Clément Bénesse

On quantum error-correction

Mémoire d'initiation à la recherche

Sous la direction de Madame **Danijela Markovic**

Cycle **p**luridisciplinaire d'**é**tudes **s**upérieures
Troisième année (L3)

Abstract

The aim is to make a description of different types of quantum correction codes, starting from the basics in quantum theory and finishing with some implementations. This paper is firstly addressed to undergraduate students and sources will be given to those more curious.

1 Quantum Mechanic

Here, we will introduce the main ideas in quantum mechanics. Most of them will be needed in order to understand what quantum correction is about. Anyone familiar with it should recognise easily every part of this section. At the end, however, we will be interested in information theory , exploring the quantum concept of entropy, as von Neumann formulated it.

1.1 Generalities

- As in classical physic, we can associate any quantum isolated system to a Hilbert complex space (ie the phase space $\mathcal{H}$) in which the system will be entirely described by an unitary vector (norm equal to 1). This vector of $\mathcal{H}$ will be noted, accordingly to Dirac's notation $|\psi\rangle$ (it is a *ket*) and its dual,which is a vector of $\mathcal{H}^\star$ will be noted $\langle\psi|$ (it is a *bra*). The action of a bra $\langle\phi|$ on a ket $|\psi\rangle$ is the bracket $\langle\phi|\psi\rangle = \langle\phi|\,(|\psi\rangle)$, which is the classical scalar product. We can remark that the orthogonal projector on the state $|\psi\rangle$ is $|\psi\rangle\langle\psi|$.
  In the case of $\mathcal{H} = \mathbb{C}^2$, we can denote by $|0\rangle$ and $|1\rangle$ two orthogonal vector of the unit sphere. For example, we can take:
  $$|0\rangle = \begin{pmatrix}1\\0\end{pmatrix};|1\rangle = \begin{pmatrix}0\\1\end{pmatrix}$$
  Therefore, we have a basis of $\mathcal{H}$ and we can write every state $|\psi\rangle = \alpha\,|0\rangle + \beta\,|1\rangle = \begin{pmatrix}\alpha\\\beta\end{pmatrix}$ with the condition of normalization: $\alpha^2 + \beta^2 = 1$. This is a qubit and we can see it as a superposition of bits 0 and 1 with probabilities, as we will see later.

  We can define operators to which we can apply vectors (i.e. $A\,|\psi\rangle$). The expectation value of the operator A in the state $|\psi\rangle$ is $\langle A\rangle = \langle\psi|A|\psi\rangle$. Moreover, the adjoint of the operator A, noted $A^\dagger$, is defined to be the transposition of the conjugate of A and therefore, $\langle\phi|A^\dagger|\psi\rangle = \langle\psi|A|\phi\rangle^*$. Some adjoints can have nice proprieties such as being hermitian. A is hermitian if and only if $A = A^\dagger$ and in that case, the expected value is real (for $< a >= \langle\phi|A|\phi\rangle = \langle\phi|A^\dagger|\phi\rangle =< a >^*$).Such an operator, if it is diagonalisable, is called an observable.

- There are two types of evolution for a system.
  - The first kind of evolution is an unitary evolution which preserves the norm of the states :
    If, at a time t, the system is in the state $|\psi\rangle$ and, at a time t', in a state $|\psi'\rangle$, then there exists an unitary operator U such that $|\psi'\rangle = U\,|\psi\rangle$
    In the same way as in classical physic with Newton's laws, the evolution of a system follows laws. In quantum physic,the equivalent of Newton's laws is the Schrödinger' equation which describes the behaviour of a system. There are two versions depending on time dependence:
    $$i\hbar\frac{\partial}{\partial t}\,|\psi(t)\rangle = H\,|\psi(t)\rangle$$
    $$E\,|\psi\rangle = H\,|\psi\rangle$$
    where H is the Hamiltonian operator and E the energy of the system. Whereas the second one tell us that energy is an eigenvalue of the Hamiltonian, the first one, from a differential equation result, tell us that $|\psi(t))\rangle = |\psi_0\rangle \exp(-iHt/\hbar) = |\psi_0\rangle\,U(t)$. This inverse of this operator $U(t)$ is its adjoint $U^\dagger$

If we are looking for another example, we can take the case of a qubit, as defined previously ($|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$), and see what happens with the matrix :

$$X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

In this case, $X|\psi\rangle = \begin{pmatrix} \beta \\ \alpha \end{pmatrix}$. This is called a bit flip and it changes the roles of $|0\rangle$ and $|1\rangle$. Such an evolution is reversible and in this particular case, the inverse is X itself.This kind of evolution can be unwanted depending on the result we want to obtain, and we might want to erase it, as we will see in the part on error-correction.

