Alma Mater Studiorum · Università di Bologna

Scuola di Scienze
Dipartimento di Fisica e Astronomia
Corso di Laurea in Fisica

Entanglement and Tolman paradox

Relatore:
Prof. Francesco Ravanini

Presentata da:
Gabriel Fenati

Anno Accademico 2021/2022

Abstract

Nella presente tesi è discusso il fenomeno dell'entanglement quantistico, che si manifesta come correlazione nei risultati delle misure di sistemi aperti, in precedenza localmente interagenti, indipendentemente da quella che è la loro distanza spaziale. Il celebre teorema di Bell esclude la possibilità che l'entanglement possa essere descritto mediante una teoria di variabili nascoste locale. Si indaga allora l'ipotesi che particelle che si muovono a velocità super-luminali (tachioni), non vietate dalla Relatività, possano intervenire nella spiegazione fisica dell' entanglement. In particolare, si presenta un' interpretazione alternativa del noto paradosso di Tolman che riguarda i tachioni, tramite la quale è possibile ricostruire la sovrapposizione quantistica di stati della coppia entangled EPR, come associata a loop di tachioni scambiati tra i due sistemi.

Abstract

In this thesis we discuss the phenomenon of quantum entanglement, which consists in the correlation between the results of measurements performed on open systems, previously locally interacting, regardless of their spatial distance. Bell's famous theorem excludes the possibility of a non-local hidden variable theory accounting for entanglement. Therefore, we inspect the hypothesis that faster than light particles (tachyons), not forbidden by Relativity, could have a role in the physical explanation of the entanglement. In particular, we present an alternative interpretation of Tolman paradox regarding tachyons, by which we are able to recover the quantum superposition of states of the EPR entangled pair as linked to loops of tachyons exchanged between the two systems.

Ai miei genitori

Contents

Introduction

≪If two separated bodies, each by itself known maximally, enter a situation in which they influence each other, and separate again, then there occurs [...] entanglement of our knowledge of the two bodies. [...] The knowledge of the two systems cannot be separated into the logical sum of knowledge about two bodies.≫
With these words Erwin Schrödinger, in his famous *cat* paper [1], first introduced the term *entanglement* to describe the peculiar quantum correlations between two systems that had been shown manifestly by Albert Einstein, Boris Podolsky and Nathan Rosen earlier in 1935 [2]. Both Einstein and Schrödinger denied the possibility of a physical connection between measurements performed on separated systems (i.e. far apart in space): ≪external influence on A has no direct influence on B; this is known as the principle of contiguity [i.e. separability]≫, in Einstein's words [3]. According to Einstein, entanglement correlations brought to light the incompleteness of (standard) quantum theory, since they would require a non-local collapse of the wave-function of the entangled system, something which clearly worried the father of Relativity. Hence, he believed in the existence of a local deterministic theory of hidden-variables (because not known) which had to stand behind those correlations. However, in 1964 John S. Bell [4] proved his famous theorem which states that any local realist theory is incompatible with the predictions of quantum mechanics, these latter experimentally well verified. So, Bell showed that non-locality, or perhaps non-separability, lies at the heart of our quantum world. Is this ≪spooky action at a distance≫ [3] in conflict with Special Relativity? No, in fact one cannot exploit entanglement to communicate (i.e. convey controlled information) with another party faster than classically permitted. Nevertheless, entanglement is a powerful resource for many important applications, as first noticed by Richard Feynman [5]: for example it allows secure communication links (quantum cryptography) for the exchange of messages; also, it can be used to transfer an (undisturbed) quantum state, by means of classical bits of information, from a laboratory to another (quantum teleportation); moreover, through the transfer of quantum qubits, it is possible to transmit more bits of information with respect to the classical case (dense coding). These outstanding techniques can only be achieved by (locally) pre-established entanglement between two parties and thanks to the fact that entanglement correlations can outperform classical ones (either local deterministic or stochastic).

But a fundamental question is still unanswered, namely, how two entangled systems, regardless of their distance, can effectively influence each other. Could faster than light particles be involved?
In the second half of the last century several works, for example those of George Sudarshan [6] and of Gerald Feinberg [7], contributed to the building of a theory of faster than light particles, named *tachyons*, whose existence is not prohibited by Einstein's theory. Indeed, the speed of light c is a double limit: for tachyons, a lower one! This fact by itself is a valid reason to study these particles. Many issues arise when one considers the motion of tachyons; for example, there may exist some observers who see the story of the tachyon reversed in time and measure a negative energy for these particles; but these alarming questions can be disarmed thanks to a reinterpretation principle of the absorption-emission of tachyons. A really serious problem appears when one considers the exchange of tachyons between two detectors. In fact, these particles are found to live in a limbo situation of existence/non-existence, since they are trapped in a causal loop. This odd situation is known as the *Tolman paradox* [8] and has been regarded as the main prove against the possible existence of faster than light particles. The causal loops in the exchanging of tachyons between two parties turn out to be a possible way to explain how quantum entanglement correlations could establish: this is the idea presented by Moses Fayngold [9]. In particular, the loop of tachyons which occurs in Tolman paradox, might be connected to the quantum superposition of EPR pair states. Besides giving to entanglement an intriguing physical interpretation, this approach also allows to see quantum measure as a breaking of this loop, through the interaction of a tachyon with a detector. The disruption of the loop translates in the determination of a single story of the tachyon from the previous limbo and the picking of the related states of the two quantum subsystems, which are exactly the eigenstates of the observable wherein they are respectively found by the measurement.

In Chapter 1, after recalling the postulates of quantum mechanics, we focus on the theory of open quantum systems, with particular detail to the study of density matrices and their properties. This leads us to give a formal definition of entangled states of a bipartite system, through the Schmidt rank. We also introduce the qubit.
In Chapter 2, the EPR paradox is described and the realist position compared to the Copenhagen interpretation. Then, we present Bell's theorem with a detailed demonstration and we show how Bell's inequality is violated by any local hidden-variable theory model. We also show another example of Bell's inequalities, in the CHSH framework, and we finally consider a tripartite entangled state: GHZ.
In Chapter 3, after briefly recalling Special Relativity, we will see that it does not, by itself, forbid the existence of faster than light particles.
In Chapter 4, we treat the motion of tachyons and we discuss Tolman paradox.
Finally, in Chapter 5, we present a simple situation in which entangled states appear to arise if we interpret Tolman paradox, not as a paradox, but as the mechanism behind quantum superposition of the EPR pair's states.

Part I

Entanglement

Chapter 1

Quantum Mechanics

Anyone who is not shocked by
quantum theory has not
understood it.

Niels Bohr

1.1 Postulates for closed quantum systems

A closed system is an ideal system which is completely isolated from all the universe and thus cannot communicate with other systems. We shall begin recalling the postulates of quantum mechanics for closed systems. According to the formal structure of quantum theory due to Dirac [10], we have what follows:

- **States** A state of the system (one of its physical modes of being) is a ray in an Hilbert space $\mathcal{H}$. A ray is a normalized ket, $|\psi\rangle$, defined up to a phase factor (i.e. a complex constant of unitary module):

$$e^{i\theta}|\psi\rangle\,, \quad \theta \in (0, 2\pi) \tag{1.1}$$

  represents the same state as $|\psi\rangle$.

- **Observables** An observable of the system (one of its quantifiable properties) is a self-adjoint (or Hermitian[1]) operator acting on $\mathcal{H}$. This means that observable A, 1) is real:

[1]Since we will only deal with finite dimensional Hilbert spaces, in which self-adjointness and hermiticity are equivalent, we will use them as synonyms.

$$A^\dagger = A \iff \langle\psi|A|\phi\rangle = \langle\psi|A|\phi\rangle^* \tag{1.2}$$

$\forall\ |\psi\rangle$ and $|\phi\rangle$; 2) it is diagonalized by an orthonormal basis of its eigenkets (this fundamental result of algebra is known as the *spectral theorem*):

$$A = \sum_j a_j |j\rangle\langle j| \quad , \quad A|j\rangle = a_j|j\rangle \tag{1.3}$$

Note that, since $\langle i|j\rangle = \delta_{ij}$, $|j\rangle\langle j|$ is the orthogonal projector along the span of $|j\rangle$ and a_j is real. We may call $P_j := |j\rangle\langle j|$ and briefly recall the properties of orthogonal projection:

1. $P_i P_j = \delta_{ij} P_i$
2. $P_i^\dagger = P_i = P_i^2$
3. $\sum_j P_j = I$ (I is the identity operator)

the latter is called *completeness relation*.

- **Measurements** The outcome of a measurement of the observable A performed on a system in an arbitrary state $|\psi\rangle$ is probabilistic. The probability distribution is given by Born rule:

$$p(a_j) = |\langle j|\psi\rangle|^2 = \langle\psi|P_j|\psi\rangle \tag{1.4}$$

and the expectation value is therefore:

$$\langle A\rangle = \sum_j a_j p(a_j) = \langle\psi|A|\psi\rangle \tag{1.5}$$

What does this mean? According to the Copenhagen (or orthodox) interpretation, unless the system is initially in an eigenstate of A, the value of the observable is undefined. When a measurement is performed, the "reduction of the state" (also called "collapse of wave function") occurs: the system switches from its initial state to one of the eigenstates of A, say $|j\rangle$, with probability $p(a_j)$. Thus, the post-measurement state, $|\psi'\rangle$, is defined (i.e. a successive measurement of A results, with certainty, in that same state $|\psi'\rangle$):

$$|\psi'\rangle = \frac{P_j|\psi\rangle}{\sqrt{p(a_j)}} \tag{1.6}$$

Another aspect of quantum measurement regards the simultaneous measure of different observables. To be possible with arbitrary precision, common eigenstates of the set of observables are needed, since a simultaneous measure would reduce the state of the system into one of them; but these do not necessarily exist, so that in general interference between measurements occurs, i.e. uncertainty relations subsist among non-commuting observables ($[A, B] = AB - BA \neq 0$). This makes the quote at the beginning of this chapter rather proper.

- **Time evolution** The time evolution of a closed system is governed by Schrödinger equation:

$$i\hbar \frac{d\,|\psi(t)\rangle}{dt} = H\,|\psi(t)\rangle \tag{1.7}$$

where $\hbar$ is the reduced Planck costant and H is the Hamiltonian operator ($H^\dagger = H$). The evolution is unitary (i.e. rotation in the Hilbert space):

$$|\psi(t+dt)\rangle = (I - \frac{i}{\hbar} H dt)\,|\psi(t)\rangle \simeq e^{-\frac{i}{\hbar} H dt}\,|\psi(t)\rangle \equiv U(t+dt, t)\,|\psi(t)\rangle \tag{1.8}$$

where $U(t', t)$ is the unitary ($U^\dagger = U^{-1}$) evolution operator from time t to t'.

Note that Schrödinger equation is linear and deterministically evolves the state of the system. The reduction of the state, which occurs when a measure is performed, *cannot* be described by Schrödinger equation. So, there is an odd role regarding the observation of a system, which consists in the non-unitary collapse of the wave function and the associated becoming-defined of observables: this is the so called *measurement problem*[2].

