z - u_z v_y \\ u_z v_x - u_x v_z \\ u_x v_y - u_y v_x \end{pmatrix}$$

Effectuant le produit des matrices associées à \hat{U} et \hat{V} , on obtient d'après ces relations :

$$\begin{pmatrix} u_z & u_x - i u_y \\ u_x + i u_y & -u_z \end{pmatrix} \begin{pmatrix} v_z & v_x - i v_y \\ v_x + i v_y & -v_z \end{pmatrix} = \begin{pmatrix} \cos \frac{\theta}{2} + i n_z \sin \frac{\theta}{2} & i(n_x - i n_y) \sin \frac{\theta}{2} \\ i(n_x + i n_y) \sin \frac{\theta}{2} & \cos \frac{\theta}{2} - i n_z \sin \frac{\theta}{2} \end{pmatrix}$$

$$\Rightarrow (\hat{\sigma} \cdot \hat{U})(\hat{\sigma} \cdot \hat{V}) = \hat{U} \cdot \hat{V} + i \hat{\sigma}(\hat{U} \times \hat{V}) = \sigma_0 \cos \frac{\theta}{2} + i \hat{n} \hat{\sigma} \sin \frac{\theta}{2}$$

L'interprétation géométrique du produit de deux matrices associées à \hat{U} et \hat{V} est qu'elles représentent chacune une symétrie par rapport à un plan orienté. Le produit de ces deux symétries est une rotation $\mathcal{R}(\hat{n}, \theta)$, qui est donc représentée par l'opérateur :

$$\mathbf{R}(\hat{n}, \theta) = \sigma_0 \cos \frac{\theta}{2} + i \hat{n} \hat{\sigma} \sin \frac{\theta}{2} = e^{i \frac{\theta}{2} \hat{n} \hat{\sigma}}$$

En effet, en développant l'exponentielle, et en remarquant que $\sigma_i^2 = \sigma_0$, il vient :

$$e^{i \lambda \sigma_i} = \sum_k \frac{(i \lambda)^k}{k!} \sigma_i^k = \sum_{2p} \frac{(i \lambda)^{2p}}{(2p)!} \sigma_i^{2p} + \sum_{2p+1} \frac{(i \lambda)^{2p+1}}{(2p+1)!} \sigma_i^{2p+1} = \sigma_0 \cos \lambda + i \sigma_i \sin \lambda$$

On a donc par rapport à chaque axe Ox , Oy , Oz les rotations :

$$\mathbf{R}_x(\theta) = e^{i \frac{\theta}{2} \sigma_x} = \begin{pmatrix} \cos \frac{\theta}{2} & i \sin \frac{\theta}{2} \\ i \sin \frac{\theta}{2} & \cos \frac{\theta}{2} \end{pmatrix} \quad n_x = 1 \quad n_y = n_z = 0$$

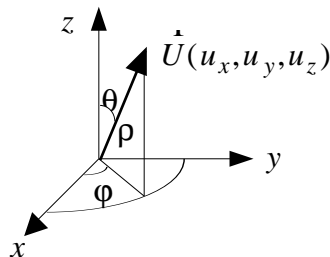
$$\mathbf{R}_y(\theta) = e^{i\frac{\theta}{2}\sigma_y} = \begin{pmatrix} \cos\frac{\theta}{2} & \sin\frac{\theta}{2} \\ -\sin\frac{\theta}{2} & \cos\frac{\theta}{2} \end{pmatrix} \quad n_y = 1 \quad n_x = n_z = 0$$

$$\mathbf{R}_z(\theta) = e^{i\frac{\theta}{2}\sigma_z} = \begin{pmatrix} e^{i\frac{\theta}{2}} & 0 \\ 0 & e^{-i\frac{\theta}{2}} \end{pmatrix} \quad n_z = 1 \quad n_x = n_y = 0$$

• *Représentation géométrique du bit quantique*

$\mathbf{R}(\hat{n}, \theta)$ est la forme la plus générale des matrices unitaires 2x2. Cela conduit à représenter le bit quantique comme l'état d'un objet de spin $1/2$ (un électron par exemple). Toute transformation unitaire agissant sur l'état (mis à part une rotation globale de phase) est une rotation du spin. Les vecteurs $|0\rangle$ et $|1\rangle$ sont les états de spin *up* ($|\uparrow\rangle$) et *down* ($|\downarrow\rangle$) le long d'un axe particulier, par exemple l'axe z . Les modules des deux nombres complexes a et b caractérisant le bit quantique décrivent l'orientation du spin dans l'espace à trois dimensions (l'angle polaire θ et l'azimuth φ).

Remarque : il faut imprimer deux "tours" à l'objet pour retrouver celui-ci dans son état initial, car $\mathbf{R}(\hat{n}, 2\pi) = -\sigma_0$ et $\mathbf{R}(\hat{n}, 4\pi) = \sigma_0$. La représentation par une figure géométrique du bit quantique est donc en toute rigueur impossible. Mais nous donnerons celle-ci à partir de la représentation usuelle d'un repère orthonormé de E^3 en coordonnées polaires (ρ , θ , φ).



• *Figure 4.2* : repère orthonormé

Une rotation d'un angle θ autour de l'axe z se traduit par :

$$\mathbf{R}(\hat{n}, \theta) = \begin{pmatrix} \cos\frac{\theta}{2} & e^{-i\varphi} \sin\frac{\theta}{2} \\ e^{i\varphi} \sin\frac{\theta}{2} & \cos\frac{\theta}{2} \end{pmatrix} \quad n_x \pm in_y = e^{\pm i\varphi} \quad n_z = 0$$

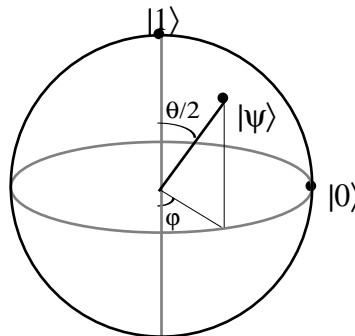
ce qui permet de calculer un état propre, avec le vecteur propre $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ de σ_z :

$$\mathbf{R}(\hat{n}, \theta) \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \cos \frac{\theta}{2} \\ e^{i\varphi} \sin \frac{\theta}{2} \end{pmatrix} = |\psi(\theta, \varphi)\rangle$$

L'état d'un bit quantique est donc représenté par la fonction d'onde :

$$|\psi(\theta, \varphi)\rangle = \cos \frac{\theta}{2} |0\rangle + e^{i\varphi} \sin \frac{\theta}{2} |1\rangle \text{ avec } |0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}; |1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

Sur la boule unité (dite "sphère de Bloch"), sa représentation est :



• Figure 4.3 : sphère de représentation de l'état d'un bit quantique



Le symbole binaire quantique est donc matérialisé par un état physique $|\psi\rangle$ qui, bien qu'unitaire, non sécable, "atomique", est beaucoup plus élaboré qu'un état binaire ordinaire (par ex. ouvert/fermé, 0/5V, ...), même si, au bout du compte, la mesure de $|\psi\rangle$ ramène au cas classique.

Conséquence

La bibliothèque Ξ d'une source binaire quantique Σ se réduit à $\{|0\rangle, |1\rangle\}$ si le bit quantique est dans un état simple. Mais s'il est dans un état superposé, on ne peut plus faire la distinction a priori entre les deux symboles qui composent la bibliothèque. Celle-ci est constituée du continuum $\{|\psi(\theta, \varphi)\rangle = \alpha|0\rangle + \beta|1\rangle\}$ qui ne devient séparable, comme on va le voir, que par le biais d'une opération de mesure.

S'il y a une mesure, on a par exemple :

$$\Xi = \{|0\rangle, |1\rangle\} \text{ avec } p(|0\rangle) = \cos^2 \frac{\theta}{2}; p(|1\rangle) = \sin^2 \frac{\theta}{2}; p(|0\rangle) + p(|1\rangle) = 1$$

Alors que dans le cas classique on a :

$$\Xi = \{0, 1\} \text{ avec } p(0) + p(1) = 1$$

La différence est évidente : la structure des probabilités quantiques est connue, prévisible, grâce à la connaissance de l'existence du paramètre θ (mais pas nécessairement la connaissance de la valeur de ce paramètre), alors que la connaissance que l'on peut avoir des probabilités classiques est soit une connaissance a priori (par exemple $p(0) = p(1) = 1/2$) soit une connaissance statistique.

Une mesure reflète en quelque sorte "le point de vue du récepteur", comme évoqué au §3.6. Mais en amont de toute mesure, "le point de vue de l'émetteur" ne considère que la bibliothèque :

$$\Xi = \{|\psi\rangle\}$$

Ceci permet de préciser la notion d'opérateur de distinction appliqué à un message de longueur unité : la distinction entre OUI et NON n'a de sens qu'au niveau du récepteur.

4.1.3. Paires corrélées de bits quantiques

4.1.3.1. États quantiques d'une paire corrélée. Base de Bell.

Nous nous placerons ici d'un point de vue algébrique, préférentiellement aux aspects purement physiques de cette question.

On considère deux bits quantiques dont les bases orthonormées sont respectivement $\{|0\rangle_A, |1\rangle_A\}$ et $\{|0\rangle_B, |1\rangle_B\}$, et soit $|\psi\rangle_{AB}$ l'état quantique de la paire AB (\otimes : produit tensoriel. cf § 4.2.1) :

$$|\psi\rangle_{AB} = a |0\rangle_A \otimes |0\rangle_B + b |1\rangle_A \otimes |1\rangle_B$$

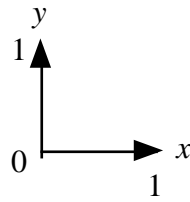
On dit que la corrélation entre A et B est maximum quand les amplitudes de probabilités sont telles que $|a|^2 = |b|^2 = 1$, soit $a = \pm 1$, $b = \pm 1$. On écrit :

$$|\psi\rangle_{AB} = \frac{1}{\sqrt{2}}(|00\rangle \pm |11\rangle)$$

Le facteur $1/\sqrt{2}$ assure la normalisation du vecteur. On choisit $a = 1$, le vecteur étant défini à une rotation de phase globale près ($= \pi$). On note :

$$|\Phi^+\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$$

Dans l'espace $\mathcal{H}_{A \text{ ou } B} = \mathbb{C}^2$ de description de chaque bit, la base orthonormée $\{|0\rangle, |1\rangle\}$ correspond aux matrices :



$$|0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad |1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

Dans l'espace $\mathcal{H}_{AB} = \mathbb{C}^4$ de description de la paire, il vient :

$$|00\rangle + |11\rangle = |0\rangle_A \otimes |0\rangle_B + |1\rangle_A \otimes |1\rangle_B = \begin{pmatrix} 1 \\ 0 \end{pmatrix}_A \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix}_B + \begin{pmatrix} 0 \\ 1 \end{pmatrix}_A \otimes \begin{pmatrix} 0 \\ 1 \end{pmatrix}_B = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

$$\Rightarrow |\Phi^+\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

$|\Phi^+\rangle$ est un des quatre vecteurs de la base (appelée *base de Bell*) des quatre états mutuellement orthogonaux d'une paire de bits quantiques, corrélés de façon maximale. On accède aux autres états en effectuant une des quatre transformations unitaires possibles (éventuellement à une rotation globale de phase près) sur un des éléments de la paire, par exemple B :

(i) $\mathbf{1}$ (ne rien faire)

$$|\Phi^+\rangle \rightarrow |\Phi^+\rangle$$

(ii) σ_x , équivalent à une rotation de 180° autour de l'axe Ox , ou encore à l'opérateur NON

:

$$\sigma_x|0\rangle = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} = |1\rangle \quad \text{et} \quad \sigma_x|1\rangle = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} = |0\rangle$$

$$|\Phi^+\rangle \rightarrow |\Psi^+\rangle = \frac{1}{\sqrt{2}}(|01\rangle + |10\rangle) = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix}$$

(iii) σ_y , équivalent à une rotation de 180° autour de l'axe Oy , ou encore, à $-\pi/2$ près,

$-i\sigma_y$:

$$-i\sigma_y|0\rangle = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} = |1\rangle \quad \text{et} \quad -i\sigma_y|1\rangle = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -1 \\ 0 \end{pmatrix} = -|0\rangle$$

$$|\Phi^+\rangle \rightarrow |\Psi^-\rangle = \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle) = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \\ -1 \\ 0 \end{pmatrix}$$

(iv) σ_z , équivalent à une rotation de 180° autour de l'axe Oz :

$$\sigma_z|0\rangle = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} = |0\rangle \quad \text{et} \quad \sigma_z|1\rangle = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ -1 \end{pmatrix} = -|1\rangle$$

$$|\Phi^+\rangle \rightarrow |\Phi^-\rangle = \frac{1}{\sqrt{2}}(|00\rangle - |11\rangle) = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ 0 \\ -1 \end{pmatrix}$$

4.1.3.2. Création d'une paire corrélée

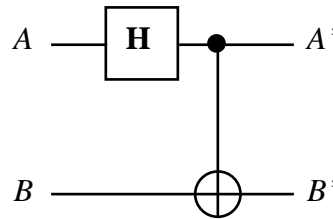
On veut créer les états corrélés $|*A*B\rangle \pm |*A*B\rangle$ (avec "*" = 0 ou 1) à partir de deux états séparés $|*A*B\rangle$. Un moyen classique d'y parvenir est de construire un dispositif dont la "table de vérité" est la suivante :

00	$\frac{1}{\sqrt{2}}(00\rangle + 11\rangle) = \Phi^+\rangle$
01	$\frac{1}{\sqrt{2}}(01\rangle - 10\rangle) = \Psi^-\rangle$
10	$\frac{1}{\sqrt{2}}(00\rangle - 11\rangle) = \Phi^-\rangle$
11	$\frac{1}{\sqrt{2}}(01\rangle + 10\rangle) = \Psi^+\rangle$

• Tableau 4.1 : "table de vérité" d'un circuit quantique corrélé.

Cette table est usuellement obtenue à l'aide de deux transformations, selon un schéma de calcul que nous reprendrons en section 6. Ce calcul consiste à appliquer au bit A de la paire AB , une première transformation unitaire, la transformation de Hadamard \mathbf{H} . Puis à l'ensemble une transformation de type "OU Exclusif" (**CNOT**, pour *Controlled Not*).

