Entanglement and nonlocality

(collected notes)

Fabio Grazioso
2015-06-24 23:06 EDT

Contents

Chapter 1

Introduction

1.1 foreword

These notes are a collection of different material, coming from different sources. In particular, two lectures from Prof Binney, University of Oxford (youtube online videos) and the "Quantum Information Theory" class of prof Patrick Heyden, McGill University. I have also used the original articles, such as [Bel64] and [CHSH69].

1.2 general picture

EPR At first, the *problem of entanglement* has been risen by Einstein Podolsky and Rosen [EPR35], and was written as a demonstration that Quantum Theory was not complete. In the EPR paper, the idea of *hidden variables* was suggested as a solution to the "quantum oddity".

A good place where the EPR paper is well explained is [WM08, §13.1, p. 247-248],

Bell In '64 John Bell managed [Bel64] to express the problem in a way which was *experimentally testable*. Indeed, EPR had formulated the problem in a conceptual way. On one hand there was the *hidden variable theory* of which EPR had suggested the existence; on the other hand, the Quantum Theory was presented as an alternative description, the conceptual difference being the probabilistic or deterministic description of reality. The contribution of Bell was to find a *point of contact* between the two descriptions, *i.e.* a specific quantity, which can be formulated using the two different descriptions, but which has a different, unreconcilable value in the models. Moreover, this quantity can be experimentally measured, and in general a result will be compatible with only one of the two models.

One important remark is that the theorem discriminates between Quantum Theory

(QT) and a *Local* Hidden Variables Theory (LHVT). This means that a Non-Local Hidden Variable Theory is compatible with the description of QT.

Entanglement [...]

Non-local boxes [...] (another way to describe these things)

Teleportation [...] (show the key role of entanglement)

Chapter 2

Background theory

2.1 No-cloning theorem

Theorem 2.1.1 (no-cloning thorem).
Quantum information can not be cloned.
More formally, given two non ortogonal, non parallel vectors i.e. two vectors $|\phi_1\rangle$, $|\phi_2\rangle$ such that

$$0 < |\langle\phi_1|\phi_2\rangle| < 1 \tag{2.1}$$

then no unitary transformation $\hat{U}$ can implement the following

$$\forall j,\ \hat{U}\,|\phi_j\rangle_A\,|0\rangle_B\,|0\rangle_C = |\phi_j\rangle_A\,|\phi_j\rangle_B\,|\omega\rangle_C \tag{2.2}$$

remarks

- In closed systems, the only transformations allowed are unitary transformations.
- the unitary transformations do not change the inner product.
- the value of j is unknown to the transformation. In other words the task is to clone an unknown state. If the state is known, it is possible to clone it.
- A is the system *the state of which* is to be copied.
- B is the system *to which* we want to copy the state of A. The starting state is an arbitrary *known* state. Here $|0\rangle$ is used just as a dummy ket. If we want to start with another dummy ket, we can always apply an unitary transformation from $|0\rangle$ to the desired state, and "absorb" this unitary transformation in $\hat{U}$.

proof of the no-cloning theorem.

We want to compare the inner product between the two possible initial states, and the inner product between the two corresponding final states. Here *initial* and *final* means before and after the application of the transformation U. The inner product before the transformation is:

$$
\begin{aligned}
&|(\langle\phi_1|_A \langle 0|_B \langle 0|_C)(|\phi_2\rangle_A |0\rangle_B |0\rangle_C)| = && (2.3a)\\
&= |\langle\phi_1|\phi_2\rangle_A \underbrace{\langle 0|0\rangle_B}_{=1} \underbrace{\langle 0|0\rangle_C}_{=1}| && (2.3b)\\
&= |\langle\phi_1|\phi_2\rangle| && (2.3c)
\end{aligned}
$$

If we apply the transformation to both the vectors, and we compute the inner product, we have:

$$
\begin{aligned}
&\left|\left(\langle\phi_1|_A \langle 0|_B \langle 0|_C \hat{U}^\dagger\right)\left(\hat{U}\,|\phi_2\rangle_A |0\rangle_B |0\rangle_C\right)\right| = && (2.4a)\\
&= |\,(\langle\phi_1|_A \langle\phi_1|_B \langle\omega_1|_C)\,(|\phi_2\rangle_A |\phi_2\rangle_B |\omega_2\rangle_C)\,| && (2.4b)\\
&= |\langle\phi_1|\phi_2\rangle_A \langle\phi_1|\phi_2\rangle_B \langle\omega_1|\omega_2\rangle_C| && (2.4c)\\
&= |\langle\phi_1|\phi_2\rangle|^2\;|\langle\omega_1|\omega_2\rangle| && (2.4d)
\end{aligned}
$$

Now, because of the relation (2.1), it is

$$|\langle\phi_1|\phi_2\rangle|^2 < |\langle\phi_1|\phi_2\rangle| \tag{2.5}$$

with *strict* inequality. This means that the result (2.4d) is incompatible with the result (2.3c). This contradicts the property of the unitary transformations (which are supposed to conserve the inner product).
This proves that such transformation does not exist!