- **Composite system** When a closed system is a partition of n subsystems, then its Hilbert space is the Cartesian product of the n Hilbert spaces:

$$\mathcal{H} = \mathcal{H}_1 \otimes ... \otimes \mathcal{H}_n \tag{1.9}$$

Therefore: $dim(\mathcal{H}) = dim(\mathcal{H}_1) \cdot ... \cdot dim(\mathcal{H}_n)$. A basis of $\mathcal{H}$ is given by the outer product of the basis vectors of each Hilbert space:

$$|i_1\rangle \otimes ... \otimes |i_n\rangle \equiv |i_1 ... i_n\rangle \tag{1.10}$$

[2]See Appendix A.

1.2 Density Matrices

1.2.1 Mixed states

Let's suppose we do not know in which state, among an ensemble of pure states $|\psi_i\rangle$ (rays non necessarily orthogonal), our system lies in. However, we may assign (classical!) probabilities p_i, which sum to one, to each state and compute the expectation value of a generic observable A as follows:

$$\langle A \rangle = \sum_i p_i \langle \psi_i | A | \psi_i \rangle \tag{1.11}$$

where we are weighting the mean values of the observable over each possible state: both classical and quantum probabilities intervene. We can define the *density matrix* associated to our system in this way:

$$\rho = \sum_i p_i |\psi_i\rangle \langle \psi_i| \tag{1.12}$$

When more than one p_i is different from zero, the state of the system is said to be *mixed*. Eq.1.11 now becomes:

$$\langle A \rangle = Tr(\rho A) \tag{1.13}$$

(recall that $Tr(\sum()) = \sum Tr()$ and the cyclic property of trace).

Density matrices are fundamental to describe open quantum systems. Suppose two systems, A and B, are in contact with each other (i.e. they can exchange energy and/or (quantum) information); therefore they are open subsystems of a bipartite closed system, whose Hilbert space is $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$. A generic pure state of the composite system is of this form:

$$|\Psi\rangle = \sum_{i,\mu} c_{i\mu} |i\rangle \otimes |\mu\rangle \quad , \quad \sum_{i,\mu} |c_{i\mu}|^2 = 1 \tag{1.14}$$

where the $|i\rangle$ and the $|\mu\rangle$ form an orthonormal basis of, respectively, $\mathcal{H}_A$ and $\mathcal{H}_B$. Suppose Alice, who possesses subsystem A, wants to compute the expectation value of the measurement of an (arbitrary) observable T performed on her subsystem; since $|\Psi\rangle$ is a pure state, we may use Eq.1.5 with observable $T \otimes I$ (T acts on A, I, the identity, acts on B):

$$\langle T \rangle_A = \langle T \otimes I \rangle_{AB} = \langle \Psi | T \otimes I | \Psi \rangle = \sum_{i,j,\mu} c^*_{j\mu} c_{i\mu} \langle j | T | i \rangle = Tr(\rho_A T) \tag{1.15}$$

where we have defined, observing that $\rho_{AB} = |\Psi\rangle \langle \Psi|$, the density matrix for subsystem A as

$$\rho_A = \sum_{i,j,\mu} c^*_{j\mu} c_{i\mu} |i\rangle \langle j| = Tr_{\mathcal{H}_B}(\rho_{AB}) \tag{1.16}$$

The density matrix characterizes a quantum open system since it encodes the probability distribution for the outcomes of any observable of that system. As we shall see, it represents indeed the state of an open system. Note that even if the composite system lies in a pure state, subsystems may lie in mixed states. If this is the case, the overall system AB is said to be *entangled*; otherwise it is called *separable*.

1.2.2 Properties

Density matrices have the following interesting properties:

- **Self-adjointness**: $\rho^\dagger = \rho$
- **Non-negativity**: $\langle\phi|\rho|\phi\rangle \geq 0 \quad , \quad \forall \phi \in \mathcal{H}$
- **Unit-trace**: $Tr(\rho) = 1$

In general, every operator who behaves in this way is a density operator. Notice that, from these properties, it is justified to interpret ρ as a probability distribution; in fact, if we consider its eigenvectors $|a\rangle$:

$$\rho |a\rangle = p_a |a\rangle \tag{1.17}$$

it follows that the eigenvalues p_a are all real, $0 \leq p_a$ and $\sum_a p_a = 1$. Indeed, in the orthonormal basis of its eigenkets, the density operator can be expanded as (particular form of Eq.1.12, where we are making use of spectral theorem, Eq.1.3):

$$\rho = \sum_a p_a |a\rangle \langle a| \tag{1.18}$$

Thus, if only one p_a differs from zero, the state is pure ($\rho^2 = \rho$); otherwise it is mixed. Note that, if the density matrix is associated as before to subsystem A, Alice cannot distinguish whether the probability distribution originates from her lack of knowledge about A (i.e. ensemble of possible pure states of A), or because of A is an open system (i.e. is part of a joint entangled system AB).

1.2.3 Schmidt Decomposition

There is a preferred basis for $\mathcal{H}_A \otimes \mathcal{H}_B$ in which to describe a bipartite pure state: the Schmidt basis. Let's consider again a generic pure state for the joint system AB, but this time, since we have learned the properties of density matrices, we choose for system A the orthonormal basis $|a\rangle$ in which ρ_A is diagonalized (Eq.1.18):

$$|\Psi\rangle = \sum_{a,\mu} c_{a\mu} |a\rangle \otimes |\mu\rangle \tag{1.19}$$

where, again, the $|\mu\rangle$ form a basis for $\mathcal{H}_B$ and $\sum_{a,\mu} |c_{a\mu}|^2 = 1$. We define the vectors (a priori neither necessarily normalized nor orthogonal):

$$|\beta_a\rangle \equiv \sum_{\mu} c_{a\mu} |\mu\rangle \tag{1.20}$$

Then, recalling Eq.1.16, we obtain:

$$\rho_A = \sum_{aa'} |a\rangle \langle a'| \, Tr_{\mathcal{H}_B}(|\beta_a\rangle \langle \beta_{a'}|) = \sum_{aa'} \langle \beta_{a'}|\beta_a\rangle |a\rangle \langle a'| \tag{1.21}$$

and, since it must equal Eq.1.18, therefore the $|\beta_a\rangle$ are indeed orthogonal:

$$\langle \beta_{a'}|\beta_a\rangle = p_a \delta_{aa'} \tag{1.22}$$

Thus, we can define the following orthonormal basis for subsystem B:

$$|b_a\rangle = \frac{1}{\sqrt{p_a}} \beta_a \tag{1.23}$$

Note that there is no issue here about dividing by zero: vectors with $p_a = 0$ just will not appear in Eq.1.21.
We have found the *Schmidt form* for $|\Psi\rangle$:

$$|\Psi\rangle = \sum_{a} \sqrt{p_a} \, |a\rangle \otimes |b_a\rangle \tag{1.24}$$

where the $|b_a\rangle$ are the orthonormal eigenkets that diagonalize the density matrix of system B since:

$$\rho_B = Tr_{\mathcal{H}_A}(|\Psi\rangle \langle \Psi|) = \sum_{a} p_a |b_a\rangle \langle b_a| \equiv \sum_{b} p_b |b\rangle \langle b| \tag{1.25}$$

where in the last passage we have changed the index of summation and we have left the subscript in the kets in order to recover Eq.1.18 for ρ_B. It should be pointed out that, whereas it exists a Schmidt decomposition for each given pure composite state, the basis $|a\rangle \otimes |b_a\rangle$ does depend on the specific state.
We observe an interesting thing: although $\mathcal{H}_A$ and $\mathcal{H}_B$ might not have the same dimension, ρ_A and ρ_B have indeed the same non-zero eigenvalues. The number of these common probabilities is called the *Schmidt rank* ($=: R_S$). It allows us to make a fundamental distinction among bipartite pure states of system AB:

- $\boldsymbol{R_S = 1}$: ***separable*** (or product) ***state***.
The most general separable pure state of AB can be written as an outer product of two pure states, $|\phi\rangle$ for A and $|\chi\rangle$ for B:

$$|\Psi\rangle = |\phi\rangle \otimes |\chi\rangle \tag{1.26}$$

that is: the Schmidt decomposition of the state has a single term, with unitary probability; the density matrices of the subsystems are respectively $\rho_A = |\phi\rangle\langle\phi|$ and $\rho_B = |\chi\rangle\langle\chi|$; the density matrix of AB is $\rho_{AB} = |\phi\rangle\langle\phi| \otimes |\chi\rangle\langle\chi|$.
If the joint state AB is separable, the correlations between A and B are entirely classical. Suppose that A and B are far away from each other; if $|\phi\rangle$ is the state of A and $|\chi\rangle$ the state of B, we can always consider the composite system AB, whose state will be evidently $|\phi\rangle \otimes |\chi\rangle$.[3]

- $\boldsymbol{R_S > 1}$: ***entangled state***
The generic Schmidt form of Eq.1.24, with more than one $p_a \neq 0$, represents an entangled state of AB. Correlations among systems through entanglement, as we shall widely see, are profoundly different from the classical ones, in many ways. First of all, while classically we can always interpret probabilities as emerging from ignorance, we cannot do the same for quantum (entanglement) probabilities, they are in some way intrinsic. Further, whereas a separable state can always be created locally, i.e. Alice and Bob act on their respective system, perhaps far from each other, it is impossible to create entanglement in such way:

$$U_A \otimes U_B \tag{1.27}$$

(U_A and U_B are unitary transformation acting respectively on A and B) cannot increase the Schmidt rank of the system. The only way we can create entanglement is by bringing the two systems together and make them interact opportunely.
We finally write how a maximally entangled state of a N dimensional system looks like:

$$|\Psi\rangle = \frac{1}{\sqrt{N}} \sum_{a=0}^{N-1} |a\rangle \otimes |b_a\rangle \tag{1.28}$$

[3] We may also have separable bipartite mixed states: they can be expressed as $\rho_{AB} = \sum_i p_i (|\phi_i\rangle\langle\phi_i|) \otimes (|\chi_i\rangle\langle\chi_i|)$ and represent an ensemble of pure bipartite states, each weighted with its probability. Again, there is no odd correlation between A and B, in the sense that we can interpret the p_i just as ignorance on the actual state of AB.