Conventionnellement, on représente ce schéma de calcul ainsi :



• Figure 4.4 : circuit quantique pour la création d'une paire EPR

avec :

$$\mathbf{H} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} = \frac{1}{\sqrt{2}} (\sigma_x + \sigma_z)$$

On vérifie que cette transformation est une rotation d'un angle $\theta = \pi$ autour de l'axe

$\vec{n} = \frac{1}{\sqrt{2}} (\vec{n}_x + \vec{n}_z)$ qui fait passer de l'axe Ox à l'axe Oz et vice-versa, à une rotation globale de

phase près :

$$\mathbf{R}(\vec{n}, \pi) = \sigma_0 \cos \frac{\pi}{2} + i \vec{n} \cdot \vec{\sigma} \sin \frac{\pi}{2} = i \frac{1}{\sqrt{2}} (\sigma_x + \sigma_z) = i\mathbf{H}$$

Quant à l'opérateur OU Exclusif, noté $a \oplus b$, il s'agit de l'addition modulo 2 classique.

On a :

$$\mathbf{CNOT} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

avec $\mathbf{CNOT}^2 = \mathbf{1}$. Le bit a est le bit de contrôle, b est la variable (*target bit*).

Matériellement, on peut implémenter un tel opérateur au moyen par exemple de manipulations de spectroscopie dites "de double résonance" [DiVincenzo, 1995].

Remarque

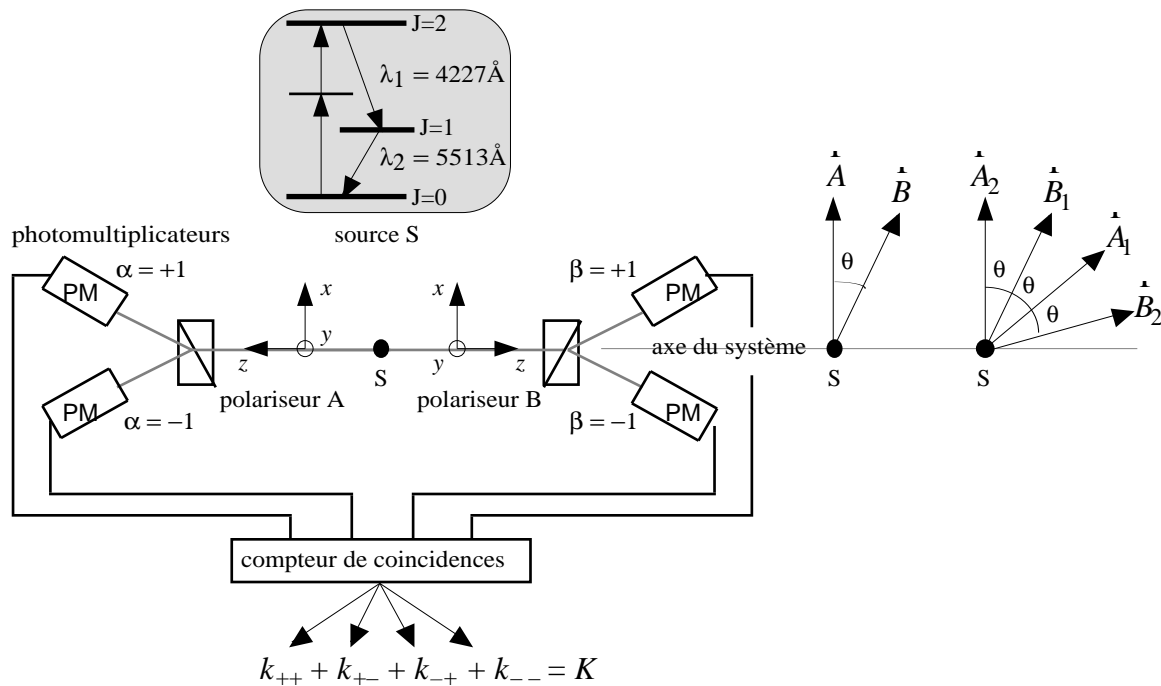
Toutes les opérations quantiques étant réversibles, le nombre de lignes entrantes dans le schéma de calcul est obligatoirement égal au nombre de lignes sortantes. Parmi les nombreuses et importantes conséquences que cela implique, il faut noter que la transformation ci-dessus est réversible, et que le schéma peut donc se lire de gauche à droite (création d'une paire corrélée) ou de droite à gauche (opération appelée "mesure de Bell", qui entraîne une rotation de la base corrélée $\{|\Phi^\pm\rangle, |\Psi^\pm\rangle\}$ vers la base séparée $\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$).

4.1.3.3. Exemple : photons corrélés

Sans les détailler, rappelons brièvement les conséquences de la corrélation d'une paire de photons d'après l'expérience bien connue de Aspect *et al* [Aspect,]

(remarque : s'agissant de particules de spin 1, les calculs diffèrent légèrement des calculs présentés ci-dessus).

Les détails techniques sont exposés dans la légende de la figure ci-dessous :



- *Figures 4.5.a* [d'après Ruhla, 1985] *4.5.b* *4.5.c*
- La figure 4.5.a est une vue perpendiculaire à l'axe du système. La source lumineuse est obtenue par l'excitation par laser et la désexcitation d'un atome de calcium qui émet deux photons. Soient \vec{A} et \vec{B} les axes d'analyse des détecteurs.
- La figure 4.5.b, qui est une vue dans l'axe, précise que les directions d'analyse de A et B font entre elles un angle θ .
- La figure 4.5.c, qui est également une vue dans l'axe, représente les choix expérimentaux faits par Aspect et son équipe : chaque détecteur a deux orientations possibles \vec{A}_1 et \vec{A}_2 d'une part, \vec{B}_1 et \vec{B}_2 d'autre part.
- Soient $\alpha_i = \pm 1$ ($i = 1, 2$) la réponse de A dans l'orientation \vec{A}_i et $\beta_j = \pm 1$ la réponse de B dans l'orientation \vec{B}_j .
- Sur K comptages simultanés pour chaque paire d'orientation (α_i, β_j) , en un temps donné, on calcule en fonction des fréquences k_{\pm} observées la valeur moyenne expérimentale $\langle \alpha_i \beta_j \rangle = (k_{++} - k_{+-} - k_{-+} + k_{--})/K = p_{++} - p_{+-} - p_{-+} + p_{--}$, qui est telle que : $-1 \leq \langle \alpha_i \beta_j \rangle \leq +1$.
- On en déduit la quantité $\langle \gamma \rangle = \langle \alpha_1 \beta_1 \rangle + \langle \alpha_1 \beta_2 \rangle + \langle \alpha_2 \beta_1 \rangle - \langle \alpha_2 \beta_2 \rangle$.

Théorèmes préliminaires :

- - Soient quatre nombres $\alpha_1, \alpha_2, \beta_1, \beta_2$ ne pouvant prendre chacun que deux valeurs $+1$ et -1 . Alors la quantité $\gamma = \alpha_1\beta_1 + \alpha_1\beta_2 + \alpha_2\beta_1 - \alpha_2\beta_2$ ne peut prendre que deux valeurs : $+2$ et -2 .

Une table de vérité à 16 entrées permet de vérifier facilement cette affirmation.

- Sur un grand nombre d'ensembles de valeurs $\alpha_1, \alpha_2, \beta_1, \beta_2$, on a nécessairement :

$$-2 \leq \langle \gamma \rangle = \langle \alpha_1\beta_1 \rangle + \langle \alpha_1\beta_2 \rangle + \langle \alpha_2\beta_1 \rangle - \langle \alpha_2\beta_2 \rangle \leq +2.$$

Cette relation est appelée "inégalité BCHSH" du nom de ses cinq auteurs Bell, Clauser, Horne, Shimony et Holt [Bell, 1964, Clauser *et al.*, 1969].



Hypothèse H_B : dans le cadre d'une théorie séparable, la réponse du détecteur A dans l'orientation $\overset{\cdot}{A}_1$ doit être indépendante de l'orientation $\overset{\cdot}{B}_1$ ou $\overset{\cdot}{B}_2$ du détecteur B.

Donc la réponse $\alpha'_1\beta_2$ qui correspondrait à $\overset{\cdot}{A}_1\overset{\cdot}{B}_2$ doit être la même que la réponse $\alpha_1\beta_1$ correspondant à $\overset{\cdot}{A}_1\overset{\cdot}{B}_1$: autrement dit, $\alpha'_1 \equiv \alpha_1$. Sur les huit possibilités (deux réponses possibles pour quatre paires d'orientations), il n'y a donc que quatre réponses distinctes $\alpha_1, \alpha_2, \beta_1, \beta_2$, chacune pouvant être égale à $+1$ ou -1 . D'où :

$$-2 \leq \gamma \leq +2$$

Le schéma de ce raisonnement est analogue à celui de la théorie statistique de l'information : d'une part le cadre de l'expérience est un cadre de pensée a priori, qui organise les variables en les classant séparément dans telle ou telle catégorie. Ensuite on affecte à ces variables des valeurs de probabilités mesurées a posteriori par le biais de fréquences constatées expérimentalement.

Mais dans le cadre de la mécanique quantique, il en va tout autrement. De façon similaire à ce qui a été constaté dans la conjonction de sources par transmission, la théorie quantique apporte ici un élément nouveau : elle précise d'emblée la valeur des probabilités car elle tient compte d'une variable supplémentaire ignorée du raisonnement précédent, l'angle θ . La catégorisation des éléments du réel est repoussée plus en amont, la connaissance de θ fait que l'on est en mesure de prédire la valeurs des probabilités et donc celle du paramètre γ par la fonction $\gamma(\theta)$. Détaillons ce calcul.

Dans le référentiel du laboratoire $Oxyz$, l'état du photon a traversant A est représenté par le vecteur :

$$|\Psi_A\rangle = \frac{1}{\sqrt{2}}(|x_A\rangle + |y_A\rangle)$$

Une mesure de la polarisation de a consiste à projeter ce vecteur sur la base $\{|X_A\rangle, |Y_A\rangle\}$ associée au détecteur A. φ_A est l'angle que font les deux repères (Ox, Oy) et (OX, OY) :

$$|x_A\rangle = \cos\varphi_A |X_A\rangle - \sin\varphi_A |Y_A\rangle$$

$$|y_A\rangle = \sin\varphi_A |X_A\rangle + \cos\varphi_A |Y_A\rangle$$

De même pour b . La paire de photons, préalablement à toute mesure, forme un tout inséparable représenté par :

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

Le calcul du vecteur d'état dans la représentation φ conduit à :

$$|\Phi\rangle = \frac{1}{\sqrt{2}} [\cos(\varphi_B - \varphi_A) |X_A, X_B\rangle - \sin(\varphi_B - \varphi_A) |X_A, Y_B\rangle + \sin(\varphi_B - \varphi_A) |Y_A, X_B\rangle + \cos(\varphi_B - \varphi_A) |Y_A, Y_B\rangle]$$

On en déduit par exemple la probabilité de détecter simultanément a polarisé sous l'angle φ_A et b polarisé sous l'angle φ_B :

$$p_{++} = 1/2 \cos^2(\varphi_B - \varphi_A) \quad (\text{réponses } |X_A\rangle, |X_B\rangle : \alpha = +1, \beta = +1 \Rightarrow \alpha\beta = +1)$$

De même :

$$p_{+-} = 1/2 \sin^2(\varphi_B - \varphi_A) \quad (\text{réponses } |X_A\rangle, |Y_B\rangle : \alpha = +1, \beta = -1 \Rightarrow \alpha\beta = -1)$$

etc. Il vient, en répétant le calcul pour chaque orientation :

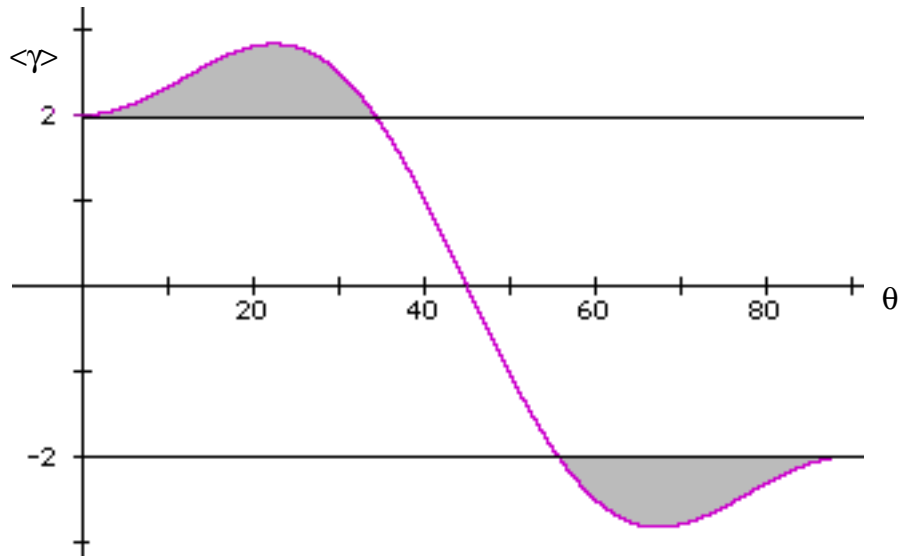
$$\langle \alpha\beta \rangle = p_{++} - p_{+-} - p_{-+} + p_{--} = \cos [2(\varphi_B - \varphi_A)]$$

$$\Rightarrow \langle \alpha_1\beta_1 \rangle = \langle \alpha_1\beta_2 \rangle = \langle \alpha_2\beta_1 \rangle = \cos 2\theta \quad \text{et} \quad \langle \alpha_2\beta_2 \rangle = \cos 6\theta$$

D'où :

$$\langle \gamma \rangle = \langle \alpha_1\beta_1 \rangle + \langle \alpha_1\beta_2 \rangle + \langle \alpha_2\beta_1 \rangle - \langle \alpha_2\beta_2 \rangle = 3 \cos 2\theta - \cos 6\theta$$

La courbe ci-dessous illustre les zones où il y a violation de l'inégalité de Bell (maximum pour $\theta = \pi/4 = 22,5^\circ$) :



- *Figure 4.6* : résultats de l'expérience d'Aspect. Les zones en grisé indiquent les valeurs de θ pour lesquelles la corrélation entre les photons a et b apparaît. Pour ces valeurs, l'hypothèse H_B (voir texte) est fausse.

