Part I

Quantum entanglement and non-locality

1

Non-locality beyond quantum mechanics

Sandu Popescu

For Abner Shimony. Your influence on me goes well beyond physics. Knowing you and being close to you is one of the greatest privileges and pleasures in my life.

1.1 Introduction

Quantum mechanics is, without any doubt, a tremendously successful theory: it started by explaining black-body radiation and the photoelectric effect, it explained the spectra of atoms, and then went on to explain chemical bonds, the structure of atoms and of the atomic nucleus, the properties of crystals and the elementary particles, and a myriad of other phenomena. Yet it is safe to say that we still lack a deep understanding of quantum mechanics – surprising and even puzzling new effects continue to be discovered with regularity. That we are surprised and puzzled is the best sign that we still don't understand; however, the veil over the mysteries of quantum mechanics is starting to lift a little.

One of the strangest things microscopic particles do is to follow non-local dynamics and to yield non-local correlations. That particles follow non-local equations of motion was discovered by Aharonov and Bohm [1], while non-local correlations – which are the subject of this chapter – were discovered by John Bell [2] and first cast in a form that has physical meaning, i.e., that can be experimentally tested, by Clauser, Horne, Shimony, and Holt [3]. When they were discovered, both phenomena seemed to be quite exotic and at the fringe of quantum mechanics. By now we understand that they are some of the most important aspects of quantum-mechanical behavior.

Consider two experimentalists, Alice and Bob, situated on different planets, far from each other. They perform experiments on particles that come from a common source and are prepared in a so-called entangled state. The experiments are "space-like separated." They take a short time compared with the time required for light, and, according to Einstein, for any other signal, to propagate from Alice to Bob.

Philosophy of Quantum Information and Entanglement, ed. A. Bokulich and G. Jaeger. Published by Cambridge University Press.

Furthermore, by pre-arrangement, the experiments are timed in such a way that Alice's experiment finishes before she could receive any signal from Bob about the experiment he performed (what experiment he did and what the result was) and similarly all information from Alice about the experiment she performed can reach Bob only after he has finished his experiment. Nevertheless, their results turn out to be correlated (although this can be found out only later, when Alice and Bob are able to compare their results). The fact that the results are correlated is not a great surprise – after all, the particles came from a common source. What is astonishing, however, is that they are correlated in such a way that, if we want to establish such correlations with any classical devices, they have to communicate with each other. Owing to the timing of the experiments, this communication has to be superluminal. We refer to such correlations as non-local.

That non-local correlations can exist at all, and not lead immediately to conflict with Einstein's relativity, is possible only because the outcomes of the measurements are probabilistic. It is the fact that quantum mechanics is fundamentally indeterministic that opens an umbrella under which non-locality can "peacefully coexist," as Abner Shimony said, with relativity [4]. This is true not only for non-local correlations but also for the non-local equations of motion discussed by Aharonov and Bohm.

While non-locality requires indeterminacy, one can easily imagine indeterministic theories that do not present non-locality. This led Aharonov [5] and Shimony [6] independently to suggest that non-locality is a deeper aspect than indeterminism, and may be the reason why quantum mechanics is what it is. In effect, they suggested that relativity and the existence of non-locality could be the axioms that determine quantum mechanics.

Following these suggestions, Daniel Rohrlich and I asked whether quantum mechanics is the only possible theory that allows the coexistence of non-locality and relativity [7]. As a first step we asked whether there could be non-local correlations that do not violate relativity (are "non-signaling") but cannot be obtained from quantum mechanics. To our surprise we found that such correlations are theoretically possible. But are there such correlations in nature? If yes, where? And if not, why not?

While, after asking the question, finding out about the theoretical possibility of non-local correlations stronger than those arising from quantum mechanics proved to be relatively easy, understanding the significance of this discovery is far more difficult. Here I will describe some steps toward this goal. The story presented here is just a small part of the research in this direction. After its discovery, the idea of non-locality beyond quantum mechanics lay dormant for more than a decade (despite a breakthrough by van Dam [8], which was not published and hence went largely unnoticed). The subject was revived by Barrett *et al.* [9], who established a

framework for describing these correlations. At present the study of these correlations is a very active research area [10].

