University Paris-Saclay - IQUPS

Optical Quantum Engineering: From fundamentals to applications

Philippe Grangier,
Institut d'Optique, CNRS, Ecole Polytechnique.

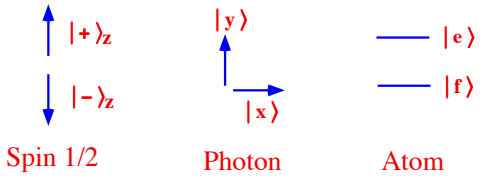
- Lecture 1 (7 March, 9:15-10:45) :
  Qubits, entanglement and Bell's inequalities.
- Lecture 2 (14 March,11:00-12:30) :
  From QND measurements to quantum gates and quantum information.
- Lecture 3 (21 March, 9:15-10:45) :
  Quantum optics with discrete and continuous variables.
- Lecture 4 (28 March, 11:10-12:30) :
  Quantum cryptography and optical quantum networks.

Exemple of Hilbert spaces with dimension equal to two.

* Spin 1/2 particle

* Polarized photon : states with linear or circular polarisation; mathematical structure very close to a spin 1/2 (factor 2 on angles, see below).

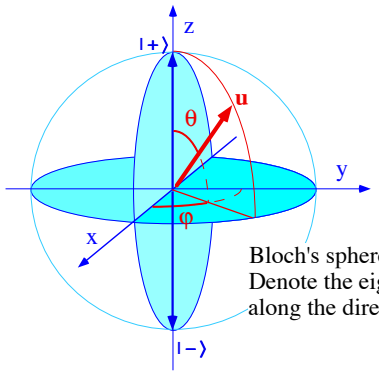
* "Two-level atom" (attention ! spontaneous emission).



* Superconducting circuit : anharmonic quantum oscillator due to a Josephson junction, allows one to isolate two energy levels

These systems are various implementations of a "quantum bit" (qubit).

Bloch's sphere.



Bloch's sphere (spin 1/2)
Denote the eigenstate
along the direction **u**(θ, φ)

Normalized vector $\vec{u}$

$u_x = \cos(\phi)\sin(\theta)$,
$u_y = \sin(\phi)\sin(\theta)$,
$u_z = \cos(\theta)$.

$$\vec{S}.\vec{u} = \frac{\hbar}{2}\,\vec{\sigma}.\vec{u}$$

$$\vec{\sigma}.\vec{u} = \begin{pmatrix} \cos(\theta) & \sin(\theta)e^{-i\phi} \\ \sin(\theta)e^{i\phi} & -\cos(\theta) \end{pmatrix}$$

Eigenvalues of $\vec{\sigma}.\vec{u}$: ± 1, eigenstates of $\vec{S}.\vec{u}$ = eigenstates of $\vec{\sigma}.\vec{u}$:

$$|+_{\vec{u}}\rangle = \cos(\theta/2)e^{-i\phi/2}\,|+_z\rangle + \sin(\theta/2)e^{i\phi/2}\,|-_z\rangle$$
$$|-_{\vec{u}}\rangle = -\sin(\theta/2)e^{-i\phi/2}\,|+_z\rangle + \cos(\theta/2)e^{i\phi/2}\,|-_z\rangle$$

A few considerations on entangled systems.

* Within classical physics, correlations between measurements carried out on separated subsystems are explained by attributing to each subsystems some properties which are correlated to properties of the other subsystem.

* If one tries to reproduce quantum correlations using such a model, Bell's inequalities show that these properties must be non-local, i.e. must contradict relativistic causality $\rightarrow$ **inacceptable**.

* Quantum mechanics remains in perfect agreement with relativistic causality, but there is a price : it is impossible to attribute a "local physical reality" to the state of each subsystem.

"EPR Paradox" (Einstein Podolsky Rosen, 1935)
"Quantum non-separability"

* We will see now that entanglement plays an essential role in quantum mechanics in general, and especially in quantum information...

Lecture 2 : Entanglement in a Quantum Measurement Process : from QND measurements to quantum gates.

1. Direct and indirect measurements in Quantum Mechanics

2. Analysis of a quantum measurement process, no-cloning theorem

3. From quantum measurement to quantum gates.

4. From Shannon to quantum cryptography.

Direct and indirect measurements in Quantum Mechanics

How can we measure the state of a qubit ? (spin 1/2, photon, atom...)