1.2.4 Open quantum systems

We can now make some generalisations for open systems of concepts seen in Sec.1.1. Firstly, the state of an open quantum system is a density matrix ρ. Further, Born rule (Eq.1.4) may now be written by using Eq.1.13 (recall that P_j is an orthogonal measurement operator):

$$p_j = Tr(\rho P_j) \tag{1.29}$$

The post-measurement state (Eq.1.6) will therefore be:

$$\rho' = \frac{P_j \rho P_j}{p_j} \tag{1.30}$$

note that this is indeed a density matrix.
If we perform a measurement but we do not record the outcome (i.e. it is inaccessible) then the correct description of the post-measurement state is given by weighting all the possible outcomes with the respective probabilities:

$$\rho' = \sum_i p_i \rho'_i = \sum_i P_i \rho P_i \tag{1.31}$$

The previous equation is an example of a quantum channel $\rho' = \epsilon(\rho)$, which is a linear map of density operators to density operators. It describes the evolution of an open quantum system. As it can be easily seen, it is *not* unitary, in general, since it maps initial pure states into mixed ones. This peculiar evolution of open systems is called *decoherence*, and it's obviously an irreversible process. How can we interpret it? Consider a system, initially for example in a pure state, which interacts with its environment, this latter being inaccessible to the observer. Hence, this interaction causes leaks of information (noise) regarding the state of the system, and because it is irrecoverable, we are obliged to describe it as an ensemble of possible pure states.[4]

1.3 Qubit and Bloch Ball

1.3.1 Pauli matrices and spin

In this last section of the first chapter we introduce a system which will accompany us throughout the following sections: the qubit. What is a qubit? Basically, it is a 2-level

[4]Sometimes orthogonal measurements are too restrictive; in fact generalized measurements might be performed on a system, for example this is more realistic for the description of the interaction between the system and its environment. These generalized measurements are a set of operators $\{E_m\}$ with the properties of hermiticity, non-negativity and completeness. Since they are not orthogonal, after a first measurement is done, a second measure will not result in the same outcome, in general.

quantum system, like the familiar spin-$\frac{1}{2}$ particle. Thus, its pure states are rays $|\psi\rangle$ in the Hilbert space $\mathcal{H} = \mathbb{C}^2$. The operator algebra is 4-dimensional and it is generated by this basis:

$$I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad \sigma_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \quad \sigma_2 = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \quad \sigma_3 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \tag{1.32}$$

where $\sigma = (\sigma_1, \sigma_2, \sigma_3)$ is the Pauli (vector) operator and the Pauli matrices have the following properties:

1. self-adjoint: $\sigma_j^\dagger = \sigma_j$
2. $\sigma_j^2 = I \qquad$ (with 1. $\implies$ unitary: $\sigma_j^\dagger = \sigma_j^{-1}$)
3. traceless: $\mathrm{Tr}(\sigma_j) = 0$
4. $\sigma_i\sigma_j + \sigma_j\sigma_i = 2\delta_{ij}I$
5. $[\sigma_i, \sigma_j] = 2i\epsilon_{ijk}\sigma_k$

where in the commutation relation ϵ_{ijk} is the Levi-Civita symbol and we make use of Einstein's convention (sum over repeated indices). From properties 1. and 2. follows that each Pauli matrix admits (by Eq.1.3) an orthonormal basis of its eigenstates, with eigenvalues respectively +1 and -1.
Moreover, properties 4. and 5. are equivalent to the following:

$$\sigma_i\sigma_j = \delta_{ij}I + i\sum_k \epsilon_{ijk}\sigma_k \tag{1.33}$$

Another useful relation is the Pauli exponential formula, which expands a generic unitary transformation U as ($\hat{n}$ is an arbitrary unit vector of $\mathbb{R}^3$ and α is an angle):

$$U(\hat{n}, \alpha) = \exp(i\alpha\hat{n} \cdot \sigma) = \cos(\alpha)I + i\sin(\alpha)\hat{n} \cdot \sigma \tag{1.34}$$

A generic Hermitian operator A can be written as ($a_0 \in \mathbb{R}$, $\vec{a} \in \mathbb{R}^3$):

$$A = a_0 I + \vec{a} \cdot \sigma \tag{1.35}$$

How can we write the generic density operator of a qubit? A 2x2 matrix with properties of Sec.1.2.2 can be decomposed in the form of Eq.1.35 as:

$$\rho(\vec{a}) = \frac{1}{2}(I + \vec{a} \cdot \sigma) = \frac{1}{2}\begin{bmatrix} 1 + a_3 & a_1 - ia_2 \\ a_1 + ia_2 & 1 - a_3 \end{bmatrix} \tag{1.36}$$

where $\vec{a}$ is known as the *polarisation* of the qubit. For the non-negativity we must require:

$$\det \rho(\vec{a}) = \frac{1}{4}(1 - |\vec{a}|^2) \geq 0 \implies |\vec{a}| \leq 1 \tag{1.37}$$

so that one's has ($\text{Tr}\rho = 1$) two real non-negative eigenvalues. Therefore, the density matrices of a qubit (i.e. a spin-$\frac{1}{2}$ particle) can be represented as points of a 3D unit ball: the Bloch Ball (Fig.1.1).

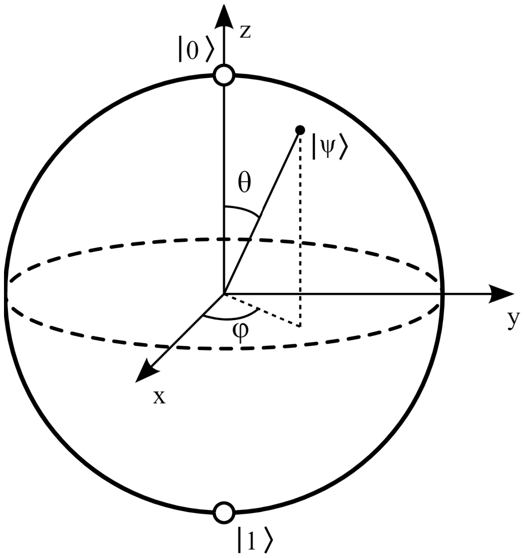

Figure 1.1: *The Bloch Ball.*

Pure states of a qubit are on the boundary (the sphere): $|\vec{a}| = 1$, so $\vec{a} \equiv \hat{a}$. The associated density matrix has one eigenvalue equal to 0 and the other is 1. If we introduce the spherical angles $\theta \in [0, \pi]$, $\phi \in [0, 2\pi]$, we can write $\hat{a} = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta)$ and the density matrix becomes:

$$\rho(\hat{a}) = \frac{1}{2}(I + \hat{a}\cdot\sigma) = \frac{1}{2}\begin{bmatrix} 1+\cos\theta & \sin\theta e^{-i\phi} \\ \sin\theta e^{-i\phi} & 1-\cos\theta \end{bmatrix} = |\psi(\theta,\phi)\rangle\langle\psi(\theta,\phi)| \tag{1.38}$$

where

$$|\psi(\theta,\phi)\rangle = \begin{pmatrix} e^{-i\frac{\phi}{2}}\cos\left(\frac{\theta}{2}\right) \\ e^{+i\frac{\phi}{2}}\sin\left(\frac{\theta}{2}\right) \end{pmatrix} \tag{1.39}$$

allows us, given the angles θ,ϕ and so a point on the Bloch sphere, to individuate the state of the qubit in $\mathcal{H}$; note however that if we consider a rotation of $\phi = 2\pi$, we get $|\psi(\theta,\phi+2\pi)\rangle = -|\psi(\theta,\phi)\rangle$: the vectors $|\psi(\theta,\phi)\rangle$ are spinors, since we need a rotation of 4π on the sphere to re-obtain the same vector; however, this does not make a problem, since we know from Eq.1.1 that states are rays, and here we just have a variation of the overall phase. Note also that mutual orthogonal states are antipodal: $\langle\psi(\theta',\phi')|\psi(\theta,\phi)\rangle = 0 \implies \theta' = \pi - \theta \ , \ \phi' = \phi + \pi$. Moreover, the $|\psi(\theta,\phi)\rangle$ is an eigenstate of $\hat{a}\cdot\sigma$ with eigenvalue $+1$:

$$\hat{a}\cdot\sigma\,|\psi(\theta,\phi)\rangle = \begin{bmatrix} \cos\theta & e^{-i\phi}\sin\theta \\ e^{i\phi}\sin\theta & -\cos\theta \end{bmatrix} |\psi(\theta,\phi)\rangle = |\psi(\theta,\phi)\rangle \tag{1.40}$$

and so can be indeed seen as a spin pointing along $\hat{a}$. An arbitrary unitary transformation $U(\hat{n},\alpha)$ (Eq.1.34) acting on the system, rotates its spin of an angle α with respect to axis $\hat{n}$ and results in the eigenstate of eigenvalue $+1$ of the operator:

$$\hat{b}\cdot\sigma = e^{i\frac{\alpha}{2}\hat{n}\cdot\sigma}(\hat{a}\cdot\sigma)e^{-i\frac{\alpha}{2}\hat{n}\cdot\sigma} \tag{1.41}$$

We now define the standard basis of $\mathbb{C}^2$ as

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix} \equiv |0\rangle \equiv |{\uparrow_z}\rangle \quad , \quad \begin{pmatrix} 0 \\ 1 \end{pmatrix} \equiv |1\rangle \equiv |{\downarrow_z}\rangle \tag{1.42}$$

where we have identified $\hat{z}$ and $-\hat{z}$ as the directions of the corresponding spins. These kets are indeed eigenstates of σ_3, with eigenvalues respectively $+1$ and -1. We can derive Eq.1.39 by applying at $|0\rangle$ a rotation of angle $-\theta$ around axis $\hat{n} = (-\sin\phi, \cos\phi, 0)$:

$$e^{-i\frac{\theta}{2}\hat{n}\cdot\sigma}\,|0\rangle = e^{\frac{\theta}{2}\begin{bmatrix} 0 & -e^{-i\phi} \\ e^{i\phi} & 0 \end{bmatrix}} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{bmatrix} \cos\frac{\theta}{2} & -e^{-i\phi}\sin\frac{\theta}{2} \\ e^{i\phi}\sin\frac{\theta}{2} & \cos\frac{\theta}{2} \end{bmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \tag{1.43}$$

Often we will also use these equivalent notations:

$$|\psi\rangle = e^{-i\frac{\phi}{2}}\cos\frac{\theta}{2}\,|0\rangle + e^{i\frac{\phi}{2}}\sin\frac{\theta}{2}\,|1\rangle \equiv a\,|0\rangle + b\,|1\rangle \tag{1.44}$$

with $a, b \in \mathbb{C}$ and $|a|^2 + |b|^2 = 1$.
Let's analyze the action of the Pauli matrices on a generic state:

- $\sigma_1\,|0\rangle = |1\rangle \quad , \quad \sigma_1\,|1\rangle = |0\rangle$ $\qquad \sigma_1$ is a bit-flip operator
- $\sigma_3\,|0\rangle = |0\rangle \quad , \quad \sigma_3\,|1\rangle = -\,|1\rangle$ $\qquad \sigma_3$ is a phase-flip operator
- $\sigma_2\,|0\rangle = i\,|1\rangle \quad , \quad \sigma_2\,|1\rangle = -i\,|0\rangle$ $\qquad \sigma_2$ is both a bit&phase-flip operator

The eigenstates of σ_1 are:

$$|{\uparrow_x}\rangle \equiv |+\rangle \equiv \frac{|0\rangle + |1\rangle}{\sqrt{2}} \quad , \quad |{\downarrow_x}\rangle \equiv |-\rangle \equiv \frac{|0\rangle - |1\rangle}{\sqrt{2}} \tag{1.45}$$

In general, we will call $|\uparrow_{\hat{n}}\rangle$ and $|\downarrow_{\hat{n}}\rangle$ respectively the eigenstates of $\hat{n}\cdot\sigma$ with eigenvalues ± 1 ($(\hat{n}\cdot\sigma)^2 = I$).

