Foundations of Quantum Information

Contents

CHAPTER 1

The model

1. The first postulate

POSTULATE 1. *Associated to any isolated physical system is a complex Hilbert space known as the* state space *of the system. The system is completely described by its* state vector*, which is a unit vector in the state space.*

The simplest quantum mechanichal system is the *qubit*. It is the system whose associated vector space is a two dimensional Hilbert space.

We will use the notation $\{|0\rangle, |1\rangle\}$ for the canonical basis of $\mathbb{C}^2$.

Therefore, an arbitrary state for a qubit is a vector

$$|\varphi\rangle = a|0\rangle + b|1\rangle$$

with $a, b \in \mathbb{C}$, $|a^2| + |b^2| = 1$.

We will often think of a qubit as a system that can be in the situations $|0\rangle$ or $|1\rangle$. In that case, when $a \neq 0 \neq b$ we will say that the state is in a *superposition* of both situations.

2. The second postulate

POSTULATE 2. *The evolution of an isolated physical system (with associated Hilbert space H) is described by an* unitary transformation*. That is, if the state of the system at time t_1 is described by $|\varphi_1\rangle$ and the state of the system at $t_2 > t_1$ is described by $|\varphi_2\rangle$, then there exist a unitary operator $U \in B(H)$ such that*

$$|\varphi_2\rangle = U|\varphi_1\rangle.$$

3. The third postulate

POSTULATE 3. *In a given physical system with associated Hilbert space H,* quantum measurements *are described by a collection $\{M_n\}_n \subset B(H)$ of* measurement operators*. The index n refers to the measurement outcomes that may occur in the experiment. If the state of the quantum system is $|\psi\rangle$ immediately before the measurement then the probability that result n occurs is given by*

$$p(n) = \langle\psi|M_n^\dagger M_n|\psi\rangle,$$

and the state of the system after the measurement is

$$\frac{M_n|\psi\rangle}{\sqrt{\langle\psi|M_n^\dagger M_n|\psi\rangle}}.$$

Measurement operators satisfy

$$\sum_m M_n^\dagger M_n = I,$$

needed for the probabilities to sum to one.

One of the simplest examples is the measurement of a qubit in the computational basis. This is defined by the measurement operators

$$M_0 = |0\rangle\langle 0| \quad \text{and } M_1 = |1\rangle\langle 1|.$$

It is easy to see that both operators are selfadjoint, $M_i^\dagger M_i = M_i^2 = M_i$ and $M_0 + M_1 = \mathbb{1}$.

It is also easy to check that, when we measure the state $|\varphi\rangle = a_0|0\rangle + a_1|1\rangle$ the probability of obtaining measurement i is $|a_i|^2$ and the state after measurement in that case is $\frac{a}{|a|}|i\rangle$. We will see later that this is operationally equivalent to $|i\rangle$.

3.1. Distinguishability. One of the typical problems in quantum information will be to distinguish two (or more) quantum states from each other. That is, we have a particle in one of several possible states and we want to find out in which of them the particle actually is.

We will study this problem now in the simplest case: distinguish between two possible states.

Let us first assume that the states we want to distinguish, $|\varphi_1\rangle$ and $|\varphi_2\rangle$ are orthogonal. Then we can choose the measurement operators $M_i = |\varphi_i\rangle\langle\varphi_i|$ $(i = 1, 2)$ and $M_0 = \mathbb{1} - \sum_i M_i$. Note that all of these operators are projections. Then, if $|\varphi\rangle$ is prepared in the state $|\varphi_i\rangle$ then

$$p(i) = \langle\varphi|M_i|\varphi\rangle = 1.$$

Therefore, both states can be unambiguously distinguished.

Suppose now that we want to distinguish two states $|\varphi_1\rangle$ and $|\varphi_2\rangle$ which are not orthogonal. Let us prove that there is no way we can do that:

Assume there is a measurement $\{M_n\}_{n\in I}$ capable of distinguishing both states. In that case we must be able to decompose $I = I_1 \cup I_2$ disjointly so that we can decide that the state is $|\varphi_i\rangle$ if the result of the measurement is $n_0 \in I_i$. Consider then the operators $E_i = \sum_{n\in I_i} M_n^\dagger M_n$.

We must have

$$\langle\varphi_i|E_i|\varphi_i\rangle = 1.$$

Since $E_1 + E_2 = \mathbb{1}$, we get $\langle\varphi_1|E_2|\varphi_1\rangle = 0$. Since E_2 is positive, we can write

$$0 = \langle\varphi_1|E_2|\varphi_1\rangle = \langle\varphi_1|\sqrt{E_2}\sqrt{E_2}|\varphi_1\rangle,$$

hence $\sqrt{E_2}|\varphi_1\rangle = 0$.

Since $|\varphi_1\rangle$ and $|\varphi_2\rangle$ are not orthogonal, we know that there exist $\alpha \neq 0 \neq \beta$ and ψ orthogonal to φ_1 such that $|\alpha|^2 + |\beta|^2 = 1$ and

$$|\varphi_2\rangle = \alpha|\varphi_1\rangle + \beta|\psi\rangle.$$

Then we must have

$$\sqrt{E_2}|\varphi_2\rangle = \alpha\sqrt{E_2}|\varphi_1\rangle + \beta\sqrt{E_2}|\psi\rangle = \beta\sqrt{E_2}|\psi\rangle,$$

but this is a contradiction since

$$|\beta\sqrt{E_2}|\psi\rangle| \leq |\beta| < 1$$

and

$$|\sqrt{E_2}|\varphi_2\rangle| = 1.$$

3.2. POVM Measurements. In many chances, we will not be interested in the post-measurement state of our particle, but only in the probabilities of the different possible measurement outcomes. In these cases, it is often more convenient to follow the formalism of the so called *Positive Operator Valued Measurements* (POVM's):

Suppose a measurement $\{M_n\}_n$ defined as in Postulate 3. Then we can define the *positive* operators $E_n = M_n^\dagger M_n$. We have that $\sum_n E_n = \mathbb{1}$ and that the probability of obtaining outcome m is

$$p(m) = \langle\varphi|E_m|\varphi\rangle.$$

Conversely, whenever we have a collection of positive operators $\{E_n\}_n$ such that $\sum_n E_n = \mathbb{1}$ we can define the measurement $\{M_n\}_n$ where $M_n = \sqrt{E_n}$.

3.3. Projective Measurements. In many applications, we will be very interested in a special case of measurements called *projective measurements*. These are measurements $\{M_n\}_n$ as in Postulate 3 with the additional property that each the M_n's are orthogonal projections; that is, they are selfadjoint and verify

$$M_n M_m = \delta_{mn} M_n.$$

In this case, we can define an *observable* M as the Hermitian operator

$$M = \sum_n n M_n$$

With this notation, the average value of the measurement is

$$\sum_n np(n) = \sum_n n\langle\varphi|M_n^\dagger M_n|\varphi\rangle = \sum_n n\langle\varphi|M_n|\varphi\rangle = \langle\varphi|M|\varphi\rangle.$$

Conversely, if we consider a Hermitian operator M, we can consider its espectral decomposition and write it like

$$M = \sum_n \lambda_n P_n,$$

where each P_n is a projection onto an eigenspace.

3.4. Heisenberg uncertainty principle. Recall that for a Hermitian operator $M = \sum_n \lambda_n P_n$ considered as a measurement acting on a state φ, its average is

$$\langle M\rangle = \langle\varphi|M|\varphi\rangle.$$

Recall also that the standard deviation of this measurement is defined like $\Delta(M)$, where

$$(\Delta(M))^2 = \langle(M - \langle M\rangle)^2\rangle = \langle M^2 + \langle M\rangle^2 - 2M\langle M\rangle\rangle = \langle M^2\rangle - \langle M\rangle^2$$

Suppose A, B are selfadjoint operators in $B(H)$ and $|\varphi\rangle$ is a quantum state. Suppose that A, B do not commute and that we have

$$\langle\varphi|AB|\varphi\rangle = a + \imath b.$$

Then

$$\langle\varphi|[A,B]|\varphi\rangle = 2\imath b$$

and

$$\langle\varphi|\{A,B\}|\varphi\rangle = 2a.$$

Hence

$$|\langle\varphi|[A,B]|\varphi\rangle|^2 + |\langle\varphi|\{A,B\}|\varphi\rangle|^2 = 4|\langle\varphi|AB|\varphi\rangle|^2.$$

Moreover, by the Cauchy-Schwarz inequality we have

$$|\langle\varphi|AB|\varphi\rangle|^2 = |\langle\varphi A|B\varphi\rangle|^2 \leq \langle\varphi A|A\varphi\rangle\langle\varphi B|B\varphi\rangle|^2 = \langle\varphi|A^2|\varphi\rangle\langle\varphi|B^2\varphi\rangle.$$

Therefore we get

$$|\langle\varphi|[A,B]|\varphi\rangle|^2 \leq 4\langle\varphi|A^2|\varphi\rangle\langle\varphi|B^2\varphi\rangle.$$

If we consider now two observables C, D and we replace A, B for $C - \langle C\rangle$, $D - \langle D\rangle$ we get

$$\Delta C\Delta D \geq \frac{|\langle\varphi|[C,D]|\varphi\rangle|}{2}.$$

So, if C and D do not commute, the standard deviations of the probability distributions that we obtain when we measure many copies of a state φ with them can not be simultaneously arbitrarily small.

3.5. Global phase. We claimed previously that the states $|\varphi\rangle$ and $e^{\imath\theta}|\varphi\rangle$ were essentially equal. We can justify now that statement.

Assume that we measure both states with a measurement $\{M_n\}_n$. Then, the probability of outcome n is, in the second case,

$$\langle\varphi e^{-\imath\theta}|M_n^\dagger M_n|e^{\imath\theta}\varphi\rangle = \langle\varphi|M_n^\dagger M_n|\varphi\rangle,$$

therefore both states are operationally identical.

4. Fourth Postulate

POSTULATE 4. *The state space of a composite physical system is the tensor product of the state spaces of the component physical systems. Moreover, if system number i is prepared in the state $|\varphi_i\rangle$ then the composite system is in the state $|\varphi_1\rangle \otimes \cdots \otimes |\varphi_n\rangle$*

We will see later that this postulate allows us to modelize *entanglement*, a behavior that seems to be at the root of many of the most surprising phenomena in quantum mechanics.

For the moment, with the aid of this postulate, we prove that projective measurements are as universal as general measurements, for as long as we allow for the use of *ancilla systems*

Suppose we have a physical system with state space H, and we want to perform a measurement $\{M_n\}_{n\in I}$ in it. To do this only with projective measurements, we introduce an auxiliary system (ancilla system) with state space K, where K is a $|I|$-dimensional system with orthogonal basis $(|n\rangle)_{n\in I}$.

Let $|0\rangle$ be a fixed state of K. Let

$$U : H \otimes [|0\rangle] \longrightarrow H \otimes K$$

be defined by

$$U(\varphi 0\rangle) = \sum_{n\in I} M_n|\varphi n\rangle.$$

Let us see that U preserves inner products on $H \otimes [|0\rangle]$. Take $|\varphi\rangle$, $|\psi\rangle \in H$. Then

$$\langle\varphi 0|U^\dagger U|\psi 0\rangle = \sum_{i,j}\langle\varphi i|M_i^\dagger M_j|\psi j\rangle = \sum_i \langle\varphi i|M_i^\dagger M_i|\psi i\rangle =$$

$$= \sum_i \langle \varphi | M_i^\dagger M_i | \psi \rangle = \langle \varphi | \psi \rangle = \langle \varphi 0 | \psi 0 \rangle$$

It is an easy exercise now (do it!!) to see that in that case U can be extended to an unitary operator (which we also call U)

$$U : H \otimes K \longrightarrow H \otimes K$$

Now we consider the projective measurement in the composite system $H \otimes K$ given by the projections $P_n = \mathbb{1}_H \otimes |n\rangle\langle n|$. The state we consider for the composite system is $U|\varphi 0\rangle = \sum_n M_n \varphi n\rangle$.

In that case, the probability of outcome n taking place is

$$p(n) = \langle \varphi 0 | U^\dagger P_n U | \varphi 0 \rangle = \sum_{i,j} \langle \varphi M_i^\dagger \otimes i | \mathbb{1} \otimes |n\rangle\langle n | M_j \varphi \otimes j \rangle = \langle \varphi | M_n^\dagger M_n | \varphi \rangle$$

and the post-measurement state in that case is

$$\frac{P_n U |\varphi 0\rangle}{\sqrt{\langle \varphi 0 | U^\dagger P_n U | \varphi 0 \rangle}} = \frac{M_n |\varphi \times n\rangle}{\sqrt{\langle \varphi | M_n^\dagger M_n \varphi \rangle}},$$

both results are exactly the same as if we would have considered the measurement $\{M_n\}_n$ in the system H acting on the state $|\varphi\rangle$.

4.1. Joint measurements. In many cases, we will have a state in a joint system formed by two or more parties and the parties will separately measure their part of their state. The following exercise should help to understand the mathematical description of this situation.

EXERCISE 4.1. *Consider a composite system $H_A \otimes H_B$. Let $\{P_i^A\}_i$ be a measurement system in H_A and let $\{Q_j^B\}_j$ be as measurement system in H_B. Prove that $\{P_i^A \otimes \mathbb{1}_B\}_i$ $\{\mathbb{1}_A \otimes Q_j^B\}_j$ and $\{P_i^A \otimes Q_j^B\}_{i,j}$ are measurement systems in the joint system $H_A \otimes H_B$. The first of them describes the situation when Alice measures with $\{P_i^A\}_i$ and Bob does nothing, the second describes the situation when Bob measures with $\{Q_j^B\}_j$ and Alice does nothing and the third one describes the situation when Alice measures with $\{P_i^A\}_i$ and Bob measures with $\{Q_j^B\}_j$*

Now, describe mathematically the situation when first *Alice measures with $\{P_i^A\}_i$ and* then *Bob measures with $\{Q_j^B\}_j$, and viceversa.*

5. Density operator formalism

Before we introduce the density operator formalism, we will recall some notions from linear algebra and operator algebras.

5.1. Trace. let us recall the definition of trace of a matrix, and some of its properties. We assume our Hilbert space H is finite dimensional, with dimension d. The main ideas of the following reasonings extend to infinite dimensional spaces, but we will not see this in this course.

We recall first some notations and basic notions from linear algebra. We use M_d for the space of the $d \times d$-dimensional matrices.

DEFINITION 5.1. *Given $A \in M_d$, we define $A^\dagger = \overline{A^T}$, that is, the adjoint of the traspose matrix.*

DEFINITION 5.2. *We say that a matrix $U \in M_d$ is* unitary *if $UU^\dagger = U^\dagger U = \mathbb{1}$, where $\mathbb{1}$ denotes the identity matrix.*

Recall that the matrices associated to a change of basis (from one orthonormal basis to another) are unitary matrices.

We recall now the definition of trace of a matrix.