□

This proof is shorter, and uses explicitly the state that is to be cloned. In [Wil11][§3.5.4] there is a different, longer proof, which works with the unknown $|\phi_j\rangle$ state.

2.2 Side on the tensor product

The inner product defined on the tensor-product-space (compound system) is the product of the inner products defined on each sub-system.
The intuition on the inner product is that it measures the indistinguishability between two vectors (quantum states), zero (perpendicular states) meaning the highest distinguishability. Then the distinguishability of a compound system is the product from

each subsystem. Taking the product is consistent with the meaning of the overall distinguishability, because if the two states of a subsystem are perpendicular, their inner product is zero, and this makes zero the all inner product, *i.e.* makes the whole state of the compound system, regardless whether the other states of the other subsystems are indistinguishable or not.

Part I

Entanglement in physics

Chapter 3

Bell's inequalities

(Note to self: look for "Tsirelson's bound" [Cir80])

3.1 Bell theorem

[Bel64]
As anticipated in the introductory chapter 1.2, Since Quantum Theory yields only *expected values*, the quantity chosen to investigate the difference between LHVT and QT is an expectation value.

3.1.1 Definitions

We will study a compound system, made of two subsystems. We will describe the outcomes of measurements done on the two subsystems, in two separate spatial locations. The measurement apparatuses are supposed to be identical. Moreover, the measurement is specified by a *measurement parameter*. We will call the two outcomes

$$A(\vec{a}); B(\vec{b}) \tag{3.1}$$

respectively, where a and b are the measurement parameters, which in general they can be vectors. To fix the ideas we can imagine the measurement being the projection of a physical quantity (*e.g.* the angular momentum) along a certain direction, the direction's vector being the measurement's parameter. The measurement outcomes are limited to two possible values:

$$\forall a, A(\vec{a}), B(\vec{a}) \in \{+1, -1\}. \tag{3.2}$$

Finally, we assume that there exist a "physical correlation" between the two outcomes of the two measurements (on the two subsystems) in the case when the two measurement parameters are chosen to be equal:

$$\forall a A(\vec{a}) = -B(\vec{a}). \tag{3.3}$$

We can imagine a realistic physical situation where this is the case: if the measurement performed is the angular momentum of the two subsystem, a total angular momentum conservation will provide for this desired correlation. We introduce also the *correlation function*, defined as the *expectation value* (average over several measurements) of the product of the two outcomes:

$$P(\vec{a}, \vec{b}) \equiv \langle A(\vec{a}) B(\vec{b}) \rangle. \tag{3.4}$$

Comments:
- the correlation in the case of coincident measurement parameters is assumed to be 1 as "physical hypothesis"
- in the generic case of different measurement parameters, the *expectation value of the product* is a good expression of the correlation: if the two outcomes are uncorrelated, the different values of the product will be averaged out, and the expectation value will be zero. The more the *correlation* (same outcomes from the two measurements on the two subsystems, either both $+1$ or both -1), the closer the average will be to 1. The more the *anti-correlation* (opposite outcomes) the closer the average will be to -1.

In the two following sections we will consider the difference of two correlations for two different couples of measurements parameters value, one of the two values being fixed:

$$P(\vec{a}, \vec{b}) - P(\vec{a}, \vec{c}) = \langle A(\vec{a}) B(\vec{b}) - A(\vec{a}) B(\vec{c}) \rangle \tag{3.5}$$

and we will compute this difference using two different "models".

3.1.2 Quantum mechanical model

Applying the formalism of quantum mechanics [...] we have:

$$\langle A(\vec{a}) B(\vec{b}) \rangle = -\vec{a} \cdot \vec{b} \tag{3.6}$$

$$\boxed{\langle A(\vec{a}) B(\vec{a}) \rangle = -\vec{a} \cdot \vec{b}} \tag{3.7}$$

3.1.3 Local Hidden Variables model

In this section we want to compute the correlation difference $\langle P(\vec{a}, \vec{b}) - P(\vec{a}, \vec{c}) \rangle$ using the following model. We consider a set of unknown variables, denoted collectively with λ.

In this model the outcomes A and B depend not only on the measurement parameter, but also on these hidden, unknown variables:

$$A = A(\vec{a}, \lambda);\;\; B = B(\vec{b}, \lambda) \tag{3.8}$$

The outcomes values are deterministically dependent on those two values (the measurement parameter and the hidden variable). We observe unpredictable fluctuations because the hidden variables fluctuate, and we don't know their value. If we call Γ the set of all the possible values of the hidden variables, the expectation values will be found integrating λ over Γ:

$$P(\vec{a}, \vec{b}) = \int_\Gamma A(\vec{a}, \lambda) B(\vec{b}, \lambda) \rho(\lambda) \mathrm{d}\lambda \tag{3.9}$$

where $\rho(\lambda)$ is the probability distribution of λ.

derivation of the inequality

From now on, to make the notation more legible, we will omit the arrow on the vectors representing the measurement parameter.