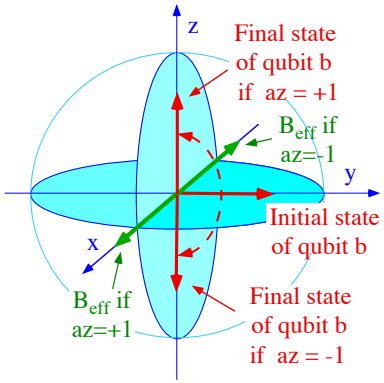
* Direct measurement process :
Spin 1/2 : Stern-Gerlach magnet $\rightarrow$ spatial splitting as a function of the value of the spin component parallel to the gradient of magnetic field.
Photon : Polarizer $\rightarrow$ spatial splitting as a function of the polarization...
The measurement process "demolishes" the qubit, which is no more available after the measurement.

* Indirect measurement process :
One "reads" the state of the qubit by coupling it to another qubit. One can thus realize an ideal measurement, including the state preparation stage. This is called a "Quantum Non Demolition" (QND) measurement.
Attention : QND does not mean that there is no effect on the system's state ! A QND measurement is a quantum measurement, so the system's state is changed, unless it is already in an eigenstate of the measured observable.

QND measurement of a spin component.

One wants to perform a QND measurement of $\hat{\sigma}_z$ on a qubit "a" : if the qubit is a spin 1/2 particle, one gets the spin "a" to interact with another spin "b" during a time τ, and read out the result on spin "b".

An appropriate interaction Hamiltonian is : $H_m = \hbar g\, \hat{\sigma}_{az}\, \hat{\sigma}_{bx}/2$



Everything happens as if qubit a creates on qubit b an effective magnetic field, aligned along Ox, with a sign depending on the state $|\pm\rangle_{az}$ (see exercise !).

$$|+\rangle_{az}\otimes|+\rangle_{by} \longrightarrow |+\rangle_{az}\otimes|+\rangle_{bz}$$
$$|-\rangle_{az}\otimes|+\rangle_{by} \longrightarrow i|-\rangle_{az}\otimes|-\rangle_{bz}$$

QND measurement of a spin component : conclusion.

In the general case the initial state is

$$|\psi(0)\rangle = (\alpha|+\rangle_{az} + \beta|-\rangle_{az}) \otimes |+\rangle_{by}$$

and from the superposition principle :

$$|\psi(\tau)\rangle = (\alpha\, |+\rangle_{az}\, |+\rangle_{bz} + i\beta\, |-\rangle_{az}\, |-\rangle_{bz})$$

- This is an entangled state like the EPR state seen before : a measurement on qubit b gives $+1$ with probability $|\alpha|^2$ and -1 with probability $|\beta|^2$.
- For each result, the state of qubit a is perfectly known after the measurement ("reduction of the wave packet").
- The quantum measurement of σ_{az} is done by an "indirect measurement", called a QND measurement : qubit a is still there for further action !
- Attention : the final state is **not** a duplication of an arbitrary initial state, which would be $(\alpha|+\rangle_{az}+\beta|-\rangle_{az})\otimes(\alpha|+\rangle_{bz}+\beta|-\rangle_{bz})$: **no-cloning !**

The no-cloning theorem (1).

"Quantum cloning" means to duplicate an arbitrary and unknown initial quantum state onto another system with an initial state $|\psi_0\rangle$ ("white sheet"), keeping the original intact in order to have two identical copies.

Considering two different initial states to be copied $|\phi_1\rangle$ et $|\phi_2\rangle$ one wants :

$$|\phi_1\rangle \otimes |\psi_0\rangle \rightarrow |\phi_1\rangle \otimes |\psi_1\rangle$$
$$|\phi_2\rangle \otimes |\psi_0\rangle \rightarrow |\phi_2\rangle \otimes |\psi_2\rangle$$

The copies are denoted as $|\psi_1\rangle$ et $|\psi_2\rangle$ because they may be made on different physical systems : e.g. "cloning" a polarized photon onto a spin 1/2 particle.