For every matrix $A \in M_d$ we can define its *trace* $tr(A)$ as the sum of the elements of the diagonal of A, that is,

$$tr(A) = \sum_{i=1}^{d} a_{ii}$$

.

EXERCISE 5.3. *Show that trace is linear and cyclic. That is, show that for any two matrices $A, B \in M_d$ and for any two $\alpha, \beta \in \mathbb{C}$,*

(1) $tr(\alpha A + \beta B) = \alpha tr(A) + \beta tr(B)$
(2) $tr(AB) = tr(BA)$

Note that it follows from cyclicity that the trace of a matrix is invariant by unitary transformations. That is, if $U \in M_d$ is an unitary matrix and $A, B \in M_d$ are two matrices such that $B = U^\dagger AU$, then

$$tr(B) = tr(U^\dagger AU) = tr(UU^\dagger A) = tr(\mathbb{1}A) = tr(A)$$

If we now consider an operator $T : H \longrightarrow H$ and A, B are the representing matrices of T with respect to two different orthogonal basis, we know from linear algebra that there exists an unitary matrix U such that $B = U^\dagger AU$.

Since, it that case, $tr(A) = tr(B)$, it follows that we can define the trace of T as the trace of any of its representing matrices. It is now very simple to see that, given a basis $\{|i\rangle\}_{i=1}^d$ of H,

$$tr(T) = \sum_{i=1}^{d} \langle i|A|i\rangle.$$

We will often use the following property of the trace. Let $|\varphi\rangle \in H$ be a norm one vector and consider the rank one operator $|\varphi\rangle\langle\varphi| : H \longrightarrow H$. Note that this operator is the projection onto the direction of $|\varphi\rangle$. Let $T \in B(H)$ be an arbitrary operator. We want to evaluate $tr(T|\varphi\rangle\langle\varphi|)$. To do this, first we extend $|\varphi\rangle$ to a basis $\{|i\rangle\}$ of H, where $|\varphi\rangle = |1\rangle$. Then

$$tr(T|\varphi\rangle\langle\varphi|) = \sum_i \langle i|T|\varphi\rangle\langle\varphi|i\rangle = \langle\varphi|T|\varphi\rangle,$$

where the last equality follows from the orthogonality of the basis.

5.2. Singular value decomposition and trace norm. (You do not need this subsection!!)

We recall the singular value decomposition (SVD) of a square matrix A.

Consider an $n \times n$ square matrix A. Then, the matrix $AA^\dagger$ is semidefinite positive. Therefore, there exists a unitary matrix U and a positive diagonal matrix Λ such that

$$AA^\dagger = U\Lambda U^\dagger$$

The elements $\lambda_1, \ldots, \lambda_n$ in the diagonal of Λ are the eigenvalues of $AA^\dagger$, and they verify $\lambda_i \geq 0$ for every $1 \leq i \leq n$. Then, we can define the diagonal matrix $\Sigma = \sqrt{\Lambda}$, that is, the diagonal matrix formed by the elements $\sigma_i = \sqrt{\lambda_i}$.

These elements $\sigma_1, \ldots, \sigma_n$ are called the *singular values* of A. We define a matrix $V^\dagger$ by $\Sigma^{-1}U^\dagger A$. Note that $V^\dagger$ is unitary, since

$$V^\dagger V = \Sigma^{-1}U^\dagger AA^\dagger U\Sigma^{-1} = \Sigma^{-1}U^\dagger U\Lambda U^\dagger U\Sigma^{-1} = \Sigma^{-1}\Lambda\Sigma^{-1} = I$$

The SVD decomposition of A is then

$$A = U\Sigma V^\dagger$$

.

Given $A = U\Sigma V^\dagger$ as above, we define the *trace norm* of A by

$$\|A\|_1 = \sum_i \sigma_i$$

EXERCISE 5.4. *Check that the trace norm is indeed a norm. Not totally obvious the triangle inequality. One possible path is the following. Prove that*

$$\|A\|_1 = \sup\{|tr(BA)|;\ B \in M_n(\mathbb{C}),\ \|B\| \leq 1\} \tag{1}$$

To do this, show that $|tr(BA)| \leq \|B\||tr(A)\| \leq \|B\|\|A\|_1$. *For the other inequality, choose* $B = VU^\dagger$, *where* $A = U\Sigma V^\dagger$ *is the SVD of* A.

Once (1) is proved, the triangle inequality follows fast.

DEFINITION 5.5. *Given a Hilbert space H, finite dimensional for simplicity, we define $S_1(H)$ as the normed space of the operators $T : H \longrightarrow H$, endowed with the trace norm.*

5.3. Positive operators. We recall the definition of positive operators.

DEFINITION 5.6. *Given a Hilbert space H, an operator $T : H \longrightarrow H$ is said to be* positive *if, for every $|\varphi\rangle \in H$, one has*

$$\langle\varphi|T|\varphi\rangle \geq 0$$

EXERCISE 5.7. *Given a finite dimensional Hilbert space H, an operator $T : H \longrightarrow H$ is positive if and only if its associated matrix (in any given basis) is semidefinite positive.*

EXERCISE 5.8. *Given a finite dimensional Hilbert space H, and an operator $T : H \longrightarrow H$, if T is positive then T is self-adjoint (also called Hermitian).* Hints*: Decompose $T = A + \imath B$, with A, B Hermitian. Prove that, for Hermitian operators $\langle\varphi|A|\varphi\rangle \in \mathbb{R}$ for every $|\varphi\rangle \in H$, and same for B. Now, use this property and the decomposition of T to show that $B = 0$ and, hence, $A = T$.*

We will use often that Hermitian operators (and, in particular, positive operators) are diagonalizable.

5.4. Density operator formalism. So far we have described the state of a physical system as a unit vector in the Hilbert space H. There is an equivalent description where states are no longer elements in the Hilbert space but trace class operators on it. This last description offers advantages in certain problems, specially (but not only) when dealing with real experiments and systems where noise is always present. We describe next this formalism.

The situation we often face is that we will not know that our system is in a state $|\varphi\rangle$, but rather we will know that our system is in one of the states $|\varphi_i\rangle$, with probability p_i respectively.

REMARK 5.9. *When you begin studying quantum information, it is very easy to confuse this notion with the previous notion of quantum superposition. Please observe that both notions are essentially different. If you have questions regarding this, please ask!*

Therefore, we would like to consider something like the "state"

$$\sum_i p_i|\varphi_i\rangle$$

but the problem is that this is not a state anymore, since it does not have norm one.

A way to circumvent this difficulty is to associate each state $|\varphi\rangle$ to the operator $|\varphi\rangle\langle\varphi| \in S_1(H)$, where $S_1(H)$ is the set of the positive operators $\rho : H \longrightarrow H$ with trace one. Following the previous notation, we will call *states* to these kind of operators $\rho : H \longrightarrow H$.

Now, if we have a state as in the previous situation we can describe it with the positive trace 1 operator

$$\rho = \sum_i p_i |\varphi_i\rangle\langle\varphi_i|.$$

ρ is called the *density operator* or *density matrix*.

To see that ρ is indeed positive note that, for any $\psi \in H$,

$$\langle\psi|\rho|\psi\rangle = \sum_i p_i \langle\psi|\varphi_i\rangle\langle\varphi_i|\psi\rangle = \sum_i p_i |\langle\varphi_i|\psi\rangle|^2 \geq 0$$

Conversely, assume we have a positive trace one operator $\rho \in B(H)$. Being positive, ρ is diagonalizable. That is, it admits an spectral decomposition

$$\rho = \sum_j \lambda_j |j\rangle\langle j|,$$

where the eigenvectors $|j\rangle$ are orthogonal, with norm 1, and the eigenvalues λ_j are real, positive and verify $\sum_j \lambda_j = 1$ (because ρ has trace 1 and trace is unitarily invariant).

Therefore, we can see the numbers λ_j as the probabilities of our system being in the state $|j\rangle$.

Note also that if we now have our system with probability q_j described by the density operator $\rho_j = \sum_i p_{ij} |\varphi_{ij}\rangle\langle\varphi_{ij}|$ then

$$\rho = \sum_j q_j \rho_j = \sum_{ij} q_j p_{ij} |\varphi_{ij}\rangle\langle\varphi_{ij}|$$

is again a density operator, and it describes our system.

Therefore, Postulate 1 now says

POSTULATE' 1. *Associated to any isolated physical is a complex Hilbert space known as the* state space *of the system. The state of the system is completely described by its* density operator*, which is a positive operator $\rho \in S_1(H) \subset B(H)$ with trace one. If the system is in the state ρ_i with probability p_i, then the density operator for the system is $\sum_i p_i \rho_i$.*

We will often use the notation *pure states* for the states of the form $|\varphi\rangle$ and *mixed states* for states of the form $\rho = \sum_i p_i |\varphi_i\rangle\langle\varphi_i|$.

The evolution of a system H is, like before, given by unitaries on H. To see how they act on $\rho = \sum_i p_i|\varphi_i\rangle\langle\varphi_i|$, just notice that if the system initially is in the state $|\varphi_i\rangle$ with probability p_i, then after the evolution given by U it will be in state $U|\varphi_i\rangle$ with probability p_i, hence the associated density operator will be

$$\sum_i p_i|\varphi_i U\rangle\langle U^\dagger\varphi_i| = U\left(\sum_i p_i|\varphi_i\rangle\langle\varphi_i|\right)U^\dagger = U\rho U^\dagger.$$

Therefore, the second postulate now says

POSTULATE' 2. *The evolution of an isolated physical system (with associated Hilbert space H) is described by an* unitary transformation. *That is, if the state of the system at time t_1 is described by ρ_1 and the state of the system at $t_2 > t_1$ is described by ρ_2, then there exist a unitary operator $U \in B(H)$ such that*

$$\rho_2 = U\rho_1 U^\dagger.$$

As for the measurements, suppose we measure with a measurement $\{M_n\}$ a mixed state $\rho = \sum_i p_i|\varphi_i\rangle\langle\varphi_i|$. If the initial state is $|\varphi_i\rangle$ then the probability of outcome n taking place is

$$p(n|i) = \langle\varphi_i|M_n^\dagger M_n|\varphi_i\rangle = tr(M_n^\dagger M_n|\varphi_i\rangle\langle\varphi_i|).$$

Therefore, the total probability of outcome n is

$$p(n) = \sum_i p(n|i)p_i = \sum_i p_i tr(M_n^\dagger M_n|\varphi_i\rangle\langle\varphi_i|) = tr(M_n^\dagger M_n\rho)$$

With similar reasonings, we can see that the post-measurement state of the system when outcome n has taken place is

$$\frac{M_n\rho M_n^\dagger}{tr(M_n\rho M_n^\dagger)}$$

. That is, our third postulate in this formalism reads

POSTULATE' 3. *In a given physical system with associated Hilbert space H,* quantum measurements *are described by a collection $\{M_n\}_n \subset B(H)$ of* measurement operators. *The index n refers to the measurement outcomes that may occur in the experiment. If the state of the quantum system is ρ immediately before the measurement then the probability that result n occurs is given by*

$$p(n) = tr(M_n^\dagger M_n\rho),$$

and the state of the system after the measurement is

$$\frac{M_n\rho M_n^\dagger}{tr(M_n\rho M_n^\dagger)}.$$

Measurement operators satisfy

$$\sum_m M_n^\dagger M_n = I,$$

needed for the probabilities to sum to one.

A simple consequence of the separate linearity of tensor products is

POSTULATE' 4. *The state space of a composite physical system is the tensor product of the state spaces of the component physical systems. Moreover, if system number i is prepared in the state ρ_i then the composite system is in the state $\rho_1 \otimes \cdots \otimes \rho_n$*

EXERCISE 5.10. *Suppose we have a composite system $H_A \otimes H_B$. Suppose system A is prepared in the state $\rho_A = \sum_i p_i \varphi_i\rangle\langle\varphi_i$ and system B is prepared in state $\rho_B = \sum_j q_j \psi_j\rangle\langle\psi_j$. Check that*

$$\rho_A \otimes \rho_B = \sum_{i,j} p_i q_j (\varphi_i \otimes \psi_j)\rangle\langle(\varphi_i \otimes \psi_j)$$

This is essentially trivial.

6. Partial trace

Suppose we have a composite physical system made up of the subsystems H_A and H_B. Then the Hilbert space associated to the global system is $H_A \otimes H_B$ and the state of the system is described by the density operator ρ_{AB}. Sometimes we need to describe the "A part" of our state. For that we define the partial trace tr_B as the linear operator

$$tr_B : S_1(H_A \otimes H_B) \longrightarrow S_1(H_A)$$

defined on elementary tensors by

$$tr_B(|a_1\rangle\langle a_2| \otimes |b_1\rangle\langle b_2|) = |a_1\rangle\langle a_2| tr(|b_1\rangle\langle b_2|),$$

where $tr(|b_1\rangle\langle b_2|)$ is the usual trace in H_B, hence equal to $\langle b_2|b_1\rangle$.

Let us see some examples of the action of the partial trace.

EXERCISE 6.1. *Suppose the simplest case, where $\rho_{AB} = \rho_A \otimes \rho_B$ and ρ_A, ρ_B are as in Exercise 0.5. Check that, in that case,*

$$tr_B(\rho_{AB}) = \rho_A tr(\rho_B) = \rho_A$$

Note that this is the result we would expect: if Alice and Bob have a non-entangled pair $\rho_A \otimes \rho_B$, then Alice's part must be ρ_A and Bob's part must be ρ_B

To see that things are in general not so simple, consider the EPR state

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

Its associated density operator is

$$\rho_{AB} = \left(\frac{|00\rangle + |11\rangle}{\sqrt{2}}\right)\left(\frac{\langle 00| + \langle 11|}{\sqrt{2}}\right) = \frac{|00\rangle\langle 00| + |11\rangle\langle 00| + |00\rangle\langle 11| + |11\rangle\langle 11|}{2}$$

Then

$$\rho_A = tr_B(\rho_{AB}) = \frac{tr_B(|00\rangle\langle 00|) + tr_B(|11\rangle\langle 00|) + tr_B(|00\rangle\langle 11|) + tr_B(|11\rangle\langle 11|)}{2} =$$

$$= \frac{|0\rangle\langle 0|\langle 0|0\rangle + |1\rangle\langle 0|\langle 0|1\rangle + |0\rangle\langle 1|\langle 1|0\rangle + |1\rangle\langle 1|\langle 1|1\rangle}{2} = \frac{|0\rangle\langle 0| + |1\rangle\langle 1|}{2} = \frac{\mathbb{1}}{2}$$

Note that in this case the original state ρ_{AB} is a pure state, about which we have maximal knowledge, whereas ρ_A is a mixed state, and indeed the identity, which is equivalent to not knowing anything!!