$$P(a, b) - P(a, c) = \int_\Gamma \left[A(a, \lambda)B(b, \lambda) - A(a, \lambda)B(c, \lambda)\right] \rho(\lambda)\mathrm{d}\lambda \tag{3.10a}$$

we can use the "physical correlation" hypothesis: $\forall a, \lambda A(a, \lambda) = -B(a, \lambda)$ and have

$$P(a, b) - P(a, c) = -\int_\Gamma \left[A(a, \lambda)A(b, \lambda) - A(a, \lambda)A(c, \lambda)\right] \rho(\lambda)\mathrm{d}\lambda \tag{3.10b}$$

Then we use the hypothesis: $\forall b, A(b, \lambda) \in \{\pm 1\} \Rightarrow A^2(b, \lambda) = 1$

$$\begin{aligned}
P(a, b) - P(a, c) &= -\int_\Gamma \left[A(a, \lambda)A(b, \lambda) - A(a, \lambda)A^2(b, \lambda)A(c, \lambda)\right] \rho(\lambda)\mathrm{d}\lambda & (3.10c)\\
&= -\int_\Gamma A(a, \lambda)A(b, \lambda)\left[1 - A(b, \lambda)A(c, \lambda)\right] \rho(\lambda)\mathrm{d}\lambda & (3.10d)\\
&= \int_\Gamma A(a, \lambda)A(b, \lambda)\left[A(b, \lambda)A(c, \lambda) - 1\right] \rho(\lambda)\mathrm{d}\lambda. & (3.10e)
\end{aligned}$$

Now, if we consider the absolute value of both terms, we have:

$$\begin{aligned}
|P(a, b) - P(a, c)| &= \left|\int_\Gamma A(a, \lambda)A(b, \lambda)\left[A(b, \lambda)A(c, \lambda) - 1\right] \rho(\lambda)\mathrm{d}\lambda\right| & (3.10f)\\
&= \left|\int_\Gamma A(a, \lambda)A(b, \lambda)\left[1 - A(b, \lambda)A(c, \lambda)\right] \rho(\lambda)\mathrm{d}\lambda\right|. & (3.10g)
\end{aligned}$$

Then, we consider that the content of the square bracket can be either 0 or 2, so it is non negative; and on the other hand the product $A(a,\lambda)A(b,\lambda)$ can be either $+1$ or -1, and then $[A(a,\lambda)A(b,\lambda)] \leq 1$.

So, we can write:

$$
\begin{aligned}
|P(a,b) - P(a,c)| \leq & \left| \int_\Gamma [1 - A(b,\lambda)A(c,\lambda)]\, \rho(\lambda) \mathrm{d}\lambda \right| && (3.10h) \\
= & \left| \int_\Gamma \rho(\lambda)\mathrm{d}\lambda - \int_\Gamma A(b,\lambda)A(c,\lambda)\rho(\lambda)\mathrm{d}\lambda \right| && (3.10i) \\
\leq & \left| \int_\Gamma \rho(\lambda)\mathrm{d}\lambda \right| + \left| \int_\Gamma A(b,\lambda)A(c,\lambda)\rho(\lambda)\mathrm{d}\lambda \right| && (3.10j) \\
= & 1 + P(b,c) && (3.10k)
\end{aligned}
$$

where we have used the normalization of the probability distribution: $\int_\Gamma \rho(\lambda)\mathrm{d}\lambda = 1$. Finally, we obtain the Bell inequality

$$
\boxed{|P(a,b) - P(a,c)| \leq 1 + P(b,c)} \qquad \text{(Bell's inequality)} \qquad (3.11)
$$

3.1.4 Contradiction

Once we have computed, in the two previous subsections, an expression of the same quantity (the difference of two correlation functions computed with "three measurement parameter settings"), using to different models, we want now to compare the two results. The claim is that the two results are incompatible. More precisely, the claim is that the quantum mechanical result can not satisfy equation (3.11).

first argument, semi-rigorous

If we look at the left hand side of equation (3.11), we can perform a Taylor (actually Mac Laurin) expansion of the two functions $P(a,b)$ and $P(a,c)$, with respect to $(a-b)$ and $(a-c)$ respectively, around the value 0 *i.e.* around $(a=b)$ and $(a=c)$ respectively. Because of the minus sign, the terms in a will cancel out. We can conclude that the left hand side is of the order of $|b-c|$. This means that for $b \to c$ the right hand side is approaching the value 0 linearly, and because of the absolute value, this minimum can not be a stationary point (technically, it is a point where the left and right derivative are different: they are the slope of the line). See figure 3.1.

So we have that $P(b,c)$ is lower bounded by such a function.