4.2. Entropies quantiques d'une source unique

4.2.1. Matrices de densité

La théorie quantique imprime une deuxième particularité aux sources symboliques : de même qu'un symbole unique présente une structure "interne" complexe, un groupe de symboles est plus qu'une simple concaténation, car la "présence" d'un symbole ξ_2 associé au symbole ξ_1 modifient les propriétés de chacun d'eux.

Un bit quantique simple est un vecteur dans l'espace de Hilbert \mathbb{C}^2 . Pour caractériser les interactions de deux bits quantiques, il faut construire un nouvel espace de Hilbert.

Si \hat{U} et \hat{V} sont des vecteurs sur les bases $\{|i\rangle_U\} = \{a_1, \dots, a_i, \dots, a_n\}$, $\{|j\rangle_V\} = \{b_1, \dots, b_j, \dots, b_m\}$, le produit tensoriel $\hat{U} \otimes \hat{V}$ est un vecteur d'un espace de dimension nm dont les éléments $u \otimes v$ (appelés tenseurs élémentaires), bilinéaires, obéissent aux relations :

$$\alpha(u \otimes v) = \alpha u \otimes v = u \otimes \alpha v$$

$$u \otimes v + w \otimes v = (u+w) \otimes v \quad u \otimes v + u \otimes w = u \otimes (v+w)$$

Sur la base $a_i \otimes b_j$ le produit tensoriel est de la forme :

$$\hat{U} \otimes \hat{V} = \sum_{i,j} \alpha_{ij} a_i \otimes b_j \quad (1 \leq i \leq n, 1 \leq j \leq m)$$

Cette définition s'étend aux produits tensoriels comprenant plus de deux termes.

L'espace produit est aussi un espace de Hilbert qui hérite du produit scalaire $(u \otimes v)(u \otimes v)$.

Un système de deux bits quantiques est représenté par un vecteur unitaire dans l'espace de Hilbert $C^2 \otimes C^2$ (isomorphe à C^4) dont la base est :

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

Notation : on écrit souvent $|0\rangle|0\rangle$ ou $|00\rangle$ pour $|0\rangle \otimes |0\rangle$

Un message de n symboles est donc représenté dans un espace de Hilbert à 2^n dimensions formé à partir de n copies de C , alors que la représentation géométrique du même message dans le cas classique est un point d'un espace à n dimensions : on constate à nouveau que la symbolique quantique est beaucoup plus riche que son homologue classique.

En outre, ces n symboles quantiques ne sont pas séparés, indépendants les uns des autres comme dans le cas classique, mais forment un tout : ils sont *corrélés*. L'information portée par le message est *partagée* par l'ensemble des symboles qui le composent, même si ceux-ci ne sont pas physiquement réunis en un même lieu : par exemple $|\xi\rangle$ peut représenter un photon occupant la position A et $|\eta\rangle$ un autre photon situé en un point B distant de A : $|\xi\rangle|\eta\rangle$ représente l'état de la paire {photon en A et photon en B}.

Remarque : distinguer ce vecteur $|\xi\rangle|\eta\rangle$ du vecteur $a|\xi\rangle + b|\eta\rangle$, superposition linéaire des vecteurs représentant l'état d'un seul photon dont la position se répartit entre deux points A et B séparés.

On voit apparaître une différence fondamentale dans la structure des messages entre le cas classique et le cas quantique :

- Dans le cas classique, rappelons qu'une source Σ comprend une bibliothèque de symboles $\Xi = \{\chi_1, \dots, \chi_i, \dots, \chi_N\}$. Un message σ est une concaténation de signes $\xi_1, \dots, \xi_k, \dots, \xi_m$ où $\xi_k \in \Xi \forall k$. L'opération de concaténation est une règle minimale, qui ne suppose aucune connaissance particulière concernant la formation des messages. La théorie classique ne fait pas la distinction entre : a) un message $\sigma = \xi\eta$ formé par la concaténation de deux signes d'une même source dont l'entropie moyenne par symbole est H sh/symbole, soit une quantité d'information portée par σ égale à $\theta = 2H$ sh/message ; b) un message $\sigma = \xi\eta$ formé par la concaténation de deux signes appartenant à deux sources distinctes mais aux propriétés identiques (même bibliothèque, même entropie H sh/symbole) portant la quantité d'information totale $\theta = H+H$ sh/message. Autrement dit, le rôle de l'environnement sur la

structure supposée du message, rôle qui consiste à mesurer ce message, n'est pas traité par la théorie, qui ne dit rien des conventions implicites sous-jacentes à la notion d'information contenue dans un message.

- Dans le cas quantique, "un plus un état" n'est pas égal à "deux états", car, grossièrement parlant, l'addition de deux informations n'est plus une addition scalaire mais une addition vectorielle. L'ambiguïté sur le rôle de l'environnement de la source est interdite, car celui-ci intervient de façon explicite dans la structure du message à travers l'accord nécessaire entre la source et son environnement concernant la préparation des états symboliques. Deux cas de figure extrêmes se présentent, mais tous les cas intermédiaires sont possibles : a) on forme le message $\sigma = \xi\eta$ en 1°) préparant l'état $|\xi\rangle$, puis 2°) en faisant interagir $|\xi\rangle$ avec l'environnement (qui joue donc le rôle nécessaire parce qu'inévitable de deuxième source, réceptrice), mesure qui détruit la cohérence globale du système, puis 3°) en préparant l'état $|\eta\rangle$, 4°) suivi d'une seconde mesure. La concaténation est ici une succession temporelle d'événements indépendants. Ou bien, b), on forme le message $\sigma = \xi\eta$ sur une paire d'états corrélés $|\xi\rangle \otimes |\eta\rangle$, que l'environnement mesure ensuite globalement. En fait la présence de l'environnement influe directement sur la structure du message par le type de mesure qu'il est sensé effectuer sur ce dernier, le concept (implicite dans la théorie classique) de source isolée étant ici un concept vide de sens.

Nous allons envisager différents cas de figure, par ordre de complexité croissante :

(i) Soient deux bits quantiques A et B dont les bases orthonormées sont respectivement $\{|0\rangle_A, |1\rangle_A\}$ et $\{|0\rangle_B, |1\rangle_B\}$, et soit $|\psi\rangle_{AB}$ l'état quantique du message AB :

$$|\psi\rangle_{AB} = a |0\rangle_A \otimes |0\rangle_B + b |1\rangle_A \otimes |1\rangle_B$$

où a et b sont des amplitudes de probabilité. Une mesure de A selon une projection sur la base $\{|0\rangle_A, |1\rangle_A\}$ donne :

- le résultat $|0\rangle_A$ avec la probabilité $|a|^2$, en laissant le système dans l'état $|0\rangle_A \otimes |0\rangle_B$, ce qui implique que l'on trouvera $|0\rangle_B$ avec une probabilité de 1 si l'on mesure B après la mesure de A .

- le résultat $|1\rangle_A$ avec la probabilité $|b|^2$, en laissant le système dans l'état $|1\rangle_A \otimes |1\rangle_B$, ce qui implique que l'on trouvera $|1\rangle_B$ avec une probabilité de 1 si l'on mesure B après la mesure de A .

Donc A et B sont parfaitement corrélés dans l'état $|\psi\rangle_{AB}$. Le message AB se réduit au message A .

(ii) Soit une mesure effectuée sur A par l'opérateur auto-adjoint \mathbf{M}_A , associée à l'opérateur identité (matrice identité) \mathbf{I}_B affecté à B . Les vecteurs $|0\rangle_B$ et $|1\rangle_B$ étant orthogonaux, la valeur attendue de l'observable dans l'état normalisé $|\psi\rangle_{AB}$ est :

$$\begin{aligned} & {}_{AB}\langle\psi| \mathbf{M}_A \otimes \mathbf{I}_B |\psi\rangle_{AB} \\ &= a^* {}_A\langle 0| \otimes {}_B\langle 0| + b^* {}_A\langle 1| \otimes {}_B\langle 1| \quad (\mathbf{M}_A \otimes \mathbf{I}_B) \quad (a |0\rangle_A \otimes |0\rangle_B + b |1\rangle_A \otimes |1\rangle_B) \\ &= |a|^2 {}_A\langle 0| \mathbf{M}_A |0\rangle_A + |b|^2 {}_A\langle 1| \mathbf{M}_A |1\rangle_A \\ &= \text{tr}(\mathbf{M}_A \rho_A) \quad \text{avec} \quad \rho_A = |a|^2 |0\rangle_A \langle 0| + |b|^2 |1\rangle_A \langle 1| \end{aligned}$$

où $\text{tr}(\mathbf{M})$ est la trace de la matrice \mathbf{M} , et où $p_0 = |a|^2$ et $p_1 = |b|^2$ sont les probabilités respectives des états mélangés $|0\rangle_A$ et $|1\rangle_A$.

L'opérateur

$$\rho_A = \sum_i p(i) |\psi_i\rangle \langle \psi_i|$$

est appelé *opérateur de densité*, ou *matrice de densité* (de probabilité) du bit quantique A . On montre que cet opérateur est auto-adjoint, positif, de trace unité.

Exemple 1 :

□ Retour sur un bit quantique unique (voir paragraphe précédent) : le calcul de la densité de probabilité $\rho(\theta, \varphi)$ conduit au résultat suivant :

$$\rho(\theta, \varphi) = |\psi(\theta, \varphi)\rangle \langle \psi(\theta, \varphi)|$$

$$\rho(\theta, \varphi) = \begin{pmatrix} \cos^2 \frac{\theta}{2} & e^{-i\varphi} \cos \frac{\theta}{2} \sin \frac{\theta}{2} \\ e^{i\varphi} \cos \frac{\theta}{2} \sin \frac{\theta}{2} & \sin^2 \frac{\theta}{2} \end{pmatrix}$$

$$\rho(\theta, \varphi) = \frac{1}{2} \sigma_0 + \frac{1}{2} \begin{pmatrix} \cos \theta & e^{-i\varphi} \sin \theta \\ e^{i\varphi} \sin \theta & -\cos \theta \end{pmatrix}$$

$$\rho(\theta, \varphi) = \frac{1}{2} (\sigma_0 + \mathbf{n} \cdot \boldsymbol{\sigma})$$



Exemple 2 :

□ Appelons $p(i)$ les probabilités de N états mélangés $|\chi_i\rangle$, constituant une bibliothèque Ξ , espace de Hilbert à N dimensions. La matrice de densité qui représente ce mélange est :

$$\rho = \sum_{i=1}^N p(i) |\chi_i\rangle \langle \chi_i| \quad \text{avec} \quad \sum_{i=1}^N p(i) = 1$$

Ce mélange est la superposition non cohérente des états $|\chi_i\rangle$, et l'on a : $\rho^2 \neq \rho$.

(Remarque : dans le cas contraire, si les états ne sont pas mélangés (états "purs"), cette matrice se réduit au projecteur $\rho = \sum_i |\chi_i\rangle \langle \chi_i|$, qui est tel que, par définition d'une opération de projection sur un sous-espace : $\rho^2 = \rho$).

Effectuons une mesure (selon une projection dans Ξ) par :

$$\mathbf{M} = |\mu\rangle \langle \mu|$$

Alors :

$$\mathbf{M}\rho = \left(\sum_i p(i) |\chi_i\rangle \langle \chi_i| \right) |\mu\rangle \langle \mu|$$

$$\mathbf{M}\rho = \sum_i p(i) |\chi_i\rangle \langle \chi_i|\mu\rangle \langle \mu|$$

$$\mathbf{M}\rho = \sum_i p(i) \langle \chi_i|\mu\rangle |\chi_i\rangle \langle \mu|$$

$$\Rightarrow \text{tr}(\mathbf{M}\rho) = \sum_i p(i) \langle \chi_i|\mu\rangle \langle \mu|\chi_i\rangle$$

$$\text{tr}(\mathbf{M}\rho) = \sum_i p(i) |\langle \chi_i|\mu\rangle|^2$$

Cette expression contient un mélange de probabilités classiques $p(i)$ et de probabilités quantiques $q(i) = |\langle \chi_i|\mu\rangle|^2$.

□

(iii) Ce calcul se généralise une seconde fois à l'état $|\psi\rangle_{AB}$ d'un système de deux bits quantiques corrélés, où $\{|i\rangle_A\}$, $\{|j\rangle_B\}$, $\{|i\rangle_A \otimes |j\rangle_B\}$ sont respectivement les bases orthogonales des espaces de Hilbert \mathcal{H}_A , \mathcal{H}_B , $\mathcal{H}_A \otimes \mathcal{H}_B$:

$$|\psi\rangle_{AB} = \sum_{i,j} \alpha_{ij} |i\rangle_A \otimes |j\rangle_B \quad \text{avec} \quad \sum_{i,j} |\alpha_{ij}|^2 = 1$$

L'observable $\mathbf{M}_A \otimes \mathbf{I}_B$ agissant sur A a pour valeur :

$$\begin{aligned}
& {}_{AB}\langle \psi | \mathbf{M}_A \otimes \mathbf{I}_B | \psi \rangle_{AB} \\
&= \sum_{j,k} \alpha_{jk}^* ({}_A\langle j | \otimes {}_B\langle k |) (\mathbf{M}_A \otimes \mathbf{I}_B) \sum_{i,k} \alpha_{ik} (|i\rangle_A \otimes |k\rangle_B) \\
&= \sum_{i,j,k} \alpha_{jk}^* \alpha_{ik} {}_A\langle j | \mathbf{M}_A |i\rangle_A \\
&= \text{tr}(\mathbf{M}_A \rho_A) \\
&= \langle \mathbf{M}_A \rangle, \text{ valeur moyenne de l'opérateur } \mathbf{M}_A.
\end{aligned}$$

où

$$\rho_A = \text{tr}_B (|\psi_{AB}\rangle \langle \psi_{AB}|) = \sum_{i,j,k} \alpha_{ik} \alpha_{jk}^* |i\rangle_A \langle j|$$

Dans ce cas, on dit que l'opérateur de densité ρ_A est obtenu en calculant la trace partielle sur le sous-système B de la matrice de densité du système global AB, avec les propriétés :

- ρ_A est auto-adjoint : $\rho_A = \rho_A^\dagger$
- ρ_A est positive : ${}_A\langle \psi | \rho_A | \psi \rangle_A = \sum_j \left| \sum_i \alpha_{ik} {}_A\langle \psi | i \rangle_A \right|^2 \geq 0$
- $\text{tr}(\rho_A) = \sum_{i,j} |\alpha_{ij}|^2 = 1$

(iv) Enfin, la théorie de la matrice de densité se généralise à un ensemble d'opérateurs de mesure \mathbf{M}_a , qui peuvent être quelconques (ce ne sont pas de simples projections), mais doivent être positifs (i.e. dont les valeurs propres sont positives ou nulles) et tels que $\sum_a \mathbf{M}_a = \mathbf{1}_A$ pour que l'espace des résultats reste probabilisable (i.e. somme des probabilités = 1). Alors chaque résultat de mesure a apparaît avec la probabilité :

$$p(a) = \text{tr} \rho \mathbf{M}_a$$

Conclusions préliminaire :

- Une matrice de densité associe probabilités classiques et probabilités quantiques. Le formalisme quantique entraîne donc la nécessité de définir deux formes distinctes d'entropies :

- Entropie "classique quantique" H : on applique la théorie statistique classique à des sources constituées d'états quantiques mélangés, auxquels sont attribuées des probabilités classiques. On évalue des quantités d'information classiques sur des systèmes quantiques. Par exemple, on calcule le nombre de bits classiques transmis par un canal quantique.