Let us show that the partial trace gives the "correct" statistics. Suppose we have a system A where we measure an observable $M = \sum_m mP_m \in B(H)$. If we consider now a composite system $A \otimes B$, then the corresponding observable in this system is $M \otimes \mathbb{1}_B = \sum_m mP_m \otimes \mathbb{1}_B$, in the sense that, for any $|\varphi\langle\otimes|\psi\rangle \in A \otimes B$, the probability of obtaining outcome m is

$$p(m) = \langle \varphi \otimes \psi | P_m \otimes \mathbb{1}_B | \varphi \otimes \psi \rangle = \langle \varphi | P_m | \varphi \rangle \langle \psi | \mathbb{1}_B | \psi \rangle = \langle \varphi | P_m | \varphi \rangle$$

which coincides with the probability of obtaining outcome m if we measure only in the system A.

Consider now again both systems A and B, and an observable M in the system A. Suppose we have the system in the state ρ_{AB} and we want to find a state ρ_A for system A which verifies that the average value we obtain when measure ρ_A with M coincides with the average value we obtain when we measure ρ_{AB} with $M \otimes \mathbb{1}_B$. That is, we want

$$tr(M\rho_A) = tr(M \otimes \mathbb{1}_{AB}\rho_{AB})$$

Note that, if $\rho_{AB} = |a_1\rangle\langle a_2| \otimes |b_1\rangle\langle b_2$, then $tr_B(\rho_{AB}) = |a_1\rangle\langle a_2|\langle b_2|b_1\rangle$ and

$$tr(M \otimes \mathbb{1}_{AB}\rho_{AB}) = tr(M \otimes \mathbb{1}_{AB}|a_1\rangle\langle a_2| \otimes |b_1\rangle\langle b_2) = tr_A(M|a_1\rangle\langle a_2|)tr_B(|b_1\rangle\langle b_2),$$

which is what we wanted.

It can also be proved that the partial trace is the only function verifying this.

7. Pauli matrices

We will often make use of the following matrices, which play a prominent role when dealing with qbits.

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

$$Y = \sigma_Y = \begin{pmatrix} 0 & -\imath \\ \imath & 0 \end{pmatrix}$$

$$Z = \sigma_Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

Note that they are all self adjoint and they all verify that their eigenvalues are ± 1. Therefore, they can be seen as the *observable* associated to the a ± 1 valued measurement in the corresponding basis of eigenvalues.

8. No-cloning

In classical information, one takes for granted that information can be copied, and this is an important fact in many areas (criptography, memories, transmission, etc). We show now that, in general, quantum information can not be copied. The result is called the no-cloning theorem.

THEOREM 8.1 (No-cloning). *There is no quantum operation that we can perform on a system (of dimension at least 2) that will duplicate the (quantum) state of that system.*

PROOF. We give two proofs. For the first proof, suppose such quantum operation exists. Call H_A the system whose state is going to be duplicated in another system H_B, with potentially the help of an ancilla system H_E. Then our quantum operation $U : H_A \otimes H_B \otimes H_E \longrightarrow H_A \otimes H_B \otimes H_E$ starts out initializing the system H_B and the ancilla in two states $|R\rangle$, $|M\rangle$, and, *for every* choice of a state $|\varphi\rangle \in H_A$ verifies

$$U(|\varphi RM\rangle) = |\varphi\varphi M(\varphi)\rangle$$

In that case, consider two orthogonal different states $|0\rangle$, $|1\rangle \in H_A$, and a superposition state $|\varphi\rangle = \alpha|0\rangle + \beta|1\rangle$. Then we have

$$\begin{aligned} U(|0RM\rangle) =& |00M(0)\rangle \\ U(|1RM\rangle) =& |11M(1)\rangle \end{aligned}$$

In that case, $U(|\varphi RM\rangle)$ must verify, on the one hand, being $|\varphi\rangle$ itself a state that can be cloned,

$$\begin{aligned} U(|\varphi RM\rangle) =&|\varphi\varphi M(\varphi)\rangle = (\alpha|0\rangle + \beta|1\rangle)(\alpha|0\rangle + \beta|1\rangle M(\varphi)\rangle = \\ &(\alpha^2|00\rangle + \alpha\beta(|01\rangle + |10\rangle) + \beta^2|11\rangle)M(|\varphi\rangle) \end{aligned}$$

and, on the other hand, it follows from linearity that

$$U(|\varphi RM\rangle) = \alpha U(|0RM\rangle) + \beta U(|0RM\rangle) = \alpha|00M(0)\rangle + \beta|11M(1)\rangle,$$

which clearly can not coincide with the expression above.

For a different proof, note that unitary operators are isometries, and, therefore, preserve scalar products. Therefore, for every $|\varphi\rangle, |\psi\rangle \in H_A$, we have

$$\begin{aligned} &\langle\varphi RM|\psi RM\rangle = \langle\varphi|\psi\rangle\langle R|R\rangle\langle M|M\rangle = \langle\varphi|\psi\rangle = \\ &\langle U(\langle\varphi RM)|U(\psi RM\rangle)\rangle = \langle\varphi|\psi\rangle^2\langle M(\varphi)|M(\psi)\rangle. \end{aligned}$$

which is not possible (think of the modulus of the corresponding complex numbers). □

9. Teleportation

Even though quantum states can not be cloned, they can be *teleported* without physically moving the particles on which they are supported.

The situation is as follows. Suppose Alice's system is a 2-dimensional Hilbert space H_A (a q-bit) prepared in the state $|\phi\rangle = a|0\rangle + b|1\rangle$, where $|0\rangle, |1\rangle$ is a fixed orthonormal basis.

Alice wants to *teleport* her state to Bob. That is, she wants Bob to have a state in exactly the same superposition between $|0\rangle$ and $|1\rangle$ that she has. In order to do this, she can *not* measure the state, since measuring it would destroy it.

Rather than this, we assume that Alice and Bob share an EPR pair $|EPR\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$. So, the whole system is $\mathbb{C}^2 \otimes \mathbb{C}^2 \otimes \mathbb{C}^2$ and is originally in the state

$$\frac{1}{\sqrt{2}}\left(a|0\rangle(|00\rangle + |11\rangle) + b|1\rangle(|00\rangle + |11\rangle)\right)$$

TO BE CONTINUED!

10. Bell inequalities

10.1. Correlations in EPR. Consider the state

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

Assume we give the first qubit to Alice and the second qubit to Bob. Suppose both Alice and Bob measure in the computational basis in their respective particles. Call P_0, P_1 to the projections in Alice's lab, and Q_0, Q_1 to the projections in Bob's lab. Then the projections for the composite system are

$$\sum_{i,j} P_i \otimes Q_j = \sum_{i,j} (P_i \otimes \mathbb{1}_B)(\mathbb{1}_A \otimes Q_j).$$

Then, necessarily Alice and Bob get the same result in their measurements.

EXERCISE 10.1. *Prove this last statement. That is, prove that if Alice and Bob share the state* $|\varphi\rangle = \frac{|00\rangle+|11\rangle}{\sqrt{2}}$ *and they each measure separately in the computatioal basis, then their outcomes will coincide with probability 1.*

To do this, first read Subsection 4.1 and do Exercise 0.1.

We forget now for a moment about quantum mechanics and we perform the following thought experiment. Charlie prepares two particles, in whatever way he wants, and he sends one of these particles to Alice and the other one to Bob. Upon receiving her particle, Alice flips a coin. If she gets heads, she measures property Q of the particle. We assume that this measurement can only take the two values ± 1. If she gets tails, then she measures property R, getting again a result of ± 1. Bob does the same with his particle, and let us call S, T to the properties he measures, again with possible outcomes ± 1. We assume that Alice and Bob can perform their measurements in a causally disconnected manner (that is, sufficiently simultaneously and far apart that the outcome of Alice's measurement can not influence Bob's measurement and viceversa). We assume Charlie can prepare similar pairs of particles once and again.

We consider now the number

$$RS + QS + RT - QT = (R + Q)S + (R - Q)T.$$

Clearly either $(Q + R)$ or $(R - Q)$ is 0, and therefore

$$QS + RS + RT - QT = \pm 2.$$

Call $p(q,r,s,t)$ to the probability (in a possible probability space of "hidden variables") that, for a given preparation of the pair of particles $Q=q$, $R=r$, etc. Then

$$\mathbf{E}(QS)+\mathbf{E}(RS)+\mathbf{E}(RT)-\mathbf{E}(QT)=\mathbf{E}(QS+RS+RT-QT)=$$

$$=\sum_{q,r,s,t}p(q,r,s,t)(qs+rs+rt-qt)\leq 2\sum_{q,r,s,t}p(q,r,s,t)=2$$

In order to describe the above situation in the quantum mechanical setting, let us first detail the calculus of bipartite correlations with observables when Alice and Bob each measure with a ± 1 valued measure. Call A^+,A^-,B^+,B^- to the corresponding hermitian positive operators verifying

$$A^++A^-=B^++B^-=\mathbb{1}$$

We have that $A^\pm\in B(H_A)$ and $B^\pm\in H_B$. To reflect the action of each of these measurements in the joint system, we must actually consider the operators $A^\pm\otimes\mathbb{1}_B$, $\mathbb{1}_A\otimes B^\pm\in B(H_A\otimes H_B$.

Then, the probability of Alice and Bob each obtaining $+1$ in their measurement is

$$\langle\varphi|A^+\otimes\mathbb{1}|\mathbb{1}\otimes B^+|\varphi\rangle=\langle\varphi|A^+\otimes B^+|\varphi\rangle,$$

and similarly for the other possible outcomes. Therefore, the expectation of the product of their outcomes is

$$\begin{aligned}\mathbb{E}(AB)&=\mathbb{P}(1,1)+\mathbb{P}(-1,-1)-\mathbb{P}(-1,1)-\mathbb{P}(1,-1)=\\&=\langle\varphi|A^+\otimes B^+|\varphi\rangle+\langle\varphi|A^-\otimes B^-|\varphi\rangle-\\&-\langle\varphi|A^-\otimes B^+|\varphi\rangle-\langle\varphi|A^+\otimes B^-|\varphi\rangle=\\&=\langle\varphi|A\otimes B|\varphi\rangle\end{aligned}$$

We now go back to quantum mechanics. First, we describe the above experiment with quantum mechanical resources. Alice's and Bob's joint system is described by a Hilbert space $H_A\otimes H_B$. The particles Charlie sends them are prepared in a state $|\varphi\rangle\in H_A\otimes H_B$.

When receiving their particle, Alice and Bob will measure the corresponding ± 1 valued property. This corresponds to choosing measurements Q^+,Q^-, etc

Assume Charlie prepares both particles in the state

$$|\varphi\rangle=\frac{|01\rangle-|10\rangle}{\sqrt{2}}$$

The first qubit goes to Alice, the second to Bob.

Alice then measures with the observables $Q = Z$, $R = X$ and Bob measures with observables $S = \frac{-Z-X}{\sqrt{2}}$, $T = \frac{Z_2-X_2}{\sqrt{2}}$. Then

$$\langle QS\rangle = \frac{1}{\sqrt{2}},\ \langle RS\rangle = \frac{1}{\sqrt{2}},\ \langle RT\rangle = \frac{1}{\sqrt{2}},\ \langle QT\rangle = -\frac{1}{\sqrt{2}},$$

hence

$$\langle QS\rangle + \langle RS\rangle + \langle RT\rangle - \langle QT\rangle = 2\sqrt{2}$$

Uppss!!!!!!!!!!!This is what is called a *violation of a Bell inequality.* It has been experimentally verified. Therefore, the universe can not be simultaneously *local* and *realistic.*

EXERCISE 10.2. *Check all the calculations above in full detail. In particular, check that the observables Q, R, S, T above defined are indeed ± 1 valued measurement systems, and check that the probabilities obtained are indeed the ones above stated.*

CHAPTER 2

Quantum criptography

Modern criptography is strongly based on RSA and other *public key* cryptosystems. Their big advantage is that they work over a public channel. Their main disadvantege is that their security is based on the computational weakness of the adversary. Most likely, the messages encoded with RSA and sent today will be easy to decipher with the computers that will exist in ten or twenty years.

Before public key criptosystems were invented, criptography used to be based on *private keys*.

The simplest, and most secure, private key criptosystem is the *one-time pad*, or Vernam cipher.

In this system, Alice and Bob share an n-bit secret key random string. Alices encodes her n-bit message adding it to the key. Then she sends it over a public channel and Bob decodes it substracting the key from the message.

If the key is known to be totally secret and it is used only once, then this protocol is provably totally secure.

The main difficulty with this protocol is the distribution of the keys.

In particular, since the one-time pad is secure only as long as the key is used just once, Alice and Bob need to share as many bits as the length of the message. Before the development of quantum criptography, the only known way to do this was that Alice and Bob meet in advance, share the key, and keep it for later use.

Quantum Key Distribution, QKD, allows Alice and Bob the creation of a *provably secure* key over a *public channel*. Therefore, it aims to keep the best of both worlds: provable security over a public channel. The main disadvantage of QKD systems so far is that they do not work (with security) over long distances and that, so far, they are slow creating the key. Both limitations should improve as technology does.

1. BB84

The protocol BB84 was designed by Bennett and Brassard in 1984. Using it, Alice and Bob can, using only a public (quantum) channel,

generate a private key, and they can prove that the mutual information between their key and a possible eavesdropper is arbitrarily small.

The only requirement will be that they can communicate qbits over a public channel with a low error rate, below certain threshold.

This is the main obstacle so far for practical implementations, since transmission of quantum bits along long distances is not easy.

In this notes, we will present the protocol BB84. We will show how it works and we will proceed as far as we can with the proof of its security.

The proof of security is not simple. When Bennet and Brassard presented their protocol, they proved it secure against a limited amount of attacks. Next proofs of security took more than 10 years to appear, and the most general proof have to wait over 20 years since the appearance of the protocol.

This means, that will not be able in this course to go through a full formal proof. But we will present rigorously several of the needed steps, and most of the relevant ideas.

The main idea behind the security of BB84, and QKD in general, is that *information gain implies disturbance*. That is, if our favorite eavesdropper, Eve, interacts with the sent qbits, either she gains no information or she modifies the qbits. In this last case, Alice and Bob will be able to detect this modification and will abort the protocol.

Before we prove formally that information gain implies disturbance, we remark that both the proof and the idea are very much related to the no-cloning theorem.

Note that the no-cloning theorem tells us that Eve cannot just intercept the communication, copy the state, keep it and send the original one to Bob. This could only be done if the states were known to be all in an orthogonal family. That is, if they were actually "only" classical information.

1.1. Information gain implies disturbance. Let $\phi, \psi \in H$ be two non-orthogonal quantum states about which Eve is trying to obtain information. That is, there will be a setting, which we could think that is going to be repeated many times, where one of ϕ, ψ will arrive to Eve, without she knowing a priori which of them will be. She can do with the received states any operation allowed by quantum mechanics and, afterwards, she has to send again the state. She also wants to extract some information about which state was, the probability of one of them appearing, or any other information. We will prove that she can not do this without modifying one, or both, of ϕ, ψ.