1.2 Non-local correlations beyond quantum mechanics

What allows one to derive general statements about the nature of correlations is the fact that experiments such as those of Alice and Bob can be described in a very general, model-independent way. For our present purpose, a very convenient way to view experiments is as input/output devices. Alice has a black box that accepts as input a number x and yields as output a number a. Similarly, Bob has a black box that accepts an input y and yields the output b. One can think of these black boxes as entire automated laboratories, containing particles, measuring devices, computers, etc. The laboratories are pre-arranged, ready to perform a number of different experiments. The inputs x and y simply indicate which experiment is to be performed, while the outputs a and b are the results of the experiments.

We consider that, as a matter of principle, Alice and Bob cannot look inside the labs and see exactly how the outputs are obtained. Furthermore, all Alice and Bob can do is to give the inputs – they do not control the outputs. In particular, when for a given input different outputs are possible, Alice and Bob cannot force a particular output. This rule mimics the quantum-mechanical behavior – when an experiment can yield different outcomes, the experimentalist cannot control which particular outcome will be obtained. We say that such boxes exhibit "fundamental indeterminacy."

Given the above setting, the entire physics is encapsulated in $P(a, b|x, y; t_x, t_y)$, the conditional joint probabilities of obtaining the outcomes a and b when the inputs are x, given at time t_x, and y, given at time t_y.

Throughout this text we are interested only in boxes that are consistent with relativity. In our context this implies two things. First, that as long as the inputs are outside the light-cone of each other, the probabilities are independent of the exact timing, hence they reduce to $P(a, b|x, y)$. Second, that they obey non-signaling constraints, namely that the probability that Alice obtains a particular outcome a is independent from Bob's input and vice versa:

$$\begin{aligned} \sum_b P(a, b|x, y) &= P(a|x), \\ \sum_a P(a, b|x, y) &= P(b|y). \end{aligned} \tag{1.1}$$

The situation is illustrated in Figure 1.1.

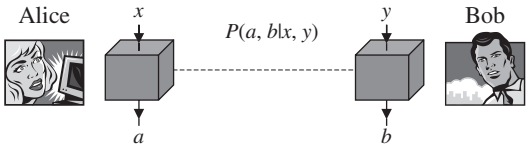

Fig. 1.1 An experimental setup viewed in terms of black boxes.

The case famously considered by Clauser, Horne, Shimony, and Holt is the simplest non-trivial case, in which all the inputs and outputs are binary. For our purposes we find it convenient to associate with x, y, a, and b the values of 0 and 1. Suppose Alice and Bob input at random, with equal probability, the values 0 and 1. It is easy to see that, in this language, what CHSH did was to analyze the probability with which the boxes succeed in yielding outputs such that

$$a \oplus b = xy, \tag{1.2}$$

where $\oplus$ denotes addition modulo 2. In other words, when x and y are both equal to 1 we succeed if the outputs are different, while in all other cases success is defined by the outputs being the same. The CHSH inequality tells us that, if the boxes work according to classical physics, the probability of success is bounded by

$$P_{\text{success}}^{\text{classical}}(a \oplus b = xy) \leq \frac{3}{4}. \tag{1.3}$$

On the other hand, if the boxes work according to quantum mechanics they can yield a larger success probability,

$$P_{\text{success}}^{\text{quantum}}(a \oplus b = xy) \leq \frac{2+\sqrt{2}}{4}. \tag{1.4}$$

(All one needs to do to find the above expressions is to convert the expectation values in the standard form of the CHSH inequality into probabilities and to note that each pair of inputs occurs with probability 1/4.)

The question Daniel Rohrlich and I asked was whether an even larger success probability could be obtained, that would be consistent with non-signaling. In the above language the answer is trivial, non-signaling doesn't constrain the maximal value of the success probability at all:

$$P_{\text{success}}^{\text{super-quantum}}(a \oplus b = xy) \leq 1. \tag{1.5}$$

Indeed, note that to each value of the product xy there correspond two different solutions for a and b: for $xy = 0$, the two solutions are $a = 0$, $b = 0$ and $a = 1$, $b = 1$, while for $xy = 1$ we have $a = 0$, $b = 1$ and $a = 1$, $b = 0$. As long as the devices yield each of the two solutions with equal probabilities, the local

probabilities for all outcomes are $1/2$ regardless of the inputs, so the correlations are non-signaling.