We can remark that the evolution can be applied to $\langle\psi|$ and not to $|\psi\rangle$. In this case, due to the fact that we work in a complex space, we have the following propriety: $\langle\psi| A = A^\dagger |\psi\rangle$. The convention is different of the one used by mathematicians as the scalar product is "antilinear on the left" here, and not "on the right".

- The second type is a non-unitary evolution. An example is a measurement.. Measurements are the only way to have some information on the current state of the system. However, the result of a measure is probabilist and modifies the state of the system. The formal way to introduce a measure is mathematical and hard to work with:

  A family $(M_m)_m$ of operators of $\mathcal{H}$ defines a measure of the quantum system if it satisfies the relation $\sum_m M_m^\star M_m = Id_{\mathcal{H}}$.

  On a more practical point of view, one can see a measure as a projection of the system on a subspace of $\mathcal{H}$. Let's take back our example: we have the vector $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$. If we decide to "measure it" on the basis $(|0\rangle, |1\rangle)$,then the result of the measure will be $|0\rangle$ with probability $|\alpha|^2$ or $|1\rangle$ with probability $|\beta|^2 = 1 - |\alpha|^2$. An important propriety of a measure is that it leaves the system changed! Which means, in the previous example, that if we observe $|1\rangle$ and we decide to remeasure right after, we will obtain $|1\rangle$ with probability 1! To finish with this example, one can ask himself how to measure a state. In this case, the answer was given earlier: the projector on $|0\rangle$ is $|0\rangle\langle 0|$. Applied to the vector $|\psi\rangle$, it will give the result described a few lines before.

- We will now introduce the concept of quantum entanglement.

A system can be composed of smaller systems (for example, we have two electrons with spin up or down, we can study each one separately or combined). If we have states $|\psi\rangle \in \mathcal{H}_1$ and $|\phi\rangle \in \mathcal{H}_2$, then the global system will be described by the tensor product $|\psi\rangle \otimes |\phi\rangle$, noted $|\psi\phi\rangle$. If $\mathcal{H} = \mathbb{C}^2$, we use the basis $(|0\rangle, |1\rangle)$ and an orthonormal basis of $\mathcal{H}^{\otimes n}$ is $(|x_1 x_2 ... x_n\rangle)_{x_i \in \mathbb{F}^2}$. We said that a vector of $\mathcal{H}$ can be seen as a qubit; therefore a vector of $\mathcal{H}^{\otimes n}$ can be seen as n qubits. We encode a binary message on qubits but during the evolution errors can occur. It is these errors that we will want to correct.An example of those "product states" is the state $|a\rangle = (|00\rangle + |01\rangle)/\sqrt{2}$

Nevertheless, all states of $\mathcal{H}^{\otimes n}$ cannot be "factorized", cannot be explained as a tensor product of states of smaller systems. The better way to see it is through an example:consider the vector of $\mathbb{C}^2 \otimes \mathbb{C}^2$

$$\frac{1}{\sqrt{2}}(|01\rangle + |10\rangle)$$

It is impossible to express this state as one could "factorize" $|a\rangle$. Indeed , we can write $|a\rangle$ in a more condensed way as $|0\rangle(|0\rangle + |1\rangle)$, times a constant for normalisation. But it is impossible to do such a factorisation for $(|01\rangle + |10\rangle)/\sqrt{2}$ .In this case, we speak of entangled state. A particularity of this kind of state is that we need less measure to fully know it.Lets say , in this case, we measure on $\mathcal{H}_1$; it is enough in order to determine the second state, whereas it would not be possible if the vector could be factorized (if we observe $|0\rangle$, we know that the entangled state was been projected on $|01\rangle$ so the system $\mathcal{H}_2$ is on the state $|1\rangle$)

- We define the Pauli matrices X,Y and Z as the following:

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

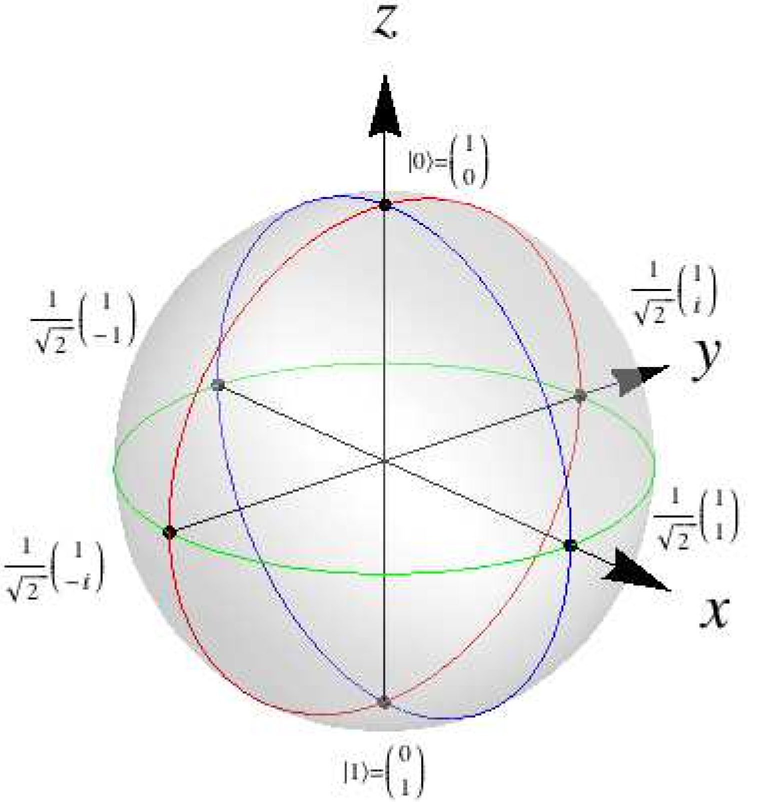
  These matrices can be considered as a single vector $\sigma = (X, Y, Z)$ for the Bloch sphere, as we will see in the next part.Each one is hermitian and, together with the identity matrix, the Pauli matrices span the full vector space of 2x2 hermitian matrices We already saw the X matrix in the first example and said it represent a bit-flip (the roles of $|0\rangle$ and $|1\rangle$ are swapped).In a similar case, the Z matrix is associated to a flip of the phase sign.

1.2 Bloch sphere

The Bloch sphere is a geometrical representation for quantum states $|\psi\rangle$ of a two level quantum system (ie. $\mathbb{C}^2$),which can be written as a superposition of the basis $(|0\rangle, |1\rangle)$. As only the relative phase between the two coefficients has a physical meaning, we can take one of the coefficient real and non negative. Therefore,if, we have chosen the coefficient of $|0\rangle$, we can write the state as following:

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

Therefore, we can represent a state on the unit sphere of $\mathbb{R}^3$ with these spherical coordinates uniquely.For those interested in the mathematics, the Bloch sphere is isomorphic to the unit sphere of quaternions, which we quotient by the antipodal relation. An algebra result proves that it is isomorphic to SO_3. One can notice that antipodal points match orthogonal state vectors.Another way to give a vector of the Bloch Spere is to use the Pauli matrices defined earlier and a straightforward computation is enough to see that a vector can be expressed as $\rho = \frac{1}{2}(I + \vec{a} \cdot \vec{\sigma})$, where $\vec{a}$ is the Bloch vector. .The points on the surface of the sphere correspond to the pure states of the system, whereas the interior points correspond to the mixed states associated to the density operator defined in the following section is ρ.


z
$|0\rangle=\begin{pmatrix}1\\0\end{pmatrix}$
$\frac{1}{\sqrt{2}}\begin{pmatrix}1\\-1\end{pmatrix}$
$\frac{1}{\sqrt{2}}\begin{pmatrix}1\\i\end{pmatrix}$
y
$\frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\end{pmatrix}$
$\frac{1}{\sqrt{2}}\begin{pmatrix}1\\-i\end{pmatrix}$
x
$|1\rangle=\begin{pmatrix}0\\1\end{pmatrix}$

1.3 Density operator

The density operator, often noted ρ, is a matrix which describes a quantum system , may it be mixed or pure. We can define it by $\rho(t) = |\phi(t)\rangle \langle\phi(t)|$.

In the case of a pure state, where we can write $|\phi(t)\rangle = \sum c_n(t)\,|u_n\rangle$ in a well chosen basis, a quick computation gives the coefficients of the matrix $\rho_{i,j} := \langle u_i|\rho|u_j\rangle = c_j^* c_i$. If we take back our example ($|\phi\rangle = \cos(\theta/2)\,|0\rangle + e^{i\phi}\sin(\theta/2)\,|1\rangle$),the density operator is given by the following matrix:

$$\begin{pmatrix} \cos^2(\theta/2) & \cos(\theta/2)\sin(\theta/2)e^{-i\phi} \\ \cos(\theta/2)\sin(\theta/2)e^{i\phi} & \sin^2(\theta/2) \end{pmatrix}$$

From a similar computation, the mean value of an operator can be given thanks to the density operator as well by the relation $< A >= Tr(\rho(t)A)$. This goes along with a set of proprieties:

$$Tr\rho(t) = 1$$

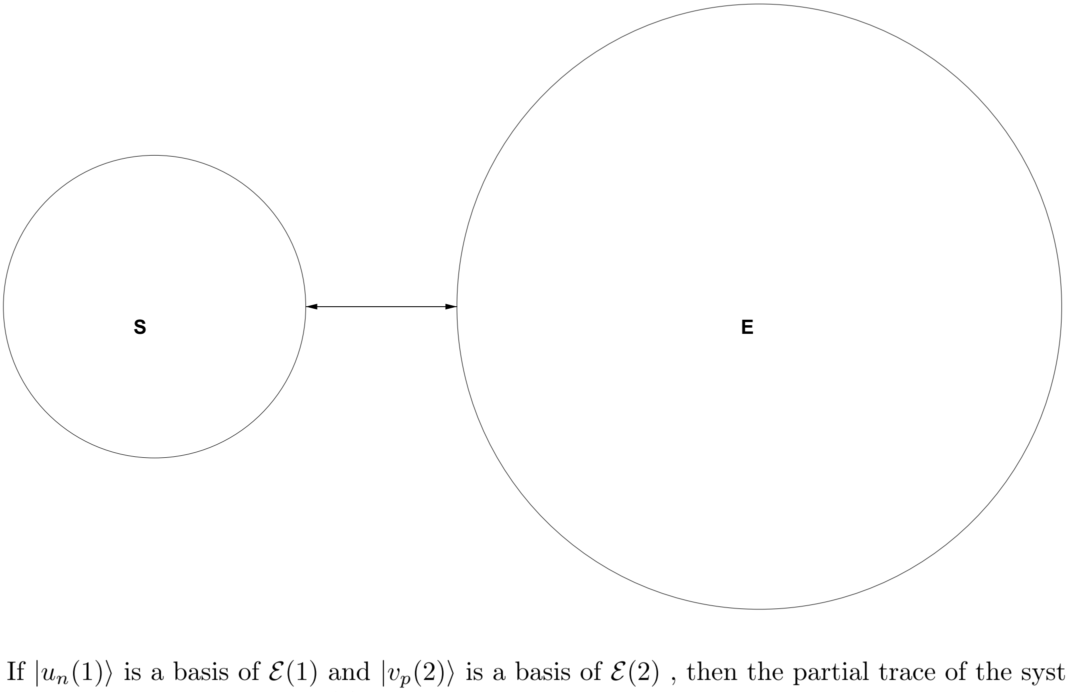
$$Tr(A\rho) = Tr(\rho A)$$

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

$$\rho^\dagger = \rho = \rho^2$$

Now, let's see what happens in the case of mixed states. Be careful, a mixed state is not the same thing as a quantum superposition of pure states! Indeed , a quantum superposition of pure state is still a pure state (cf $|\phi\rangle = (|01\rangle + |10\rangle)/\sqrt{2}$). A mixed state can be used for example to describe the spin of an electron which has been in contact with the environment or if we study the polarization of unpolarized light. In that case, the Hilbert space of the system is the tensor product of two Hilbert spaces $\mathcal{E} = \mathcal{E}(1) \otimes \mathcal{E}(2)$. One of the main tools used for mixed state of partial trace.

Interaction between the system and its environment:



If $|u_n(1)\rangle$ is a basis of $\mathcal{E}(1)$ and $|v_p(2)\rangle$ is a basis of $\mathcal{E}(2)$, then the partial trace of the system (1) with respect to the system (2) can be defined as:

$$\rho(1) = Tr_2\rho_{12} = \sum_p \langle v_p(2)|\rho|v_p(2)\rangle$$

We can also use the fact that $tr_B(|a_1\rangle\langle a_2| \otimes |b_1\rangle\langle b_2|) = |a_1\rangle\langle a_2| \, tr(|b_1\rangle\langle b_2|)$

As an example, we can study the Bell state $(|00\rangle + |11\rangle)/\sqrt{2})$. The density operator will be :

$$\begin{aligned} \rho &= \left(\frac{|00\rangle + |11\rangle}{\sqrt{2}}\right)\left(\frac{\langle 00| + \langle 11|}{\sqrt{2}}\right) \\ &= \frac{|00\rangle\langle 00| + |11\rangle\langle 11| + |11\rangle\langle 00| + |00\rangle\langle 11|}{2} \end{aligned}$$

Tracing out the second qubit and using the propriety of the partial trace, we obtain:

$$\begin{aligned} \rho_1 &= tr_2(\rho) \\ &= \frac{|0\rangle\langle 0|\langle 0|0\rangle + |1\rangle\langle 0|\langle 0|1\rangle + |0\rangle\langle 1|\langle 1|0\rangle + |1\rangle\langle 1|\langle 1|1\rangle}{2} \\ &= \frac{I}{2} \end{aligned}$$

This object is not a number but an operator on the space (1). But why is it so useful?If we now try to compute the mean value of an other operator $A = A(1) \otimes I(2)$ where I is the identity, one can remark that

$$< A >= Tr(\rho_{12}A) = Tr_1(\rho(1)A(1)$$

In fact, in order to know the mean value of this kind of operator, we do not need to know the whole system but only the reduced density matrix. All remarks about density matrices are valid, except for the last one: $\rho^2(1) \neq \rho(1)$, the reduced matrix density is not a projector. We said earlier that , on the Bloch sphere, pure states were on the surface and mixed states were inside the sphere. That can be expressed with the density operator as $tr(\rho) = 1$ and $\rho^2 = \rho$ only for pure states.