What could she do to obtain information?. The most general action she could do (see Section 8.2 in [**1**]) would be:

- Prepare an ancillary system in a state $|u\rangle \in E$. We denote by E the ancillary system.
- Unitarily evolve the joint system. That is, implement an unitary operator $U : H \otimes E \longrightarrow H \otimes E$. This unitary evolution will verify

$$U(|\psi\rangle \otimes |u\rangle) = |\psi'\rangle \otimes |u_\psi\rangle$$

$$U(|\phi\rangle \otimes |u\rangle) = |\phi'\rangle \otimes |u_\phi\rangle$$

- Then, Eve will send $|u_\psi\rangle$ or $|u_\phi\rangle$ to Bob and she keeps $|\psi'\rangle$, $|\phi'\rangle$ to gain information from them.

In order not to disturb the states, she needs $|u_\psi\rangle = |\psi\rangle$ and $|u_\phi\rangle = |\phi\rangle$. But unitary evolutions are isometries. Therefore, they preserve inner products. This means that the product

$$(\langle\phi| \otimes \langle u|)|(|\psi\rangle \otimes |u\rangle) = \langle\phi|\psi\rangle\langle u|u\rangle = \langle\phi|\psi\rangle$$

must coincide with

$$((\langle\phi'| \otimes \langle u_\phi|)|(|\psi'\rangle \otimes |u_\psi)\rangle = \langle\phi'|\psi'\rangle\langle u_\phi|u_\psi\rangle$$

If we use $|u_\psi\rangle = |\psi\rangle$ and $|u_\phi\rangle = |\phi\rangle$, we obtain

$$\langle\phi|\psi\rangle = \langle\phi'|\psi'\rangle\langle\phi|\psi\rangle. \tag{2}$$

Now, if $|\phi\rangle, |\psi\rangle$ are orthogonal, then $\langle\phi|\psi\rangle = 0$ and Equation 2 is always true.

But if $|\phi\rangle, |\psi\rangle$ are not orthogonal, Equation 2 holds if and only if $\langle\phi'|\psi'\rangle = 1$, which in turn implies that $|\phi'\rangle = |\psi'\rangle$ and, therefore, Eve can not obtain from them different information for the case $|\phi\rangle$ and $|\psi\rangle$.

Or, read the other way, if Eve wants to obtain information, she will need $|\phi'\rangle \neq |\psi'\rangle$, and she can only achieve this disturbing the states, that is, making $|u_\psi\rangle \neq |\psi\rangle$ or $|u_\phi\rangle \neq |\phi\rangle$.

Note that the reasonings above work whenver $|\phi\rangle, |\psi\rangle$ are not orthogonal. BB84, and, in general, QKD protocols, will rely on that fact. Alice will send non orthogonal states so that, if Eve tampers with them, she will disturb the states and Alice and Bob will be able to detect this and they abort the protocol.

Otherwise, that is, if the states have not been tampered with, they know no one has eavesdropped.

1.2. BB84 protocol. We start now the main part of this course, the study of the BB84 protocol.

First of all, we describe the protocol.

(1) Alice has two random bit strings $\mathbf{a}, \mathbf{b} \in \{0,1\}^{(4+\delta)n}$. We denote the i^{th} bits of $\mathbf{a}, \mathbf{b}$ by a_i, b_i. Alice selects $\mathbf{a}, \mathbf{b}$ in such way that each of the bits a_i, b_j is independent from the rest, and each has probability $\frac{1}{2}$ of being 0 or 1. In this expression, $\delta > 0$ is a parameter which we will fix later.

(2) She encodes the strings as a state

$$|\psi\rangle = \otimes_{i=1}^{(4+\delta)n} |\psi_{a_i,b_i}\rangle \in \otimes_{i=1}^{(4+\delta)n} \mathbb{C}^2$$

So $|\psi\rangle$ is a state made of $(4+\delta)n$ qbits. Each of these qbits is defined by

$$|\psi_{a_i,b_i}\rangle = \begin{cases} |0\rangle \text{ if } a_1 = 0,\ b_i = 0 \\ |1\rangle \text{ if } a_1 = 1,\ b_i = 0 \\ |+\rangle \text{ if } a_1 = 0,\ b_i = 1 \\ |-\rangle \text{ if } a_1 = 1,\ b_i = 1 \end{cases}$$

Remember that $|+\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$ and $|-\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$

What is Alice doing? She looks at the bit b_i to decide whether to use the canonical basis $\{|0\rangle, |1\rangle\}$ or the $\{|+\rangle, |-\rangle\}$ basis. In either case, she codifies bit a_i in the basis indicated by b_i

EXERCISE 1.1. *Check that* $\{|+\rangle, |-\rangle\}$ *is an orthogonal basis of* $\mathbb{C}^2$

EXERCISE 1.2. *Check that* $|\psi_{0,0}\rangle$, $|\psi_{1,0}\rangle$ *are orthogonal to each other and so are* $|\psi_{0,1}\rangle$, $|\psi_{1,1}\rangle$, *but* $|\psi_{i,0}\rangle$ *is never orthogonal to* $|\psi_{j,1}\rangle$.

(3) Alice sends $|\psi\rangle = \otimes_{i=1}^{(4+\delta)n} |\psi_{a_i,b_i}\rangle$ to Bob over a public (quantum) channel $\mathcal{E}$.

(4) Bob receives $\mathcal{E}(|\psi\rangle)$ and announces reception. Note that $\mathcal{E}(|\psi\rangle)$ could a priori be different from $|\psi\rangle$, because of noise in the channel and/or tampering from Eve.

Note, since Bob does not know in which basis to measure, he, or Eve had she received the bits, can not yet do much with the received qbits.

In the next two exercises, suppose for the sake of simplification that Bob (or Eve) received exactly $|\psi\rangle$.

EXERCISE 1.3. *Check that if Bob measures qbit $|\psi_i\rangle$ in the basis indicated by b_i, he obtains bit a_i with probability 1. Check also that if he measures $|\psi_i\rangle$ in the basis not indicated by b_i, then he obtains a_i with probability $\frac{1}{2}$ (and $a_i \oplus 1$ with probability $\frac{1}{2}$).*

This exercise is very simple and totally crucial for an understanding of BB84. Do it with full detail. If you have any doubt, please ask!!

(5) Bob measures the state he received. To do this he generates first a uniformly random bit chain $\mathbf{b}' \in \{0,1\}^{(4+\delta)n}$, and then he measures each of the qbits $|\psi_i\rangle$ in the basis indicated by b_i'. Let us call $\mathbf{a}' \in \{0,1\}^{(4+\delta)n}$ to the bit chain Bob obtains.

(6) Bob announces he has finished measuring. He does not send $\mathbf{a}'$.

(7) Now, Alice announces $\mathbf{b}$ and Bob announces $\mathbf{b}'$. In (aproximately) half of the positions $b_i \neq b_i'$. The qbits and measurements corresponding to those positions are discarded and they are not used any more in the rest of the protocol.

They keep the other half of the bits. In those, if there were no transmission errors and no tampering from Eve, we would have always $a_i = a_i'$. That is, Alice and Bob would be sharing a random $2n$-bit chain.

But:

-Typically quantum channels induce errors

-Eve might have eavesdropped, and she would have perturbed the bit chain.

The goal of Alice and Bob now is to remove errors (this is called *error correction*) and to reduce the mutual information between Eve and the shared chain of bits they will effectively use. This is called *privacy amplification*.

Assume the length of the not removed part of the chain is at least $2n$. We will see later that δ is chosen to make sure this happens with high enough probability. Alice and Bob now want to test how much noise, or tampering, they have in their useful bits.

(8) Alice chooses n out of the $2n$ bits, and announces which ones she chose.

(9) Alice and Bob both announce over a public channel the bits they have in those positions. If there are more than t not coincident bits, they abort. The parameter t will be chosen later.

(10) If they did not abort, Alice and Bob do *information reconciliation* and *privacy amplification* on the remaning n bits to obtain a string of $m \leq n$ bits of shared key for which the mutual information with Eve is as small as desired.

Let us see an example of a possible transmission and attack of Eve.

EXAMPLE 1.4. *Suppose Alice generates* $\mathbf{a} = 01010100$ *and* $\mathbf{b} = 11011001$. *Suppose Eve intercepts the communication, measures each of the qbits in a given basis. Say the string of Eve's basis is* $\mathbf{l} = 11000011$. *She measures each qbit with the corresponding basis and sends along the postmeasurement state. Bob, receives the qbits resent by Eve and he measures with the basis indicated by the random chain* $\mathbf{b}' = 11000101$.

Let us analyze the results. Whenever Eve measures in the same basis as Alice used to encode the qbit, she obtains the encoded qbit with probability 1 and sends along that same qbit. Therefore, her presence goes undetected. But, whenever she measures in the wrong basis, she obtains the right bit with probability $\frac{1}{2}$, *and sends along a qbit in this wrong basis.*

We will use the notation $(0,1)$ *(respectively* $(+,-)$*) to denote that an agent obtains* $|0\rangle, |1\rangle$ *(resp.* $|+\rangle$, $|-\rangle$*) with probability* $\frac{1}{2}$.

Encode $\mathbf{a} = 01010100$ *with* $\mathbf{b} = 11011001$ *and note that Alice will be sending*

$$|+\rangle|-\rangle|0\rangle|-\rangle|+\rangle|1\rangle|0\rangle|+\rangle$$

When Eve measures in the basis indicated by $\mathbf{l} = 11000011$, *she will get*

$$|+\rangle|-\rangle|0\rangle(0,1)(0,1)|1\rangle(+,-)|+\rangle$$

and will send this along to Bob.

When Bob now measures with the basis indicated by $\mathbf{b}' = 11000101$ *he will get*

$$|+\rangle|-\rangle|0\rangle(0,1)(0,1)|1\rangle(0,1)|+\rangle$$

You can see that in position 7, Bob would have got $|0\rangle$ *with probability 1, but the action of Eve causes him to obtain* $|0\rangle, |1\rangle$ *with probability* $\frac{1}{2}$.

In general, it is easy to see (do it!) that if Eve follows this strategy (measure with a randomly selected basis the intercepted qbits and send to Bob the postmeasurement state), the the probability of Eve choosing the wrong basis is $\frac{1}{2}$, the probability of Eve introducing an error in Bob's measurement is $\frac{1}{4}$ and the probability of Eve learning a correct

bit is $\frac{1}{2}$. So, she introduces one error for each two bits she learns, and that way she can be detected.

This is the basic idea. The formalization of the fact that, not matter what strategy Eve follows, this same idea will remain essentially true, is not easy. In our search for a proof, we will visit important notions in Quantum Information Theory, together with some facts on (classical) linear error correcting codes.

We leave as an exercise the theoretical justification for step (7) in the protocol. It is a non trivial probability theory exercise.

EXERCISE 1.5. *Consider two strings* $\mathbf{a}$, $\mathbf{b}$ *of* $2n$ *bits. Suppose that, for each position in the string, the probability that* $a_i = b_i$ *is the same. We choose randomly* n *bits of the string and we find that they differ in* μn *bits, with* $0 < \mu < 1$. *Prove that the probability that the remaining* n *bits differ in more than* $(\mu + \epsilon)n$ *bits is smaller than ??*

2. BB84 a la EPR

In the previous section we have seen the BB84 protocol. In it, the roles of Alice and Bob are apparently different, and do not seem interchangeable. In this section we present another protocol, based on shared EPR pairs, which is totally equivalent to BB84 but the roles of Alice and Bob are now clearly symmetric.

We describe the protocol.

(1) In this new protocol, there is a third party, Charlie, which prepares $(4+\delta)n$ EPR pairs $\frac{|00\rangle+|11\rangle}{\sqrt{2}}$ and, for each pair, sends one of the particles to Alice and the other one to Bob.
(2) Alice and Bob receive their particles and they announce so.
(3) They could have a longer chain (say $8n$ long) and use half of the chain to check that they maximally violate CHSH. This guarantees that they do indeed have entangled EPR pairs. We will not see this in this course.
(4) Each of them chooses randomly a $(4+\delta)n$ bit string. We will call, as before, $\mathbf{b}$ to Alice's string and $\mathbf{b}'$ to Bob's string.
(5) Now, each of them measures his/her part of the EPR states. In the ith pair, each of them measures with the basis indicated by b_i and b'_i. That is, Alice measures her ith qbit in the $\{|0\rangle, |1\rangle\}$ basis if $b_i = 0$ and she measures it in the $\{|+\rangle, |-\rangle\}$ basis if $b_i = 1$. Bob does the same with b'_i.
(6) After measuring, they both announce they have measured. Next, they both announce their bit strings $\mathbf{b}$ and $\mathbf{b}'$. The positions i where $b_i \neq b_i$ are discarded and those bits are not used any more during the protocol

(7) In the rest of the positions, if there was no noise in the channel and no eavesdropper, they share exactly the same bits.

EXERCISE 2.1. *In Exercise 0.8 you proved that if Alice and Bob share an EPR state and they each measure separately in the computational basis, then their outcomes will coincide with probability 1. Prove now that this is also the case if they mesure separately in the* $\{|+\rangle, |-\rangle\}$ *basis. Prove also that actually the result is also true no matter in which orthogonal basis they measure (for as long as they both use the same one). Or, put in another way, prove that when Alice measures with an orthogonal basis and obtains a result, Bob's state collapses to that same state. This phenomenom is called* steering*: Alice's measurement steers Bob's state.*

Note that, in this point, the situation is indistisguishable from the situation after point (7) in the BB84 protocol.

CHAPTER 3

Classical error correcting codes

Alice and Bob will have to do error correction in their transmission. It is not obvious that quantum error correction can be done. We will see that it indeed can be done. In order to understand how, we need to understand first classical error correction, in particular classical linear error correction.

Let us forget for the moment about quantum mechanics, and consider the situation of classical communication. We all know that transmission channels are physically imperfect and can induce errors in the communication. To deal with those, there is a large amount of error correcting techniques, adapted to the different transmissions.

How one can correct errors?. First option is redundance. This is essentially brute force: Alice wants to send 000 and she sends 000000. If B receives 000010, for instance, he knows there was a mistake, but he can not correct it.

To solve this, we could add more redundance: Alice sends now 000000000. If Bob receives, for instance, 000010000, he knows there was a mistake and he can correct it (assuming that we know that, at most, there was one mistake).

This is brute force. Since error correction is very necessary, more efficient methods have been developed. In our case, we will be interested in linear error correcting codes, since they can be adapted to quantum error correction.

1. Classical linear codes

.

We will be interested now in *linear* error correcting codes. Linearity will arise from a vector space structure.

Typically in undergrad linear algebra courses you have studied vector spaces over the field of the real numbers, or maybe the complex numbers. Here, our scalar field is going to be $\mathbb{Z}_2 = \{0, 1\}$ with the usual product and addition mod 2.

An n-dimensional vector space over $\mathbb{Z}_2$ will always be (isomorphic to) $\mathbb{Z}_2^n$. Think of $\mathbb{Z}_2$ as a bit, and an element (vector) in $\mathbb{Z}_2^n$ is an n-bit word.