- Entropie "quantique quantique" S , encore appelée "entropie de Von Neumann". Il s'agit alors d'un formalisme spécifiquement quantique. On évalue des quantités d'information

quantiques sur des systèmes quantiques. Par exemple, on calcule le nombre de bits quantiques transmis par un canal lui-même quantique.

- Dans l'absolu, un système isolé (par exemple un bit quantique), dont l'évolution temporelle est dirigée par l'équation de Schrödinger, est une forme (en ce sens que le bit quantique possède une structure interne, d'ailleurs complexe), mais n'est pas une information, ni un message. En pratique, "s'informer" sur cette source consiste à associer le système et son observateur, et c'est au niveau seulement de cette association que réside "l'information" .

Comme le montre l'exemple 2 du point (ii), lorsqu'un observateur O effectue une mesure sur un système A de probabilités $p(i)$, il réalise une sorte "d'échelon informationnel" qui fait passer instantanément du système A au système OA, ajoutant aux probabilités $p(i)$ de A seul les probabilités $q(i) = |\langle \chi_i | \mu \rangle|^2$ de OA. L'information contenue dans A est devenue l'information *partagée* par A étant donné O. Ce schéma se généralise à tout couple AB de systèmes A et B, ou l'influence (réciproque) exercée par B sur A est équivalent à une mesure de A par B : l'information contenue dans A change instantanément de nature en devenant l'information *partagée* par A étant donné B.

4.2.2. Entropie de Von Neumann

Considérons une source symbolique Σ préparant un message constitué de signes ξ pris dans une bibliothèque de symboles $\Xi = \{\chi_i\}$ formée d'une base orthonormée d'états quantiques $|\chi_i\rangle$, chacun apparaissant avec une probabilité $p(i)$. L'ensemble a une matrice de densité :

$$\rho = \sum_i p(i) |\chi_i\rangle \langle \chi_i|$$

La valeur moyenne attendue de l'application de tout observable \mathbf{M} agissant sur Σ est :

$$\langle \mathbf{M} \rangle = \text{tr} (\mathbf{M}\rho) = \sum_i p(i) \langle \chi_i | \mathbf{M} | \chi_i \rangle$$

L'entropie (classique) de l'ensemble est :

$$H(\xi) = - \text{tr} (\rho \log \rho)$$

Etendue au cas (non classique) où les états ne commutent pas, cette quantité est appelée *entropie de Von Neumann* :

$$S(\rho) = - \text{tr} (\rho \log \rho)$$

Soit un message σ formé de m signes ξ_k choisis dans la bibliothèque Ξ . Le message a la

matrice de densité :

$$\rho^m = \rho \otimes \dots \otimes \rho$$

Soit $\dim \mathcal{H}$ le nombre *minimum* de dimensions que doit posséder l'espace de Hilbert \mathcal{H} dans lequel le message est produit. On montre [Schumacher, 1995] que :

$$\log (\dim \mathcal{H}) = m S(\rho)$$

Ce résultat, qui est le pendant quantique du premier théorème de Shannon, montre que l'entropie de Von Neumann est le nombre moyen de bits porté par chaque état du système. Encore faudrait-il faire dans le cas quantique la même distinction que celle qui est faite dans le cas classique entre "shannon" (unité de mesure d'une quantité d'information) et "bit" (support symbolique binaire de cet information) : le bit *quantique* est le support binaire (dans le sens où il est décrit dans \mathbb{C}^2 , espace de Hilbert à deux dimensions) d'une quantité d'information *classique* qui s'exprime comme il se doit en shannons/états.

4.2.3. Propriétés de l'entropie de Von Neumann

Les preuves des principales propriétés de S sont données dans [Wehrl, 1978] :

(i) En l'absence de superposition :

$$\rho = |\xi\rangle \langle \xi| \Rightarrow S(\rho) = 0$$

(ii) S invariante dans un changement de base (S ne dépend que des valeurs propres de ρ):

$$S(\mathbf{U}\rho\mathbf{U}^{-1}) = S(\rho)$$

(iii) Maximum :

$$S(\rho) \leq \log N$$

(si ρ a N valeurs propres non dégénérées)

(iv) Concavité : pour $\lambda_i \geq 0 \forall i$ et $\sum_{i=1}^n \lambda_i = 1$, la concavité de la fonction $-x \log x$ entraîne :

$$S\left(\sum_{k=1}^n \lambda_k \rho_k\right) \geq \sum_{k=1}^n \lambda_k S(\rho_k)$$

Comme son homologue classique, l'entropie augmente si l'ignorance de l'observateur est

plus grande concernant la façon avec laquelle les états ont été préparés.

(v) Entropie d'une mesure : Si dans l'état ρ on mesure l'observable

$$\mathbf{M} = \sum_i |\chi_i\rangle \chi_i \langle \chi_i|,$$

alors le symbole χ_i est observé avec une probabilité :

$$p(i) = \langle \chi_i | \rho | \chi_i \rangle.$$

L'entropie $H(\eta)$, avec $\eta \in \{\chi_i\}$, est telle que

$$H(\eta) \geq S(\rho)$$

l'égalité étant vraie si \mathbf{M} et ρ commutent (ce qui est le cas d'un "bon" observable)

(vi) Entropie d'une source symbolique : A partir d'un ensemble $\Xi = \{\xi : |\chi_i\rangle, p(i), i = 1 \dots N\}$ dont la matrice de densité est

$$\rho = \sum_i p(i) |\chi_i\rangle \langle \chi_i|,$$

on a :

$$H(\xi) \geq S(\rho)$$

l'égalité étant vraie si les états $|\chi_i\rangle$ sont orthogonaux, ce qui est une condition de discernabilité.

Bien que les points (v) et (vi) soient similaires dans leurs résultats, la théorie quantique apporte une précision importante concernant le concept de source : alors que dans le cas classique il n'y a pas lieu de distinguer entre source émettrice et source réceptrice (les calculs de probabilités se font de la même façon dans les deux cas), il n'en est plus de même au niveau quantique : la source et la mesure de la source font appel à deux calculs différents.

(vii) Sous-additivité :

$$S(\rho_{\xi\eta}) \leq S(\rho_\xi) + S(\rho_\eta)$$

$$\text{où } \rho_\xi = \text{tr}_\eta \rho_{\xi\eta} \text{ et } \rho_\eta = \text{tr}_\xi \rho_{\xi\eta}$$

l'égalité étant vraie si $\rho_{\xi\eta} = \rho_\xi \otimes \rho_\eta$ (cas de systèmes non corrélés)

(viii) Sous-additivité "forte"

$$S(\rho_{\xi\eta\gamma}) + S(\rho_\eta) \leq S(\rho_{\xi\eta}) + S(\rho_{\eta\gamma})$$

Bien que triviale dans le cas classique, la sous-additivité "forte" est délicate à démontrer dans le cas quantique. Elle est toutefois d'une grande importance en ce qui concerne les propriétés de l'information quantique.

(ix) Inégalité du triangle

$$S(\rho_{\xi\eta}) \geq |S(\rho_{\xi}) - S(\rho_{\eta})|$$

Autrement dit : dans le cas classique, l'entropie globale d'un système est supérieure à l'entropie de chacune de ses parties. Ce n'est plus vrai dans le cas quantique. A la limite, dans le cas d'un système de deux états corrélés purs tels que $S(\rho_{\xi}) = S(\rho_{\eta})$, on a $S(\rho_{\xi\eta}) = 0$. Nous obtenons un critère fort de distinction entre ces deux types informationnels (échangé *versus* partagé) : le message quantique composite a bien été préparé de façon définie, mais sa mesure donne des résultats imprédictibles (de probabilité $1/2$). On ne peut observer les deux sous-systèmes (c'est-à-dire les deux symboles composant le message) séparément pour en déduire leurs états respectifs de préparation initiale : l'information est partagée dans la corrélation quantique (non-locale) entre les deux signes ξ et η .

4.2.4. Théorème de Holevo

Considérons un ensemble d'états pouvant être employés pour préparer un bit quantique unique, dont toute l'information doit pouvoir être contenue dans un espace de Hilbert bi-dimensionnel. Les états constituent une bibliothèque symbolique $|\chi_1\rangle, |\chi_2\rangle, \dots, |\chi_i\rangle, \dots, |\chi_N\rangle$ de probabilités $p(i)$ et sont mélangés au sein d'un message de telle façon que :

-la matrice de densité de chaque symbole est :

$$\rho_i = |\chi_i\rangle\langle\chi_i| \text{ avec } \text{tr} \rho_i = 1$$

-la matrice de densité de la source est :

$$\rho = \sum_i p(i) \rho_i$$

De la relation :

$$S(\rho) = -\text{tr} (\rho \log \rho) = -\sum_i \lambda_i \log \lambda_i$$

où les λ_i sont les valeurs propres de ρ , on tire, sachant que $\text{tr} \rho_i = 1 \quad \forall i$:

$$S(\rho) = -\sum_i \text{tr} (p(i)\rho_i) \log (p(i)\rho_i)$$

$$S(\rho) = -\sum_i p(i) \log p(i) - \sum_i p(i) \text{tr} \rho_i \log \rho_i$$

$$S(\rho) = I_H(\xi) + \sum_i p(i) S(\rho_i)$$

La quantité $I_H(\xi)$ est appelée *information de Holevo*. Cette quantité, valable pour un ensemble d'états mélangés mutuellement orthogonaux, a la même forme qu'une quantité mutuelle d'information ($I_H(\xi)$ se réduisant à $S(\rho)$ pour un ensemble d'états non mélangés):

$$I_H(\xi:\eta) = H(\xi) - H(\xi|\eta) = S(\rho) - \sum_i p(i) S(\rho_i)$$

De même que l'information mutuelle classique exprime que l'entropie sur ξ est diminuée de la connaissance que l'on a sur η , l'information de Holevo exprime que l'entropie de Von Neumann sur un ensemble d'états est réduite *si l'on connaît le mode de préparation du mélange*.

Comme son homologue classique, l'information de Holevo est toujours positive ou nulle : ceci est une conséquence de la concavité de $S(\rho)$ (propriété iv du § précédent). Mais on montre que cette quantité représente une limite (appelée *limite de Holevo*) : dans le cas d'états mélangés non orthogonaux, l'information que l'on peut retirer de la mesure d'un bit quantique concernant l'identité de ces états est toujours moindre :

$$H(\xi:\eta) \leq I_H(\xi:\eta)$$

Nous examinerons dans le paragraphe 4.4 un exemple illustrant l'existence de cette limite de la capacité classique d'un canal quantique.

4.3. Sources symboliques quantiques : entropie classique quantique

4.3.1. Exemple

Cet exemple est adapté d'après Preskill et Landahl [Preskill, Landahl, chap.5, 1998].

On considère deux sources binaires Σ et Γ , respectivement émetteur et récepteur, et un message échangé par ces deux sources constitué d'un unique bit quantique. La bibliothèque de Σ est constituée de deux états non orthogonaux préparés de façon équiprobable :

$$\xi \in \Xi = \{\chi_1 = |0\rangle, \chi_2 = |1\rangle\}, \quad \text{avec :}$$

$$|0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} ; \quad |1\rangle = \begin{pmatrix} \cos \frac{\theta}{2} \\ \sin \frac{\theta}{2} \end{pmatrix} \quad \text{avec} \quad 0 < \theta < \pi \quad \text{et} \quad p(0) = p(1) = \frac{1}{2}$$

La lecture du message par Γ est effectuée en appliquant différents opérateurs de mesure

notés $\mathbf{M}_j, j = 1, 2, 3, \dots$. La bibliothèque de Γ est l'ensemble *dénombrable* de ces opérateurs :

$$\eta \in \Pi = \{ \pi_1 = \mathbf{M}_1, \pi_2 = \mathbf{M}_2, \dots, \pi_j = \mathbf{M}_j, \dots \}$$

On note respectivement $H(\xi)$ et $H(\eta)$ les entropies des deux sources, $p(\eta|\xi)$ les différentes probabilités conditionnelles, calculées par :

$$p(j|\xi) = | \langle \xi | \mathbf{M}_j | \xi \rangle |$$

On calcule la quantité d'information échangée par une des relations habituelles :

$$H(\xi:\eta) = H(\eta) - H(\eta|\xi)$$

(On pourrait tout aussi bien utiliser l'autre forme : $H(\xi:\eta) = H(\xi) - H(\xi|\eta)$). Les entropies sont calculées par :

$$H(\eta) = - \sum_j p(j) \log p(j) \quad \text{avec} \quad p(j) = p(0) \cdot p(j|0) + p(1) \cdot p(j|1)$$

$$H(\eta|\xi) = - p(0) \sum_j p(j|0) \log p(j|0) - p(1) \sum_j p(j|1) \log p(j|1)$$

On note H_2 la fonction $-x \log x - (1-x) \log (1-x)$.