Theorem : Perfect cloning of an arbitrary unknown quantum state is forbidden both by linearity and by unitarity of quantum mechanics.

The no-cloning theorem (2).

1. Demonstration based on linearity : let us consider $|\phi_3\rangle = \frac{1}{\sqrt{2}}(|\phi_1\rangle + |\phi_2\rangle)$

$$|\phi_1\rangle \otimes |\psi_0\rangle \rightarrow |\phi_1\rangle \otimes |\psi_1\rangle$$
$$|\phi_2\rangle \otimes |\psi_0\rangle \rightarrow |\phi_2\rangle \otimes |\psi_2\rangle$$
$$|\phi_3\rangle \otimes |\psi_0\rangle \rightarrow \frac{1}{\sqrt{2}}(|\phi_1\rangle \otimes |\psi_1\rangle + |\phi_2\rangle \otimes |\psi_2\rangle) \neq |\phi_3\rangle \otimes |\psi_3\rangle$$

The sum of the copies is not the copy of the sum !

2. Demonstration based on unitarity : conservation of the scalar product.

$$\langle\phi_1|\phi_2\rangle \ \langle\psi_0|\psi_0\rangle = \langle\phi_1|\phi_2\rangle\langle\psi_1|\psi_2\rangle$$
$$\langle\phi_1|\phi_2\rangle \ (1 - \langle\psi_1|\psi_2\rangle) = 0$$

$\langle\phi_1|\phi_2\rangle = 0$: orthogonal states $\quad \langle\psi_1|\psi_2\rangle = 1$: identical copies.

Conclusion : one can clone within a known set of orthogonal states, but it is not possible to clone (perfectly) within a set of non-orthogonal states.

Generalization : "CNOT" quantum gate

The interaction between two qubits seen previously implements a logical operation between the qubits : $|0,\ 0\rangle \longrightarrow |0,\ 0\rangle$, et $|1,\ 0\rangle \longrightarrow i\,|1,\ 1\rangle$.

Generalizing, one defines a "C-NOT" (Controlled-NOT) gate, which does the following (denoting $|i\rangle_a \otimes |j\rangle_b = |i, j\rangle$):

$$|0,\ 0\rangle \longrightarrow |0,\ 0\rangle$$
$$|0,\ 1\rangle \longrightarrow |0,\ 1\rangle$$
$$|1,\ 0\rangle \longrightarrow |1,\ 1\rangle$$
$$|1,\ 1\rangle \longrightarrow |1,\ 0\rangle$$

The first qubit (control qubit) is unchanged.
The second qubit (target qubit) is inverted if the first qubit value is one.

This logical gate (applied to many qubits...) is a building block to implement quantum computation.

Simple application : preparation and measurement of Bell's states

$$|0,0\rangle \longrightarrow \boxed{?} \longrightarrow \frac{1}{\sqrt{2}}(|0,0\rangle + |1,1\rangle)$$

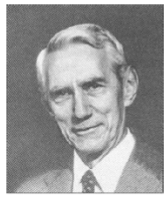
Simple application : preparation and measurement of Bell's states

The control qubit is driven in a superposition of the computational states $|0\rangle$ et $|1\rangle$ ($\pi/2$ rotation of Bloch's vector), then the CNOT gate is applied:

$$
\begin{aligned}
|0,0\rangle &\longrightarrow \frac{1}{\sqrt{2}}(|0\rangle+|1\rangle)|0\rangle &&\longrightarrow \frac{1}{\sqrt{2}}(|0,0\rangle+|1,1\rangle)\\
|0,1\rangle &\longrightarrow \frac{1}{\sqrt{2}}(|0\rangle+|1\rangle)|1\rangle &&\longrightarrow \frac{1}{\sqrt{2}}(|0,1\rangle+|1,0\rangle)\\
|1,0\rangle &\longrightarrow \frac{1}{\sqrt{2}}(|0\rangle-|1\rangle)|0\rangle &&\longrightarrow \frac{1}{\sqrt{2}}(|0,0\rangle-|1,1\rangle)\\
|1,1\rangle &\longrightarrow \frac{1}{\sqrt{2}}(|0\rangle-|1\rangle)|1\rangle &&\longrightarrow \frac{1}{\sqrt{2}}(|0,1\rangle-|1,0\rangle)
\end{aligned}
$$

- One gets 4 orthogonal entangled states forming a basis of the Hilbert space of the two qubits (with dimension 4) : "Bell's states".
- This transformation is reversible : starting from Bell's states, on can get the four factorized basis states, which are easy to identify from a direct measurement of the two qubits : Bell's states measurement.