We state some definitions and notation: A code of length n is a subset C of $\mathbb{Z}_2^n$. The elements in $\mathbb{Z}_2^n$ will be called *words* (of length n) and the elements in C are the *code words*.

If C is a code using n-bit words to codify every possible k-bit words ($k \leq n$) we say that C is an $[n, k]$ code.

If C is a vector subspace of $\mathbb{Z}_2^n$, we say the code is *linear*.

From now on, we will only be interested in linear codes.

A brief remark on words and dimensions: $\mathbb{Z}_2^n$ has 2^n elements, and dimension n. If C is going to codify k-bit words, C will have 2^k elements, which results in dimension k.

EXAMPLE 1.1. *The binary repetition code of length 6. In order to send* 0*, Alice sends* 000000*. To send* 1*, she sends* 111111*. Therefore,* $C = \{000000, 111111\}$*, which is a linear subspace (check it!). Therefore, it is a* $[6, 1]$ *code. In this case, error correction is majority vote: If Bob receives* 010010 *it decodes it as* 0*. It is clear that* C *can decode correctly up to 2 errors per word.*

We see now a much more interesting linear code, which will serve as model to understand much of what follows.

EXAMPLE 1.2. *The* $[7, 4]$*-Hamming code. It will use words of length 7 to codify all words of length 4. The code is*

$$C = \{(a_1, a_2, \ldots, a_7) \in \mathbb{Z}_2^7 \text{ that verify the equations below }\}$$

$$\begin{aligned} a_1 + a_2 + a_3 + a_5 &= 0 \\ a_1 + a_2 + a_4 + a_6 &= 0 \\ a_2 + a_3 + a_4 + a_7 &= 0 \end{aligned}$$

What this does, it codifies the 4 bit strings in the first 4 bits, and the three bits left are parity checks.

C is clearly linear. Why? Because its defined via linear restrictions over a vector space.

We apply now linear algebra results to our linear codes. We will work in the canonical basis $\{e_1, \ldots, e_n\}$ of $\mathbb{Z}_2^n$. Since $C \subset \mathbb{Z}_2^n$ is a k-dimensional subspace, it admits a basis $\{\mathbf{r_1}, \ldots, \mathbf{r_k}\}$, where, for every i, $\mathbf{r_i} = (r_{i,1}, \ldots, r_{i,n})$.

We consider the matrix

$$G = \begin{pmatrix} r_{1,1} & r_{1,2} & \cdots & r_{k,1} \\ r_{1,2} & r_{2,2} & \cdots & r_{k,2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1,n} & r_{2,n} & \cdots & r_{k,n} \end{pmatrix}$$

This is the matrix generating the code, in the following sense. Consider G as describing an operator

$$G : \mathbb{Z}_2^k \longrightarrow \mathbb{Z}_2^n$$

acting by columns. Note that, for every $1 \leq i \leq k$, $G(e_i) = \mathbf{r_i}$. Then, if we want to codify a word $\mathbf{x} \in \mathbb{Z}_2^k$, we calculate $G(\mathbf{x})$

EXERCISE 1.3. *Prove that the previous mapping sends each word in $\mathbb{Z}_2^k$ to a code word in $C \subset \mathbb{Z}_2^n$ in a bijective way. In particular, different words go to different words.*

EXERCISE 1.4. *Prove that for the binary repetition code of length 3 (sending 0 to 000 and 1 to 111), the generating matrix is*

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

EXERCISE 1.5. *Calculate the generating matrix G for the $[7, 4]$-Hamming code.*

As a side remark, note that the generating matrix allows us to describe the code with only the kn bits of G, as opposed to the $n2^k$ bits that could be needed in general.

Also, codifying is very efficient: only $O(kn)$ operations are needed to codify a k-bit message.

But the real reason for us to be interested in linear codes comes from the error correction mechanism.

C is a linear subspace of $\mathbb{Z}_2^n$. Recall from your linear algebra courses that in order to think of a linear subspace you can consider a generating set, equivalently, you can think of C as the image of a linear mapping. This approach leads to the generating matrix G. But you can also think of C as the subspace verifying a set of linear restrictions, equivalently, as the kernel of a linear mapping

$$H : \mathbb{Z}_2^n \longrightarrow \mathbb{Z}_2^{n-k}$$

In this case, C is described by $n - k$ linear restrictions, and can be written as

$$C = Ker(H) = \{\mathbf{x} \in \mathbb{Z}_2^n \text{ such that } H(\mathbf{x}) = 0\}$$

Therefore, we have two descriptions of the linear code C: as the image of G or as the kernel of H. One passes from one description to the other exactly the same way as you used to do in undergrad courses:

If you know H and you want to find G, you need k linearly independent vectors $\mathbf{r_1}, \ldots, \mathbf{r_k}$ such that $Ker(H) = [\mathbf{r_1}, \ldots, \mathbf{r_k}]$. They are the columns of the matrix G.

Coversely, if you know G, its columns are the vectors $\mathbf{r_1}, \ldots, \mathbf{r_k}$ generating C. We now look for $\mathbf{s_1}, \ldots, \mathbf{s_{n-k}}$ orthogonal to them, in the sense that $\langle \mathbf{r_i} | \mathbf{s_j} \rangle = 0$. Then, $\mathbf{s_1}, \ldots, \mathbf{s_{n-k}}$ are the rows of H.

EXERCISE 1.6. *Check this last paragraphs*

EXERCISE 1.7. *Calculate H for the $[7, 4]$-Hamming code.*

1.1. Error correction using H. Let us see how we use H for error correction.

Alice encodes her message x as $y = G(x)$, and sends it to Bob. There is noise and/or eavesdroppers in the channel, and Bob receives y', which we may write as

$$y' = y + (y' - y) = y + e$$

e is the error. Note that Bob does not know y nor e, he only knows y'.

Then, Bob applies H to y'. Since $y \in C = ker(H)$, Bob obtains

$$H(y') = H(y) + H(e) = H(e)$$

We call $H(e)$ the error syndrome. It is the footprint that tells us that there was an error.

Now, we want to use $H(y') = H(e)$ to calculate e and, therefore, to calculate $y = y' - e$.

To understand the situation, let us first consider the simplest case: imagine that for some reason we can assume that there was at most an error in one bit. Then, if there was no mistake, $H(y') = H(y) = 0$. If there was a mistake in bit j, then we have that $y' = y + e_j$, where $e_j = (0, \ldots, 0, 1, 0, \ldots, 0)$ with the 1 in the jth position. In that case $H(y') = H(e_j)$. If we chose our code in such way that, for every $i \neq j$, $e_j - e_i \notin C$, then $H(e_j - e_i) \neq 0$ and, therefore, $H(e_i) \neq H(e_j)$. So, if we look at $H(y') = H(e_j)$ we can recover e_j and, hence, y.

This was just a simple analysis of a very simple situation. For the general case, we need to introduce the notion of Hamming distance.

DEFINITION 1.8. *Given two n-bit strings* $\mathbf{x}, \mathbf{y}$*, we define its Hamming distance as*

$$d_H(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^{n} |x_i - y_i|$$

We define the Hamming weight of $\mathbf{x}$ *as*

$$W_H(\mathbf{x}) = d_H(\mathbf{x}, \mathbf{0})$$

Note that, since we are working in $\mathbb{Z}_2^n$, we do not actually need the absolute value. The Hamming distance measures how many corresponding bits differ in $\mathbf{x}$ and $\mathbf{y}$. The Hamming weight tells how many 1's are there in $\mathbf{x}$.

EXERCISE 1.9. *Show that*

$$d_H(\mathbf{x}, \mathbf{y}) = W_H(\mathbf{x} + \mathbf{y})$$

For the following, note that, in general, we may always assume that the probability of a bit flip is less than $1/2$.

EXERCISE 1.10. *Why??*

Very often we can also assume that the probabilities of bit flips in different bits are independent from each other.

EXERCISE 1.11. *Suppose your (classical) channel is such that the probability of a bit flip in each separate bit is less than* $1/2$ *and that the the probabilities of bit flips in different bits are independent from each other. Prove that, in that case, if Bob receives* y'*, then the most likely word Alice sent is the* y *minimizing*

$$W_H(e) = W_H(y - y') = d_H(y, y').$$

In general, calculating y can be not computationally demanding. We will see later that the special choice of Hamming codes makes it easier, and useful for quantum error correction purposes.

We will need to define one important parameter associated to our code C.

DEFINITION 1.12. *We define the* distance *of* C *by*

$$d(C) = \min_{\mathbf{x},\mathbf{y} \in C, \mathbf{x} \neq \mathbf{y}} d_H(\mathbf{x}, \mathbf{y}) = \min_{\mathbf{x},\mathbf{y} \in C, \mathbf{x} \neq \mathbf{y}} W_H(\mathbf{x} + \mathbf{y}) = \min_{\mathbf{z} \in C, \mathbf{z} \neq 0} W_H(\mathbf{z})$$

In an $[n, k]$ code has distance d, we say it is an $[n, k, d]$ code.

EXERCISE 1.13. *Prove that a code with distance* d *can*

(1) *Detect* $d - 1$ *errors.*
(2) *Correct* $\frac{d-1}{2}$ *errors.*

EXERCISE 1.14. *Let C be a linear code, H the matrix such that $C = ker(H)$. Prove that if C has d columns which are linearly dependent but every choice of $d-1$ column are linearly independent, then $d(C) = d$.*

1.2. Hamming Codes. Given $r \geq 2$ (and consider the special case r=3) define the $[2^r - r - 1, 2^r - 1]$ Hamming code ($[4, 7]$ if $r = 3$) by means of its matrix H:

H is the $(r \times 2^r - 1)$ matrix whose columns are all the r-bit strings not constantly 0.

In the case $r = 3$, the matrix is

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

EXERCISE 1.15. *Prove that the distance of every Hamming code is 3.* Hint*: Use Exercise 1.14.*

Therefore, a Hamming code can detect two errors and correct 1.

Suppose we are using a Hamming code and there was en error in bit j. Then, the syndrome $H(e_j)$ is the binary representation of j.

1.3. Dual code. Given a linear code C, we define its dual code $C^\perp$ as the set of vectors perpendicular to all codewords in C,

$$C^\perp = \{\mathbf{v} \in \mathbb{Z}_2^n \text{ such that } \mathbf{v} \cdot \mathbf{c} = 0 \text{ for every } \mathbf{c} \in C\}$$

Here, and in the following, we will use $\cdot$ to refer to the inner product in $\mathbb{Z}_2^n$, to distinguish from the inner product $\langle | \rangle$ in $\mathbb{C}^n$.

EXERCISE 1.16. *Show that $C^\perp$ is a linear subspace of $\mathbb{Z}_2^n$ with $dim(C^\perp) = n - k$, where $k = dim(C)$.*

CHAPTER 4

Quantum error correcting codes

In this chapter, we use of the classical linear error correcting codes of the previous chapter to implement quantum error correction.

1. Quantum codes

One classical bit is an element of $\mathbb{Z}_2$, whereas a qbit is an element of $\mathbb{C}_2$. An n bit string is an element in $\mathbb{Z}_2^n$ and n-qbits is an element in $\mathbb{C}_2^{\otimes n} = \mathbb{C}^2 \otimes \ldots \mathbb{C}$. We are going to use the notation $\mathbb{C}_2^{\otimes n} = \mathbb{C}_2^n$. Lead by the analogies above, our road map for the construction of quantum codes is to replace $\mathbb{Z}_2$, $\mathbb{Z}_2^n$ in linear codes by $\mathbb{C}_2$, $\mathbb{C}_2^n$ in quantum codes.

Therefore, similarly to $[k, n]$ classical linear codes being k-dimensional subspaces C of $\mathbb{Z}_2^n$, we define a $[k, n]$ quantum code to be a 2^k-dimensional subspace of $\mathbb{C}_2^n$. This can also be seen as an isometry $Q : \mathbb{C}_2^k \longrightarrow \mathbb{C}_2^n$.

We define now the quantum codes we will be interested in. Suppose we have a linear code $C \subset \mathbb{Z}_2^n$. We define the linear code H_C as the subspace of $\mathbb{C}_2^n$ generated by each of the elements in C, when viewed in $\mathbb{C}_2^n$.

An example will help: suppose C is the subspace generated by $\{(1000), (0100), (0010)\}$. Then the codewords are

$$\{(0000), (0010), (0100), (0110), (1000), (1010), (1100), (1110)\}$$

and, then we define H_C as the 8 dimensional subspace in C_2^4 generated by

$$\{(|0\rangle|0\rangle|0\rangle|0\rangle), (|0\rangle|0\rangle|1\rangle|0\rangle), (|0\rangle|1\rangle|0\rangle|0\rangle), (|0\rangle|1\rangle|1\rangle|0\rangle),$$
$$(|1\rangle|0\rangle|0\rangle|0\rangle), (|1\rangle|0\rangle|1\rangle|0\rangle), (|1\rangle|1\rangle|0\rangle|0\rangle), (|1\rangle|1\rangle|1\rangle|0\rangle)\}$$

EXERCISE 1.1. *Prove that those states are pairwise orthogonal.*

Important: Note that this implies that $dim(H_C) = 2^{dim(C)}$. Please, note that one of the dimensions is on a vector space over the complex numbers, and the other one on a vector space over $\mathbb{Z}_2^n$.

We can also think of the code as the isometry

$$Q : \mathbb{C}_2^k \longrightarrow \mathbb{C}_2^n$$

taking the elements of the canonical basis of : $\mathbb{C}_2^k$ to the above mentioned vectors generating H_C.

Suppose G is the matrix generating C as in the previous chapter. That is $G(v)$ ranges over all the codewords in C as v ranges over all the words in $\mathbb{Z}_2^{dim(C)}$.

We will denote by $|Gv\rangle$ the quantum state associated to the codeword $G(v) \in C$. These are our *quantum codewords*. For every $w \in \mathbb{Z}_2^n$ we consider now the following superposition of the quantum codewords:

$$|c_w\rangle = \frac{\sum_{v \in \mathbb{Z}_2^{dim(C)}} (-1)^{w \cdot G(v)} |Gv\rangle}{2^{\frac{dim(C)}{2}}}$$

EXERCISE 1.2. *Prove that if $w_1 + w_2 \in C^\perp$, then $|c_{w_1}\rangle = |c_{w_2}\rangle$. Hint: Prove first that, in the hypothesis, $w_1 \cdot G(v) = w_2 \cdot G(v)$ for every $v \in \mathbb{Z}_2^{dim(C)}$*

EXERCISE 1.3. *Prove that if $w_1 + w_2 \notin C^\perp$, then $\langle c_{w_1} | c_{w_2} \rangle = 0$. Hint: Show first that, if there exists $v \in \mathbb{Z}_2^{dim(C)}$ such that $w \cdot G(v) \neq 0$, then*

$$\sum_{v \in \mathbb{Z}_2^{dim(C)}} (-1)^{w \cdot G(v)} = 0$$

Consider a k-dimensional classical linear code $C \subset \mathbb{Z}_2^n$. Consider its orthogonal complement $C^\perp$, with dimension $n - k$. Recall from linear algebra the notion of the quotient space $\mathbb{Z}_2^n / C^\perp$. The elements in $\mathbb{Z}_2^n / C^\perp$ are the equivalence classes $w + C^\perp$ in $\mathbb{Z}_2^n$. Recall also from linear algebra that $\mathbb{Z}_2^n / C^\perp$ is a linear space of dimension $n-(n-k) = k$. Therefore $\mathbb{Z}_2^n / C^\perp$ is linearly isomorphic to C. It is not difficult to explicitely construct the linear isomorphism.