(i) Mesure orthogonale, avec :

$$\mathbf{M}_1 = |0\rangle \langle 0| = \begin{pmatrix} 1 & \\ & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \quad (\text{réponse } 0)$$

$$\mathbf{M}_2 = \mathbf{1} - |0\rangle \langle 0| = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \quad (\text{réponse } \neg 0)$$

On trouve :

$$p(1|0) = \begin{pmatrix} 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = 1 \quad p(1|1) = \begin{pmatrix} \cos \frac{\theta}{2} & \sin \frac{\theta}{2} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \cos \frac{\theta}{2} \\ \sin \frac{\theta}{2} \end{pmatrix} = \cos^2 \frac{\theta}{2}$$

$$p(2|0) = \begin{pmatrix} 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = 0 \quad p(2|1) = \begin{pmatrix} \cos \frac{\theta}{2} & \sin \frac{\theta}{2} \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \cos \frac{\theta}{2} \\ \sin \frac{\theta}{2} \end{pmatrix} = \sin^2 \frac{\theta}{2}$$

$$H(\eta) = -\frac{1}{2} \left(1 + \cos^2 \frac{\theta}{2} \right) \log \left(\frac{1}{2} \left(1 + \cos^2 \frac{\theta}{2} \right) \right) - \frac{1}{2} \sin^2 \frac{\theta}{2} \log \left(\frac{1}{2} \sin^2 \frac{\theta}{2} \right)$$

$$H(\eta|\xi) = -\frac{1}{2} \cos^2 \frac{\theta}{2} \log \left(\cos^2 \frac{\theta}{2} \right) - \frac{1}{2} \sin^2 \frac{\theta}{2} \log \left(\sin^2 \frac{\theta}{2} \right)$$

$$H(\eta; \xi) = 1 - \frac{1}{2} \left(1 + \cos^2 \frac{\theta}{2} \right) H_2 \left(\frac{1}{1 + \cos^2 \frac{\theta}{2}} \right)$$

(ii) Mesure "généralisée" à trois opérateurs positifs $\mathbf{M}_{1,2,3}$ tels que $\mathbf{M}_1 + \mathbf{M}_2 + \mathbf{M}_3 = \mathbf{1}$:

$$\mathbf{M}_1 = \alpha (\mathbf{1} - |0\rangle \langle 0|) \quad (\text{réponse } \neg 0)$$

$$\mathbf{M}_2 = \alpha (\mathbf{1} - |1\rangle \langle 1|) \quad (\text{réponse } \neg 1)$$

$$\mathbf{M}_3 = (1 - 2\alpha) \mathbf{1} + \alpha (|0\rangle \langle 0| + |1\rangle \langle 1|) \quad (\text{réponse vide})$$

α est le plus grand réel positif tel que \mathbf{M}_3 soit positif. Sa valeur est donnée en cherchant les valeurs propres positives de \mathbf{M}_3 , par l'équation caractéristique :

$$\det(\mathbf{M}_3 - \lambda \mathbf{1}) = 0 \Leftrightarrow \lambda^2 - \lambda \operatorname{tr} \mathbf{M}_3 + \det \mathbf{M}_3 = 0$$

$$\Rightarrow \lambda = 1 - \alpha \pm \alpha \cos \frac{\theta}{2}$$

$$\lambda \geq 0 \Rightarrow \alpha \leq \frac{1}{1 \pm \cos \frac{\theta}{2}}$$

Pour $0 < \theta < \pi$, la plus grande valeur de α compatible avec des valeurs propres positives est donc $\alpha = \frac{1}{1 + \cos \frac{\theta}{2}}$.

On trouve :

$$p(1|0) = 0 \quad p(1|1) = \alpha \sin^2 \frac{\theta}{2}$$

$$p(2|0) = \alpha \sin^2 \frac{\theta}{2} \quad p(2|1) = 0$$

$$p(3|0) = 1 - \alpha + \alpha \cos^2 \frac{\theta}{2} \quad p(3|1) = 1 - \alpha + \alpha \cos^2 \frac{\theta}{2}$$

$$H(\eta) = -\alpha \sin^2 \frac{\theta}{2} \log \left(\frac{1}{2} \alpha \sin^2 \frac{\theta}{2} \right) - \left(1 - \alpha + \alpha \cos^2 \frac{\theta}{2} \right) \log \left(1 - \alpha + \alpha \cos^2 \frac{\theta}{2} \right)$$

$$H(\eta|\xi) = -\alpha \sin^2 \frac{\theta}{2} \log \left(\alpha \sin^2 \frac{\theta}{2} \right) - \left(1 - \alpha + \alpha \cos^2 \frac{\theta}{2} \right) \log \left(1 - \alpha + \alpha \cos^2 \frac{\theta}{2} \right)$$

$$H(\eta; \xi) = \alpha \sin^2 \frac{\theta}{2} = 2 \sin^2 \frac{\theta}{4} = 1 - \cos \frac{\theta}{2}$$

(iii) Mesure orthogonale par projection par rapport aux bissectrices de l'angle $(|0\rangle, |1\rangle)$.

$$\mathbf{M}_1 = |\chi\rangle \langle \chi|$$

$$\mathbf{M}_2 = \mathbf{1} - |\chi\rangle \langle \chi|$$

$$|\chi\rangle = \begin{pmatrix} \cos\left[\frac{1}{2}\left(\frac{\theta}{2} + \frac{\pi}{2}\right)\right] \\ \sin\left[\frac{1}{2}\left(\frac{\theta}{2} + \frac{\pi}{2}\right)\right] \end{pmatrix} = \begin{pmatrix} \cos\left[\frac{\theta + \pi}{4}\right] \\ \sin\left[\frac{\theta + \pi}{4}\right] \end{pmatrix}$$

On trouve :

$$p(1|0) = \cos^2 \frac{\theta + \pi}{4}$$

$$p(1|1) = \sin^2 \frac{\theta + \pi}{4}$$

$$p(2|0) = \sin^2 \frac{\theta + \pi}{4}$$

$$p(2|1) = \cos^2 \frac{\theta + \pi}{4}$$

$H(\eta) = 1$ (\Rightarrow cette méthode de mesure maximalise l'information contenue dans η)

$$H(\eta|\xi) = -\sin^2 \frac{\theta + \pi}{4} \log\left(\sin^2 \frac{\theta + \pi}{4}\right) - \cos^2 \frac{\theta + \pi}{4} \log\left(\cos^2 \frac{\theta + \pi}{4}\right)$$

$$H(\eta:\xi) = 1 - H_2\left(\cos^2 \frac{\theta + \pi}{4}\right)$$

(iv) Enfin, on peut calculer la limite de Holevo, plus grande quantité d'information qu'il est possible de transmettre par le canal de ce bit quantique. Puisque les états de la source Σ ne sont pas mélangés, la limite de Holevo se réduit ici à $S(\rho)$.

On calcule la matrice de densité :

$$\rho = \frac{1}{2} |0\rangle \langle 0| + \frac{1}{2} |1\rangle \langle 1| = \frac{1}{2} \begin{pmatrix} 1 + \cos^2 \frac{\theta}{2} & \cos \frac{\theta}{2} \sin \frac{\theta}{2} \\ \cos \frac{\theta}{2} \sin \frac{\theta}{2} & \sin^2 \frac{\theta}{2} \end{pmatrix}$$

dont les valeurs propres sont obtenues par l'équation caractéristique :

$$\det(\rho - \lambda \mathbf{1}) = 0 \Leftrightarrow \lambda^2 - \lambda \operatorname{tr} \rho + \det \rho = 0$$

$$\Leftrightarrow \lambda^2 - \lambda + \frac{1}{4} \left(1 + \cos^2 \frac{\theta}{2}\right) \sin^2 \frac{\theta}{2} - \frac{1}{4} \cos^2 \frac{\theta}{2} \sin^2 \frac{\theta}{2} = 0$$

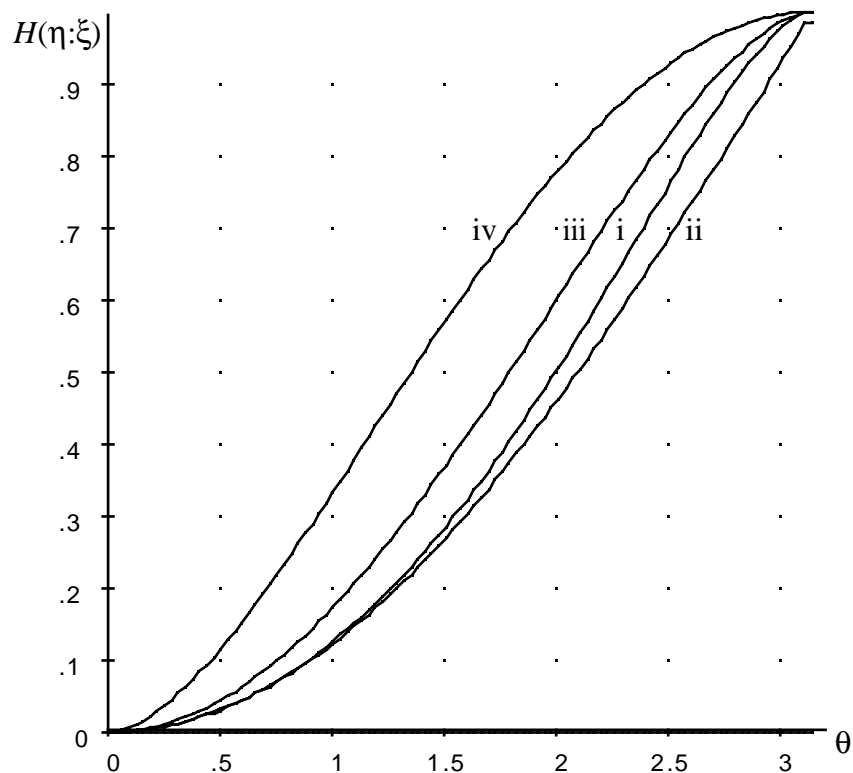
$$\Leftrightarrow \lambda = \frac{1}{2} \pm \frac{1}{2} \sqrt{1 - \sin^2 \frac{\theta}{2}} = \frac{1}{2} \pm \cos \frac{\theta}{2} = \sin^2 \frac{\theta}{4}, \quad \cos^2 \frac{\theta}{4}$$

ce qui permet de calculer ρ sous forme diagonalisée, et donc $S(\rho)$:

$$\rho = \begin{pmatrix} \sin^2 \frac{\theta}{2} & 0 \\ 0 & \cos^2 \frac{\theta}{2} \end{pmatrix}$$

$$I_H(\eta;\xi) = S(\rho) = -\text{tr}(\rho \log \rho) = -\sin^2 \frac{\theta}{4} \log \sin^2 \frac{\theta}{4} - \cos^2 \frac{\theta}{4} \log \cos^2 \frac{\theta}{4} = H_2\left(\cos^2 \frac{\theta}{4}\right)$$

(v) On peut tracer l'évolution de l'information mutuelle en fonction de θ , selon les cas i, ii, iii et iv qui viennent d'être détaillés. On constate que la méthode iii s'approche le plus de la limite de Holevo.



- Figure 4.7 : évolution de $H(\eta;\xi)$ en fonction de $\theta \in [0, \pi]$. La courbe iv est la courbe limite de Holevo. Elle correspond à la courbe $p=q$ de la figure 3.4 pour p variant de $1/2$ à 1.
- On voit que pour s'approcher le plus possible de la limite de Holevo il faut choisir la méthode iii.

En comparant la courbe iv de la figure 4.7 et la courbe $p=q$ de la figure 3.4 (canal binaire symétrique), on constate que la théorie quantique complète le raisonnement classique : dans cet exemple, dans le résultat $H(x:y)$, où x est la variable et H la quantité d'information mutuelle, la variable x n'est plus une probabilité, mais un angle.

Face à un couple de deux sources dont l'observateur cherche à connaître les lois de probabilités pour en déduire les diverses quantités d'information, celui-ci n'a le choix, dans le cadre classique, qu'entre : a) se donner des probabilités a priori. Mais il doit s'attendre à ce que celles-ci ne traduisent que les propriétés de sa propre cognition, et non celles des sources. Par exemple, il peut se donner une loi telle que $p(a) = 1/2$, $p(\neg a) = 1/2$, parce que, en l'absence de connaissances contextuelles supplémentaires, l'observateur doit se contenter de tautologies (qui accèdent certes au sommet de l'entendement humain, mais sont de piètre valeur informationnelle). Ou bien, b) mesurer des fréquences, puis à partir de ces mesures extrapoler et prédire des lois de probabilités qui lui serviront ensuite pour déterminer les diverses quantités d'information des sources.

Dans le cadre quantique, l'observateur se donne une description a priori de la structure des objets, et en tient compte dans l'évaluation des probabilités. A des lois $p(i)$ posées a priori ou évaluées de façon statistique sont substituées de lois $p(\theta)$ raisonnées, où θ est un paramètre (ici un angle) supplémentaire introduit par la théorie, elle-même issue des catégorisations que l'observateur a effectuées préalablement.

4.3.2. Structure heuristique de la théorie classique quantique

Comme précédemment, nous donnons pour finir un schéma de la structure des opérations cognitives et des calculs que la théorie met en jeu. Nous nous baserons sur l'exemple du § 4.3.1.

(i) Le noyau des calculs (calculs des entropies H à partir des probabilités p) est le même que dans le schéma 3.6. Les probabilités obéissent encore à des relations ensemblistes, les vecteurs de probabilités sont aussi les lignes et les colonnes de tableaux ordonnés horizontalement et verticalement suivant l'indexation des bibliothèques.

(ii) Toutefois, certaines probabilités ne sont plus établies axiomatiquement, ou d'après un calcul statistique. Si une probabilité est posée a priori, ou après un comptage de fréquences, il s'agit d'une probabilité classique (cas de la source Σ dans l'exemple). Si cette probabilité est calculée selon la méthode vectorielle propre au formalisme quantique, il s'agit d'une probabilité

quantique (cas de la source Γ dans l'exemple). Dans ce cas, il faut non seulement tenir compte des propriétés de l'objet, à travers le paramètre θ , mais aussi des propriétés du processus de mesure, à travers les matrices \mathbf{M}_j .

La bibliothèque de la source Γ dans l'exemple est le résultat d'une construction effectuée par l'observateur à partir de l'objet associé à Σ par l'intermédiaire de la connaissance de l'existence et de la valeur du paramètre θ . Objet et observateur sont liés dans une même activité de modélisation, qui fournit les bibliothèques des deux sources. La connaissance des probabilités p conserve une nature ensembliste, mais c'est la définition des ensembles eux-mêmes qui n'est plus triviale : comme nous le disions au début de cette section, cet acte cognitif initial et préalable à toute évaluation de l'information consiste à définir le sujet de cette évaluation. Il est double, correspondant aux deux premières étapes évoquées en introduction : (a) étape opérationnelle : préparation P (dont le résultat est la connaissance de θ) des micro-états étiquetés $|u\rangle$ et $|v\rangle$. (b) définition des observables \mathbf{M}_j .