Back to Eq.1.37, when $|\vec{a}| < 1$, in the interior of the ball, we have mixed states: both the eigenvalues of ρ are in $(0,1)$. We may compute:

$$\langle \hat{n}\cdot\sigma\rangle = Tr(\rho(\vec{a})(\hat{n}\cdot\sigma)) = \frac{1}{2}\sum_{ij} Tr(n_i\sigma_i(I + a_j\sigma_j)) = \hat{n}\cdot\vec{a} \tag{1.46}$$

this shows us the physical significance of the polarisation: it individuates the direction in which the expectation value relative to the measure of the spin is maximized. When $|\vec{a}| = 0$ the state is unpolarised and from Eq.1.36 we see that the density matrix takes the form:

$$\rho = \frac{1}{2}I \tag{1.47}$$

and, by Eq.1.46, we conclude that a measure of the spin, along an arbitrary axis, results ± 1 with the same probability. This state is called the *maximally mixed* state.

1.3.2 Convex set of density operators

In general, density operators live in a convex (finite dimensional) set, i.e. a generic density operator $\rho(\lambda)$ can be expressed as a convex combination:

$$\rho(\lambda) = \lambda\rho_1 + (1-\lambda)\rho_2 \qquad 0 \le \lambda \le 1 \tag{1.48}$$

where ρ_1 and ρ_2 are also density operators of the set. This can be geometrically view in Fig.1.2, where every point along the line between ρ_1 and ρ_2 is a possible density operator of the system. Recalling the ensemble interpretation (Sec.1.2.1), this means that every density operator can be prepared by suitable sampling from other probability distributions.

Mixed states (ρ has more than one eigenvalue different from zero) can thus be prepared in infinite different ways; the same density matrix can for example describe the probability distribution associated to two different ensemble of (non necessarily orthogonal) pure states, $\{|\phi_i\rangle\}$ and $\{|\chi_\mu\rangle\}$:

$$\rho = \sum_i p_i |\phi_i\rangle\langle\phi_i| = \sum_\mu p_\mu |\chi_\mu\rangle\langle\chi_\mu| \tag{1.49}$$

Instead, pure states are extremal: they can be prepared in unique ways, in the sense that they cannot be obtained as a mixture of other states.
Back to the qubit, we see that the convex set of 2x2 density operators is really the Bloch Ball. We also observe that every mixed state can here be expressed in a unique

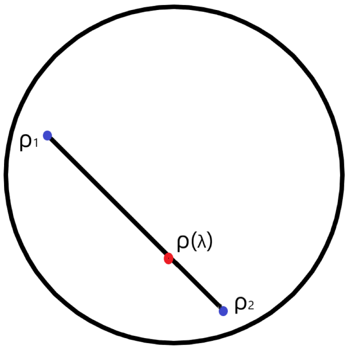

Figure 1.2: *The convex set of density matrices.*

way by two mutually orthogonal pure states, represented by antipodal points, since (for every point different from the centre) there is just one diameter which passes through the point of a mixed state (Fig.1.3-left). This is due to the non-degeneracy of eigenvalues. An obvious exception is the centre of the ball ($p_1 = p_2 = \frac{1}{2}$), which is the maximally mixed state of Eq.1.47: since through the centre of a sphere an infinite number of diameters passes, we can express this state by infinitely many combinations of two orthogonal pure states (Fig.1.3-right):

$$\rho = \frac{1}{2}I = \frac{1}{2}(|\uparrow_{\hat{n}}\rangle\langle\uparrow_{\hat{n}}| + |\downarrow_{\hat{n}}\rangle\langle\downarrow_{\hat{n}}|) \tag{1.50}$$

for every $\hat{n}$. In this sense, this is also called the state of maximum ignorance. We will make use of this state in abundance in the next sections.

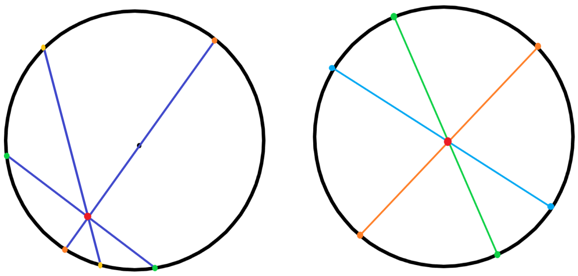

Figure 1.3: *Left: A generic mixed state of a qubit. Right: The maximally mixed state.*

1.3.3 Separable and entangled states of qubits

We conclude this chapter with an overview of the possible composite states of two qubits. Consider Alice and Bob who respectively have got qubits A and B. By recalling Sec.1.2.3, we may write an example of a separable state of the two qubits as:

$$|\uparrow_{\hat{n}}\rangle \otimes |\uparrow_{\hat{m}}\rangle \equiv |\uparrow_{\hat{n}}\rangle\,|\uparrow_{\hat{m}}\rangle \equiv |\uparrow_{\hat{n}}\uparrow_{\hat{m}}\rangle \tag{1.51}$$

where we have indicated equivalent notations which we shall frequently use.
One important example of entangled state is (note Schmidt rank $R_S = 2$):

$$|\Psi\rangle = \frac{1}{\sqrt{2}}(|\uparrow_z\rangle\,|\uparrow_z\rangle + |\downarrow_z\rangle\,|\downarrow_z\rangle) \tag{1.52}$$

This is the so called maximally entangled pure state of AB. If we compute ρ_A and ρ_B by Eq.1.16 we get:

$$\rho_A = Tr_{\mathcal{H}_B}(\rho_{AB}) = \frac{1}{2}I = Tr_{\mathcal{H}_A}(\rho_{AB}) = \rho_B \tag{1.53}$$

so both density matrices are of the type of Eq.1.47. This implies that the Schmidt decomposition is ambiguous, i.e. we can obtain the same bipartite state in infinite ways:

$$|\Psi\rangle = \frac{1}{\sqrt{2}}(|\uparrow_{\hat{n}}\rangle\,|\uparrow_{\hat{n}}\rangle + |\downarrow_{\hat{n}}\rangle\,|\downarrow_{\hat{n}}\rangle) \tag{1.54}$$

$\forall \hat{n}$. Suppose Alice performs a measure of σ_3 on her qubit A: she will obtain with probability $\frac{1}{2}$ the value $+1$ ($|\uparrow_z\rangle$) and with same probability the outcome -1 ($|\downarrow_z\rangle$). If Bob then measures the same observable on his qubit he will find, with certainty, the same outcome Alice obtained. This occurs independently of the distance between A and B, so there is no classical communication between the two qubits that occurs. Further, the situation is perfectly symmetric, as Bob could be the first to perform the measure and in fact he is the first one for certain frame of references, if the two measurement events are space-like. Thus, if A and B are entangled, they are correlated in a very peculiar and subtle way. In fact, there is no way that one party can communicate to the other through entanglement: first of all, because of Eq.1.54, independently from which measurement Alice does perform, Bob may measure along any axis $\hat{n}$; they find the same outcomes only if they measure the same observable (i.e. along the same axis, for Eq.1.35). In addition, supposing they agree in advance on the measure to perform, they cannot convey controlled information one to the other, since the outcomes of the measurement are probabilistic. So, Alice and Bob are obliged to compare the results of their measurements and this evidently occurs through classical communication.
Besides, note that Alice and Bob cannot exploit entanglement in order to violate the commutation relations, i.e. to simultaneously measure incompatible observables (Sec.1.1). In fact, suppose again Alice measures σ_3 on her qubit A and finds $|\uparrow_z\rangle$; as we have seen:

$$|\uparrow_z\rangle_A \implies |\uparrow_z\rangle_B \qquad , \qquad then \qquad |\Psi\rangle = |\uparrow_z\uparrow_z\rangle \tag{1.55}$$

She might think: "If now Bob measures σ_1 on his qubit, since our qubits are entangled, I would also know the spin along x of my qubit and I also (already) know how is the spin along z!" However, Bob's measurement erases the information about qubit A encoded in B. In fact, if the outcome is $|\uparrow_x\rangle_B$, then:

$$|\uparrow_x\rangle_B \implies |\uparrow_x\rangle_A = \frac{1}{\sqrt{2}}(|\uparrow_z\rangle_A + |\downarrow_z\rangle_A) \tag{1.56}$$

and so his measurement causes the transition of A's state into a coherent superposition of $|\uparrow_z\rangle, |\downarrow_z\rangle$ and so the spin along z becomes indefinite. This process is also called *quantum erasure*. This is consistent with Prop.5. of Pauli matrices, i.e. none party can measure simultaneously (even with the described strategy that makes use of the entanglement) non-commutative observables.
We will deepen in detail these ideas in what follows.

Chapter 2

EPR paradox and Bell's Theorem

> I would not call [entanglement] one but rather the characteristic trait of quantum mechanics, the one that enforces its entire departure from classical lines of thought.
>
> Erwin Schrödinger

2.1 EPR paradox

2.1.1 Copenhagen interpretation and local realism

One may arise several issues regarding how to interpret entanglement in the framework of quantum mechanics. We have already seen in Sec.1.3.3 that a subtle correlation subsists among entangled systems, which has no classical counterpart. These correlations appear to be *instantaneous* and *non-local*, although, as we previously discussed, they cannot be exploited to communicate (super-luminary) from one party to the other. But an irrepressible question arises: what is really going on?

According to the Copenhagen interpretation (recall Sec.1.1), when a measurement of the spin is performed on a qubit, the reduction of the state of the composite system (instantaneously) comes about. For example, consider again the state of Eq.1.52, and suppose a party measures the spin along z:

$$\frac{1}{\sqrt{2}}(|\uparrow_z\rangle\,|\uparrow_z\rangle + |\downarrow_z\rangle\,|\downarrow_z\rangle) \quad \Longrightarrow \quad |\uparrow_z\uparrow_z\rangle \tag{2.1}$$

consequently, the other party will find the same result by measuring along z. In other words, the orthodox position attributes the entanglement correlations to the collapse of the common wave-function of the two systems caused by the measure. So, this interpretation involves such an "ethereal" influence and also a rather peculiar role of the measurement process, since it determines the emergence of the definite value of the observable from indeterminacy. This position is historically due to two giants of quantum physics: Niels Bohr and Werner Heisenberg.
Speaking of giants, Albert Einstein, together with Boris Podolsky e Nathan Rosen, proposed in 1935 the famous EPR paradox [2] in order to refute the Copenhagen interpretation. Einstein's position is known as *local realism*.
The principle of locality is the assumption that causal influence cannot propagate faster than light. So, the measurement of the spin of one qubit cannot determine the outcome of the measure on the other, for example through a superluminal signal [1].
The realist viewpoint believes that the observables of a system are always defined, also non-commuting ones, although not being simultaneously measurable. Therefore, the act of measurement has nothing special and is deterministic, since it simply acknowledges the value of a prior-defined observable. Einstein's picture embraces also the concept of *separability*, which in his own words means ≪the real factual situation of the system S_2 is independent of what is done with the system S_1, which is spatially separated from the former≫, i.e. each system possesses its own physical state.
Local realism is a quite comforting vision, since it does not conceive any instantaneous non-local collapse (i.e. state reduction), or ≪spooky action at a distance≫ in Einstein's words. However, this implies that quantum mechanics, which gives us only the probabilities of measurements outcomes, is not a complete theory, in the sense that it not fully explains the state of the system. Indeed, it must exist some local hidden-variable theory that accounts for the correct and precise prediction of measure's results. "Hidden" because such hypothetical theory is not known and this (our ignorance) is the reason behind the quantum probabilities and uncertainty: could we attain the complete theory, they would vanish.