On the other hand, we remember that the quantum mechanical model predicts for the same quantity, the result $-b \cdot c = -\cos\theta$, where θ is the angle between b and c. As b

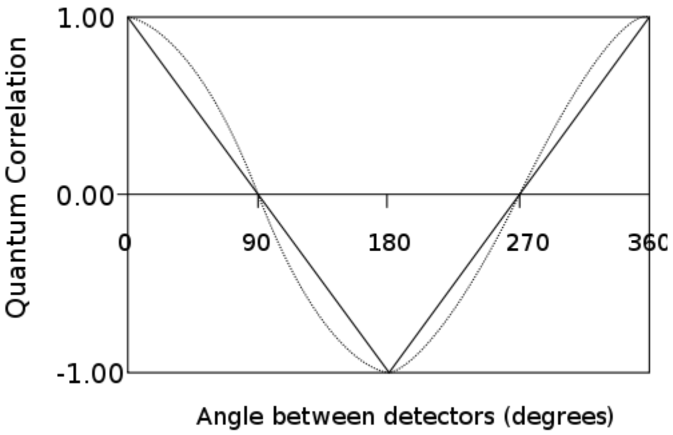

Figure 3.1: solid line= prediction of LHV model. Dotted line=prediction of the QM model

approaches c, *i.e.* as θ approaches zero, the function $\cos\theta$ has the same derivative on both sides. In other words the function has a stationary point in zero.

Since the LHV model's prediction is lower bounded by the linear function, which is strictly greater than the QM prediction in a neighborhood of 0, this means that the two predictions are incompatible.

Detailed proof

[...]

3.2 CHSH inequality

The CHSH [CHSH69] inequality is a generalization of the Bell's inequality. The meaning is similar, but it applies to a more general situation.

In particular, the limitation of Bell's hypotheses, that CHSH want to address, is the "visibility" of the anti-correlated states measurements. Let's explain. In a Bell-like experiment, we have a couple of particles, in a singlet state. If the two measurements are done along the same direction, according to the theory the data should be perfectly correlated (or anti-correlated), *i.e.* the correlation function should be ±1. But because of experimental errors, it is likely that the results are not exactly ±1.

Now, let's go into the details.

In the same experimental situation as for Bell's theorem, and using the same definition (3.4) of correlation $P(a, b)$, we consider the following absolute value of the difference between two correlations:

$$|P(a,b)-P(a,c)| = \left|\int_\Gamma \left[A(a,\lambda)B(b,\lambda)-A(a,\lambda)B(c,\lambda)\right]\rho(\lambda)\mathrm{d}\lambda\right| \tag{3.12}$$

Now, we use the relation:

$$\left|\int f(x)\mathrm{d}x\right| \leq \int |f(x)|\,\mathrm{d}x \tag{3.13}$$

(see [Fio95, vol I, §11.II, pag 426, eq. (11)]) Now, using the fact that

$$\forall a, A(a,\lambda) \in \{+1,-1\} \tag{3.14}$$

we can write:

$$\begin{align}
|P(a,b)-P(a,c)| &\leq \int_\Gamma |A(a,\lambda)B(b,\lambda)-A(a,\lambda)B(c,\lambda)|\,\rho(\lambda)\mathrm{d}\lambda \tag{3.15a}\\
&= \int_\Gamma \left|A(a,\lambda)B(b,\lambda)-A(a,\lambda)\underbrace{B^2(b,\lambda)}_{=1}B(c,\lambda)\right|\rho(\lambda)\mathrm{d}\lambda \tag{3.15b}\\
&= \int_\Gamma \left|A(a,\lambda)B(b,\lambda)\underbrace{[1-B(b,\lambda)B(c,\lambda)]}_{=\{0\text{ or }2\}\geq 0}\right|\rho(\lambda)\mathrm{d}\lambda \tag{3.15c}\\
&= \int_\Gamma \underbrace{|A(a,\lambda)B(b,\lambda)|}_{=1}\left[1-B(b,\lambda)B(c,\lambda)\right]\rho(\lambda)\mathrm{d}\lambda \tag{3.15d}\\
&= \int_\Gamma \left[1-B(b,\lambda)B(c,\lambda)\right]\rho(\lambda)\mathrm{d}\lambda \tag{3.15e}\\
&= \underbrace{\int_\Gamma \rho(\lambda)\mathrm{d}\lambda}_{} = 1-\int_\Gamma B(b,\lambda)B(c,\lambda)\rho(\lambda)\mathrm{d}\lambda \tag{3.15f}\\
&= 1-\int_\Gamma B(b,\lambda)B(c,\lambda)\rho(\lambda)\mathrm{d}\lambda \tag{3.15g}
\end{align}$$

summarizing:

$$|P(a,b)-P(a,c)| \leq 1-\int_\Gamma B(b,\lambda)B(c,\lambda)\rho(\lambda)\mathrm{d}\lambda \tag{3.16}$$