Since $dim(\mathbb{Z}_2^n / C^\perp) = k$, this means that there are 2^k different elements in $\mathbb{Z}_2^n / C^\perp$, equivalently, 2^k different equivalence classes $w + C^\perp$ in $\mathbb{Z}_2^n$.

We saw before that $dim(H_C) = 2^{dim(C)} = 2^k$.

With all of this at our disposal, do the following exercise.

EXERCISE 1.4. *For every equivalence class $w + C^\perp$ in $\mathbb{Z}_2^n$, choose one representative w. Prove now that the collection of vectors $\{|c_w\rangle\}_w$ thus generated is a basis for H_C. Hint: Use Exercises 1.2 and 1.3, and count vectors.*

2. Quantum error correction

In this section I want to present, mostly without proofs, the ideas that will allow us to use our quantum codes to correct quantum errors. The ideas are the following:

- First of all: If we correct bit flips and phase flips, we correct everything.
- In the canonical basis, the $|c_w\rangle$ states are superpositions of the basis vectors $|v\rangle$, with $v \in C_1^\perp$. This will allow to correct $t = \frac{d-1}{2}$ bit flips.
- Later we will perform a change of basis in each qbit. The new basis will be called the s basis. We will see that in the s basis, the $|s_w\rangle$ states are superpositions of the basis vectors $|v\rangle$, with $v \in C_2^\perp$. This will allow us to correct t bit flips in the s basis.
- But bit flips in the s basis are phase errors in the canonical basis (and vice versa).
- Error correction in the canonical basis does not interfere with error correction in the s basis. Therefore, both types of errors can be corrected.

Let us see how far we can go in understanding each of them.
We go first with the first point.

2.1. Quantum errors. In order to understand how to correct quantum errors, we must have some understanding of quantum channels.

We want to describe what is a quantum channel sending quantum states from an n-dimensional system H_n to an m-dimensional system H_m. If we call $S_1(H_n), S_1(H_m)$ to the operators in H_n, H_m, then the states are the positive, trace one operators there. Then, a quantum channel is an operation $\mathcal{E} : S_1(H_n) \longrightarrow S_1(H_m)$. This operation must be linear and trace preserving. It must also send positive operators to positive operators. But this is not enough.

If the dimensions n, m do not coincide, then $\mathcal{E}$ can never be an unitary mapping, and we know quantum evolutions are described by unitary mappings. But there is more. Both systems can be paired to the environment. And $\mathcal{E}$ should be an unitary mapping also in the system Alice, Bob, Environment. For this, we require the mapping to be linear, trace preserving and *completely positive*. We will not go into the definition of complete positivity in this course.

We skip the math, and we just describe, without proof, the conclusion we are interested in.

Given a quantum channel between two spaces of the same dimension $\mathcal{E} : S_1(H_n) \longrightarrow S_1(H_n)$, there exists a number $k \in \mathbb{N}$ and operators $E_k : H_n \longrightarrow H_n$ such that, for every state ρ,

$$\mathcal{E}(\rho) = \sum_k E_k \rho E_k^\perp,$$

and the operators E_k verify

$$\sum E_k^{\perp} E_k = \mathbb{1}$$

.

These operators E_k are often called Kraus operators.

We consider now a channel sending just one qbit. Since the Pauli matrices are a basis for the space of 2×2 matrices, we can decompose each of the E_k by

$$E_k = a_{k0}\mathbb{1} + a_{k1}X + a_{k2}Z + a_{k3}XZ$$

Our quantum error correction methods, are, like all quantum operations, linear. This implies, (non trivially) that if we are able to fix errors associated to the operators X, Z and XZ then we can fix all errors on a qbit. If you want to read this in more detail, see [**1**, Chapter 10].

3. Our quantum code

We choose now two linear codes C_1, C_2 in $\mathbb{Z}_2^n$, such that

$$\{0\} \subsetneq C_2 \subsetneq C_1 \subsetneq \mathbb{Z}_2^n.$$

We choose them such that both C_1 and $C_2^{\perp}$ can correct t errors.

To simplify writing, we fix $dim(C_1) = k_1$, $dim(C_2) = k_2$.

Our working example is C_1 equal to the $[7, 4]$ Hamming code, and $C_2 = C_1^{\perp}$. In that case $C_2^{\perp} = C_1$, and the previous chapter tells us that they have distance 3, and, hence, can correct 1 error.

EXERCISE 3.1. *If C_1 is the $[7, 4]$ Hamming code, check that $C_1^{\perp} \subset C_1$.*

Given any bit sting $x \in \mathbb{Z}_2^n$, we associate it a quantum state $|x\rangle$ with the same natural identification as before, that is, if $x = (x_i)_{i=1}^n \in \mathbb{Z}_2^n$, then

$$|x\rangle = |x_i\rangle \otimes \cdots |x_n\rangle.$$

We will often use, probably without further mention that if $x \neq x'$, then $|x\rangle$ and $|x'\rangle$ are orthogonal. Check this!

Then, given $x \in C_1$, we define the quantum state $|x + C_2\rangle$ by

$$|x + C_2\rangle = \frac{1}{2^{\frac{k_2}{2}}} \sum_{y \in C_2} |x + y\rangle$$

EXERCISE 3.2. *Prove that if $x \pm x' \in C_2$, then $|x + C_2\rangle = |x' + C_2\rangle$.*

Therefore, the quantum state $|x + C_2\rangle$ depends only on the class of C_1/C_2 in which x lies, and not on the particular representative chosen.

EXERCISE 3.3. *Prove that if $x-x' \notin C_2$, then $|x+C_2\rangle$ and $|x'+C_2\rangle$ are orthonormal. Hint: if $x - x' \notin C_2$, then, for every $y, y' \in C_2$, $x + y \neq x' + y'$.*

Now, we define our quantum code Q_{C_1,C_2} by

$$Q_{C_1,C_2} = \{|x + C_2\rangle;\ x \in C_1\}$$

Put together, the two previous exercises show that the dimension of our quantum code Q_{C_1,C_2} is the cardinal of C_1/C_2, which is $2^{dim(C_1/C_2)}$ which is $2^{k_1-k_2}$.

In our working example, this dimension is $2^{4-3} = 2$.

The elements of the $[7, 4]$ Hamming code are

$$\begin{gathered}0000000,\ 0001011,\ 0010101,\ 0011110,\\ 0100111,\ 0101100,\ 0110010,\ 0111001,\\ 1000110,\ 1001101,\ 1010011,\ 1011000,\\ 1100001,\ 1101010,\ 1110100,\ 1111111\end{gathered}$$

EXERCISE 3.4. *If C_1 is the $[7, 4]$ Hamming code, enumerate the $8 = 2^3$ elements of $C_2 = C_1^{\perp}$.*

In this case, C_1/C_2 has dimension 1, and, therefore, just two elements. We must find a representative for each of the class. Take one of the representatives as the vector $x_0 = 0000000$. Then, to find a representative for the other class, we just need a vector x_1 such that $0000000 + x_1 \notin C_2$. Thus, we can take $x_1 = 1111111$, because $x_1 + x_0 = x_1 \notin C_2$. Thus, this code allows us to encode 1 qbit in 7 qbits. The encoding procedure sends $|0\rangle$ to

$$|x_0 + C_2\rangle = \frac{1}{2^{\frac{k_2}{2}}} \sum_{y \in C_2} |0000000 + y\rangle$$

and $|1\rangle$ to

$$|x_1 + C_2\rangle = \frac{1}{2^{\frac{k_2}{2}}} \sum_{y \in C_2} |1111111 + y\rangle$$

In general, if our quantum code code has dimension $2^{k_1-k_2}$, the encoding procedure is as follows:

Choose $2^{k_1-k_2}$ different classical n-bit strings, one in each of the different classes C_1/C_2. Call x_j to each of this classical bit strings. Associated to each of them, define the quantum state $|x_j + C_2\rangle$ as before.

Then, our quantum code will encode $\mathbb{C}_2^{k_1-k_2}$, which is a space of dimension $2^{k_1-k_2}$ into $\mathbb{C}_2^n$. To do this, it will send the j^{th} qbit of the canonical basis of $\mathbb{C}_2^{k_1-k_2}$ to $|x_j + C_2\rangle$.

Let us see how our code allows us to correct up to t bit flip errors and t phase flip errors.

Suppose we send the quantum state

$$|x + C_2\rangle = \frac{1}{2^{\frac{k_2}{2}}} \sum_{y \in C_2} |x + y\rangle$$

and there were bit flip and phase flip errors. Let us call $e \in \mathbb{Z}_2^n$ to the bit string with a 1 in the positions where there were bit flips, and 0 in the rest. Similarly, let us call $f \in \mathbb{Z}_2^n$ to the bit string with a 1 in the positions where there were phase flips, and 0 in the rest.

Recall that a bit flip (respectively phase flip) in a qbit $|\varphi\rangle$ can be written as $X|\varphi\rangle$ (respectively where $Z|\varphi\rangle$), where X, Z are the Pauli matrices.

We will use the notation

$$X^e := X^{e_1} \otimes \cdots X^{e_n}$$

and

$$Z^f := Z^{f_1} \otimes \cdots X^{f_n},$$

where e_i, f_i are the i^{th} bits of e, f and $X^1 = X$, $X^0 = \mathbb{1}$, and similarly for Z.

Then, the received state is

$$X^e Z^f \left(|x + C_2\rangle\right) = X^e Z^f \left(\frac{1}{2^{\frac{k_2}{2}}} \sum_{y \in C_2} |x + y\rangle \right) = \tag{3}$$

$$\frac{1}{2^{\frac{k_2}{2}}} \sum_{y \in C_2} X^e Z^f \left(|x + y\rangle\right) = \frac{1}{2^{\frac{k_2}{2}}} \sum_{y \in C_2} (-1)^{(x+y)\cdot f} \left(|x + y + e\rangle\right)$$

The first two inequalities above are elementary. We show the last one.

EXERCISE 3.5. *Given two bits a, b, show that $X^a Z^b = (-1)^{ab} Z^b X^a$*

EXERCISE 3.6. *Using this, prove that*

$$X^e Z^f \left(|x + y\rangle\right) = (-1)^{e \cdot f} Z^f |x + y + e\rangle$$

Hint: Use the fact that $X|0\rangle = |1\rangle$, $X|1\rangle = |0\rangle$.

EXERCISE 3.7. *Prove that, for bits a, b, w,*

$$(-1)^{ab} Z^b |w + a\rangle = (-1)^{wb} |w + a\rangle$$

Hint: the case $b = 0$ is trivial. For the other case, use $Z|0\rangle = |0\rangle$, $Z|1\rangle = -|1\rangle$

Finally,

EXERCISE 3.8. *Prove that*

$$(-1)^{e \cdot f} Z^f |x + y + e\rangle = (-1)^{(x+y) \cdot f} |x + y + e\rangle$$

Hint: use the previous exercise, with w the corresponding bit in $x + y$.

We must show now that we can fix the errors in the received state. We need to recall the basis of quantum information manipulation: if we measure the state, say to find the syndrome, then we perturb it, and we might not be able to fix the error.

So, we have to be a little more careful. The first observation is that we can use an auxiliary register and unitary evolution to compute the syndrome without perturbing the state. First, we attach an auxiliary register, initialized in any fixed state, which we will call $|0\rangle$, to our quantum state. That is, we receive a quantum state, which we call $|x\rangle$, and we create $|x\rangle \otimes |0\rangle$.

LEMMA 3.9. *Let H be the parity check matrix of C_1, with dimension $(n - k) \times n$. Then the mapping*

$$|x\rangle \otimes |0\rangle \mapsto |x\rangle \otimes |Hx\rangle$$

can be implemented unitarily. Indeed, it only requires the use of CNOT gates, and the ancilla can be chosen to be $\mathbb{C}_2^{n-k}$.

PROOF. Recall first that a controlled-NOT gate (CNOT gate) is an unitary mapping

$$\mathbb{C}_2 \otimes \mathbb{C}_2 \longrightarrow \mathbb{C}_2 \otimes \mathbb{C}_2$$

verifying

$$|0\rangle \otimes |0\rangle \mapsto |0\rangle \otimes |0\rangle$$
$$|0\rangle \otimes |1\rangle \mapsto |0\rangle \otimes |1\rangle$$

$$|1\rangle \otimes |0\rangle \mapsto |1\rangle \otimes |1\rangle$$

$$|1\rangle \otimes |1\rangle \mapsto |1\rangle \otimes |0\rangle$$

Observe that the first qbit remains unaltered, and the second one is flipped conditioned to the first one.

We will work with an example

Assume

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

(the parity matrix of the $[7, 4]$ Hamming code).

Then, to build the desired mapping, consider your state $|x\rangle \otimes |000\rangle \in \mathbb{C}_2^7 \otimes \mathbb{C}_2^3$.

Now, for each of the rows in H, create the sequence of CNOT gates which fixes the qbit of $|x\rangle$ and flips the bit in the register indexed by the row in the corresponding cases.

If we write $\text{CNOT}(i, j)$ for a CNOT gate where the control qbit is the i^{th} bit of $|x\rangle$ and the target qbit is the j^{th} qbit of the ancilla, then the first row of H is implemented by the sequence

$$\text{CNOT}(1,1)\text{CNOT}(2,1)\text{CNOT}(3,1)\text{CNOT}(5,1)$$

The second row is implemented by the sequence

$$\text{CNOT}(1,2)\text{CNOT}(2,2)\text{CNOT}(4,2)\text{CNOT}(6,2)$$

etc.

□

EXERCISE 3.10. *Suppose that H is the $n \times n$ identity matrix. Discuss why the above does not contradict the no cloning theorem, despite the fact that the associated unitary matrix sends*

$$|x\rangle \otimes |0\rangle \mapsto |x\rangle \otimes |x\rangle$$

for every $|x\rangle$ in the canonical basis.