2.1.2 Bohm's version of the paradox

We shall describe a simplified version of the EPR paradox due to David Bohm and Yakir Aharonov [11]. Consider the decay of a pion π^0 which produces an electron-positron pair (Fig.2.1):

$$\pi^0 \quad \rightarrow \quad e^- + e^+ \tag{2.2}$$

[1]Thus, the correlations between the spin measurements can only be explained by a common cause, i.e. an event which lies in the intersection of their back light cones (see Sec.3.2).

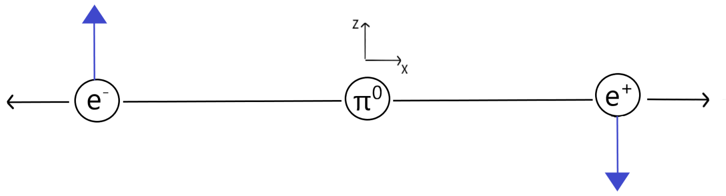

Figure 2.1: *Bohm-EPR experiment. In blue colour we represent the (anti-symmetric) spins along z of the particles of the EPR pair.*

If we are in the frame of reference in which the pion is initially at rest, e^- and e^+ will move in opposite directions with the same velocity, for momentum conservation. Moreover, in this frame the total angular momentum is just the spin angular momentum; thus, since π^0 has zero spin, e^- and e^+ must have opposite spins (± 1 [2]), along any direction. So, the quantum state of the electron-positron pair is the singlet configuration:

$$|EPR\rangle = \frac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle) \tag{2.3}$$

which is also called *EPR pair*, a maximal entangled state. The density matrix of the EPR pair is thus:

$$\rho = \frac{1}{2}(|\uparrow\downarrow\rangle\langle\uparrow\downarrow| + |\downarrow\uparrow\rangle\langle\downarrow\uparrow| - |\uparrow\downarrow\rangle\langle\downarrow\uparrow| - |\downarrow\uparrow\rangle\langle\uparrow\downarrow|) \tag{2.4}$$

whereas the density matrices of both the e^- and the e^+ are (by Eq.1.16):

$$\rho_- = \frac{1}{2}I = \rho_+ \tag{2.5}$$

Comparing Eq.2.4 with Eq.2.5, we see that while the density matrix of the whole system knows about the entanglement correlations, those of each single member of the pair do not: they represent maximum ignorance states. This is indeed in agreement with the ensemble interpretation of Sec.1.2.1. Now suppose that a measurement of the spin of, say, the e^- is performed along, for example, the z axis. As we have already widely commented, this may result, with same probability, in ± 1. But imagine that we don't record the outcome (i.e. it is unknown); the density matrix of the EPR pair then becomes the one of a mixed state:

$$\rho_u = \frac{1}{2}(|\uparrow\downarrow\rangle\langle\uparrow\downarrow| + |\downarrow\uparrow\rangle\langle\downarrow\uparrow|) \tag{2.6}$$

[2] A measure of any component of the spin, for spin-$\frac{1}{2}$ particles like e^- and e^+, can return $+\frac{\hbar}{2}$ or $-\frac{\hbar}{2}$; so just suppose that our detectors record in units of $\frac{\hbar}{2}$.

where the subscript stands for unrecorded. See that it differs from Eq.2.4; instead, the density matrices of e^- and e^+ are still given by Eq.2.5. So, nothing changes in the state of a system when a measurement is done on the other: this perfectly agrees with Einstein locality. But, if a particular outcome of a measure is known (suppose we obtained $|\uparrow\rangle$), then the density matrices do change:

$$\rho' = |\uparrow\downarrow\rangle\langle\uparrow\downarrow| \quad , \quad \rho'_- = |\uparrow\rangle\langle\uparrow| \quad , \quad \rho'_+ = |\downarrow\rangle\langle\downarrow| \tag{2.7}$$

That is, when we acquire the spin of the electron, we immediately know in what state the positron will be found once measured, regardless of how distant are the two. In Einstein's argument, if Copenhagen interpretation is correct, this would require a non-local collapse of the wave-function which signals to the positron how to behave when measured. This represents, in EPR viewpoint, a clear and unacceptable violation of locality: this is the EPR paradox.

In their position (local realism), it exists a local hidden-variable theory, which, giving a complete description of physical reality, prevents that an action performed on one system can modify the description of the other (entangled with the first). In fact, if realism really describes nature, then entanglement is nothing more than a classical-type correlation, where probabilities emerge as ignorance in the quantum incomplete theory. We can understand this through a simple example. Alice and Bob each receive from a third party, say Charlie, a box. These may contain one of two notes: "up" is written on one, "down" on the other. So, both Alice and Bob, will affirm to have one of the two possible notes, each with probability $\frac{1}{2}$, this because of their actual ignorance. When, for example, Alice opens her box and finds "up", then, once Bob opens his, he will find "down" with certainty. Note that this 2-notes-state is indeed described by Eq.2.6, but here no odd quantum correlation occurs! In this sense local realism is a comforting vision. We are now going to see that, although comforting, it is wrong.

2.2 Bell's Theorem

In 1964 John S. Bell discovered that quantum mechanics is incompatible with any local-hidden variable theory [4]. His theorem is perhaps the most important scientific result of the second half of last century. We will now recall his proof, by showing how the predictions of quantum mechanics violate Bell's inequality and we will examine what are the consequences on our comprehension of nature.

2.2.1 Bell's inequality

Bell proof starts exactly where EPR paradox have left us: suppose a local hidden-variable theory indeed does exist. Suppose also to generalize the (Bohm-)EPR experiment of Sec.2.1.2, by allowing to rotate independently the two detectors of the spins, in order to measure the e^- spin in a direction $\hat{a}$ and the e^+ spin along $\hat{b}$, with unit vectors $\hat{a}, \hat{b} \in \mathbb{R}^3$. For a more comfortable notation we will indicate the electron as (system) A and the positron as B. We also call $\sigma^{(A)}$ the spin operator associated with A and $\sigma^{(B)}$ the analogous for system B. The locality assumption implies:

$$[\sigma^{(A)} \cdot \hat{a}\ ,\ \sigma^{(B)} \cdot \hat{b}] = 0 \tag{2.8}$$

That is, in no way the outcome of a measure of spin performed on one system can be affected by the orientation of the other detector or by the outcome of measurements on the other system.
Let us call the hidden variable(s) λ; it can be a single parameter or a collection of functions, discrete or continuous valued: it's nature is indifferent (it is after all unknown), the important thing is that it actually characterizes the complete state of the system AB. One can imagine λ as two sets (one for A and the other for B) of initial values of some dynamical variables governed by hidden laws of motion. Nevertheless, we will treat λ as a single continuous parameter, with probability distribution $\rho(\lambda)$.
Consider many π^0 decays as the one described in Eq.2.2. Bell's idea is to calculate the mean value of the product of the spins of, respectively, A and B for arbitrary orientations of detectors. If $S_{\hat{a}}^{(A)}$ is the value of the spin of A measured along $\hat{a}$ direction and $S_{\hat{b}}^{(B)}$ is the analogous for B, we call this average as $\langle S_{\hat{a}}^{(A)} S_{\hat{b}}^{(B)} \rangle$. Since we have an hidden variable theory, $S_{\hat{a}}^{(A)}$ and $S_{\hat{b}}^{(B)}$ must be some functions of λ, namely:

$$S_{\hat{a}}^{(A)}(\lambda) = \pm 1 \qquad , \qquad S_{\hat{b}}^{(B)}(\lambda) = \pm 1 \tag{2.9}$$

where we have specified the only values they can take. These functions need to have the property to be always anti-correlated when $\hat{a} = \hat{b}$ (recall that maximally entangled states like the EPR one, Eq.2.3, can be realized through orthogonal pure states along any axis $\hat{n}$, as we saw in Eq.1.54):

$$S_{\hat{a}}^{(A)}(\lambda) = -S_{\hat{a}}^{(B)}(\lambda) \tag{2.10}$$

We are now ready for Bell's inequality. The average of the products of spins can be computed in this way:

$$\langle S_{\hat{a}}^{(A)} S_{\hat{b}}^{(B)} \rangle = \int d\lambda \rho(\lambda) S_{\hat{a}}^{(A)}(\lambda) S_{\hat{b}}^{(B)}(\lambda) = -\int d\lambda \rho(\lambda) S_{\hat{a}}^{(A)}(\lambda) S_{\hat{b}}^{(A)}(\lambda) \tag{2.11}$$

where we have made use of Eq.2.10. Now let us introduce a third unit vector $\hat{c}$; then, using the previous equation we can write:

$$
\begin{aligned}
\langle S_{\hat{a}}^{(A)} S_{\hat{b}}^{(B)} \rangle - \langle S_{\hat{a}}^{(A)} S_{\hat{c}}^{(B)} \rangle = - \int d\lambda \rho(\lambda) [S_{\hat{a}}^{(A)}(\lambda) S_{\hat{b}}^{(A)}(\lambda) - S_{\hat{a}}^{(A)}(\lambda) S_{\hat{c}}^{(A)}(\lambda)] = \\
= - \int d\lambda \rho(\lambda) S_{\hat{a}}^{(A)}(\lambda) S_{\hat{b}}^{(A)}(\lambda) [1 - S_{\hat{b}}^{(A)}(\lambda) S_{\hat{c}}^{(A)}(\lambda)]
\end{aligned}
\tag{2.12}
$$

where, for Eq.2.9, $(S_{\hat{b}}^{(A)}(\lambda))^2 = 1$. By the same equation it follows also that $|S_{\hat{a}}^{(A)}(\lambda) S_{\hat{b}}^{(A)}(\lambda)| \leq 1$ and $\rho(\lambda)[1 - S_{\hat{b}}^{(A)}(\lambda) S_{\hat{c}}^{(A)}(\lambda)] \geq 0$ (remember $\rho(\lambda) \geq 0$ since it is a probability density); therefore:

$$
|\langle S_{\hat{a}}^{(A)} S_{\hat{b}}^{(B)} \rangle - \langle S_{\hat{a}}^{(A)} S_{\hat{c}}^{(B)} \rangle| \leq \int d\lambda \rho(\lambda) [1 - S_{\hat{b}}^{(A)}(\lambda) S_{\hat{c}}^{(A)}(\lambda)] \tag{2.13}
$$

Since the second term on the right is just $\langle S_{\hat{b}}^{(A)} S_{\hat{c}}^{(B)} \rangle$ and $\int d\lambda \rho(\lambda) = 1$, we finally find the *Bell's inequality*:

$$
|\langle S_{\hat{a}}^{(A)} S_{\hat{b}}^{(B)} \rangle - \langle S_{\hat{a}}^{(A)} S_{\hat{c}}^{(B)} \rangle| \leq 1 + \langle S_{\hat{b}}^{(A)} S_{\hat{c}}^{(B)} \rangle \tag{2.14}
$$

Note that Bell's reasoning is particularly well-constructed: it holds for any local hidden variable theory and, moreover, since it deals with expectation values (i.e. averages), it is directly comparable with quantum mechanics' predictions.