1.4 Von Neumann entropy,information theory

In quantum mechanics, the von Neumann entropy is the extension of the classical Gibbs entropy.It is defined by $S(\rho) = -tr(\rho \ln \rho)$ where ρ is the density matrix defined previously. But if ρ can be decomposed with eigenvalues ($\rho = \sum_j n_j |j\rangle\langle j|$), then the von Neumann entropy is given by: $S(\rho) = -\sum_j n_j \ln n_j$.

Some proprieties:

- The entropy is zero if and only if ρ is a pure state.
- $S(\rho)$ is maximal and equal to $\ln N$ for a maximally mixed state, N being the dimension of the Hilbert space.
- S is concave
- If we have two independent systems (A) and (B),the entropy of the system (A+B) is the sum of the entropies of the system (A) and of the system (B). If the systems are not independent.

2 Quantum Codes

What is the purpose of error-correction codes? What is the motivation behind the study of such things? The idea is simple: if an error (such as a typo in a word) occurs while a message is transmitted, the receiver may not understand the meaning of the message. But if the "noise" (the error) is weak, one can try to understand the message by correcting the error of the message. While the aim is to have a perfectly noiseless communication, error-correction codes are here to say if a noised message is still understandable or not.

For example, suppose we want to send a message made of 0 or 1 (typically, a bit string) through a noisy channel. The typical error can be a flip (ie. a 0 becomes a 1) with probability $p > 0$ or a swap between two bits with probability $q > 0$.A very simple way to protect the message from the noise is to repeat each bit several times:

$$0 \mapsto 000$$

$$1 \mapsto 111$$

The noise of the channel will therefore be weaken and if you observe the message 000 010 110 100, we can deduce that the original message was 0010 , modified with a flip on the fifth bit and a swap between the bits number 9 and 10. Of course, every bit could have been changed but , if p and q are small, we suppose that only two errors is much more probable than a lot. Of course, we can imagine different errors (for example, the codeword is not received entirely) and other ways to discover them, even if the main idea is repetition.

2.1 Generalities about linear codes

After this heuristic and before studying quantum error-correction codes, lets see what is a linear code from the mathematical point of view.The motivation is double: not only is it better to have a few examples of classic correction codes in mind before studying quantum ones but some ideas will be important in the following of the paper. Of course,given the variety of technological applications, the theory of such codes is quite developed and the more curious can read [1].

A linear code C encoding k bits of information into an n bit code is associated to an n by k generator matrix G whose entries are elements of $\mathbb{F}_2$. The k bit message x is therefore encoded as Gx.We can remark that all operations are done modulo 2 as we work with $\mathbb{F}_2$.A code encoding k bits in n is a $[n, k]$ code.

The first example of such a code will be the repetition code described previously Recall that $0 \mapsto 000$ and $1 \mapsto 111$. In this case, the generator matrix will be the 3×1 matrix:

$$G = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

We can check that $G(0) = (0, 0, 0)^t$ and $G(1) = (1, 1, 1)^t$ We could similarly write the expression of a $[rk, k]$ linear code using r repetition of each bit.

This code can also be represented by an alternative formulation: the parity check matrix H. In this case, a $[n, k]$ code is all n-elements vectors x such that

$$Hx = 0$$

where $H \in \mathcal{M}_{n-k \times n}()$. The code is defined to be the kernel of H. One can show that, if the parity check matrix is of the form $H = [A|I_{n-k}]$, then the corresponding matrix is

$$G = [\frac{I_k}{-A}]$$

Now, suppose that we encode a message y but an error e corrupts y (due to noise for example), giving the corrupted codeword $y' = y + e$. As $Hy = 0$, it follows that $Hy' = He$. We call Hy' the error syndrome ; it is a function of the corrupted state and contains information about the error that occurred.In order to see what type of information we have,imagine there are no errors or only a flip on one bit: the syndrome is equal to 0 when there are no errors and to He when there is only one error. By computing the error syndrome for each possible value of He_j where e_j is an error on the j-th bit, we can therefore determine which bit needs to be corrected.

For this kind of error-correction performed with linear code, we can use the concept of distance. The (Hamming) distance $d(x, y)$ between x and y is defined to be the number of places at which x and y differ (ex: $d((1, 0, 1), (1, 1, 1)) = 1$). The weight of a word $x$is defined as $wt(x) := d(x, 0)$. Note that, as we work in $\mathbb{F}_2$, $d(x, y) = wt(x + y)$.Let's suppose again that we have a word y corrupted into $y' = y + e$.Then ,provided some hypothesis, the most likely codeword to have been encoded is the codeword y which minimizes $d(y, y') = wt(2y + e) = wt(e)$. Given that, we can now define another characteristic of a code: its distance. We define the distance of a code to be the minimum distance between any two different codewords :

$$d = d(C) := min_{x,y \in C, x \neq y} d(x, y)$$

or, otherwise:

$$d(C) = min_{x \in C, x \neq 0} wt(x)$$

In that case, we talk of a $[n, k, d]$ code. We can give as a propriety that a code with minimum distance d can correct $\lfloor \frac{d-1}{2} \rfloor$ errors.