Shannon's information theory



Claude E. Shannon
1916–2001

A mathematical theory of communication, Bell System Technical Journal **27** (1948) 379-423 and 623-656.

Communication theory of secrecy systems, Bell System Technical Journal **28** (1949) 656-715.

This work has laid out the entire foundation of today's information technology era.

A few elementary notions about the classical theory of information.

Fundamental results published by Claude Shannon, 1948.

(1) By how much is it possible to compress the data of a message, assuming that there is no transmission noise ?
Answer :
the maximum compression rate is the **Shannon entropy** of the message

(2) What is the maximum transmission rate through a noisy channel ?
Answer :
the maximum transmission rate of information is equal to the **channel capacity**, to be defined below

Shannon entropy

Message : (long) string of n letters chosen within an alphabet of k letters :

$$\{a_1,\ a_2,\ \ldots\ a_k\}$$

Each letter a_x is given a probability $p(a_x)$, with $\sum_{x=1}^{k} p(a_x) = 1$.

Simple exemple : binary alphabet $\{0,1\}$ with $p(0) = 1-p$, $p(1) = p$.

If $n >> 1$, the message will contain about np bits 1 and $n(1-p)$ bits 0.

The number of such "typical messages" is

$$C_n^{np} = \frac{n!}{(np)!(n-np)!}$$

Using Stirling's formula $\log(n!) = n\log n - n + O(\log n)$ one gets :

$$\begin{aligned}\log_2(C_{np}^n) &= n\log_2(n) - np\log_2(np) - (n-np)\log_2(n-np)\\ &= nH(p)\end{aligned}$$

where we define the **(binary) Shannon entropy**

$$H(p) = -p\log_2 p - (1-p)\log_2(1-p).$$

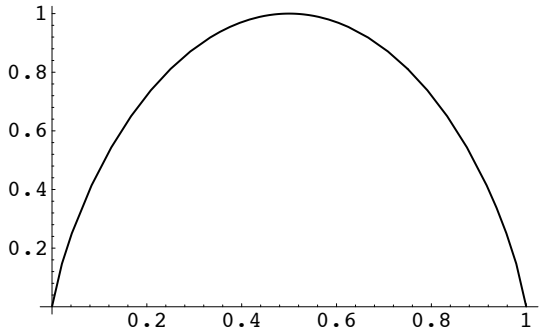
The number of typical messages is then $2^{nH(p)}$.

Main point :
When n is very large, it is sufficient to assign a "codeword" to each of these typical messages, because the "atypical" messages will almost never appear (more rigorously : the error rate due to the atypical messages will be asymptotically negligible).

In order to identify these codewords one needs $nH(p)$ bits instead of n, the compression factor is thus $H(p) \leq 1$.

Remarks :

* Shape of the binary entropy $H(p) = -p\log_2 p - (1-p)\log_2(1-p)$:



* For $p = 1/2$ one gets $H(p) = 1$: a completely random message (no redundancy) cannot be compressed.

This reasoning can be generalized for k letters, by defining

$$H(X) = -\sum_x p(x)\log_2 p(x) = -\langle \log_2 p(x)\rangle.$$

where $H(X)$ is **the Shannon entropy** of the message $X = \{x,\ p(x)\}$ (alphabet and associated probabilities).
A message with n letters can then be compressed in $nH(X)$ bits; on says also that each letter carries on average $H(X)$ bits of information.