That is, we

- receive the quantum state
- Attach a large enough ancilla initialized to $|0\rangle$ to obtain the state

$$\frac{1}{2^{\frac{k_2}{2}}} \sum_{y \in C_2} (-1)^{(x+y)\cdot f} \left(|x + y + e\rangle\right) \otimes |0\rangle =$$

$$= \frac{1}{2^{\frac{k_2}{2}}} \sum_{y \in C_2} (-1)^{(x+y)\cdot f} \left(|x + y + e\rangle \otimes |0\rangle\right)$$

- Calculate H on the ancilla, without disturbing the original state. Using the fact that H is a linear action on bits, we obtain

$$\frac{1}{2^{\frac{k_2}{2}}}\sum_{y\in C_2}(-1)^{(x+y)\cdot f}\left(|x+y+e\rangle\right)\otimes|H(x+y+e)\rangle=$$

$$\frac{1}{2^{\frac{k_2}{2}}}\sum_{y\in C_2}(-1)^{(x+y)\cdot f}\left(|x+y+e\rangle\right)\otimes|H(x+y)+H(e)\rangle=$$

$$\frac{1}{2^{\frac{k_2}{2}}}\sum_{y\in C_2}(-1)^{(x+y)\cdot f}\left(|x+y+e\rangle\right)\otimes|H(e)\rangle,$$

where the last equality follows from the fact that $x+y\in C_1$ and, hence, $H(x+y)=0$

Now, we can measure the ancilla register, to obtain $H(e)$. After measuring it, we discard the ancilla.

Now, since C_1 can correct up to t errors, and we know we have no more than that number of errors, knowing $H(e)$ we can know e. Now, to correct the bit flip errors we just do bit flips (that is, acting of the matrix X) on each of the flipped qbits. Note again that this action is just an unitary, and therefore it is physically doable.

After that, we have the state

$$\frac{1}{2^{\frac{k_2}{2}}}\sum_{y\in C_2}(-1)^{(x+y)\cdot f}\left(|x+y\rangle\right),$$

and we have removed the bit flip errors!.

We still have to remove the phase flip errors. We want to use know the correcting properties of C_1. For that, we will need the bit strings $x+y$ to be in C_1 whenever $x\in C_1$. Note that we have that, because the codes are linear and $C_2\subset C_1$, and here is where we use this fact.

In order to correct the phase flips, we need perform a change of basis in our qbits, to send the canonical basis to the $\{|+\rangle,|-\rangle\}$ basis. This is a unitary mapping, and we are going to see that our state

$$\frac{1}{2^{\frac{k_2}{2}}}\sum_{y\in C_2}(-1)^{(x+y)\cdot f}\left(|x+y\rangle\right),$$

turns into

$$\frac{1}{2^{\frac{n-k_2}{2}}}\sum_{z\in C_2^{\perp}}(-1)^{x\cdot z}\left(|z+f\rangle\right)$$

We will see that we can now remove f similarly to the way we removed e previously, and then, when we undo the change of basis, we recover our corrected quantum state.

To do this, let us call

$$Had : \mathbb{C}_2^n \longrightarrow \mathbb{C}_2^n$$

to unitary transformation implementing the change of basis from the canonical basis to the $\{|+\rangle, |-\rangle\}$. We call it Had because Hadamard gates are the gates doing this change in one qbit. Our transformation is then the tensor product of n Hadamard gates.

With this notation, we want to see that

$$Had\left(\frac{1}{2^{\frac{k_2}{2}}}\sum_{y\in C_2}(-1)^{(x+y)\cdot f}\left(|x+y\rangle\right)\right) = \frac{1}{2^{\frac{n-k_2}{2}}}\sum_{z\in C_2^\perp}(-1)^{x\cdot z}\left(|z+f\rangle\right).$$

This requires again several steps.

EXERCISE 3.11. *Calculate the four products* $\langle +|0\rangle$, $\langle +|1\rangle$, $\langle -|0\rangle$, $\langle -|1\rangle$.

Let now $w \in \mathbb{Z}_2^n$.

EXERCISE 3.12. *Use the previous exercise to show that*

$$\langle w|Had|(x+y)\rangle\rangle = \frac{1}{2^{\frac{n}{2}}}(-1)^{(x+y)\cdot w}$$

Therefore, we obtain that

$$\left\langle w|Had\left(\frac{1}{2^{\frac{k_2}{2}}}\sum_{y\in C_2}(-1)^{(x+y)\cdot f}|x+y\rangle\right)\right\rangle =$$

$$\frac{1}{2^{\frac{k_2}{2}}}\sum_{y\in C_2}(-1)^{(x+y)\cdot f}\left\langle w|Had\left(|x+y\rangle\right)\right\rangle =$$

$$\frac{1}{2^{\frac{k_2+n}{2}}}\sum_{y\in C_2}(-1)^{(x+y)\cdot f}(-1)^{(x+y)\cdot w} = \frac{1}{2^{\frac{k_2+n}{2}}}\sum_{y\in C_2}(-1)^{(x+y)\cdot(f+w)} =$$

$$= \frac{1}{2^{\frac{k_2+n}{2}}}(-1)^{x\cdot(f+w)}\sum_{y\in C_2}(-1)^{y\cdot(f+w)}$$

EXERCISE 3.13. *Show that*

$$\sum_{y\in C_2}(-1)^{y\cdot(f+w)} = \begin{cases} 0 \text{ if } f+w \notin C_2^\perp \\ 2^{k_2} \text{ if } f+w \in C_2^\perp \end{cases}$$

Using this exercise, we have that

$$\left\langle w|Had\left(\frac{1}{2^{\frac{k_2}{2}}}\sum_{y\in C_2}(-1)^{(x+y)\cdot f}|x+y\rangle\right)\right\rangle =$$

$$= \begin{cases} 0 & \text{if } f+w \notin C_2^\perp \\ \frac{1}{2^{\frac{n-k_2}{2}}}(-1)^{x\cdot(f+w)} & \text{if } f+w \in C_2^\perp \end{cases}$$

We now use the fact that the canonical basis $\{|w\rangle; w \in \mathbb{Z}_2^n\}$ is a basis of $\mathbb{C}_2^n$. This means that every vector $|\varphi\rangle$ can be written as

$$|\varphi\rangle = \sum_{w\in\mathbb{Z}_2^n} \langle w|\varphi\rangle |w\rangle$$

If we apply this to our vector $|\varphi\rangle = Had\left(\frac{1}{2^{\frac{k_2}{2}}}\sum_{y\in C_2}(-1)^{(x+y)\cdot f}|x+y\rangle\right)$, we obtain

$$Had\left(\frac{1}{2^{\frac{k_2}{2}}}\sum_{y\in C_2}(-1)^{(x+y)\cdot f}|x+y\rangle\right) =$$
$$\sum_{w\in\mathbb{Z}_2^n}\left\langle w|Had\left(\frac{1}{2^{\frac{k_2}{2}}}\sum_{y\in C_2}(-1)^{(x+y)\cdot f}|x+y\rangle\right)\right\rangle|w\rangle =$$
$$\frac{1}{2^{\frac{n-k_2}{2}}}\sum_{w;w+f\in C_2^\perp}(-1)^{x\cdot(f+w)}|w\rangle = \frac{1}{2^{\frac{n-k_2}{2}}}\sum_{z\in C_2^\perp}(-1)^{x\cdot z}|z+f\rangle,$$

where in the last equality we have renamed $w+f=z$.

Let us now complete our quantum error correction.

Note that the expression of our state now looks very much like the the expression in Eq (3). That is, our phase flip error looks exactly like a bit flip when we have changed the basis. So, to fix the phase flip error, we do the same as before:

Calling H' to the the parity check matrix of $C_2^\perp$, we attach an ancilla initialized in $|0\rangle$ and perform the unitary taking

$$\frac{1}{2^{\frac{n-k_2}{2}}}\sum_{z\in C_2^\perp}(-1)^{x\cdot z}|z+f\rangle\otimes|0\rangle \mapsto \frac{1}{2^{\frac{n-k_2}{2}}}\sum_{z\in C_2^\perp}(-1)^{x\cdot z}|z+f\rangle\otimes|H'(z+f)\rangle =$$

$$= \frac{1}{2^{\frac{n-k_2}{2}}}\sum_{z\in C_2^\perp}(-1)^{x\cdot z}|z+f\rangle\otimes|H'(f)\rangle$$

As before, we measure the ancilla to find out $H'(f)$ and, hence, f. Once we know the syndrome, we apply bit flips (X matrices) in the

proper qbit. This makes f dissapear, that is, we have now

$$\frac{1}{2^{\frac{n-k_2}{2}}} \sum_{z \in C_2^\perp} (-1)^{x \cdot z} |z\rangle.$$

Finally, we reverse the change of basis. Note that, to do this, we just need to apply the same product of Hadamard gates again. As a final exercise,

EXERCISE 3.14. *Check that*

$$Had\left(\frac{1}{2^{\frac{n-k_2}{2}}} \sum_{z \in C_2^\perp} (-1)^{x \cdot z} |z\rangle\right) = \frac{1}{2^{\frac{k_2}{2}}} \sum_{y \in C_2} |x + y\rangle$$

That is, we have corrected the errors!!

CHAPTER 5

Solution to exercises

EXERCISE 0.1. *Consider a composite system $H_A \otimes H_B$. Let $\{P_i^A\}_i$ be a measurement system in H_A and let $\{Q_j^B\}_j$ be as measurement system in H_B. Prove that $\{P_i^A \otimes \mathbb{1}_B\}_i$ $\{\mathbb{1}_A \otimes Q_j^B\}_j$ and $\{P_i^A \otimes Q_j^B\}_{i,j}$ are measurement systems in the joint system $H_A \otimes H_B$. The first of them describes the situation when Alice measures with $\{P_i^A\}_i$ and Bob does nothing, the second describes the situation when Bob measures with $\{Q_j^B\}_j$ and Alice does nothing and the third one describes the situation when Alice measures with $\{P_i^A\}_i$ and Bob measures with $\{Q_j^B\}_j$*

Now, describe mathematically the situation when first *Alice measures with $\{P_i^A\}_i$ and* then *Bob measures with $\{Q_j^B\}_j$, and viceversa.*

Solution: We do it in the case that the system $\{P_i^A\}_i$ and $\{Q_j^B\}_j$ are projective measurements, since that is the case we will use most times in this course. The general case is done similarly. We prove first that, in that case, $\{P_i^A \otimes \mathbb{1}_B\}_i$ is also a projective measurement:

(1) For every i, $P_i^A \otimes \mathbb{1}_B$ is a projection:

$$(P_i^A \otimes \mathbb{1}_B)(P_i^A \otimes \mathbb{1}_B) = P_i^A P_i^A \otimes \mathbb{1}_B \mathbb{1}_B = P_i^A \otimes \mathbb{1}_B$$

(2) For every $i \neq j$, $P_i^A \otimes \mathbb{1}_B$ and $P_j^A \otimes \mathbb{1}_B$ are orthogonal:

$$(P_i^A \otimes \mathbb{1}_B)(P_j^A \otimes \mathbb{1}_B) = P_i^A P_j^A \otimes \mathbb{1}_B \mathbb{1}_B = 0 \otimes \mathbb{1}_B = 0$$

(3) $\sum_i P_i^A \otimes \mathbb{1}_B = \mathbb{1}_{H_A \otimes H_B}$:

$$\sum_i P_i^A \otimes \mathbb{1}_B = \left(\sum_i P_i^A \right) \otimes \mathbb{1}_B = \mathbb{1}_A \otimes \mathbb{1}_B = \mathbb{1}_{H_A \otimes H_B}$$

The proof for $\{\mathbb{1}_A \otimes Q_j^B\}_j$ is totally similar.
We do now the case $\{P_i^A \otimes Q_j^B\}_{i,j}$:

(1) For every i, j, $P_i^A \otimes Q_j^B$ is a projection:

$$(P_i^A \otimes Q_j^B)(P_i^A \otimes Q_j^B) = P_i^A P_i^A \otimes Q_j^B Q_j^B = P_i^A \otimes Q_j^B$$

(2) For every $(i, j) \neq (i', j')$, $P_i^A \otimes Q_j^B$ and $P_j^A \otimes Q_j^B$ are orthogonal:

$$(P_i^A \otimes Q_j^B)(P_{i'}^A \otimes Q_{j'}^B) = P_i^A P_{i'}^A \otimes P_i^A P_{i'}^A = 0,$$

where the last equality follows from the fact that either $P_i^A P_{i'}^A$ or $P_i^A P_{i'}^A$ must be 0 if $(i,j) \neq (i',j')$

(3) $\sum_{i,j} P_i^A \otimes Q_j^B = \mathbb{1}_{H_A \otimes H_B}$:

$$\sum_{i,j} P_i^A \otimes Q_j^B = \left(\sum_i P_i^A\right) \otimes \left(\sum_j Q_j^B\right) = \mathbb{1}_A \otimes \mathbb{1}_B = \mathbb{1}_{H_A \otimes H_B}$$

For the final question, when *first* Alice measures with $\{P_i^A\}_i$ and *then* Bob measures with $\{Q_j^B\}_j$, that is acting with the operators

$$(\mathbb{1}_A \otimes Q_j^B)(P_i^A \otimes \mathbb{1}_B)$$

, the other case being

$$(P_i^A \otimes \mathbb{1}_B)(\mathbb{1}_A \otimes Q_j^B).$$

But note that $(P_i^A \otimes \mathbb{1}_B)$ and $(\mathbb{1}_A \otimes Q_j^B)$ commute and

$$(P_i^A \otimes \mathbb{1}_B)(\mathbb{1}_A \otimes Q_j^B) = (\mathbb{1}_A \otimes Q_j^B)(P_i^A \otimes \mathbb{1}_B) = P_i^A \otimes Q_j^B$$

Therefore, the situations where Alice measures first, Bob measures first or they both measure simultaneously are indistinguishable

EXERCISE 0.2. *Show that trace is linear and cyclic. That is, show that for any two matrices $A, B \in M_d$ and for any two $\alpha, \beta \in \mathbb{C}$,*

(1) $tr(\alpha A + \beta B) = \alpha tr(A) + \beta tr(B)$
(2) $tr(AB) = tr(BA)$

Solution: Easy enough.

EXERCISE 0.3. *Given a finite dimensional Hilbert space H, an operator $T : H \longrightarrow H$ is positive if and only if its associated matrix (in any given basis) is semidefinite positive.*

Solution: This is essentially the definition of semidefinite positive matrix. Recall that a matrix $A \in M_n$ is positive semidefinite if and only if, for every vector $|\varphi\rangle \in \mathbb{C}^n$,

$$\langle\varphi|A|\varphi\rangle \geq 0.$$

Compare with the definition of positive operator.

EXERCISE 0.4. *Given a finite dimensional Hilbert space H, and an operator $T : H \longrightarrow H$, if T is positive then T is self-adjoint (also called Hermitian).* Hints*: Decompose $T = A + \imath B$, with A, B Hermitian. Prove that, for Hermitian operators $\langle\varphi|A|\varphi\rangle \in \mathbb{R}$ for every $|\varphi\rangle \in H$, and same for B. Now, use this property and the decomposition of T to show that $B = 0$ and, hence, $A = T$.*

Solution: Given T (not necessarily positive nor Hermitian) consider the operators

$$A = \frac{T + T^\dagger}{2}$$

and

$$B = \frac{T - T^\dagger}{2\imath}$$

It is trivial to check that $A = A^\dagger$, $B = B^\dagger$ and $T = A + \imath B$.