2.2.2 Quantum prediction

Consider again the quantum (maximally) entangled state of AB:

$$
|\Psi\rangle = \frac{1}{\sqrt{2}} (|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle) \tag{2.15}
$$

where now we have replaced $|EPR\rangle$ with $|\Psi\rangle$ in Eq.2.3, for the sake of a clearer notation. The observable measured on A and B are respectively:

$$
\sigma^{(A)} \cdot \hat{a} \qquad , \qquad \sigma^{(B)} \cdot \hat{b} \tag{2.16}
$$

which have both eigenvalues ± 1 (recall Sec.1.3.1). Since the state $|\Psi\rangle$ has total vanishing angular momentum, we have:

$$
(\sigma^{(A)} + \sigma^{(B)}) \, |\Psi\rangle = 0 \tag{2.17}
$$

where we have left the common (arbitrary) axis of measurement ($\hat{n}$). By using Eq.1.5 we can compute the quantum expectation value of the products of the two spins (again, in general along two different axis for A and B).

$$
\begin{aligned}
\langle S_{\hat{a}}^{(A)} S_{\hat{b}}^{(B)} \rangle = \langle \Psi | (\sigma^{(A)} \cdot \hat{a}) \otimes (\sigma^{(B)} \cdot \hat{b}) | \Psi \rangle = - \langle \Psi | (\sigma^{(A)} \cdot \hat{a})(\sigma^{(A)} \cdot \hat{b}) \otimes I | \Psi \rangle = \\
= -Tr[(\sigma^{(A)} \cdot \hat{a})(\sigma^{(A)} \cdot \hat{b})\rho_A] = \\
= -\frac{1}{2} Tr[(\sigma^{(A)} \cdot \hat{a})(\sigma^{(A)} \cdot \hat{b})] = -\frac{1}{2} a_i b_j Tr(\sigma_i \sigma_j) = -\frac{1}{2} a_i b_j (2\delta_{ij}) = \\
= -\hat{a} \cdot \hat{b} = -\cos\theta
\end{aligned}
\tag{2.18}
$$

where, in sequence, we made use of Eq.2.17, 1.15, $\rho_A = \frac{1}{2} I$, $\sigma^{(A)} \cdot \hat{a} = a_i \sigma_i$, $\sigma^{(A)} \cdot \hat{b} = b_j \sigma_j$, Eq.1.33 and finally we called θ the angle between the two directions $\hat{a}$ and $\hat{b}$.

It is now straightforward to show that Bell's inequality (Eq.2.14) is incompatible with quantum mechanics. In fact, we just have to provide an example in which Bell's inequality is violated. Let's consider $\hat{a}, \hat{b}, \hat{c}$ co-planar and with the first two forming an angle of 90, while the third bisects, as shown in Fig.2.2; then, using Eq.2.18:

$$
\langle S_{\hat{a}}^{(A)} S_{\hat{b}}^{(B)} \rangle = 0 \qquad , \qquad \langle S_{\hat{a}}^{(A)} S_{\hat{c}}^{(B)} \rangle = \langle S_{\hat{b}}^{(A)} S_{\hat{c}}^{(B)} \rangle = -\frac{\sqrt{2}}{2} \tag{2.19}
$$

But this is inconsistent with Bell's inequality which would prescribe:

$$
\frac{\sqrt{2}}{2} \leq 1 - \frac{\sqrt{2}}{2} \tag{2.20}
$$

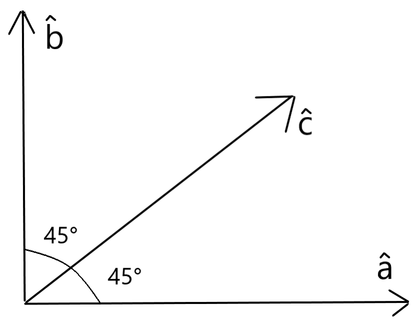

Figure 2.2: *Orientations of detectors which leads to a violation of Bell's inequality.*

So, Bell's theorem states the incompatibility between quantum mechanics and any local hidden variable theory one could imagine.
But how do we know which one is wrong? We must ask directly to nature through experiments. And in fact many of them have been settled in the decades after Bell's article. In particular Aspect's experiment (1982) was the first which addressed the so called locality loophole, which was raised against former experiments in which the two measurement events (on A and B) were not space-like, and so Eq.2.8 (which is an

hypothesis in Bell's derivation!) might not have been satisfied. Instead in Aspect's work, which used photons, the decision about the orientations of detectors was made while photons were in flight. Anyway, Aspect's and many other experiments confirmed quantum mechanics predictions and led to demise any attempt for local realism.

2.2.3 Consequences and subtleties

Since Bell found a violation of his inequality perpetuated by quantum mechanics, and given that he derived the inequality directly from EPR hypothesis, we must re-examine them. In particular, Bell's theorem leaves open two possible ways in front of us.
Again consider measurements on A and B to be, in general, space-like events.

- ***Non-local hidden variable theory***
  One possibility is to refuse the principle of locality, i.e. Eq.2.8 does not hold. The consequence is that the (type of) measurement on one system actually exerts an influence on the outcome of measures performed on the other. This unveils a subtle trait of quantum mechanics, which occurs with entanglement, that is indeed the *quantum non-locality* (note, however, that as we have already seen (Sec.1.3.3) there is no way for faster than light communication). So, we might recover realism, but with an high price to pay: these are the so called non-local hidden variable theories, to whom Bell's theorem does not apply; a famous example is Bohm's interpretation of quantum mechanics[3].

- ***Indeterminacy, quantum probability and non-separability***
  The other possibility is to keep locality but with the demise of realism, and so to accept indeterminacy. In this case we must reconsider the hypothesis of the existence of a joint probability distribution $\rho(\lambda)$, which for Bell governed the possible outcomes of every measurements on the two systems. In fact, behind this hypothesis hide the so called *counterfactuals*: the assumption that we could assign probabilities of outcomes to measurements that could have been made, but were non actually performed. We are used to employ counterfactuals in the classical world, because probability distributions arise from ignorance; but in the quantum world this is not true anymore, because of Bell's theorem. Indeed, as Bohr's *principle of complementary* states, not only we are unable to measure non-commuting observables, but also we cannot build a simultaneous probability distribution of their outcomes. Therefore, quantum probability is deeply different from the classical one, since it is intrinsic in the measurement process and not due to ignorance. How can we explain correlations in this picture? By refusing separability, which was the other assumption in EPR. So, we must envision the two entangled systems as an inseparable entity, even if they are far distant one from each other: this is

[3]See Appendix B for a brief presentation of Bohm theory.

sometimes called *quantum holism*. In this sense, one may interpret the collapse of wave function as just a mathematical artifact to treat the two A,B as subsystems of a single joint entangled system AB.

About eighty years have passed since quantum mechanics theory was formulated. But we still lack a deep understanding of it.

2.3 CHSH inequalities

We shall consider in this section a simpler form of Bell's inequality, but that indeed preserves its essence. This work is due to Clauser, Horne, Shimony and Holt (CHSH) [12].
Let us consider again Alice and Bob, respectively with a system A and a system B. Suppose each of them can choose between two observables to measure: A_1, A_2 for Alice, B_1, B_2 for Bob. Moreover, these observables have eigenvalues ± 1, i.e. these are the possible outcomes of each measurement. We make the assumption of locality, so by Eq.2.8, we have:

$$[A_i, B_j] = 0 \quad , \quad \forall i, j \in \{1, 2\} \tag{2.21}$$

Instead, in general, we do not make the same (non-interference) assumption for A_1 e A_2 nor for B_1 e B_2. We define the following observable:

$$C = (A_1 + A_2)B_1 + (A_1 - A_2)B_2 \tag{2.22}$$

Firstly, we suppose that a local hidden variable theory does subsist. If this is the case, had we the complete description of the systems A,B, we would know with certainty the values of all the four previous observables (simultaneously), since they are always defined; we call their values, respectively, a_1, a_2, b_1, b_2. Hence, the probability distribution of the outcomes is due to ignorance. If we compute the expectation value of observable C we then obtain:

$$C = (a_1 + a_2)b_1 + (a_1 - a_2)b_2 \tag{2.23}$$

$$|\langle C \rangle| \leq 2 \tag{2.24}$$

One can easily convince himself that this is true just remembering that a_i, $b_j \in \{-1, +1\}$. We have derived the *CHSH inequality*:

$$|\langle A_1 B_1 \rangle + \langle A_2 B_1 \rangle + \langle A_1 B_2 \rangle - \langle A_2 B_2 \rangle| \leq 2 \tag{2.25}$$

We see, by the reasoning we have followed, that it is a bound substantially analogous to Bell's inequality (Eq.2.14).

Now let's go quantum, but again we require locality condition, i.e. Eq.2.21 still holds. Carefully note that now Eq.2.23 cannot be written anymore! Thus we reason with operators. First of all $A_i^2 = I = B_j^2$ and so one can calculate:

$$\begin{aligned} C^2 = (A_1 + A_2)^2 + (A_1 - A_2)^2 + (A_2A_1 - A_1A_2)B_1B_2 + (A_1A_2 - A_2A_1)B_2B_1 = \\ = 4I - [A_1, A_2][B_1, B_2] \end{aligned} \tag{2.26}$$

It is straightforward to observe that:

$$|\langle [A_1, A_2] \rangle| \leq |\langle A_1A_2 - A_2A_1 \rangle| \leq |\langle A_1A_2 \rangle| + |\langle A_2A_1 \rangle| \leq 2 \tag{2.27}$$

And so, from Eq.2.26, we find: $|\langle C^2 \rangle| \leq 8$. Since in general $\langle C \rangle^2 \leq \langle C^2 \rangle$, we have obtained the *Cirel'son inequality*:

$$|\langle C \rangle| \leq 2\sqrt{2} \tag{2.28}$$

If we make a comparison between Eq.2.25 and Eq.2.28 we see that the quantum bound is larger. These equations indeed represent limitations on the strength of, respectively, classical and quantum (entanglement) correlations between two systems. So we have showed that the quantum ones outperform the classical counterparts.
For example, we may consider again the maximally entangled state of Eq.2.3 and identify $A_i = \sigma^{(A)} \cdot \hat{a}_i$, $B_j = \sigma^{(B)} \cdot \hat{b_j}$. Given the orientations of the axis as in Fig.2.3, we can compute (recall that the expectation value of product of spins is given by Eq.2.18):

$$|\langle C \rangle| = 2\sqrt{2} > 2 \tag{2.29}$$

Hence, it violates CHSH and saturates Cirel'son inequality.