At this point we introduce two “measurement positions”, called b and b' such that:

$$P(b',b) = (1-\delta) \tag{3.17}$$

where $0 \leq \delta \leq 1$. Experimentally interesting values for *delta* are values close to, but different from, zero.
Now, we split the Γ region in two, such that:

$$\Gamma_\pm := \{\lambda : A(b,\lambda) = \pm B(b',\lambda)\} \tag{3.18}$$

And, elaborating on equation (3.17) we can write:

$$\begin{aligned}
P(b,b') &\overset{\text{def}}{=} \int_\Gamma A(b,\lambda)B(b',\lambda)\rho(\lambda)\mathrm{d}\lambda = (1-\delta) && \text{(3.19a)}\\
&= \int_{\Gamma_+} A(b,\lambda)B(b',\lambda)\rho(\lambda)\mathrm{d}\lambda + \int_{\Gamma_-} A(b,\lambda)B(b',\lambda)\rho(\lambda)\mathrm{d}\lambda && \text{(3.19b)}\\
&= \int_{\Gamma_+} A(b,\lambda)A(b,\lambda)\rho(\lambda)\mathrm{d}\lambda - \int_{\Gamma_-} A(b,\lambda)A(b,\lambda)\rho(\lambda)\mathrm{d}\lambda && \text{(3.19c)}\\
&= \int_{\Gamma_+} \underbrace{A^2(b,\lambda)}_{=1}\rho(\lambda)\mathrm{d}\lambda - \int_{\Gamma_-} \underbrace{A^2(b,\lambda)}_{=1}\rho(\lambda)\mathrm{d}\lambda && \text{(3.19d)}\\
&= \int_{\Gamma_+} \rho(\lambda)\mathrm{d}\lambda - \int_{\Gamma_-} \rho(\lambda)\mathrm{d}\lambda && \text{(3.19e)}\\
&= 1 - 2\int_{\Gamma_-} \rho(\lambda)\mathrm{d}\lambda && \text{(3.19f)}\\
(1-\delta) &= 1 - 2\int_{\Gamma_-} \rho(\lambda)\mathrm{d}\lambda && \text{(3.19g)}\\
&\Rightarrow\\
\int_{\Gamma_-} \rho(\lambda)\mathrm{d}\lambda &= \tfrac{1}{2}\delta && \text{(3.19h)}
\end{aligned}$$

where in (3.19f) we have used the relation:

$$\begin{aligned}
\int_{\Gamma_+} f(x)\mathrm{d}x - \int_{\Gamma_-} f(x)\mathrm{d}x = && \text{(3.20a)}\\
= \int_{\Gamma_+} f(x)\mathrm{d}x + \underbrace{\int_{\Gamma_-} f(x)\mathrm{d}x - \int_{\Gamma_-} f(x)\mathrm{d}x}_{=0} - \int_{\Gamma_-} f(x)\mathrm{d}x && \text{(3.20b)}\\
= \int_{\Gamma} f(x)\mathrm{d}x + \int_{\Gamma_-} f(x)\mathrm{d}x - 2\int_{\Gamma_-} f(x)\mathrm{d}x && \text{(3.20c)}\\
= 1 - 2\int_{\Gamma_-} f(x)\mathrm{d}x. && \text{(3.20d)}
\end{aligned}$$

Then, looking at the right-hand-side of equation (3.16), and using the definition of $\Gamma_\pm$, we have:

$$\int_\Gamma B(b,\lambda)B(c,\lambda)\rho(\lambda)\mathrm{d}\lambda = \int_{\Gamma_+} A(b',\lambda)B(c,\lambda)\rho(\lambda)\mathrm{d}\lambda - \int_{\Gamma_-} A(b',\lambda)B(c,\lambda)\rho(\lambda)\mathrm{d}\lambda \tag{3.21a}$$

$$= \int_\Gamma A(b',\lambda)B(c,\lambda)\rho(\lambda)\mathrm{d}\lambda - 2\int_{\Gamma_-} A(b',\lambda)B(c,\lambda)\rho(\lambda)\mathrm{d}\lambda \tag{3.21b}$$

$$= P(b',c) - 2\int_{\Gamma_-} A(b',\lambda)B(c,\lambda)\rho(\lambda)\mathrm{d}\lambda \tag{3.21c}$$

$$\geq P(b',c) - 2\int_{\Gamma_-} \underbrace{|A(b',\lambda)B(c,\lambda)|}_{=1}\rho(\lambda)\mathrm{d}\lambda \tag{3.21d}$$

$$= P(b',c) - 2\int_{\Gamma_-} \rho(\lambda)\mathrm{d}\lambda \tag{3.21e}$$

$$= P(b',c) - \delta \tag{3.21f}$$

where we have used equation (3.19h). Summarizing

$$\int_\Gamma B(b,\lambda)B(c,\lambda)\rho(\lambda)\mathrm{d}\lambda \geq P(b',c) - \delta \tag{3.22}$$

Using this last result in (3.16) we have

$$|P(a,b) - P(a,c)| \leq 1 - [P(b',c) - \delta] \tag{3.23}$$

And replacing the definition (3.17):

$$|P(a,b) - P(a,c)| \leq 1 - P(b',c) + \delta \tag{3.24a}$$

$$= 1 - P(b',c) + [1 - P(b,b')] \tag{3.24b}$$

$$= 2 - P(b',c) - P(b,b') \tag{3.24c}$$

Transporting the right terms to the left, we finally have

$$\boxed{|P(a,b) - P(a,c) + P(b',c) + P(b,b')| \leq 2} \tag{3.25}$$

which is the CHSH inequality.