Remarks :
* If the 26 letters of the usual alphabet were perfectly random, each letter would carry $\log_2(26) = 4.7$ bits. However the probabilities are not equally distributed (esainturlo...) and the letters are strongly correlated, so a letter carries about 1.1 bit (see http://www.math.ucsd.edu/ crypto/java/entropy/).
* Therefore it is enough to use $n(H(p)+\delta)$ bits to encode a message with n bits, without errors asymptotically. On the other hand, one shows that with $n(H(p)-\delta)$ bits there will necessarily be errors, because there will not be enough codewords to represent all typical sequences.

$$p(x) = \begin{cases} 1/2 & \text{if } x = \text{"A"}\\ 1/4 & \text{if } x = \text{"B"}\\ 1/8 & \text{if } x = \text{"C"}\\ 1/8 & \text{if } x = \text{"D"}\end{cases}$$

Simple code:
"A" $\rightarrow$ 00
"B" $\rightarrow$ 01
"C" $\rightarrow$ 10
"D" $\rightarrow$ 11

$L = 2$ bits/symbol

Average code length $L \geq H(X)$

$H(X) = \frac{1}{2} + \frac{2}{4} + \frac{3}{8} + \frac{3}{8} = \frac{7}{4}$ bits < 2 bits

Mutual information.

In a message is transmitted through a noisy channel, the received message will be in general different from the sent message. So one should answer the question : what do we know about a message drawn in the ensemble X^n, if one knows a message drawn in the ensemble Y^n ?
Let us define the conditional probability $p(y|x)$ to receive y if x was sent. The ensemble of $p(y|x)$ caracterizes the channel, and using Bayes' formula one can calculate $p(x|y)$:

$$p(x|y) = p(x,y)/p(y) = p(y|x)p(x)/\sum p(y|x)p(x)$$

From the $p(x|y)$ one defines the conditional entropy $H(X|Y)$:

$$H(X|Y) = \langle -\log_2 p(x|y)\rangle = \langle -\log_2 p(x,y)\rangle + \langle \log_2 p(y)\rangle.$$

One has thus :

$$H(X|Y) = H(X,Y) - H(Y)$$

and also

$$H(Y|X) = H(X,Y) - H(X)$$

Mutual information.

One defines then the **mutual information** $I(X;Y)$:

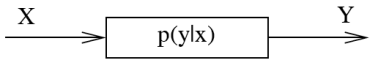
$$\begin{aligned} I(X;Y) &= H(X) - H(X|Y) \\ &= H(X) + H(Y) - H(X,Y) \\ &= H(Y) - H(Y|X) \end{aligned}$$

The mutual information $I(X;Y)$ is symmetrical with respect to X and Y, and is zero if X and Y are uncorrelated. It is the fundamental quantity measuring the information exchanged through the channel. The unit of $I(X;Y)$ ("bit per symbol") is the same as the one used for the entropies $H(X)$.

Shannon theorem for the transmission through a noisy channel :
The maximum number of bits per symbol that can be transmitted through a noisy channel is its capacity C, defined by :

$$C = \overset{Max}{\{p(x)\}} I(X;Y)$$

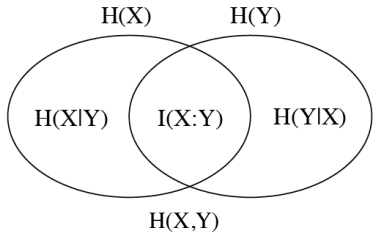
Taking the maximum over $\{p(x)\}$ eliminates the role of the message, and therefore C is a property of the channel itself (in fact, of the $p(y|x)$).



$H(X|Y) = H(X,Y) - H(Y) =$ loss of information

perfect channel $H(X|Y) = 0$
random channel $H(X|Y) = H(X)$

$I(X{:}Y) \equiv H(X) - H(X|Y) =$ mutual information
$= H(X) + H(Y) - H(X,Y)$



The density matrix (1)

One can (re)formulate quantum mechanics by replacing the state vector $|\psi(t)\rangle$ by the projector onto $|\psi(t)\rangle$:

$$|\psi(t)\rangle \quad \rightarrow \quad \hat{\rho}(t) = |\psi(t)\rangle\langle\psi(t)|$$

Then a statistical distribution of states $|\psi_i\rangle$ with a probability Π_i (**classical statistics**) will be described by the density matrix :

$$\hat{\rho} = \sum_i \Pi_i \, |\psi_i\rangle\langle\psi_i|$$

As a result the double average (classical and quantum) of an observable $\hat{A}$ can be written:

$$\langle\hat{A}\rangle_{stat} = \sum_i \Pi_i \, \langle\psi_i|\hat{A}|\psi_i\rangle = \sum_i \Pi_i \, \mathsf{Tr}(|\psi_i\rangle \, \langle\psi_i| \, \hat{A}) = \mathsf{Tr}(\hat{\rho} \, \hat{A})$$