1.3 Communication complexity

Abner Shimony pointed out to me, and I fully agreed, that the correlations Daniel Rohrlich and I discovered are just a stand-alone example, and in order for their meaning to be evaluated they need to be integrated into a full theory that will explain all the known phenomena, along with instances in which stronger-than quantum correlations would appear. While this is certainly true, little did we know just how much one can milk this extremely simple example. The shock came for me when Wim van Dam showed me his results on communication complexity.

Consider again Alice and Bob. One day, overwhelmed by their "passion at a distance," they decide to meet for the first time. But they are very busy, so finding a day when they are both free is a difficult task. To make things more fun, they decide that, instead of trying to find a good day directly, they should first find out whether the total number of convenient days when they are both free this year is even or odd. Suppose furthermore that it is only Bob who will send information to Alice, and it is Alice who has the task of finding the result.

Of course, the task can be accomplished if Bob sends Alice his entire schedule. But this is very redundant: all Alice wants is one bit of information, i.e., "even" or "odd," but Bob has to send N bits of information, a "free" or "busy" for each of the N days of the year. Of course, they could try some other communication strategy, for example Bob could first tell Alice whether the total number of his free days is even or odd, then some other information, etc., etc. Unfortunately, it is quite easy to see that there is no method that requires less communication than sending his entire schedule.

In mathematical terms, any communication problem in which the answer is a single bit (i.e., a variable that can take only the value 0 or 1) can be formulated as follows. Alice and Bob each have N bits; Alice has $x_1, \ldots, x_N$ and Bob has $y_1, \ldots, y_N$; Alice doesn't know Bob's bits and Bob doesn't know Alice's. Let $f(\mathbf{x}, \mathbf{y})$ be a function of $\mathbf{x}$ and $\mathbf{y}$, where f can take only the value 0 or 1 and where by $\mathbf{x}$ and $\mathbf{y}$ we denote the sets of N bits $\mathbf{x} = \{x_1, \ldots, x_N\}$ and $\mathbf{y} = \{y_1, \ldots, y_N\}$. Alice and Bob know in advance the function f, and the task is for them to collaborate such that in the end Alice will find out the value of $f(\mathbf{x}, \mathbf{y})$.

To solve the problem, in general, Alice and Bob will have to communicate. We assume that they agree in advance on a specific communication protocol.

There are different problems we may consider – we are interested here in a one-way communication problem, in which it is only Bob who sends information to Alice.

Obviously, any such task can be accomplished if Bob tells Alice all his N bits. But, depending on the specific form of the function f, he might not need to send so much information. For example, if f is independent of $\mathbf{y}$ then Bob doesn't need to tell Alice anything; if f depends only on p, the parity of Bob's bits ($p = y_1 \oplus y_2 \oplus \ldots \oplus y_N$, where by $\oplus$ we denote addition modulo 2), then Bob need only send p, a single bit of information. The basic question is to find the most efficient protocol, i.e., the protocol in which Bob needs to communicate the minimum number of bits. This minimum number represents the *complexity of the communication*. Incidentally, the absolute minimum is 1 bit of communication for all cases except when the function is independent of $\mathbf{y}$. Indeed, since Alice doesn't have all the information to start with, she needs to learn at least 1 bit. Furthermore, since Alice is interested only in finding out a single bit – the value of f – any communication of more than 1 bit is redundant.[1] Finding the best protocol for an arbitrary function is, in general, a very difficult task, and it is a problem considered in computer science.

In the particular dating problem with which we started our discussion Alice and Bob need to evaluate the so-called *inner product* of $\mathbf{x}$ and $\mathbf{y}$, that is

$$f(\mathbf{x}, \mathbf{y}) = \mathbf{xy} = x_1 y_1 \oplus x_2 y_2 \oplus \ldots \oplus x_N y_N, \tag{1.6}$$

where x_i describes whether Alice's day i is busy ($x_i = 0$) or free ($x_i = 1$) and similarly y_i describes Bob's days. Day i is convenient for both Alice and Bob only if the product $x_i y_i = 1$.