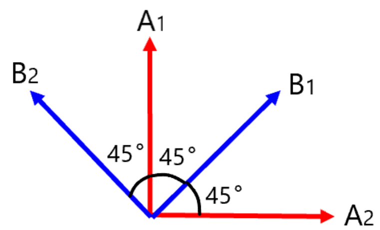

Figure 2.3: *Maximal violation of CHSH.*

2.4 GHZ state

We now consider a 3-qubit entangled state which is due to Greenberger, Horne and Zeilinger [13]:

$$|GHZ\rangle = \frac{1}{\sqrt{2}}(|000\rangle + |111\rangle) \tag{2.30}$$

Just like before, the three systems (A, B, C) are distributed respectively to Alice, Bob and Charlie. Suppose that two of them measure the spin along y and the third one along x. Again we assume locality, i.e. $[\sigma_i^{(A)}, \sigma_j^{(B)}] = 0$ and analogously for A,C and B,C. Recalling the actions of Pauli matrices on basis vector (Sec.1.3.1), one sees that:

$$\begin{aligned}\sigma_2^{(A)} \otimes \sigma_2^{(B)} \otimes \sigma_1^{(C)} |GHZ\rangle = \sigma_1^{(A)} \otimes \sigma_2^{(B)} \otimes \sigma_2^{(C)} |GHZ\rangle = \sigma_2^{(A)} \otimes \sigma_1^{(B)} \otimes \sigma_2^{(C)} |GHZ\rangle = \\ = -|GHZ\rangle\end{aligned} \tag{2.31}$$

Now, if there is an hidden variable theory, then we can simultaneously consider all the possible values of the spin for each qubit and, from Eq.2.31, we may compute:

$$S_y^A S_y^B S_x^C = S_x^A S_y^B S_y^C = S_y^A S_x^B S_y^C = -1 \tag{2.32}$$

The above are numbers, so we can multiply them and obtain:

$$(S_y^A S_y^B S_y^C)^2 S_x^A S_x^B S_x^C = -1 \quad \Longrightarrow \quad S_x^A S_x^B S_x^C = -1 \tag{2.33}$$

So this is the prediction of the products of spins which we make using counterfactuals. But what about quantum mechanics? $|GHZ\rangle$ is indeed an eigenstate of $\sigma_1^{(A)} \otimes \sigma_1^{(B)} \otimes \sigma_1^{(C)}$, but of course of eigenvalue $+1$:

$$\sigma_1^{(A)} \otimes \sigma_1^{(B)} \otimes \sigma_1^{(C)} |GHZ\rangle = +|GHZ\rangle \tag{2.34}$$

Therefore, again, Bohr's principle of complementarity forbids to reason with counterfactuals: since σ_1 and σ_2 (for the single qubit here!) do not commute, they are incompatible, and so we cannot infer the outcome of measurement of $\sigma_1^{(A)} \otimes \sigma_1^{(B)} \otimes \sigma_1^{(C)}$ from the postulation of an outcome of $\sigma_2^{(A)} \otimes \sigma_2^{(B)} \otimes \sigma_2^{(C)}$, as we implicitly did in Eq.2.33.
Further, note that here the inconsistency between local realism and quantum mechanics occurs at any single trial, whereas the violation of Bell's inequality (Eq.2.14) as we have seen in Sec.2.2.1 reveals only when one considers averages.

Part II

Tachyons

Chapter 3

Special Relativity

There is no such thing as
simultaneity of distant events.

Albert Einstein

3.1 Principles

In 1905 Einstein published his Theory of Special Relativity, which was based on these two postulates:

- **Principle of Special Relativity**
  The laws of physics are the same for all inertial observers.

- **Postulate of invariance of the speed of light**
  The speed of light in vacuum has the same value c in all inertial frames.

An inertial frame is a frame in which a body not subjected to forces is in uniform rectilinear motion[1]. The first principle expresses a physical equivalence among all inertial frames and states the absence of preferred ones. Consequently, measurements performed by observers in different inertial frames have the same physical dignity.
The invariance of the speed of light, moreover, entails renewed concepts of space and time, with respect to the Newtonian absolutes ones (i.e. invariance of the spatial distance and of the time duration), which result in (boost) Lorentz transformations of coordinates between two inertial frames of reference with relative velocity $\vec{v} = v\hat{x}$:

[1]Einstein resumed the Galilean/Newtonian definition which is actually a loophole, since (the effect of) a force is defined by Newton's law $\vec{F} = \frac{d\vec{p}}{dt}$, but this latter is only true for inertial observers. This ambiguity would have vanished thanks to General Relativity.

$$\begin{cases} x' = \frac{x-vt}{\sqrt{1-\frac{v^2}{c^2}}} \\ y' = y \\ z' = z \\ t' = \frac{t-\frac{v}{c^2}x}{\sqrt{1-\frac{v^2}{c^2}}} \end{cases} \tag{3.1}$$

Eq.3.1 can be derived simply by considering two inertial frame of references S and S', and assuming, in addition to the above postulates, linearity (and invertibility) of the transformation, space homogeneity and isotropy (no point nor direction in space is privileged) which allow us to orient the axis of the two frames in parallel and make their origins coincident at $t = t' = 0$, since also time homogeneity holds (no instant of time is privileged), and finally the requirement that this coordinates change recovers Galileian transformations in the limit $v << c$.
Perhaps the most important consequence of the principles stated by Einstein is the relativity of simultaneity. Let's consider the following situation. S (station) and S' (train) are two inertial frames and v is their relative velocity; O and O' are observers localized at the origins of these frames. Suppose that, at $t = t' = 0$, O and O' coincide and two light pulses are emitted by equidistant sources in the frame S, placed respectively at $x = \pm L$. Evidently, for observer O the two pulses will arrive simultaneously, namely after $\Delta t = \frac{L}{c}$. Instead, O' will be reached before by the pulse emitted from the source the train is moving towards. Now, since this difference in the reception of the signals is not due (as it was in the Galileian Relativity) to the different relative velocities of the two light pulses with respect to O', who instead always measures c as the speed of light, this observer must conclude that the two light pulses were not emitted at the same instant of time. That is, the *invariance* of the speed of light implies the relativity of simultaneity (remember that both versions, of O and O', are of equal dignity because of the principle of relativity).
From this, it obviously follows the impossibility of clocks' synchronization in different frames. Note again that this result is a consequence of the invariance of light speed, no speed limit has been invoked.
Furthermore, from Eq.3.1 one may obtain the formulas for "contraction" of (longitudinal) length and "dilation" of time. If S is the inertial frame of reference in which a bar is at rest along x axis and its proper length is L, then another inertial frame S' will measure as the length of the bar (here the positions of the two extremes of the bar must be measured *simultaneously*):

$$L' = L\sqrt{1-\frac{v^2}{c^2}} \leq L \tag{3.2}$$

Whereas if in S two events occur at the same spatial point, but at two different instants of time separated by $\Delta\tau$ (this is called their *proper time*), then in S' their temporal distance will be measured as:

$$\Delta t' = \frac{\Delta\tau}{\sqrt{1-\frac{v^2}{c^2}}} \geq \Delta\tau \tag{3.3}$$

which also implies that they do not occur at the same position here.
In the following we shall often call $\gamma := \frac{1}{\sqrt{1-\beta^2}} \geq 1$, where $\beta := \frac{v}{c} \leq 1$.
"Contraction" and "dilation" are in quotes because they are not real dynamical effects, i.e. a body does not shrink or slow down; in fact, since velocity (apart from that of light!) is a relative quantity, it follows that if an object really contracted, there would be a violation of the principle of relativity, as just one preferred frame would accurately account for that contraction. The point here, actually, is the relativity of space and time, i.e. of the measure of distance (length) and duration, which indeed follows from the principles of Special Relativity.

3.2 Minkowski space-time

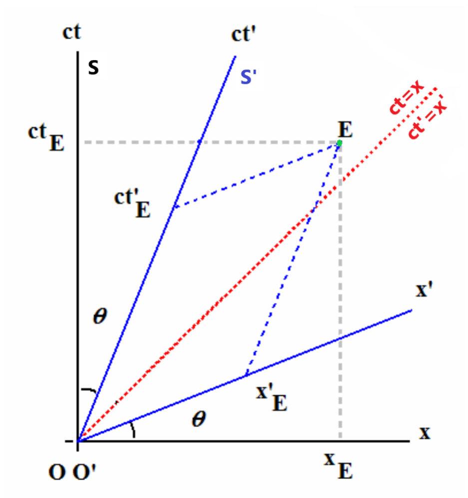

Figure 3.1: *Minkowski diagram (one spatial dimension).*

Lorentz transformations (Eq.3.1) lead us to consider space and time not more as different concepts, but rather as parts of the true fundamental structure of reality: space-time. This is a 4D geometric space, the *Minkowski space-time* (in Fig.3.1 is displayed a 1-spatial dimensional diagram), whose points are events and which is flat since the metric globally defined is the Minkowski's one[2]:

$$\eta_{\mu\nu} = diag(1, -1, -1, -1) \tag{3.4}$$

We can therefore define the space-time distance between events, which becomes in Special Relativity the true absolute quantity (i.e. Lorentz invariant), in contrast to Newtonian space and time, as:

$$ds^2 = c^2dt^2 - dx^2 - dy^2 - dz^2 \tag{3.5}$$

This invariance guarantees that all inertial frames agree on the structure of space-time. Once again, until now we have not required the speed of light to be any sort of limit. We now assume the principle of locality, which we have already encountered in Sec.2.1: the cause must precede the effect, i.e. (from Eq.3.5) no causal signal can travel faster than light. Hence, depending on the space-time distance between two given events A and B, we may classify them among three cases:

- $ds^2 < 0$: ***space-like*** separated
  The two events are in each other *elsewhere*, since they cannot be causally connected. Thus, it will exist one reference frame in which A and B are simultaneous (but differ for spatial locations), others where A precedes B and still others in which B precedes A.
  For example, if in frame S they are distanced Δx in space and Δt in time, from Eq.3.1 one finds that the inertial system S' in which the two events are simultaneous moves with velocity $v = c^2\frac{\Delta t}{\Delta x}$ with respect to S.

- $ds^2 = 0$: ***light-like*** separated
  The two events are on the light-(hyper)cone of each other; they might be causally related only by means of a light signal.

- $ds^2 > 0$: ***time-like*** separated
  One event is in the past-cone of the other and the latter is in the future cone of the former. A and B might be causally connected, through subluminal ways. Again by Eq.3.1, one can see that it exists an inertial frame in which the two events occur at the same spatial point (but necessarily at different times): it is indeed the frame of the (fictitious) signal that might connect them, that is the system which moves at $v = \frac{\Delta x}{\Delta t}$ with respect to S (where Δx and Δt are also measured in S).
  Also, it is fundamental that all inertial frames agree in the temporal sequence of time-like events: in a generic frame one finds that $\Delta t = \gamma \Delta\tau$, so they both have the same sign, the latter being the time elapsed in the frame where A and B take

[2]In the framework of General Relativity, space-time is generally not flat, but curved; however, we can always recover Minkowski space-time locally.

place at the same spatial point, which is thus the proper time we have in Eq.3.3: $\Delta\tau$ it is the minimal time lapse between two events.
In the case of a time-like path $\frac{\Delta s}{c} = \frac{1}{c}\sqrt{c^2\Delta t^2 - \Delta x^2 - \Delta y^2 - \Delta z^2}$ is the quantity measured (in a generic inertial frame) by a clock along the direction it travels; it is exactly the proper time: in fact, being invariant, $\frac{\Delta s}{c} = \Delta\tau$.