Remember that, for every operator C (Hermitian or not) and for every vectors $|\varphi\rangle, |\psi\rangle \in H$ one has

$$\langle\varphi|C\psi\rangle = \langle\varphi C^\dagger|\psi\rangle.$$

Therefore, for a Hermitian operator C and a vector $|\varphi \in H$ one has

$$\langle\varphi|C\varphi\rangle = \langle\varphi C^\dagger|\varphi\rangle = \langle\varphi C|\varphi\rangle = \overline{\langle\varphi|C\varphi\rangle}$$

and, therefore, $\langle\varphi|C\varphi\rangle \in \mathbb{R}$.

Therefore, both $\langle\varphi A\varphi\rangle$ and $\langle\varphi B\varphi\rangle$ are real numbers.

Since T is positive, for every $|\varphi\rangle$ we have

$$\langle\varphi T\varphi\rangle \geq 0,$$

and, in particular,

$$\langle\varphi T\varphi\rangle \in \mathbb{R}$$

At the same time, we have

$$\langle\varphi T\varphi\rangle = \langle\varphi A + \imath B\varphi\rangle = \langle\varphi A\varphi\rangle + \imath\langle\varphi B\varphi\rangle.$$

Since $\langle\varphi A\varphi\rangle$ and $\langle\varphi B\varphi\rangle$ are real numbers, it follows that, for every $\varphi\rangle$,

$$\langle\varphi B\varphi\rangle = 0.$$

Therefore, $B = 0$ and $T = A$.

EXERCISE 0.5. *Suppose we have a composite system $H_A \otimes H_B$. Suppose system A is prepared in the state $\rho_A = \sum_i p_i\varphi_i\rangle\langle\varphi_i$ and system B is prepared in state $\rho_B = \sum_j q_j\psi_j\rangle\langle\psi_j$. Check that*

$$\rho_A \otimes \rho_B = \sum_{i,j} p_i q_j(\varphi_i \otimes \psi_j)\rangle\langle(\varphi_i \otimes \psi_j)$$

Solution: Essentially trivial.

EXERCISE 0.6. *Suppose the simplest case, where $\rho_{AB} = \rho_A \otimes \rho_B$ and ρ_A, ρ_B are as in Exercise 0.5. Check that, in that case,*

$$tr_B(\rho_{AB}) = \rho_A tr(\rho_B) = \rho_A$$

Solution: Let $\rho_A = \sum_i p_i \varphi_i\rangle\langle\varphi_i$ and $\rho_B = \sum_j q_j \psi_j\rangle\langle\psi_j$. Using the linearity of the partial trace and the fact that, for every state $|\psi\rangle$, $tr(\psi\rangle\langle\psi) = 1$, we have

$$\begin{aligned} tr_B(\rho_{AB}) =& tr_B\left(\left(\sum_i p_i \varphi_i\rangle\langle\varphi_i\right) \otimes \left(\sum_j q_j \psi_j\rangle\langle\psi_j\right)\right) = \\ =& \sum_{i,j} p_i q_j tr_B\left((\varphi_i\rangle\langle\varphi_i) \otimes (\psi_j\rangle\langle\psi_j)\right) = \sum_{i,j} p_i q_j \varphi_i\rangle\langle\varphi_i = \\ =& \sum_j q_j \sum_i p_i \varphi_i\rangle\langle\varphi_i = \sum_i p_i \varphi_i\rangle\langle\varphi_i = \rho_A \end{aligned}$$

EXERCISE 0.7. *Prove this last statement. That is, prove that if Alice and Bob share the state* $|\varphi\rangle = \frac{|00\rangle+|11\rangle}{\sqrt{2}}$ *and they each measure separately in the computatioal basis, then their outcomes will coincide with probability 1.*

Solution: "Alice and Bob measure separately in the computational basis" means that Alice measures with $\{|0\rangle\langle 0| \otimes \mathbb{1}_B, |1\rangle\langle 1| \otimes \mathbb{1}_B\}$ and Bob measures with $\{\mathbb{1}_A \otimes |0\rangle\langle|, \mathbb{1}_A \otimes |1\rangle\langle 1|\}$. As we saw in an exercise above, we may consider any order in the measurements.

We start with Alice's measurements. Suppose Alice got $|0\rangle$ in her measurement. Then the post measurements state is

$$\frac{(|0\rangle\langle 0| \otimes \mathbb{1}_B)\,\varphi\rangle}{\|\,(|0\rangle\langle 0| \otimes \mathbb{1}_B)\,\varphi\rangle\|}$$

The denominator is

$$(|0\rangle\langle 0| \otimes \mathbb{1}_B)\,\varphi\rangle = (|0\rangle\langle 0| \otimes \mathbb{1}_B)\,\frac{|00\rangle + |11\rangle}{\sqrt{2}} =$$

$$= \frac{1}{\sqrt{2}}\left((|0\rangle\langle 0| \otimes \mathbb{1}_B)\,|00\rangle + (|0\rangle\langle 0| \otimes \mathbb{1}_B)\,|11\rangle\right) =$$

$$\frac{1}{\sqrt{2}}|0\rangle \otimes |0\rangle,$$

and, therefore,

$$\frac{(|0\rangle\langle 0| \otimes \mathbb{1}_B)\,\varphi\rangle}{\|\,(|0\rangle\langle 0| \otimes \mathbb{1}_B)\,\varphi\rangle\|} = |0\rangle \otimes |0\rangle$$

That is the state they share after Alice measured and got $|0\rangle$. It is very easy to check now that when Bob measures he will also get $|0\rangle$ with probability 1. The case when Alice gets $|1\rangle$ is totally similar.

EXERCISE 0.8. *Check that if Bob measures qbit $|\psi_i\rangle$ in the basis indicated by b_i, he obtains bit a_i with probability 1. Check also that if he measures $|\psi_i\rangle$ in the basis not indicated by b_i, then he obtains a_i with probability $\frac{1}{2}$ (and $a_i \oplus 1$ with probability $\frac{1}{2}$).*

Solution: The case when they both measure in the computational basis was done in Exercise 0.7.

Suppose now Alice measures the state $\frac{1}{\sqrt{2}}(|00\rangle+|11\rangle)$ in the $\{|+\rangle, |-\rangle\}$ basis. From the Alice-Bob point of view, that means the measurement projections are $\{|+\rangle\langle+| \otimes \mathbb{1}, |-\rangle\langle-| \otimes \mathbb{1}\}$. Suppose the outcome was $|+\rangle$. Then, the post measurement state is

$$\frac{(|+\rangle\langle+| \otimes \mathbb{1})\frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)}{\text{norm of the numerator}} = \frac{|+\rangle\langle+|0\rangle|0\rangle + |+\rangle\langle+|1\rangle|1\rangle}{\text{norm of the numerator}} =$$

$$= \frac{\frac{1}{\sqrt{2}}|+\rangle|0\rangle + \frac{1}{\sqrt{2}}|+\rangle|1\rangle}{\text{norm of the numerator}} = |+\rangle \otimes |+\rangle.$$

Therefore, when Bob measures (with $\{\mathbb{1} \otimes |+\rangle\langle+|, \mathbb{1} \otimes |-\rangle\langle-|\}$ he will obtain $|+\rangle$ with probability 1.

This finishes the case when they both measure in the same basis.

We see now the case when they each measure in a different basis. I write out one of the cases, the others being similar.

Suppose Alice measures with $\{|0\rangle\langle0| \otimes \mathbb{1}, |1\rangle\langle1| \otimes \mathbb{1}\}$, and she obtains $|0\rangle$. As we saw, the post measurement state will be $|00\rangle$. Now, when Bob measures in the $\{|+\rangle, |-\rangle\}$ basis, that is, with the operators $\{|+\rangle\langle+| \otimes \mathbb{1}, |-\rangle\langle-| \otimes \mathbb{1}\}$, the probability he obtains $|+\rangle$ is

$$\langle 00||+\rangle\langle+| \otimes \mathbb{1}|00\rangle = \langle0|+\rangle\langle0|+\rangle\langle0|0\rangle = \frac{1}{2}$$

The rest of the cases are similar.

EXERCISE 0.9. *Consider two strings* $\mathbf{a}$, $\mathbf{b}$ *of $2n$ bits. Suppose that, for each position in the string, the probability that $a_i = b_i$ is the same. We choose randomly n bits of the string and we find that they differ in μn bits, with $0 < \mu < 1$. Prove that the probability that the remaining n bits differ in more than $(\mu + \epsilon)n$ bits is smaller than ??*

EXERCISE 0.10. *In Exercises 0.7, 0.8 you proved that if Alice and Bob share an EPR state and they each measure separately in the computational basis, or in the $\{|+\rangle, |-\rangle\}$ basis. then their outcomes will coincide with probability 1. Prove now that actually the result is also true no matter in which orthogonal basis they measure (for as long as they both use the same one). Or, put in another way, prove that when Alice measures with an orthogonal basis and obtains a result, Bob's*

state collapses to that same state. This phenomenom is called steering*: Alice's measurement steers Bob's state.*

Solution: Let $\{|\alpha\rangle, |\beta\rangle\}$ be any basis of $\mathbb{C}_2$. If Alice measures $\frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$ in the $\{|\alpha\rangle, |\beta\rangle\}$ basis, that corresponds to the operators $\{|\alpha\rangle\langle\alpha| \otimes \mathbb{1},\ |\beta\rangle\langle\beta| \otimes \mathbb{1}\}$. Suppose Alice's result was $|\alpha\rangle$. Then, the post measurement state is (I write it without the normalizing denominator) is

$$(\alpha\rangle\langle\alpha| \otimes \mathbb{1})(|00\rangle + |11\rangle) = \langle\alpha|0\rangle|\alpha\rangle \otimes |0\rangle + \langle\alpha|1\rangle|\alpha\rangle \otimes |1\rangle =$$

$$= |\alpha\rangle \otimes (\langle\alpha|0\rangle|0\rangle + \langle\alpha|1\rangle|1\rangle) = |\alpha\rangle \otimes |\alpha\rangle,$$

where the last inequality follows from the fact that, since $\{|0\rangle, |1\rangle\}$ is a basis of $\mathbb{C}^2$,

$$\langle\alpha|0\rangle|0\rangle + \langle\alpha|1\rangle|1\rangle = |\alpha\rangle,$$

for every vector $|\alpha\rangle \in \mathbb{C}_2$

EXERCISE 0.11. *Prove that the previous mapping sends each word in $\mathbb{Z}_2^k$ to a code word in $C \subset \mathbb{Z}_2^n$ in a bijective way. In particular, different words go to different words.*

Solution: This is just linear algebra: Consider $x, y \in \mathbb{Z}_2^k$ such that $G(x) = G(y)$. Write $x = \sum_i x_i e_i$, $y = \sum_i y_i e_i$. Using linearity, we obtain

$$\sum_i x_i G(x_i) = \sum_i x_i \mathbf{r_i} = \sum_i y_i \mathbf{r_i} = \sum_i y_i G(y_i),$$

and, therefore,

$$\sum_i (x_i - y_i)\mathbf{r_i} = 0$$

Using the linear independence of the vectors $\mathbf{r_i}$'s we obtain that $x_i = y_i$ for every i, and, therefore, $x = y$.

EXERCISE 0.12. *Prove that for the binary repetition code of length 3 (sending 0 to 000 and 1 to 111), the generating matrix is*

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

Solution: It is enough to check that $G(0) = 0$, $G(1) = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$

EXERCISE 0.13. *Calculate the generating matrix G for the $[7, 4]$-Hamming code.*

Solution*: There are many ways to do the exercise. I choose the one I find simpler. The idea of the code is that we encode the four bits of the bit string in the first four bits of the codeword, and the rest three bits are parity check. With that in mind, the canonical basis of* $\mathbb{Z}_2^4$ *is encoded by*

$$(0001) \mapsto (0001011)$$
$$(0010) \mapsto (0010101)$$
$$(0100) \mapsto (0100111)$$
$$(1000) \mapsto (1000110)$$

Check now that the four vectors $\{(0001011), (0010101), (0100111), (1000110)\}$ *are codewords (trivial) and linearly independent. To see this, it is enough to see that the associated matrix has rank 4. Therefore,*

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

EXERCISE 0.14. *Check these last paragraphs.*

Solution*: Linear algebra!*

EXERCISE 0.15. *Calculate* H *for the* $[7, 4]$*-Hamming code*

Solution*:*

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

The proof is immediate.

EXERCISE 0.16. *Show that*

$$d_H(\mathbf{x}, \mathbf{y}) = W_H(\mathbf{x} + \mathbf{y})$$

Solution*: In* $\mathbb{Z}_2$*,* $0 = -0$ *and* $1 = -1$*. Therefore,* $-\mathbf{y} = \mathbf{y}$*. It follows that*

$$d_H(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^{n} |x_i - y_i| = \sum_{i=1}^{n} x_i - y_i = \sum_{i=1}^{n} x_i + y_i W_H(\mathbf{x} + \mathbf{y})$$

EXERCISE 0.17. *Why can we may always assume that the probability of a bit flip is less than* $1/2$*?*

Solution*: If we have a channel such that the this probability is strictly bigger than* $1/2$*, we flip the outputs of the channel on arrival. The new probability of bit flip is smaller than* $1/2$*.*

EXERCISE 0.18. *Suppose your (classical) channel is such that the probability of a bit flip in each separate bit is less than* $1/2$ *and that the the probabilities of bit flips in different bits are independent from each other. Prove that, in that case, if Bob receives* y'*, then the most likely word Alice sent is the* y *minimizing*

$$W_H(e) = W_H(y - y') = d_H(y, y').$$

Solution*: Given that we have received* y'*, we have to calculate* y *in order to maximize* $Pr(y|y')$*, where this last expression means "the probability of* y *being the sent word, conditioned to having received* y'*. This is equivalent to maximizing* $Pr(y \cap y')$*, assuming* y' *was received. We use now that the error are bitwise independent. We call* p *to the probability of bit flip, and assume* $p \leq \frac{1}{2}$*. Hence,* $\frac{1-p}{p} \geq 1$*. We call* $d = d_H(y, y')$*. We then have*

$$Pr(y \cap y') = \prod_i Pr(y_i \cap y'_i) = (1-p)^{n-d} p^d =$$

$$= \frac{(1-p)^{n-d} p^{n-d}}{p^{n-d}} p^d = \left(\frac{1-p}{p}\right)^{n-d} p^n$$

Hence, to maximize $Pr(y \cap y')$ *we must minimize* d*, the Hamming distance between* y *and* y'*.*

Bibliography

[1] Nielsen, Chuang, Quantum computation and Quantum Information
[2] https://cs.uwaterloo.ca/ watrous/LectureNotes.html
[3] https://homepages.cwi.nl/ rdewolf/qcnotes.pdf