As we mentioned before, evaluating the inner product is a very demanding task – Bob needs to send all his bits. The proof is extremely simple. Among all the possible patterns of free and busy days, a particular case is when Alice is free only on day 1, i.e., $x_1 = 1$, $x_2 = x_3 = \ldots = x_N = 0$. In that case $\mathbf{xy} = y_1$. So if Alice is to know this value she must know y_1. However, Alice might be free only on day 2 ($x_2 = 1$, $x_1 = x_3 = \ldots = x_N = 0$). In that case $\mathbf{xy} = y_2$. So if Alice is to know this value she must know y_2 and so on. But since Alice is not allowed to tell Bob anything, Bob doesn't know whether Alice needs to know y_1 or $y_2 \ldots$ or y_N, so he needs to send them all.

But what if Alice and Bob also share pairs of entangled particles? After all, entangled particles generate correlations that cannot be generated by any classical means, and the proof considered only classical manipulations. At first sight the idea that entanglement might help seems hopeless. Indeed, the first thing one learns about the non-local correlations generated by entangled particles is that they cannot be used for sending information. The reason for this is that these correlations

[1] Here, for simplicity, by bit we mean a binary digit – a "0" or a "1" – not the concept of a bit as a quantity of information, which takes into account the probabilities for the digit to be 0 or 1.

are established instantaneously, immediately when Alice and Bob perform measurements on their particles, and if the correlations could send information this would imply superluminal signaling. However, perhaps they might be useful *in conjunction with classical communication*, that is, in protocols that involve both non-local correlations and classical communication. Waiting for the classical information to arrive makes the whole protocol work at speeds slower than light. In fact Cleve and Buhrman [11] found that quantum entanglement can indeed help in some communication problems. However, it was also shown by Cleve *et al.* [12] that quantum-mechanical non-local correlations cannot help Alice and Bob in their dating task.

That quantum mechanics cannot help is unfortunate indeed, because Alice and Bob's dating problem (the "inner-product" problem) is not just a silly game. In fact every communication problem in which Alice and Bob want to learn just one single bit of information can be reduced to this particular problem. Thus, if we could succeed in reducing the redundancy in communication for the inner-product problem, we would reduce it in all communication problems.

What van Dam has shown is that, if Alice and Bob were to have access to "maximally" super-quantum non-local correlations, i.e., the correlations that reach the upper bound in (1.5), $P_{\text{success}}^{\text{super-quantum}}(a \oplus b = xy) = 1$ (so-called PR correlations or PR boxes), then they could solve their dating problem by using a single bit of communication, hence completely eliminating the redundancy in communication. Furthermore, since, as we mentioned before, any other communication problem in which Alice is interested only in learning 1 bit can be mapped into this dating problem, it implies that the existence of PR correlations would result in completely eliminating the redundancy of all communication.

Van Dam's solution is extremely simple. Recall that a pair of PR boxes are non-signaling input–output devices that obey the rule $xy = a \oplus b$, (1.5). To solve the inner-product problem, Alice and Bob use N PR-box pairs. They input x_1 and y_1 in the first pair, x_2 and y_2 in the second, and so on. Then the inner product is related to the outcomes of their boxes by

$$\mathbf{xy} = x_1 y_1 \oplus \ldots \oplus x_N y_N = a_1 \oplus b_1 \oplus \ldots \oplus a_N \oplus b_N. \tag{1.7}$$

By regrouping the terms in the last equality we obtain

$$\mathbf{xy} = a \oplus b, \tag{1.8}$$

where

$$\begin{aligned} a &= a_1 \oplus \ldots \oplus a_N, \\ b &= b_1 \oplus \ldots \oplus b_N. \end{aligned} \tag{1.9}$$

Hence all Alice needs from Bob is a single bit, the value of b.

The maximal non-local correlation used by van Dam is a very particular correlation and a very extreme case. Brassard *et al.* [13] made the next breakthrough: they suggested that perhaps every PR box pair with a probability of success above the quantum limit of $(2+\sqrt{2})/4 \approx 0.85$ leads to eliminating all the redundancy in communication. Using error-corrections methods they succeeded in showing that all PR boxes with success probability above approximatively 0.91 lead to eliminating redundancy in communication. Therefore, at present there is still a gap, from 0.85 to 0.91, about which we don't know anything.