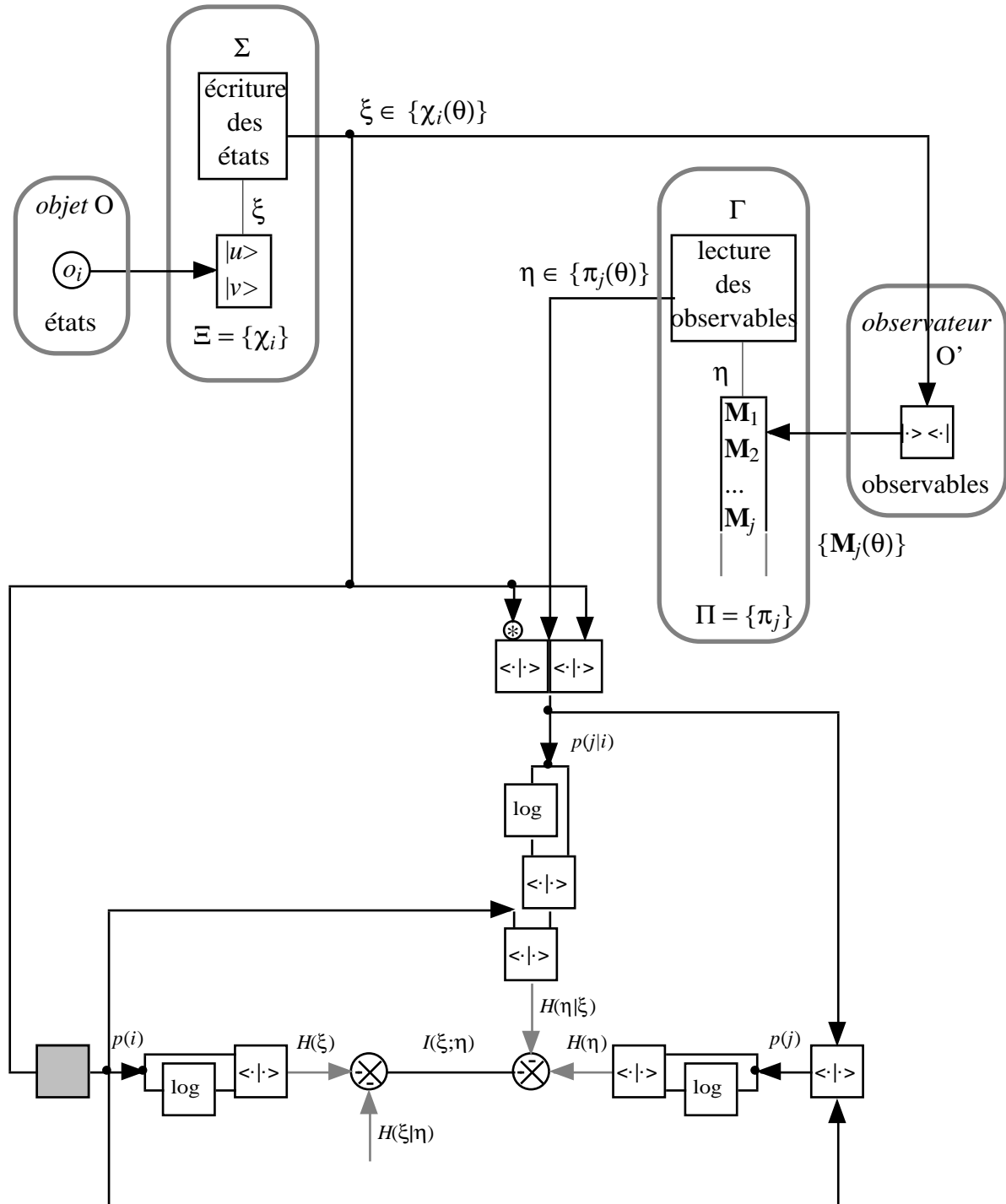
Dans le schéma, il est nécessaire de prévoir la représentation d'un nouvel opérateur : le projecteur $|\cdot\rangle\langle\cdot|$, qui permet de calculer les \mathbf{M}_j . En outre, le calcul des probabilités par $p(j|\xi) = |\langle\xi|\mathbf{M}_j|\xi\rangle|$ nécessite de faire la distinction entre un vecteur et son dual, par l'opérateur de symétrie hermitienne noté \otimes .

Enfin, notons que dans cet exemple les probabilités $p(j)$ sont calculées à partir de $p(j) = \sum_i p(i) \cdot p(j|i)$, qui est donc encore une moyenne pondérée représentée sous la forme d'un produit scalaire (selon la dimension correspondant à l'index i) de deux vecteurs de probabilité.

(iii) Les propriétés exposées au point précédent ont une autre conséquence vis-à-vis de la théorie de l'information : si un état ou un observable sont des symboles élémentaires, il n'en demeure pas moins que ceux-ci sont dotés d'une structure interne, grâce à laquelle on peut mener les calculs. Cette structure est la modélisation de l'objet et de l'observateur (c'est-à-dire une construction cognitive). On ne peut donc plus "faire l'impasse" sur les objets – plus exactement sur les modèles des objets – pour ne considérer que les sources, comme nous le supposons dans la section 2 des définitions : les propriétés statistiques des sources dépendent des modèles de constitution des objets. En ce sens, bien qu'élémentaire et non sécable, un symbole est doté d'une structure interne qui influe sur son comportement probabiliste.

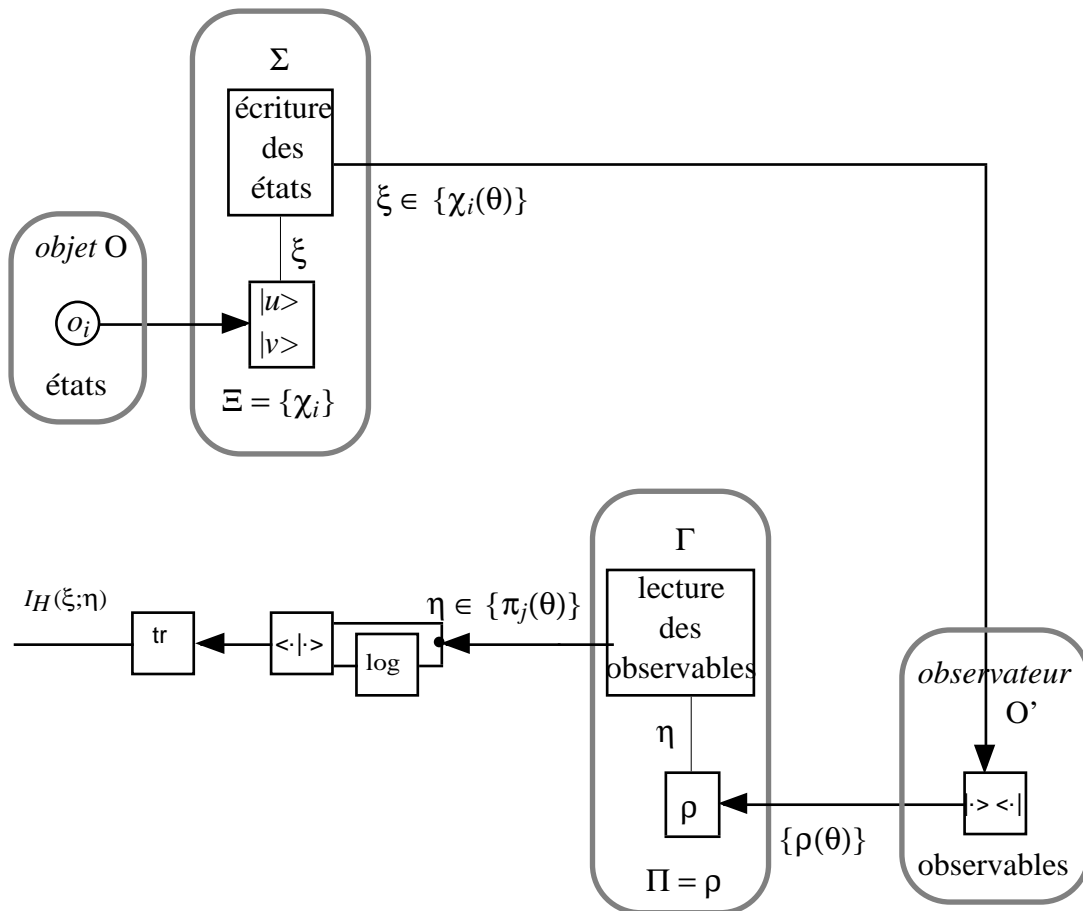
(iv) Une caractéristique essentielle de la théorie, qui dérive immédiatement du point précédent, apparaît à la lecture du schéma : objet O et observateur O' sont liés au sein d'un même processus de modélisation. Ou, dans le cas plus général d'une interaction entre deux objets, l'effet de celle-ci est que l'action d'un objet sur un autre est assimilable à un processus de mesure et de modélisation, autrement dit un processus informationnel agissant sur l'ensemble $\{O, O'\}$. Le résultat de ce processus global est la construction séparée et non symétrique de deux bibliothèques dont sont issues les deux sources : une bibliothèque de vecteurs d'états et une bibliothèque d'observables. A partir de là, on retombe dans le schéma des processus informationnels classiques.

(fig. page suivante)



- **Figure 4.8** : heuristique : représentation schématique de la structure globale des opérations cognitives et des calculs mis en jeu dans l'exemple 4.3.1. Pour la clarté du schéma n'ont été représentées que les opérations effectivement employées dans cet exemple.

(iv) Un cas extrême est celui du calcul de la limite de Holevo : en suivant les flèches qui décrivent le déroulement heuristique du calcul de I_H , on constate que "l'observateur" O' perd toute existence propre, dépendant totalement des propriétés et des caractéristiques de l'objet :



• Figure 4.9 : heuristique du calcul de la limite de Holevo.

4.4. Sources faibles quantiques : entropie quantique quantique

4.4.1. Sources corrélées

Dans le paragraphe précédent, la communication *classique* consistait à transmettre un (ou deux) bit classique en transportant d'un lieu à un autre (physiquement), ou en enregistrant puis lisant, un bit quantique choisi dans l'ensemble des états d'un système quantique : une telle manipulation ne contredit en aucun cas les lois de la physique quantique. Mais celle-ci autorise aussi une autre forme de communication, parfois dénommée "*téléportation*", issue des propriétés des systèmes quantiques corrélés (paires dites *EPR*, pour Einstein-Podolsky-Rosen).

La structure et les propriétés du bit quantique entraînent un certain nombre de conséquences que nous allons résumer.

(i) Théorème de non-duplication (No-Cloning Theorem)

Il n'existe pas de transformation unitaire U telle que $U|\psi\rangle|0\rangle = |\psi\rangle|\psi\rangle \in \forall |\psi\rangle$.

□ La démonstration se fait par l'absurde. Supposons qu'il existe U , qui serait une sorte de machine à dupliquer telle que :

$$U|\psi\rangle|0\rangle = |\psi\rangle|\psi\rangle$$

$$U|\varphi\rangle|0\rangle = |\varphi\rangle|\varphi\rangle$$

Le fait que U soit unitaire entraîne que :

$$\langle\psi|\varphi\rangle\langle 0|0\rangle = \langle\psi|\varphi\rangle\langle\psi|\varphi\rangle$$

relation qui est fautive dans le cas où $\langle\psi|\varphi\rangle \neq 1$.

□

On ne peut donc copier un bit quantique, comme on copie un bit classique.

(ii) Codage "dense" (Dense Coding)

La limite de Holevo montre que l'échange d'un bit quantique ne permet pas de transmettre plus d'un bit classique. Mais la corrélation quantique peut être mise à profit pour échanger deux bits (classiques) d'information par le biais d'un seul bit quantique. Pour cela, on considère deux sources Σ et Γ corrélées, partageant par exemple l'état $|\Phi^+\rangle_{\Sigma\Gamma}$ entre deux bits quantiques appartenant chacun à une source. Les bibliothèques des sources Σ et Γ sont :

$$\Xi = \Pi = \{ |\Phi^+\rangle_{\Sigma\Gamma}, |\Phi^-\rangle_{\Sigma\Gamma}, |\Psi^+\rangle_{\Sigma\Gamma}, |\Psi^-\rangle_{\Sigma\Gamma} \}$$

Un message constitué par un de ces quatre symboles est obtenu en appliquant à $|\Phi^+\rangle$ une des quatre transformations unitaires, respectivement : $\mathbf{1}$ (ne rien faire), σ_1 (rotation de $180^\circ / Ox$), σ_2 (rotation de $180^\circ / Oy$), σ_3 (rotation de $180^\circ / Oz$).

Ce message est alors transmis vers Γ par transfert matériel du bit quantique qui appartenait à Σ . Pour le lire, on applique une mesure orthogonale à la paire $\Sigma\Gamma$ qui projette celle-ci dans l'état correspondant à la transformation précédente et fournit donc deux bits classiques d'information.

Remarques :

- le bit quantique transmis ne contient en lui-même *aucune information* : sa matrice de densité est $\rho_{\Sigma} = \frac{1}{2}\mathbf{1}_{\Sigma}$. Le mesurer est inutile, chaque état est équiprobable ($\frac{1}{4}$). C'est sur *l'ensemble* des deux bits qu'il faut appliquer l'opérateur de mesure. Cela constitue en soi un moyen d'encryptage absolument inviolable...

- le message a été préalablement préparé en affectant à Γ un des bits de la paire corrélée, ce qui a nécessité un transfert. Il y a donc bien au total transfert de deux bits quantiques. Mais la transmission d'information proprement dite est assurée par le deuxième bit quantique. La préparation de la liaison est donc à mettre au bilan des ressources partagées, la transmission elle-même n'intervenant qu'a posteriori.

(iii) Téléportation quantique (Quantum Teleportation)

Avec le codage dense, on transmet de l'information classique par un canal quantique. L'inverse est possible : transmettre de l'information quantique à l'aide d'un canal classique.