3.2.1 Coincidences

Another merit of CHSH with respect to Bell (beside the more simple form) is to express the inequality also in terms of "coincidences of detections" instead of probabilities. This second relation is a more useful one, because the photon detectors have a quantum efficiency far from unity!
[...]

Part II

Entanglement in Information Theory

Chapter 4

Non-local boxes

(This chapter is based on the lecture notes of from Patrick Hayden, McGill University, 2013)
Let's consider a function

$$f : \underbrace{\{0,1\}^n}_{Alice's} \times \underbrace{\{0,1\}^n}_{Bob's} \to \{0,1\} \tag{4.1}$$

Definition 4.0.1 (communication complexity).
# of bits needed to be exchanged between Alice and Bob, in order to compute the function

Let's see some examples:

1. the function:

   $$f(x^n, y^n) = x_j \oplus y_j \tag{4.2}$$

   has comm. compl. = 1

   Proof:
   Alice needs to send to Bob her j-th bit, Bob already knows his j-th bit and can compute $x_j \oplus y_j$. □

2. the function

   $$f(x^n, y^n) = \bigvee_{j=1}^{n} (x_j \wedge y_j) \tag{4.3}$$

   we can give a "meaning" to this function: the binary variables mean whether in a specific day the person is free or not, and the function computes whether in the range of days considered there is at least one day where both are free. This function is called *non destroyingness.*

The comm. compl. of this function is not 1. Patrick is not sure but should be O(n)

3. the inner product

$$f(x^n, y^n) = \sum_{j=1}^{n} x_j \cdot y_j \mod 2 \tag{4.4}$$

has comm. compl. = n

Now we introduce a *fictitious device* called Non-Local Boxes (NLB)

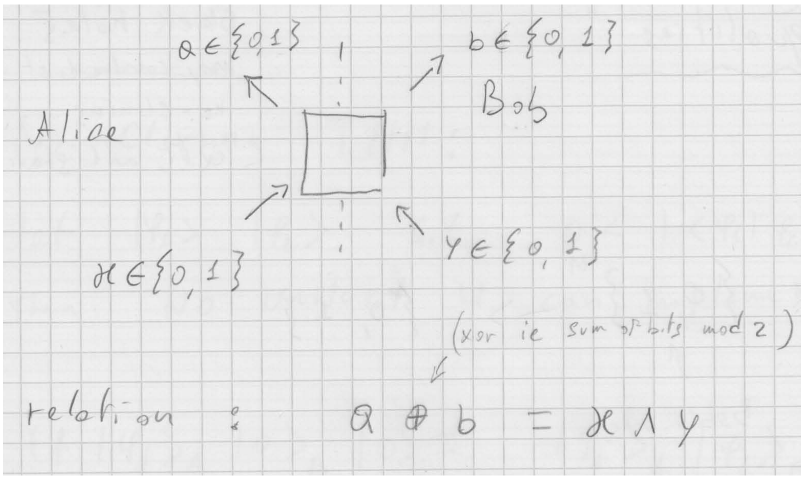

Figure 4.1: Non-Local Boxes

So, this device is made of two (non-local) parts, one for each party, who are supposed to be far apart from each other.
The device works like this:
each party inputs a bit, Alice inputs x and Bob inputs y. Then, each party receives an output of one bit, a for Alice and b for Bob. The device is such that

$$a \oplus b = x \wedge y \tag{4.5}$$

In other words, the device is such that there is a special correlation between the two outputs.
Let's point out that this device does not allow any communication between Alice and Bob. Indeed, if the device was such that the bit Bob receives as output is the bit input

by Alice, the device would allow communication. But in the case at hand there is a more subtle relation.

We can prove that the NLB do not allow communication.

Indeed, a way to implement the NLB (a way that complies with the definition (4.5)) is if a is completely random, and $b = a \oplus (x \wedge y)$.

This implementation satisfy the definition because:

$$\begin{aligned} a \oplus b &= a \oplus (a \oplus (x \wedge y)) && \text{(4.6a)} \\ &= (a \oplus a) \oplus (x \wedge y) && \text{(4.6b)} \\ &= 0 \oplus (x \wedge y) && \text{(4.6c)} \\ &= (x \wedge y) && \text{(4.6d)} \end{aligned}$$

as required.

In the last computation we have used the distributivity of the “x-or”, the result $\forall a, a \oplus a = 0$, and the result $\forall x, x \oplus 0 = x$.

This shows that, in the case a is random, b is also random, (because b is equal to a random variable xor-ed with $(x \wedge y)$), and in particular a is uncorrelated to y, and b is uncorrelated with x.

4.1 NLB collapse the communication complexity

Statement 4.1.1.