1.4 Non-local computation

If the status of super-quantum correlations with respect to communication complexity is still unknown, there is another problem, namely non-local computation [14], where the boundary between quantum and super-quantum correlations is sharp.

Consider an ordinary computation problem with N input bits, $z_1, \ldots, z_N$, and a one-bit output, $c = f(z_1, \ldots, z_N)$. To this problem there corresponds a non-local version: the computation is carried out by two devices, one at Alice and one at Bob. Each device has N bits of input and an output of one bit, $x_1, \ldots, x_N$ and a, respectively, for Alice, and $y_1, \ldots, y_N$ and b for Bob. Alice and Bob are given the input bits by some external agents, such that the parity of their inputs equals the original input

$$x_i \oplus y_i = z_i. \tag{1.10}$$

For every possible value of the original input z_i there are two possible combinations of x_i and y_i that obey (1.10): when $z_i = 0$ the two combinations are $x_i = 0$, $y_i = 0$ and $x_i = 1$, $y_i = 1$, while for $z_i = 1$ we can have $x_i = 0$, $y_i = 1$ or $x_i = 1$, $y_i = 0$. The rule of the game is that for each input bit of the original problem, z_i, we give Alice and Bob one of the two corresponding sets of inputs x_i and y_i at random, with equal probability. Consequently, seeing only their own inputs, x_i and y_i, respectively, Alice and Bob have no knowledge about the original input z_i.

The task of Alice and Bob is to output one bit each, a and b, respectively, such that their parity equals the result of the original computation

$$a \oplus b = c = f(z_1, \ldots, z_N). \tag{1.11}$$

Again, for each value of c there are two possible combinations of a and b; we do *not* impose any restriction on which combination should occur. The setup for non-local computation is illustrated in Figure 1.2.

Alice and Bob know in advance the function they have to compute and they are allowed to communicate in advance, set a common strategy, and prepare their

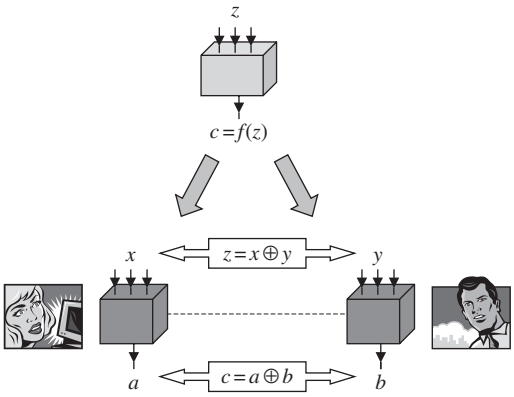

Fig. 1.2 Non-local computation.

devices in whichever way they want. However, they are no longer allowed to communicate once they have been given their inputs. To ensure this, we arrange that the whole procedure performed by Alice, from the moment when she receives her inputs until she delivers her output, is space-like separated from Bob's procedure.

There are three main cases of interest: when Alice and Bob have only classical devices, when they also use entangled quantum particles, and when they have access to super-quantum non-local correlations.

In general Alice and Bob cannot always succeed at outputting the correct answer; their task is to try to do as well as possible. There are various measures of success. A simple scenario is when Alice and Bob are given inputs at random, with equal probability, and they try to obtain the best average success probability.

In effect the whole setting we described above is a Bell-inequality experiment in which Alice can perform one out of 2^N experiments, each with a binary output, and similarly for Bob. Indeed, each of the 2^N possible combinations of, say, Alice's input bits $x_1, \ldots, x_N$ denotes an experiment she performs on her device. The 2^N experiments may all be different from each other, or some of them may be the same. The only difference from a usual Bell-inequality experiment is that Alice and Bob are given their settings by some external party, to avoid Alice and Bob cheating. The upper bound on the average probability of success when the devices are classical, $P_{\mathrm{C}}^{\max}$, is the generalized Bell inequality, whereas the upper bound on the average probability of success when the devices are quantum, $P_{\mathrm{Q}}^{\max}$, is the generalized Cirel'son inequality [15].

The astonishing result is that neither classical physics nor quantum mechanics allows non-local computation, but the very moment we go beyond the quantum-mechanical limit non-local computation becomes possible. An example suffices.