We conclude this part on classical error-correction by the *dual* construction. Suppose C is an $[n, k]$ code with generator matrix G and parity check matrix H. Then, we can define the dual of C , denoted $C^{\perp}$, to be the code with generator matrix H^T and parity check matrix G^T. Equivalently, the dual of C consists of all codewords y such that y is orthogonal to all the codewords in C . If $C \subseteq C^{\perp}$, the code is said to be weakly self dual and stricly self dual if $C == C^{\perp}$. This notion will come back in quantum error-correction, for the CSS codes for example.

2.2 Calderbank-Shor-Steane codes

An example of a large class of quantum error correcting codes is the *Calderbank - Shor - Steane* codes, usually known as CSS. They are an important subclass of stabilizer codes which will be explained later. In order to create a CSS code, we firstly need two codes.

Suppose C_1 and C_2 are $[n, k_1]$ and $[n, k_2]$ linear codes such that $C_2 \subset C_1$ and C_1 and $C_2^{\perp}$ correct both t errors. The aim is to have a code that can correct up to t errors on qubits (meaning up to t bit flips and t phase flips). The CCS code of C_1 over C_2 will be able to do that.We will construct the vector space defining the code:

For $x \in C_1$, we define the quantum state

$$|x + C_2\rangle := \frac{1}{|C_2|} \sum_{y \in C_2} |x + y\rangle$$

. It is easy to see that this state depends only upon the coset of C_1/C_2 which x is in; and we could add that different cosets lead to orthonormal states. We have therefore our vector space: $CSS(C_1, C_2)$ is defined to be the vector space spanned by the states $|x + C_2\rangle$. In this case, the code is a $[n, k_1 - k_2]$ quantum code. By a technique explained in [1] , it can be shown that the first code C_1 will be used to find and correct bit flip errors while the second code C_2 will be used to correct phase flips.

Another way to create CSS codes will be explained after the stabilizer codes are introduced. Nevertheless, we will see an example of CSS code: the *Steane* code.Its construction uses the $[7, 4, 3]$ Hamming code whose parity check is the following:

$$H = \begin{pmatrix} 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{pmatrix}$$

If we denote this code as C and set $C_1 = C$ and $C_2 = C^{\perp}$, the only hypothesis we have to check is that $C_2 \subset C$.We need to compute the parity check matrix of the code C_2 for this. From what we said on dual codes, we know that:

$$H[C_2] = G[C_1]^T = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{pmatrix}$$

The span of the rows of $H[C_2]$ contains the span of the rows of $H[C_1]$ so C_2 is contained in C_1. Therefore, we have a $[7, 1]$ CSS code which will be able to correct errors on a single qubit.If we want more informations on the code such as the space ,we need only two vectors; they can be given as

$$|0\rangle = \frac{1}{\sqrt{8}}[|0000000\rangle + |1010101\rangle + |0110011\rangle + |1100110\rangle$$
$$+ |0001111\rangle + |1011010\rangle + |0111100\rangle + |1101001\rangle]$$

and

$$|1\rangle = \frac{1}{\sqrt{8}}[|1111111\rangle + |0101010\rangle + |1001100\rangle + |0011001\rangle$$
$$+ |1110000\rangle + |0100101\rangle + |1000011\rangle + |0010110\rangle]$$

How can we find those vectors? We have 3 base vectors for C_2 given by the rows of the generator matrix (which is, here, H). Therefore, we know we have $2^3 = 8$ vectors in C_2.Now, the trick is that $|0\rangle$ is the mean of all the vectors of C_2. For $|1\rangle$, we only have to take the vectors of $|0\rangle$ and add to each an element of C_1 which is not in C_2: the element $|1111111\rangle$ does the trick.

2.3 Stabilizer codes

We said it earlier, CSS codes are a particular case of stabilizer codes. But before anything, we must introduce the stabilizer formalism.