The density matrix (2)

It is postulated that any quantum state can be described by a **density operator** $\hat{\rho}$ such that :

- $\hat{\rho}$ is hermitian with trace 1 (but $\hat{\rho}$ is not always a projector)
- All its eigenvalues are positive or zero.
- The hamiltonian evolution and measurement probabilities are given by :

$$\mathcal{P}(a_\alpha) = \mathrm{Tr}(\hat{\rho}\hat{P}_\alpha) \qquad \langle A \rangle = \mathrm{Tr}(\hat{\rho}\hat{A}).$$

$$i\hbar \frac{d\hat{\rho}}{dt} = \left[\hat{H}(t), \hat{\rho}(t)\right].$$

Remark : One can always diagonalize $\hat{\rho}$ and write : $\hat{\rho} = \sum_i \Pi_i \, |\psi_i\rangle\langle\psi_i|$. One recovers a pure state if there is only one non-zero Π_i, then

$$\hat{\rho}(t) = |\psi(t)\rangle\langle\psi(t)|, \quad \hat{\rho}^2 = \hat{\rho} \text{ and } \mathrm{Tr}(\hat{\rho}^2) = 1$$

Simple exemples of density matrices (1)

Depolarized spin 1/2 (silver atom straight from the oven) :

$$\hat{\rho}_{\text{nonpol.}} = \frac{1}{2}|+\rangle\langle+| \; + \; \frac{1}{2}|-\rangle\langle-| \; = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \frac{1}{2}\,\hat{1}.$$

to be compared to a pure state $(|+\rangle + |-\rangle)/\sqrt{2}$: $\hat{\rho}_{\text{pol.selon } x} = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$.

Qubit : two-state system $\{|e\rangle, |g\rangle\}$ ou $\{|+\rangle, |-\rangle\}$:

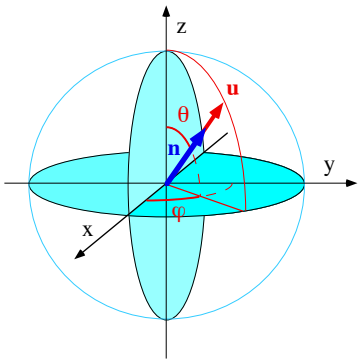
$$\hat{\rho} = \begin{pmatrix} \rho_{ee} & \rho_{eg} \\ \rho_{ge} & \rho_{gg} \end{pmatrix} = \begin{pmatrix} (1+n_z)/2 & (n_x - in_y)/2 \\ (n_x + in_y)/2 & (1-n_z)/2 \end{pmatrix} = \frac{1}{2}\left(\hat{1} + \vec{n}.\vec{\sigma}\right)$$

where we define $\vec{n} = (n_x, n_y, n_z)$ with n_i real numbers, and where we used the Pauli matrices $\vec{\sigma}$:

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

The eigenvalues of $\hat{\rho} = \frac{1}{2}\left(\hat{1} + \vec{n}.\vec{\sigma}\right)$ are equal to $\frac{1}{2}(1 \pm |\vec{n}|)$

Generalisation of Bloch sphere ("Bloch ball")



$$\hat{\rho} = \begin{pmatrix} (1+n_z)/2 & (n_x - in_y)/2 \\ (n_x + in_y)/2 & (1-n_z)/2 \end{pmatrix} = \frac{1}{2}\left(\hat{1} + \vec{n}.\vec{\sigma}\right)$$

with eigenvalues $\frac{1}{2}(1 \pm |\vec{n}|)$.