*If Alice and Bob have access to NLB, the communication complexity of *any* boolean function “collapses” to 1:*

$$\forall f, CC^{NLB}(f) \leq 1 \tag{4.7}$$

Proof.

It is possible to show that the set of the following three logic ports:

- NOT
- fan-out
- AND

is a complete set, meaning that any boolean function can be implemented using only those three ports. So, if we show that we can implement these three ports, we have shown the result for any function. Remark: what we want to implement is the non-local version of these ports. This means that we define a non-local bit, as the couple of two local bits, each held by each the two parties Alice and Bob: (a, b). The value of the non-local bit will be defined as

$$x \overset{\text{def}}{=} a \oplus b \tag{4.8}$$

non-local NOT gate

[...] it is sufficient that one of the two party negates its bit [...]

non-local fan-out

[...] it is sufficient that each party clones its bit [...]

non-local AND gate

Alice and Bob start with two non-local bits

$$x = (c \oplus d) \tag{4.9a}$$

$$y = (e \oplus f) \tag{4.9b}$$

where c and e are local to Alice, and d and f are local to Bob. The goal is to compute:

$$x \wedge y = (c \oplus d) \wedge (e \oplus f) \tag{4.10a}$$

$$= \underbrace{(c \wedge e)}_{\text{local to Alice}} \oplus (c \wedge f) \oplus (d \wedge e) \oplus \underbrace{(d \wedge f)}_{\text{local to Bob}} \tag{4.10b}$$

Since c and e are local to Alice, she can locally compute $(c \wedge e)$. In the same way, Bob can locally compute $(d \wedge f)$. To obtain $(c \wedge f)$ and $(d \wedge e)$, Alice and Bob exploit the NLB: they input their local bits in the NLB, and keep the corresponding outputs. In particular

- Alice inputs c, Bob inputs f, and then Alice gets the output g, and Bob gets the output h, such that (by definition of NLB) $g \oplus h = c \wedge f$

- Alice inputs d, Bob inputs e, and then Alice gets the output l, and Bob gets the output k, such that (by definition of NLB) $l \oplus k = d \wedge e$

At the end of this round of operations:

- Alice holds $(c \wedge e)$, g, l

- Alice holds $(d \wedge f)$, h, k

To obtain the final result what is needed is:

$$(c \wedge e) \oplus (c \wedge f) \oplus (d \wedge e) \oplus (d \wedge f) \tag{4.11a}$$

$$(c \wedge e) \oplus (g \oplus h) \oplus (l \oplus k) \oplus (d \wedge f) \tag{4.11b}$$

$$[(c \wedge e) \oplus g \oplus l] \oplus [h \oplus k \oplus (d \wedge f)] \tag{4.11c}$$

and we notice that the first square bracket is local to Alice, and the second is local to Bob.

So, to obtain the final result, it is sufficient that one of the two parties sends its partial result (the content of square brackets in (4.11c)) to the other party.

We stress that this final exchange of information is the only one necessary, and it consists of 1 bit.

In other words, all the computation has been done locally; to compute the desired "non-local bit" all is needed is that one of the two parties sends to the other its "local part" of the non-local final result (a non local bit is a couple of local bits), in order to compute the final $\oplus$.

Because of the linearity, once we have proven this result for the complete set of the three logic gates, this proves the result for any generic boolean function.

□

4.2 Classical simulation of NLB

[...]

4.3 Quantum simulation of NLB

[...]

4.4 Summary: CHSH inequality

[...]

Chapter 5

Quantum channels

5.1 Quantum channels

5.2 Superdense coding

The original idea was published in 1992 [BW92].
See also Patrick Hayden course, lecture 10 ($t1^h9'0''$)
Let's ask which is the most efficient way to transmit classical information using qubits. The trivial way is to choose a basis, encode each bit in a qubit, transmit the qubits, and measure the qubits *projecting* on the chosen basis.
We can generalize this and say that *if we want no errors* (distinguish perfectly), to transmit d distinguishable messages we need a d-dimensional quantum system. Then we choose an orthonormal basis, and encode 1 bit per qubit. This is because, if we don't want errors, we need all orthonormal projectors, and we need a d-dimensional system in order to have d orthonormal projectors. If we try to "squeeze" more information we will induce errors, because of non-distinguishable states.
However, there is a *loophole* to this, and the loophole is represented by entanglement!
Let's imagine (see figure 5.1) Alice having two bits (i.e. 4 symbols), and trying to send them to Bob. They share an entangled pair of qubits, let's say $|\Phi_+\rangle$. Then, Alice performs a unitary operation on the qubit of the entangled pair that she has, and in particular she applies one of the four "one qubit Pauli operators", depending on which of the four symbols she wants to send. In particular, she will apply one of the four "one qubit Pauli matrices" (represented in the computational basis):

$$\sigma_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}; \; \sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}; \; \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}; \; \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{5.1}$$

Then she sends her entangled qubit to Bob. Bob has now two qubits, so he can perform a projective measurement.