Consider the following state of two qubits:

$$|\phi\rangle = \frac{|00\rangle + |11\rangle}{\sqrt{2}}$$

If we recall the Pauli matrices, we can verify that $X_1X_2|\phi\rangle = Z_1Z_2|\phi\rangle = |\phi\rangle$: the state $|\phi\rangle$ is stabilized by the operators X_1X_2 and Z_1Z_2 and in fact, it is the only one, up to a global phase. Some quantum states can be more easily described by the operators that stabilize them than by the state itself and it is this idea that is behind the stabilizer formalism. The key to describe stabilizer codes is to use the Pauli group. For a single qubit, the Pauli group G_1 consists of the Pauli matrices with multiplicative factors $\pm1, \pm i$:

$$G_1 = \{\pm I, \pm iI, \pm X, \pm iX, \pm Y, \pm iY, \pm Z, \pm iZ\}$$

The general Pauli group on n qubits consist of all n tensor products of Pauli matrices, with again multiplicative factors $\pm1, \pm i$. Now, if we take S a subgroup of G_n and define V_S as all the n-qubits states stabilized by all elements of S, V_S is said to be the vector space stabilized by S and S to be the stabilizer of V_S.

For example, if $n = 3$, and $S = \{I, Z_1Z_2, Z_2Z_3, Z_1Z_3\}$, the subspace fixed by Z_1Z_2 is spanned by $|000\rangle, |001\rangle, |110\rangle$ and $|111\rangle$. For Z_2Z_3, the subspace is spanned by $|000\rangle, |100\rangle, |011\rangle, |111\rangle$. We can see what happen for Z_1Z_3 and deduce that VG_S is spanned by the states $|000\rangle$ and $|111\rangle$. Those familiar with group theory can see that used generators of S as $Z_1Z_3 = (Z_1Z_2)(Z_2Z_3)$ and $I = (Z_1Z_2)^2$.

But not just any subgroup of the Pauli group must be used as the stabilizer. Indeed, if $-I \in S$, it is obvious that the vector space will be trivial as $-|\phi\rangle = |\phi\rangle$ implies that $|\phi\rangle = 0$. In order to have a non-trivial vector space, two conditions must be met: the elements of S must commute and $-I$ must not be an element of S. If the second condition is straightforward, the first one is a bit trickier: as Pauli matrices either commute or anti-commute, any operators of the Pauli group either commute or anti-commute. If it is the latter, then we have lets say $|\phi\rangle = NM|\phi\rangle = -MN|\phi\rangle = -|\phi\rangle$.

Lets take back the Steane code, in the case of seven qubits. It turns out that six generators is enough to genrate a stabilizer for the code space of the Steane code. Those generators are listed as follow:

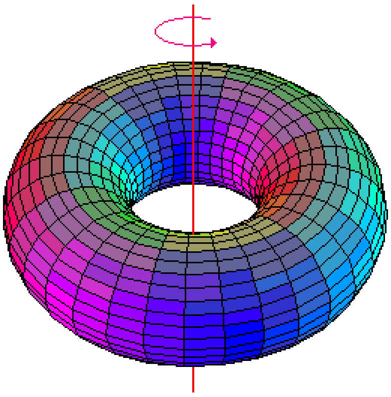
$$\begin{pmatrix} g_1 & = & IIIXXXX \\ g_2 & = & IXXIIXX \\ g_3 & = & XIXIXIX \\ g_4 & = & IIIZZZZ \\ g_5 & = & IZZIIZZ \\ g_6 & = & ZIZIZIZ \end{pmatrix}$$

Clearly, this description is way cleaner that the state vectors described previously. Also, we can see a similarity between the generators and the parity check matrices of C_1 and $C_2^\perp$ (here, $C_1 = C_2^\perp$. In fact, the first three generators of the stabilizer have Xs on the qubits corresponding to 1s in the parity check matrix for C_1 (remember, C_1 is used to correct bit flips) and the final three generators have Zs where the parity check matrix for $C_2^\perp$ had 1s (the second code is used for phase flips).

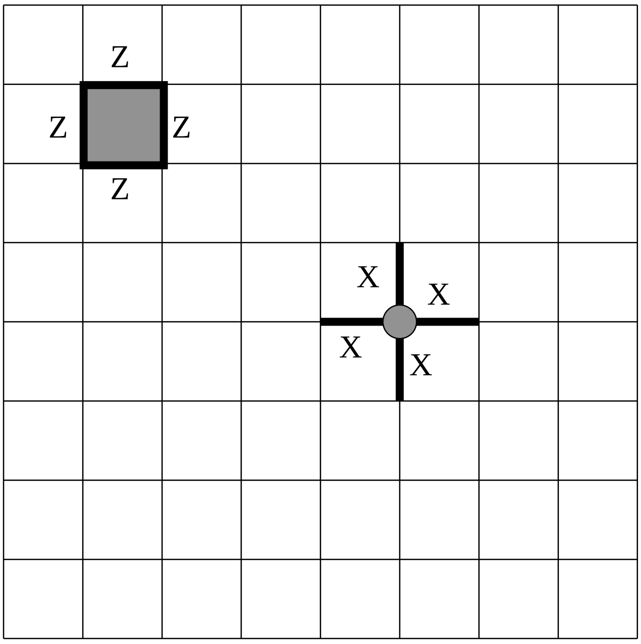
Now, we have the tools to define stabilizer codes: an $[n, k]$ stabilizer code is defined to be the vector space V_S stabilized by a subgroup S of the Pauli group G_n such that S has $n-k$ independent and commuting generators.The notion of "independent generators" is what we can guess: you can't decompose a generator as a combination of the other generators.