Using $\mathrm{Tr}(\sigma_i) = 0$, $\sigma_i^2 = \hat{1}_2$ $(i = x, y, z)$, one gets :

$$\langle \vec{\sigma} \rangle = \mathrm{Tr}(\hat{\rho}\,\vec{\sigma}) = \vec{n}$$

- if $|\vec{n}| = 1$ one can write $\vec{n} = \vec{u}$ and one retrieves a pure state. The operator $\hat{\rho}$ is then a projector on the state $|+_{\vec{u}}\rangle$, and one has $\langle \vec{\sigma} \rangle = \vec{u}$
- if $|\vec{n}| < 1$ then $\vec{n}$ is inside the Bloch ball, and one has again $\langle \vec{\sigma} \rangle = \vec{n}$
- if $|\vec{n}| = 0$ then the spin is depolarized and $\langle \vec{\sigma} \rangle = \vec{0}$.

Entangled pair of spin 1/2 particles (qubits).

In a singlet state for pour 2 spins 1/2 :

$$|\psi_{ss}\rangle = \frac{1}{\sqrt{2}}(|+,\,-\rangle - |-,\,+\rangle) \qquad \hat{\rho}_{ss} = \frac{1}{2}\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

with the ordering of the basis vectors: $\{|+,\,+\rangle,\, |+,\,-\rangle,\, |-,\,+\rangle,\, |-,\,-\rangle\}$

- Justify the above expression of $\hat{\rho}_{ss}$.
- Show that the reduced density operators are given by :

$$\hat{\rho}_A = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \hat{\rho}_B = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \qquad \hat{\rho}_A \otimes \hat{\rho}_B = \frac{1}{4}\;\hat{1}_4$$

  Conclude that the two sub-systems are completely depolarized.
- Show by using the reduced density operators that the "reduction of the wave packet" does not allow any transfer of information between Alice and Bob when doing an EPR correlation experiment.

Von Neumann entropy.

In quantum information the "symbols" may become quantum states, described by a density matrix ρ_x, and the alphabet is then an ensemble $\{\rho_x, p_x\}$. The density matrix seen by an observer is then

$$\rho = \sum_x p_x \ \rho_x.$$

One can choose an orthonormal basis $\{|a\rangle\}$ where ρ is diagonal

$$\rho = \sum_a p_a \ |a\rangle\langle a|.$$

The set $\{|a\rangle\langle a|, p_a\}$ is then équivalent to a classical alphabet, and one has:

$$H(A) = -\sum_a p_a \ \log_2(p_a) = -\text{Trace}(\rho \ \log_2 \rho) = S(\rho)$$

where $S(\rho)$ is the **Von Neumann entropy** of the density matrix ρ.

In general the symbols ρ_x are not mutually exclusive (non orthogonal states), and the Shannon and Von Neumann entropies have different properties.

Mathematical properties of Von Neumann entropy (1).

1. The entropy $S(|\phi\rangle\langle\phi|)$ of a pure state is equal to zero.

2. The entropy $S(\rho)$ is not modified by a change of basis (isometry).

3. If ρ has D non-zero eignevalues, then $S(\rho) \leq \log_2 D$.

4. For $p_i \geq 0$ and $\sum_i p_i = 1$, one has $S(\sum_i p_i \ \rho_i) \geq \sum_i p_i \ S(\rho_i)$
   (S increases when ignoring of the way by which the state was prepared).

5. Measurement : For an observable B with eigenvalues b_y one defines $Y = \{b_y, \ p(b_y)\}$ and then $H(Y) \geq S(\rho)$ (equality if $[B, \rho] = 0$).

6. Preparation : For a density matrix $\rho = \sum_x p_x \ |\phi_x\rangle\langle\phi_x|$ with $X = \{|\phi_x\rangle, \ p_x\}$ then $H(X) \geq S(\rho)$ (equality if the $\{|\phi_x\}$ are orthogonal).

Mathematical properties of Von Neumann entropy (2).

7. For a bipartite system AB one has

$$|S(\rho_A) - S(\rho_B)| \leq S(\rho_{AB}) \leq S(\rho_A) + S(\rho_B)$$

   whereas

$$H(X), H(Y) \leq H(X, Y) \leq H(X) + H(Y).$$

   Fundamental difference between Shannon and Von Neumann !
   On can get $S(\rho_A) = S(\rho_B) \neq 0$ whereas $S(\rho_{AB}) = 0$
   (entanglement ! impossible classically).

8. Strong subaddidivity : For a tripartite system ABC one has

$$S(\rho_{ABC}) + S(\rho_B) \leq S(\rho_{AB}) + S(\rho_{BC})$$

   (important, not easy to demonstrate !).