Les hypothèses et les conditions de départ sont les mêmes : on suppose que les deux sources Σ et Γ sont corrélées, en partageant par exemple l'état $|\Phi^+\rangle_{\Sigma\Gamma}$ d'une paire de bits quantiques, chacun d'eux appartenant à une source. A nouveau, la corrélation est utilisée en tant que ressource, c'est-à-dire en tant qu'information contextuelle préalable partagée par les deux sources. Dans ce contexte, on veut transmettre de Σ vers Γ un bit quantique, dont l'état est représenté par :

$$|\chi\rangle = \alpha|0_X\rangle + \beta|1_X\rangle$$

La différence avec le cas précédent est donc que la bibliothèque de Σ est :

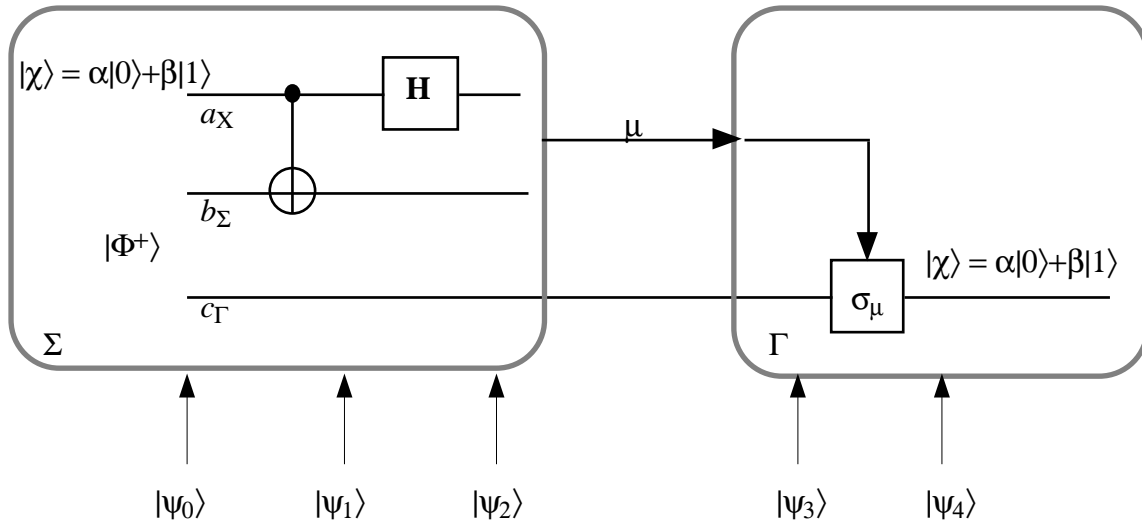
$$\Xi = \{|\chi\rangle, |\Phi^+\rangle\}$$

et la "bibliothèque" de Γ ne comporte que le symbole $|\Phi^+\rangle$, ce qui revient à dire que celle-ci est vide (une bibliothèque doit contenir au moins deux symboles distincts ou distinguables) :

$$\Pi = \{|\Phi^+\rangle\}$$

Cette fois, comme cela avait été annoncé à l'issue de l'étude de la structure d'un bit quantique (§ 4.2), par rapport à $|\chi\rangle$, la bibliothèque Ξ est un continuum *non dénombrable* qui n'est pas séparable en symboles élémentaires distincts.

Le protocole est le suivant (voit figure 4.10) [d'après Vazirani, 1997] :



• Figure 4.10 : téléportation d'un bit quantique

a) Dans Σ , on effectue la corrélation $|\psi_0\rangle$ entre $|\chi\rangle$ et $|\Phi^+\rangle_{\Sigma\Gamma}$ dans l'espace de Hilbert à huit dimensions $\mathcal{H}_{X\Sigma\Gamma} = \mathbb{C}^3$:

$$|\psi_0\rangle = (\alpha|0_X\rangle + \beta|1_X\rangle)|\Phi^+\rangle_{\Sigma\Gamma}$$

$$|\psi_0\rangle = \frac{1}{\sqrt{2}} \left[\alpha|0_X\rangle(|0_{\Sigma}0_{\Gamma}\rangle + |1_{\Sigma}1_{\Gamma}\rangle) + \beta|1_X\rangle(|0_{\Sigma}0_{\Gamma}\rangle + |1_{\Sigma}1_{\Gamma}\rangle) \right]$$

b) On applique dans Σ l'opérateur **CNOT** au bit cible b_{Σ} contrôlé par $a_X = |\chi\rangle$:

$$|\psi_1\rangle = \frac{1}{\sqrt{2}} \left[\alpha|0_X\rangle(|0_{\Sigma}0_{\Gamma}\rangle + |1_{\Sigma}1_{\Gamma}\rangle) + \beta|1_X\rangle(|1_{\Sigma}0_{\Gamma}\rangle + |0_{\Sigma}1_{\Gamma}\rangle) \right]$$

c) On applique dans Σ la transformation de Hadamard au bit a_X :

$$|\psi_2\rangle = \frac{1}{2} \left[\alpha(|0_X\rangle + |1_X\rangle)(|0_{\Sigma}0_{\Gamma}\rangle + |1_{\Sigma}1_{\Gamma}\rangle) + \beta(|0_X\rangle - |1_X\rangle)(|1_{\Sigma}0_{\Gamma}\rangle + |0_{\Sigma}1_{\Gamma}\rangle) \right]$$

En regroupant les termes :

$$|\psi_2\rangle = \frac{1}{2} \left[\alpha|0_X0_{\Sigma}\rangle + \alpha|1_X0_{\Sigma}\rangle + \alpha|0_X1_{\Sigma}\rangle + \alpha|1_X1_{\Sigma}\rangle \right. \\ \left. + \beta|0_X1_{\Sigma}0_{\Gamma}\rangle - \beta|1_X1_{\Sigma}0_{\Gamma}\rangle + \beta|0_X0_{\Sigma}1_{\Gamma}\rangle - \beta|1_X0_{\Sigma}1_{\Gamma}\rangle \right]$$

$$|\psi_2\rangle = \frac{1}{2} \begin{bmatrix} |0_X0_{\Sigma}\rangle(\alpha|0_{\Gamma}\rangle + \beta|1_{\Gamma}\rangle) \\ +|0_X1_{\Sigma}\rangle(\alpha|1_{\Gamma}\rangle + \beta|0_{\Gamma}\rangle) \\ +|1_X0_{\Sigma}\rangle(\alpha|0_{\Gamma}\rangle - \beta|1_{\Gamma}\rangle) \\ +|1_X1_{\Sigma}\rangle(\alpha|1_{\Gamma}\rangle - \beta|0_{\Gamma}\rangle) \end{bmatrix}$$

d) On effectue une mesure dans Σ de la paire corrélée $a_X b_\Sigma$, obtenant quatre résultats possibles équiprobables, et laissant le bit c_Γ dans l'état correspondant :

$$|0_X 0_\Sigma\rangle \rightarrow |\psi_3\rangle = \frac{1}{\sqrt{2}}[\alpha|0_\Gamma\rangle + \beta|1_\Gamma\rangle] \text{ et } \mu = "00"$$

$$|0_X 1_\Sigma\rangle \rightarrow |\psi_3\rangle = \frac{1}{\sqrt{2}}[\alpha|1_\Gamma\rangle + \beta|0_\Gamma\rangle] \text{ et } \mu = "01"$$

$$|1_X 0_\Sigma\rangle \rightarrow |\psi_3\rangle = \frac{1}{\sqrt{2}}[\alpha|0_\Gamma\rangle - \beta|1_\Gamma\rangle] \text{ et } \mu = "10"$$

$$|1_X 1_\Sigma\rangle \rightarrow |\psi_3\rangle = \frac{1}{\sqrt{2}}[\alpha|1_\Gamma\rangle - \beta|0_\Gamma\rangle] \text{ et } \mu = "11"$$

e) On transmet le message μ (classique) de Σ vers Γ (en obéissant aux lois de la relativité : ce message ne peut se déplacer à une vitesse supérieure à celle de la lumière).

f) En fonction du message reçu, on applique dans Γ au bit c_Γ une des quatre transformations de Bell, qui restitue en c_Γ une copie parfaite du bit a_Σ :

$$\mu = "00" \Rightarrow |\psi_4\rangle = \sigma_0 |\psi_3\rangle = |\chi\rangle$$

$$\mu = "01" \Rightarrow |\psi_4\rangle = \sigma_x |\psi_3\rangle = |\chi\rangle$$

$$\mu = "10" \Rightarrow |\psi_4\rangle = \sigma_z |\psi_3\rangle = |\chi\rangle$$

$$\mu = "11" \Rightarrow |\psi_4\rangle = -i\sigma_y |\psi_3\rangle = |\chi\rangle$$

Conclusion

Après cette opération, les bibliothèques sont devenues :

$$\Xi = \{|\psi_3\rangle\}$$

$$\Pi = \{|\chi\rangle, |\psi_3\rangle\}$$

Autrement dit, Ξ n'est plus une "bibliothèque" au sens ordinaire, puisque cette opération laisse Σ dans un état unique (un des quatre états possibles $|\psi_3\rangle$). Par contre, la bibliothèque de Γ est maintenant $|\chi\rangle$, si on laisse de côté ce même état $|\psi_3\rangle$. Ce transfert est effectué grâce à la ressource partagée qu'est l'état $|\Phi^+\rangle$, et possède plusieurs caractéristiques qui le distinguent d'une transmission classique :

- pour autant qu'il n'est pas mesuré dans Γ , le bit $|\chi\rangle$ représente une quantité non dénombrable.
- mais le "prix à payer" pour transférer une entité non dénombrable est élevé :

- la copie est impossible, ce qui physiquement a le sens suivant : une transmission au sens classique d'une entité non dénombrable supposerait un canal de capacité infinie (ou des mots de longueur infinie). Le transfert ressemble plus à un transport de matière (bien qu'il n'en soit pas !) qu'à un transport d'information (qui consiste à copier l'information contenue dans l'émetteur vers le récepteur) : après le transfert, l'émetteur est vide.

- le transfert est "aveugle" : l'émetteur *ne connaît pas* le "contenu" (l'état) du bit quantique téléporté. Pour le connaître, il faudrait le mesurer, ce qui le détruirait. Alors que le récepteur a toute latitude pour faire cela (mais s'il le fait, lui aussi détruira l'information), ou pour l'employer dans un autre processus de communication (comme au §4.3).

- la téléportation a pour résultat le transfert d'une bibliothèque d'une source vers une autre (ce n'est pas une copie, le bit initial est détruit) et non d'un message. Cette bibliothèque est non dénombrable, alors qu'un message, constitué d'une suite de résultats de mesure, est par nature dénombrable. Le bit quantique est un symbole continu, mais sa mesure est un message discret. Cette différence fait d'une source quantique corrélée une "source faible", similaire au concept (§2.5) de source faible continue, définie à partir de l'inégalité $\text{card}(\mathbf{L}) < \text{card}(\mathbf{\Xi})$. Ici, $\text{card}(\mathbf{L}) = \aleph_0$ tandis que $\text{card}(\mathbf{\Xi}) = \aleph_1$.

4.4.2. Entropies de Von Neumann de sources corrélées

Dans ces conditions, il faut s'attendre à ce que les propriétés de l'entropie de cette source faible diffèrent sensiblement des propriétés de l'entropie classique d'un canal quantique. Nous allons considérer maintenant l'entropie quantique (entropie de Von Neumann) de sources quantiques corrélées, ou entropie "quantique quantique".

On construit les entropies composées de sources conjointes en généralisant aux matrices les définitions classiques calculées à partir des probabilités [Cerf, Adami, 1995-1998] :

$$H(\xi) = - \sum_i p(i) \log p(i)$$

$$S(\xi) = - \text{tr} (\rho_\xi \log \rho_\xi)$$

$$H(\xi, \eta) = - \sum_{i,j} p(i,j) \log p(i,j)$$

$$S(\xi, \eta) = - \text{tr} (\rho_{\xi\eta} \log \rho_{\xi\eta})$$

$$H(\xi|\eta) = - \sum_{i,j} p(i,j) \log p(i|j)$$

$$S(\xi|\eta) = - \text{tr} (\rho_{\xi\eta} \log \rho_{\xi|\eta})$$

$$\text{avec } p(i|j) = p(i,j) / p(j)$$

$$\text{avec } \rho_{\xi|\eta} = \rho_{\xi\eta} / (\mathbf{1}_\xi \otimes \rho_\eta)$$

$$H(\xi:\eta) = - \sum_i p(i,j) \log p(i:j)$$

$$S(\xi:\eta) = - \text{tr} (\rho_{\xi\eta} \log \rho_{\xi:\eta})$$

$$\text{avec } p(i:j) = p(i).p(j) / p(i,j)$$

$$\text{avec } \rho_{\xi:\eta} = (\rho_\xi \otimes \rho_\eta) / \rho_{\xi\eta}$$

Ces définitions sont données pour des matrices qui commutent (Cerf et Adami en donnent une version plus générale qui s'étend aux matrices qui ne commutent pas). Notons qu'il ne faut pas confondre la définition de l'entropie conditionnelle $S(\xi|\eta)$ indiquée ici avec la définition par Von Neumann de l'entropie relative [Preskill, 1998] :

$$S(\rho|\sigma) = \text{tr } \rho (\log \rho - \log \sigma)$$

On montre que cette quantité est toujours positive, comme dans le cas classique du gain d'information.

Ces définitions conduisent aux relations :

$$S(\xi, \eta) = S(\xi) + S(\eta|\xi) = S(\eta) + S(\xi|\eta)$$

$$S(\xi:\eta) = S(\xi) + S(\eta) - S(\xi, \eta)$$

Un calcul opéré sur la trace de la quantité $\rho_{\xi|\eta}$ se réduit à son équivalent classique, somme calculée à partir des probabilités $p(\xi|\eta)$, lorsque la matrice est diagonale. L'opérateur $\rho_{\xi|\eta}$ est un opérateur hermitien positif dont le spectre est invariant sous un changement de base $U_\xi \otimes U_\eta$. Mais, alors que les probabilités conditionnelles classiques $p(i|j)$ sont comprises entre 0 et 1, la matrice $\rho_{\xi|\eta}$ n'est pas une matrice de densité, car ses valeurs propres peuvent être supérieures à un. Cela entraîne que $S(\xi|\eta)$ peut prendre des valeurs négatives, contrairement à $S(\rho|\sigma)$. Cette conséquence de la non-monotonie de l'entropie de Von Neumann rapproche celle-ci de l'entropie différentielle (par rapport à l'entropie classique discrète).