2.4 Surface codes

The easiest way to understand what is a surface code is to study one. For this, we will consider the toric code of Kitaev. Consider a torus divided in L^2 squares. We have therefore a graph (cf figure) where opposite sides are identified ("equals").



On this code, consider the edges to be qubits. There are $2L^2$ links in the lattice and hence, $2L^2$ qubits in the code block. .The faces (*plaquette*) are associated to the Pauli operator Z (a spin-flip) applied to the edges in contact and the identity for every other qubits and the vertex (*site*) to the Pauli operator X applied only to the edges in contact as well (cf Fig.).We define the dual graph as in graph theory, by swapping the roles of plaquettes and sites .

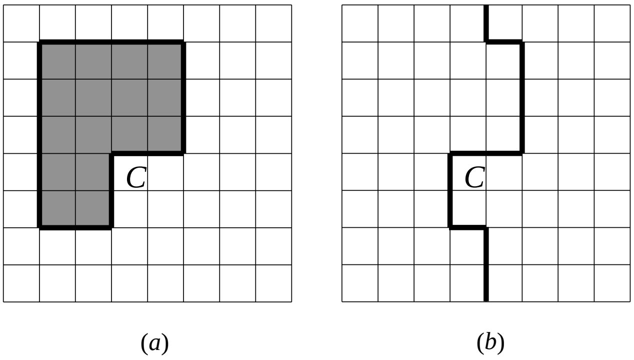


In fact, this code is a special case of a stabilizer code where a Pauli operator acting on n qubits is one of the 2^{2n} tensor product operators $\{I, X, Y, Z\}^{\otimes n}$
We can check that each elementary plaquette or site can be expressed as the product of the other $L^2 - 1$ same operators. With the fact that $X^2 = Z^2 = Id$, it is easy to be convinced that there are $2(L^2 - 1)$ independent check operators , and hence two encoded qubits.

We will need some notations here: a mapping that assigns to each link of the graph an element of $\{0,1\}$ is called a 1-chain. In an abuse of language, we will call 1-chain (or more simply chain) the links that are assign 1 by the mapping. We can similarly define 2-chains when the links are replaced by plaquettes of the graph and 0-chains when the links are replaced by sites. We can now define the boundary ∂ that takes 2-chains to 1-chains and 1-chains to 0-chains: the boundary of a plaquette is the sum of the four links touching the plaquette and the boundary of a link is the sum of the two sites of the link. If the boundary is trivial, the chain is said to be a cycle.

Now, remember that, for stabilizer codes, we want Pauli operators (which can be expressed as a tensor product of Xs and identities times a tensor product of Zs and identities) that commutes one with another. We will represent them as a mapping where qubits acted on by a Z or a X will be mapped to 1 . The condition of commutating operators is ,for example in the case of Zs, verified trivially with all of the plaquette chack operators and commutes with Xs if an even number of Zs act on the links adjacent to the site. By using the dual graph, we have the same result for the Xs. Therefore, a Pauli operator that coimmutes with the stabilizer of the code can be represented as a tensor product of Zs acting on a cycle of the lattice times a tensor product of Xs on a cycle of the dual lattice.

There are two kinds of cycles. A 1-cycle is homologically trivial if it can be expressed as the boundary of a 2-chain. The surface it defines can be tiled by plaquettes and this operator is a product of the check operators. It is contained in the code stabilizer and acts trivially (which means " no error there"). The same way, if a product of X 's acts trivially on the dual lattice, we are in the same case. Given the group structure of the stabilizer group, any element of it can be expressed as a product of Z's acting on a homological trivial cycle of the lattice, times the same thing with X's. (cf Fig)
But if a cycle is not the boundary of anything, it will still commute with the code stabilizer but will not be contained in it. Therefore, it acts non trivially on the encoded information.



We now want to know if we can correct an error made by a Pauli operator. For this, we define the weight of an operator to be the number of qubits on which non-trivial Pauli matrices act.Moreover, the distance of the toric code is the weight of the minimal-weight Pauli operator preserving the code subspace but acting non trivially within the code subspace. It is the number of lattice links contained in the shortest homologically non-trivial cycle, may it be on the graph or on the dual graph. In the case of the toric code, the dual graph is the same as the graph ; in the case of an $L \times L$ square lattice, the code distance is L .

References

[1] Nielsen and Chuang, *Quantum Computation and Quantum Information*, 2010

[2] Dalibard, *Notes de Cours de l'Ecole Polytechnique*, 2003

[3] Delfosse, *Constructions et performances de codes LDPC quantiques*, 2012