Quantum data compression.

One considers $\rho = \sum_x p_x \ |\phi_x\rangle\langle\phi_x|$, and one defines $X = \{|\phi_x\rangle, \ p_x\}$. A message with n symbols is then associated to the density matrix

$$\rho^n = \rho \otimes .... \otimes \rho.$$

What is the minimum number of **qubits** needed to encode this message ?
(Reminder : the dimension of the Hilbert space $\mathcal{E}$ for N qubits is 2^N).

It can be shown (Schumacher) that $\log_2(dim\mathcal{E}) = N = n \ S(\rho)$. The entropy $S(\rho)$ corresponds to the number of **qubits** per symbol of the message.

Principle of the demonstration :
For very long messages, the support of the density matrix ρ^n is included within a subspace of dimension $2^{nS(\rho)}$. This can be obtained from Shannon's theorem, by considering a basis where ρ is diagonal.

Quantum data compression: a few remarks.

1. Bob received qubits which are "efficiently packaged", but in general he cannot simply retrieve the classical information sent by Alice.
2. In the general case, the symbols are themselves statistical mixtures (density matrices), and not pure states. An important question is then to evaluate the maximum amount of **classical** information, that can be extracted from a quantum message.

   The answer is given by using the Holevo information :

$$\chi(\{\rho_x, p_x\}) = S(\rho) - \sum_x p_x \, S(\rho_x).$$

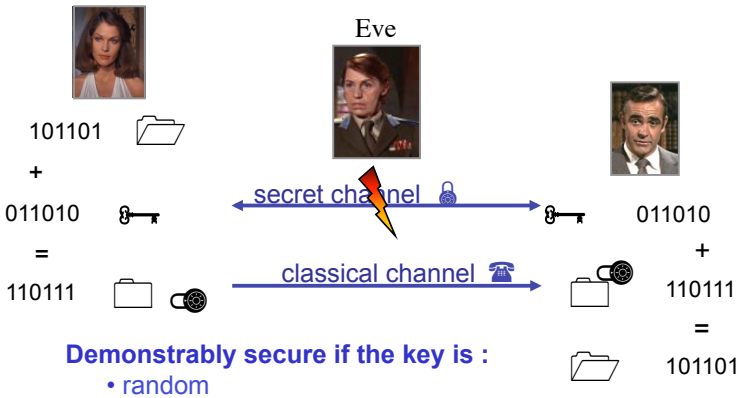
Holevo' theorem then states that the maximum accessible classical information (over all possible measurements) is bounded by $\chi(\{\rho_x, p_x\})$:

$$\underset{Mesures}{Max} I(X;Y) \leq \chi(\{\rho_x, p_x\}).$$

Demonstration : Not obvious, one has to use strong subadditivity...

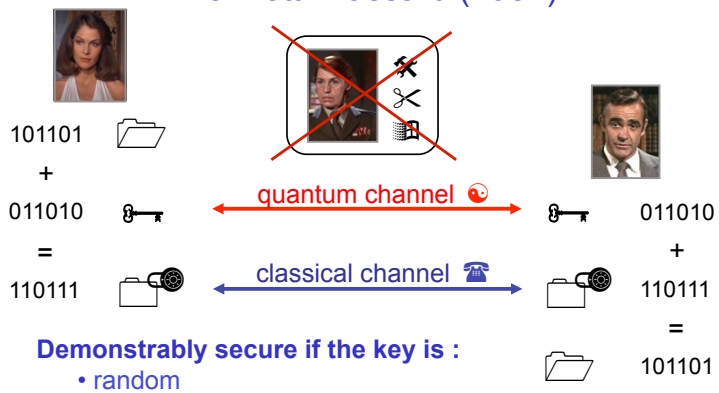
Quantum cryptography: the characters

Secret key cryptosystem : one-time pad (G. Vernam, 1917)



Demonstrably secure if the key is :
- random
- as long as the message
- used only once (Shannon)

Quantum Secret Key Cryptosystem : Bennett-Brassard (1984)



Demonstrably secure if the key is :
- random
- as long as the message
- used only once (Shannon)
- **unknown by Eve : Quantum laws !**