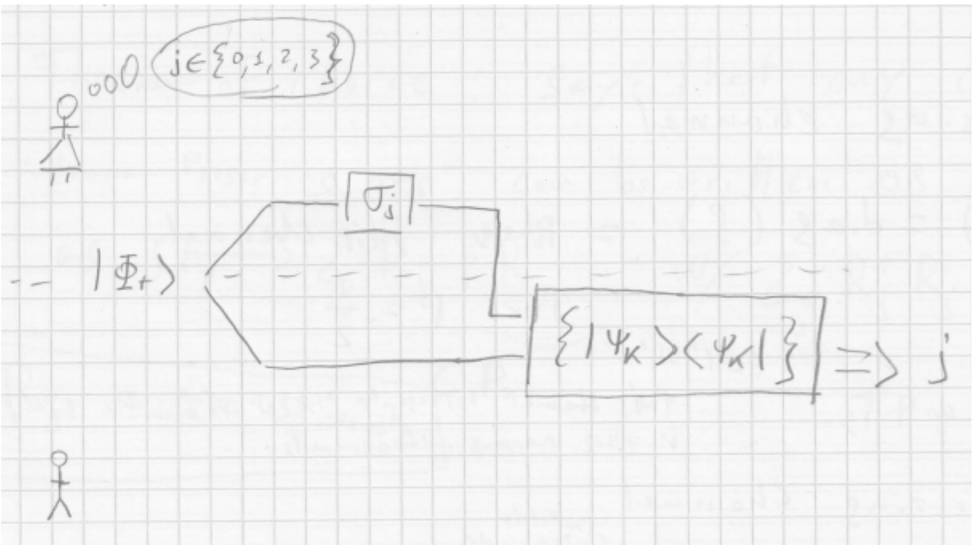

Figure 5.1: Alice has two bits (four labels), and she wants to transmit them to Bob. the dotted line separates Alice's and Bob's labs. They share an entangled state $|\Phi_+\rangle$. Then Alice sends her qubit to Bob, and Bob performs a projective measurement.

To understand which four projectors to use for this projective measurement, let's consider the four possibile two-qubits state he may have, depending on which Pauli operator Alice has applied.
The starting entangled pair is:

$$|\Phi_+\rangle = |00\rangle + |11\rangle \tag{5.2}$$

then we can represent the generic state of the entangled pair after the action of Alice's unitary operation as:

$$|\psi_k\rangle = (\sigma_k \otimes \mathbb{I})\,|\Phi_+\rangle \tag{5.3}$$

explicitly:

$$|\psi_0\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle) \tag{5.4a}$$

$$|\psi_1\rangle = \frac{1}{\sqrt{2}}(|10\rangle + |01\rangle) \tag{5.4b}$$

$$|\psi_2\rangle = \frac{1}{\sqrt{2}}(-i\,|10\rangle + i\,|01\rangle) \tag{5.4c}$$

$$|\psi_3\rangle = \frac{1}{\sqrt{2}}(|00\rangle - |11\rangle) \tag{5.4d}$$

It can be shown (by directly computing the products) that those vectors are orthonormal:

$$\langle \psi_k | \psi_j \rangle = \delta_{k,j}. \tag{5.5}$$

The net result is that:

Theorem 5.2.1 (Superdense coding)**.** *It is possible to transmit 2 bits transmitting 1 qubit, and consuming 1 ebit (1 pair of maximally entangled qubits) .*

Chapter 6

Teleportation

(note to self: read [BBC98])

Bibliography

[BBC98] Gilles Brassard, Samuel L Braunstein, and Richard Cleve. Teleportation as a quantum computation. *Physica D: Nonlinear Phenomena*, 120(1-2):43–47, 1998.

[Bel64] John S Bell. On the Einstein Podolsky Rosen paradox. *Physics*, 1(3):195–200, 1964.

[BW92] Charles H. Bennett and Stephen J. Wiesner. Communication via one- and two-particle operators on einstein-podolsky-rosen states. *Phys. Rev. Lett.*, 69:2881–2884, Nov 1992.

[CHSH69] John F. Clauser, Michael A. Horne, Abner Shimony, and Richard A. Holt. Proposed Experiment to Test Local Hidden-Variable Theories. *Phys. Rev. Lett.*, 23:880–884, October 1969.

[Cir80] B.S. Cirel'son. Quantum generalizations of Bell's inequality. *Letters in Mathematical Physics*, 4:93–100, 1980.

[EPR35] A. Einstein, B. Podolsky, and N. Rosen. Can Quantum-Mechanical Description of Physical Reality Be Considered Complete? *Physical Review*, 47(10):777–780, May 1935.

[Fio95] Renato Fiorenza. *Lezioni di Analisi Matematica*, volume I. Liguori, Napoli, 3rd edition, 1995.

[Wil11] Mark M Wilde. From Classical to Quantum Shannon Theory. *arXiv.org*, quant-ph, June 2011.

[WM08] Daniel F. Walls and Gerard J. Milburn. *Quantum optics.* Springer-Verlag, Berlin, Heidelberg, second (revised and enlarged) edition, 2008.