Exemple :

Soit une paire corrélée (ξ, η) caractérisée par l'état (exprimé dans la base $|00\rangle, |01\rangle, |10\rangle, |11\rangle$) :

$$|\Phi_{\xi, \eta}\rangle = |\Phi^+\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle) = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

Dans cette base, les matrices de densité de chaque état et de la paire sont :

$$\rho_{\xi, \eta} = |\Phi^+\rangle\langle\Phi^+| = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix} \cdot \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1/2 & 0 & 0 & 1/2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1/2 & 0 & 0 & 1/2 \end{pmatrix}$$

$$\rho_{\xi} = \rho_{\eta} = \begin{pmatrix} 1/2 & 0 \\ 0 & 1/2 \end{pmatrix} \Rightarrow \mathbf{1}_{\xi} \otimes \rho_{\eta} = \begin{pmatrix} 1/2 & 0 & 0 & 0 \\ 0 & 1/2 & 0 & 0 \\ 0 & 0 & 1/2 & 0 \\ 0 & 0 & 0 & 1/2 \end{pmatrix}$$

$$\rho_{\xi|\eta} = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix} \quad \text{on vérifie que } \rho_{\xi\eta} = \rho_{\xi|\eta} \cdot (\mathbf{1}_{\xi} \otimes \rho_{\eta})$$

Plus simplement, on peut aussi exprimer ces quantités dans la base de Bell $|\Phi^+\rangle, |\Phi^-\rangle, |\Psi^+\rangle, |\Psi^-\rangle$. Après diagonalisation des matrices 4x4 de forme générale $M = (a00a, 0000, 0000, a00a)$, on vérifie que :

$$M - \mathbf{1} \cdot \lambda = 0 \Rightarrow \lambda^3 - 2a\lambda = 0 \Rightarrow \lambda_{11} = 2a, \lambda_{22} = \lambda_{33} = \lambda_{44} = 0$$

$$\Rightarrow \rho_{\xi, \eta} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}; \rho_{\xi|\eta} = \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}; \rho_{\xi:\eta} = \begin{pmatrix} 4 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Il vient, sachant que $x \log x \xrightarrow{x \rightarrow 0} 0$:

$$S(\xi) = -\text{tr}(\rho_{\xi} \log \rho_{\xi}) = 1$$

$$S(\xi, \eta) = -\text{tr}(\rho_{\xi\eta} \log \rho_{\xi\eta}) = 0$$

$$S(\xi|\eta) = -\text{tr}(\rho_{\xi|\eta} \log \rho_{\xi|\eta}) = -1$$

$$S(\xi:\eta) = -\text{tr}(\rho_{\xi:\eta} \log \rho_{\xi:\eta}) = 2$$

On vérifie donc que :

$$S(\xi, \eta) = S(\xi) + S(\eta|\xi) = 1 - 1 = 0$$

$$S(\xi:\eta) = S(\xi) + S(\eta) - S(\xi, \eta) = 1 + 1 - 0 = 2$$

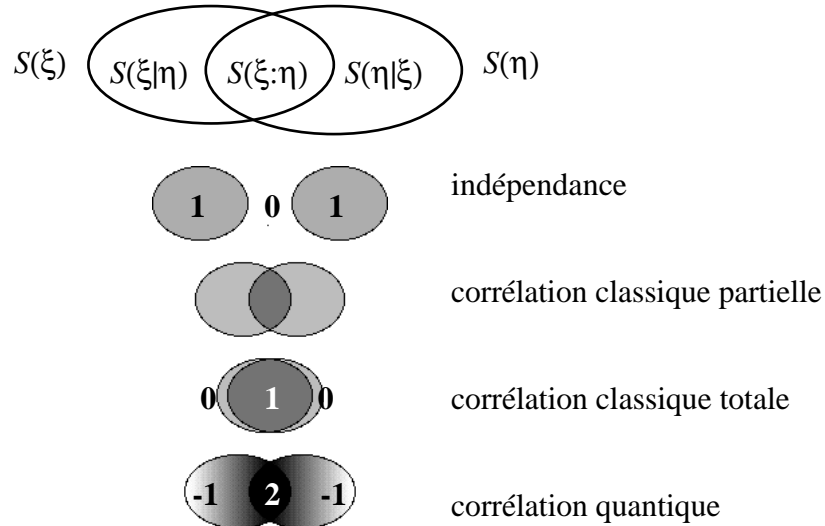
Ce qui se résume dans le diagramme de la figure 4.1.

Dans le cas classique, l'entropie mutuelle maximale est telle que :

$$H(\xi:\eta) \leq \min [H(\xi), H(\eta)]$$

Dans le cas quantique, il vient :

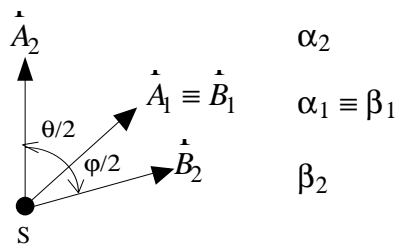
$$S(\xi:\eta) \leq 2 \min [S(\xi), S(\eta)]$$



- Figure 4.11 : différents cas de corrélation entre deux messages de un bit, classique ou quantique [d'après Cerf et Adami, 1995]. Les chiffres indiquent les valeurs respectives des quantités $S(\xi|\eta)$, $S(\xi:\eta)$, $S(\eta|\xi)$.

4.4.3. Exemple

(i) Cerf et Adami ont appliqué les considérations précédentes à une version simplifiée de l'expérience EPR présentée au § 4.1.3.3, où l'on se donne trois (au lieu de quatre) choix possibles : $\hat{A}_1 \equiv \hat{B}_1$, \hat{A}_2 et \hat{B}_2 , définissant ainsi deux angles $\theta/2$ et $\varphi/2$, compris entre 0 et π :



- Figure 4.12 :

En reprenant les mêmes notations qu'au §4.1.3.3, il vient, sachant que $\alpha_1 \equiv \beta_1$:

$$\langle \alpha_i \beta_j \rangle = p_{++} - p_{+-} - p_{-+} + p_{--}$$

$$\langle \alpha_2 \beta_1 \rangle = \cos \theta ; \langle \alpha_1 \beta_2 \rangle = -\cos \varphi ; \langle \alpha_2 \beta_2 \rangle = -\cos(\theta - \varphi)$$

Pour trois variables aléatoires dichotomiques x , y et z , ne pouvant prendre que les valeurs +1 et -1, l'inégalité de Bell s'écrit :

$$xy + xz - yz \leq 1$$

Ce qui conduit, à partir des trois alternatives précédentes, aux inéquations :

$$\langle \alpha_2 \beta_1 \rangle + \langle \alpha_1 \beta_2 \rangle - \langle \alpha_2 \beta_2 \rangle \leq 1$$

$$\langle \alpha_2 \beta_1 \rangle - \langle \alpha_1 \beta_2 \rangle + \langle \alpha_2 \beta_2 \rangle \leq 1$$

$$-\langle \alpha_2 \beta_1 \rangle + \langle \alpha_1 \beta_2 \rangle + \langle \alpha_2 \beta_2 \rangle \leq 1$$

En choisissant $\theta = \pi/4$ (valeur obtenue au §4.1.3.3. correspondant à la violation maximale), et en remplaçant les moyennes $\langle \alpha\beta \rangle$ par leurs valeurs, on constate qu'aucune de ces trois inégalités n'est respectée. Par exemple, pour la première inégalité, on trouve $\langle \alpha_2 \beta_1 \rangle + \langle \alpha_1 \beta_2 \rangle - \langle \alpha_2 \beta_2 \rangle \approx 1,47$ pour $\varphi \approx 2$ rad.

Ce résultat standard peut être étendu dans le cadre de la théorie de l'information.

Soit p la probabilité que la paire soit corrélée ($p = p_{++} + p_{--}$) et $1-p$ qu'elle soit anticorrélée ($1-p = p_{+-} + p_{-+}$) :

$$\langle \alpha_i \beta_j \rangle = p_{++} - p_{+-} - p_{-+} + p_{--} = p - (1-p) \Rightarrow p = (1 + \langle \alpha_i \beta_j \rangle) / 2$$

Comme cela a été vu au §3.3.2, l'information mutuelle entre deux sources symétriques est $H(\dot{A}_i, \dot{B}_j) = 1 - H_2(p_{ij})$ avec $H_2(p_{ij}) = -p_{ij} \log p_{ij} - (1-p_{ij}) \log (1-p_{ij})$, soit :

$$H(\dot{A}_i, \dot{B}_j) = 1 + \frac{1 + \langle \alpha_i \beta_j \rangle}{2} \log \frac{1 + \langle \alpha_i \beta_j \rangle}{2} + \left(1 - \frac{1 + \langle \alpha_i \beta_j \rangle}{2} \right) \log \left(1 - \frac{1 + \langle \alpha_i \beta_j \rangle}{2} \right)$$

Comme cela a été vu au §3.5.1, l'interprétation ensembliste de l'entropie statistique appliquée à trois ensembles A , B et C permet de traduire les inégalités de Bell sous la forme de trois inégalités portant sur les entropies des sources correspondant aux trois orientations $\dot{A}_1 \equiv \dot{B}_1$, \dot{A}_2 et \dot{B}_2 :

$$\begin{aligned} H(\dot{A}_2, \dot{B}_1) + H(\dot{A}_1, \dot{B}_2) - H(\dot{A}_2, \dot{B}_2) &\leq H(\dot{A}_1) \leq 1 \\ H(\dot{A}_2, \dot{B}_1) - H(\dot{A}_1, \dot{B}_2) + H(\dot{A}_2, \dot{B}_2) &\leq H(\dot{A}_2) \leq 1 \\ -H(\dot{A}_2, \dot{B}_1) + H(\dot{A}_1, \dot{B}_2) + H(\dot{A}_2, \dot{B}_2) &\leq H(\dot{B}_2) \leq 1 \end{aligned}$$

Le calcul numérique montre que ces inégalités sont également violées, pour des valeurs de l'angle φ qui sont toutefois différentes des résultats précédents : ainsi, pour la première inégalité, on trouve un maximum $H(\dot{A}_2, \dot{B}_1) + H(\dot{A}_1, \dot{B}_2) - H(\dot{A}_2, \dot{B}_2) \approx 1,088$ obtenu pour $\varphi \approx 2,90$ rad.

(ii) Le même principe de calcul peut être appliqué au cas général de l'inégalité BCHSH, où coexistent les quatre orientations possibles \dot{A}_1 , \dot{B}_1 , \dot{A}_2 et \dot{B}_2 (cf §4.1.3.3). En terme de quantités d'information, il vient :

$$\langle \alpha_i \beta_j \rangle = \cos[2(\varphi_B - \varphi_A)] \Rightarrow \langle \alpha_1 \beta_1 \rangle = \langle \alpha_1 \beta_2 \rangle = \langle \alpha_2 \beta_1 \rangle = \cos 2\theta$$

$$\text{et } \langle \alpha_2 \beta_2 \rangle = \cos 6\theta$$

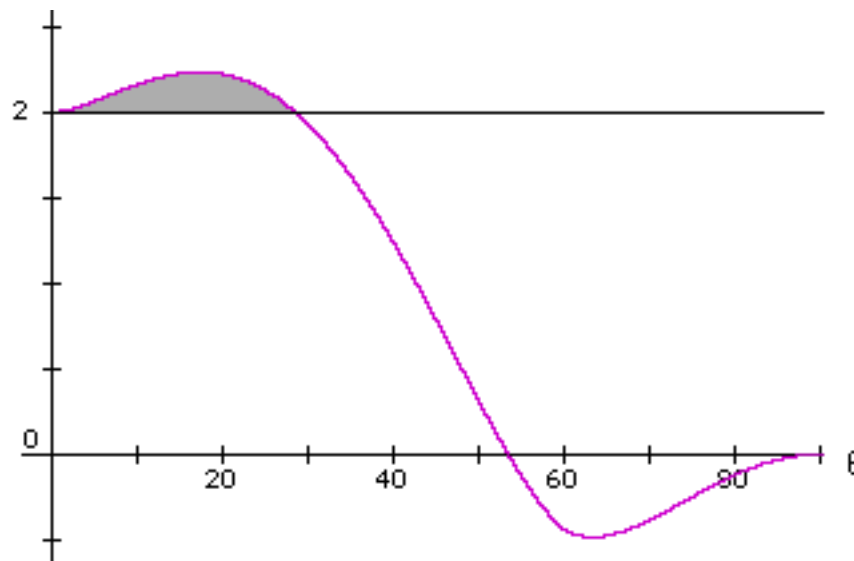
$$\langle \alpha_i \beta_j \rangle = 2p_{ij} - 1 \quad \Rightarrow \quad p_{11} = p_{12} = p_{21} = \frac{1 + \cos 2\theta}{2}$$

$$\text{et } p_{22} = \frac{1 + \cos 6\theta}{2}$$

$$H(\dot{A}_i, \dot{B}_j) = 1 - H_2(p_{ij}) \quad \text{avec } H_2(p_{ij}) = -p_{ij} \log p_{ij} - (1-p_{ij}) \log (1-p_{ij})$$

$$H(\dot{A}_1: \dot{B}_1) + H(\dot{A}_2: \dot{B}_1) + H(\dot{A}_1: \dot{B}_2) - H(\dot{A}_2: \dot{B}_2) \leq 2$$

Le calcul numérique de la partie gauche de cette inéquation conduit aux résultats suivants :



• Figure 4.13 : équivalent entropique de la figure 4.x du §4.1.3.3.

On constate que l'inéquation sur les entropies mutuelles correspondant à l'inégalité BCHSH présente également un domaine angulaire de violation, qui cependant diffère de l'original. On note en particulier que le graphe obtenu ne présente pas de symétrie : la violation a lieu pour des valeurs de θ comprises en gros entre 0 et 30° seulement. Le maximum de violation est $H(\dot{A}_1: \dot{B}_1) + H(\dot{A}_2: \dot{B}_1) + H(\dot{A}_1: \dot{B}_2) - H(\dot{A}_2: \dot{B}_2) \approx 2,22$ obtenu pour $\theta \approx 17,4^\circ$.

4.5. Conclusion : résumé des propriétés de l'entropie de Von Neumann