In the following Fig.3.2 we give a representation of the described regions of space-time with respect to a given event.

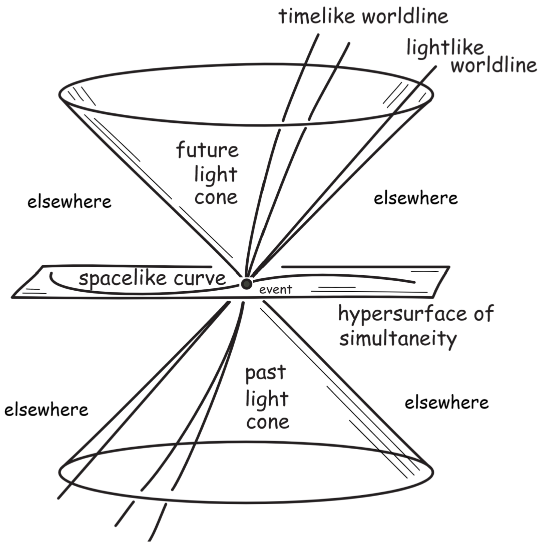

Figure 3.2: *Light-cone. This is a modified version of a picture Copyright John D. Norton.*

3.3 4-vectors and main relations

In this section we briefly recall some results of Special Relativity which will be used in what comes afterwards. The 4-position is $x^\mu \equiv (x^0, x^1, x^2, x^3) \equiv (x^0, \vec{x})$, where with index 0 we identify the time component and with $1, 2, 3$ respectively the x, y, z components, whereas $\vec{x}$ is the 3-position. We will adopt analogous notation for a generic 4-vector. So, Eq.3.5 writes as $ds^2 = \eta_{\mu\nu} dx^\mu dx^\nu = dx_\mu dx^\mu$ and $d\tau = \frac{ds}{c}$.
A Lorentz (boost) transformation (Eq.3.1), from frame S to S' can be written as $x^{\mu'} = \Lambda^{\mu'}{}_\nu x^\nu$, where:

$$\Lambda^{\mu'}{}_\nu = \begin{bmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{3.6}$$

It preserves Minkowski metric since $\lambda^T \eta \lambda = \eta$.
We can now define the 4-velocity as (we define $\gamma_p = \frac{1}{\sqrt{1-\frac{u^2}{c^2}}}$, where u is the speed of the particle in S):

$$U^\mu \equiv \frac{dx^\mu}{d\tau} = \gamma_p(c, \vec{u}) \tag{3.7}$$

whose norm is $U_\mu U^\mu = c^2$ (time-like) for any massive particle, whereas is zero (light-like) for photons. Using Eq.3.6 we can find how the components of 4-velocity Lorentz-transform. For example for spatial components we have:

$$u'_x = \frac{u_x - v}{1 - \frac{vu_x}{c^2}} \quad , \quad u'_y = \frac{u_y}{\gamma(1 - \frac{vu_x}{c^2})} \quad , \quad u'_z = \frac{u_z}{\gamma(1 - \frac{vu_x}{c^2})} \tag{3.8}$$

and its (3-)magnitude becomes:

$$u' = c\sqrt{1 - \frac{(1 - \frac{u^2}{c^2})(1 - \frac{v^2}{c^2})}{(1 - \frac{u_x v}{c^2})^2}} \tag{3.9}$$

From this latter we can note that if a particle moves with speed $u < c$ in S, it will move subluminally in any other reference frame with a speed generally $u' \neq u$ ($v < c$, in fact, from the postulate of c-invariance, it does not exist any rest frame of a photon). Instead, if the particle is a photon (i.e. a mass-less particle, moving with speed c), in any frame it will have the same speed, that is c, consistently with the postulate of invariance of the speed of light. In next sections we shall tackle the question about the possibility of particles with speed $u > c$.
Futher, if m is the proper mass of a particle (i.e. the mass measured by an observer in the rest inertial frame of that particle), its 4-momentum is:

$$P^\mu \equiv mU^\mu = \gamma_p m(c, \vec{u}) = (\frac{E}{c}, \vec{p}) \tag{3.10}$$

where we have used the famous mass-energy equivalence (for the derivation see Sec.3.4), $E = \gamma_p mc^2$, and the relativistic 3-momentum $\vec{p} = \gamma_p m\vec{u}$. Note that for a photon ($m = 0$) one can only write $P^\mu = (\frac{E}{c}, \vec{p})$ which thus is a more general relation. If we compute the norm of the 4-momentum we obtain:

$$P_\mu P^\mu = m^2c^2 = \frac{E^2}{c^2} - p^2 \tag{3.11}$$

which is the very important *energy-momentum* relation. Again, for a photon this 4-vector is light-like. The components of 4-momentum transform in this way with Eq.3.6:

$$\frac{E'}{c} = \gamma(\frac{E}{c} - \beta p_x) \quad , \quad p'_x = \gamma(p_x - \beta\frac{E}{c}) \quad , \quad p'_y = p_y \quad , \quad p'_z = p_z \tag{3.12}$$

3.4 What does Relativity actually say about faster than light motion?

It is sometimes said that Special Relativity forbids the existence of faster than light particles. This is not actually true, since the theory does not, by itself, deny the possibility that such particles do exist. In fact, what Relativity really forbids is another thing, namely the possibility that a particle which moves subluminally (in any inertial frame, as we have seen from Eq.3.9) may be accelerated to speeds equal (or higher) than the speed of light. This can be shown through energy considerations as follows.
Consider a particle of proper mass m subjected to some external work, along a path from $\vec{x}_A$ to $\vec{x}_B$. The infinitesimal work can be computed as (we now call $\vec{v}$ the particle velocity and we leave the subscript in γ_p, since there is no ambiguity here):

$$\begin{aligned} dW = \frac{d\vec{p}}{dt}\cdot d\vec{x} = \frac{d}{dt}(m\gamma(v)\vec{v}\,)\cdot d\vec{x} = m[\gamma^3(v)\frac{v}{c^2}\dot{v}\vec{v} + \gamma(v)\dot{\vec{v}}\,]\cdot\vec{v}dt = \\ = m[\gamma^3(v)\frac{v^3}{c^2}dv + \gamma(v)d\vec{v}\cdot\vec{v}\,] = m\gamma(v)vdv[\gamma^2(v)\frac{v^2}{c^2}+1] = \\ = m\gamma^3(v)vdv = d(mc^2\gamma(v)) \end{aligned} \tag{3.13}$$

where we have used $d\vec{x} = \vec{v}dt$ and $\vec{v}\cdot d\vec{v} = \frac{1}{2}d(\vec{v}\cdot\vec{v}) = vdv$. Therefore the (relativistic) kinetic energy acquired by the particle will be:

$$K = \Delta W = \int_A^B dW = \frac{mc^2}{\sqrt{1-\frac{v^2}{c^2}}} - mc^2 \tag{3.14}$$

where we assumed the particle was initially at rest in this frame. We see it has a rest energy $E_0 = mc^2$ and we have also found the total (relativistic) energy:

$$E = \gamma mc^2 = E_0 + K \tag{3.15}$$

We can observe that in the limit $v \to c$ the kinetic energy diverges: it is necessary an infinite amount of work to accelerate a massive particle to the speed of light, as one can also see in Fig.3.3.
One may argue that faster than light particles do not exist since there is already this barrier for usual particles. But note that with a similar argument one should also exclude photons, which do exist.
In the next chapter we will tackle the problem of how one can describe such particles with the relativistic theory and we will discuss related causal issues.

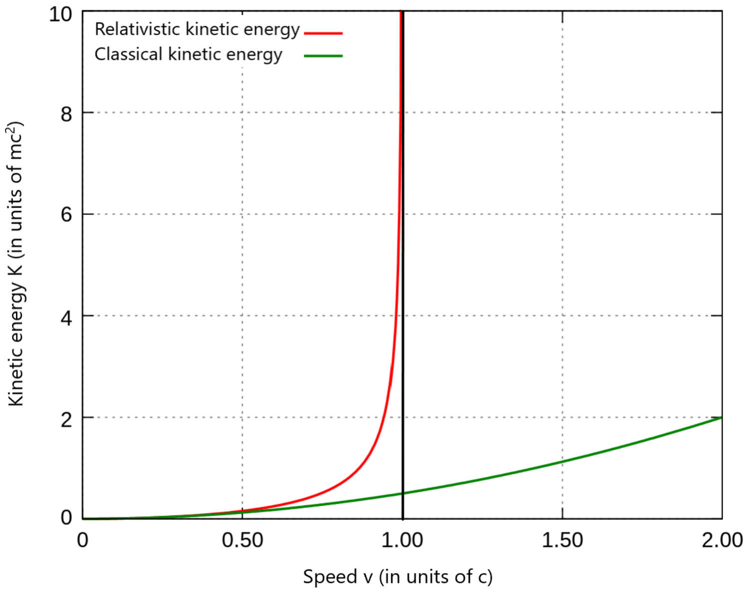

Figure 3.3: *Comparison between relativistic and classical kinetic energy as functions of speed.*

Chapter 4

Tachyons!

> The limiting velocity is c, but a limit has two sides.
>
> Gerald Feinberg

4.1 Kinematics of tachyons

Faster than light particles are commonly called *tachyons*[1]. Thus far, there is no experimental evidence for their existence. Nevertheless, as we have explained in the previous chapter, they are not prohibited by Special Relativity and a consistent theory of tachyons can be indeed formulated. Why should one investigate these quite exotic particles? Well, in a sense we follow the idea that *if something is not expressly forbidden by the laws of Nature, it possibly exists*. And also, more interestingly, tachyons might provide the non-local interaction which characterizes entangled systems (Sec.2.2.3).
Firstly, it is easy to see from Eq.3.9 that a particle which is a tachyon for a given inertial reference frame S, it will also move with a velocity $u' > c$ in any other inertial frame S' ($v < c$). This is also shown graphically in Fig.4.1, where we can note that the word line of a tachyon is in fact space-like.

One bizarre thing is that, recalling Lorentz transformation (Eq.3.1),

$$\Delta t' = \gamma \Delta t(1 - \frac{uv}{c^2}) \tag{4.1}$$

there exist inertial frames where the time ordering of (space-like, $ds^2 < 0$) events along the trajectory of the tachyon is inverted, namely when the condition

$$uv > c^2 \tag{4.2}$$

[1] From the Greek word $\tau\alpha\chi\acute{\nu}\varsigma$, which means rapid, swift.

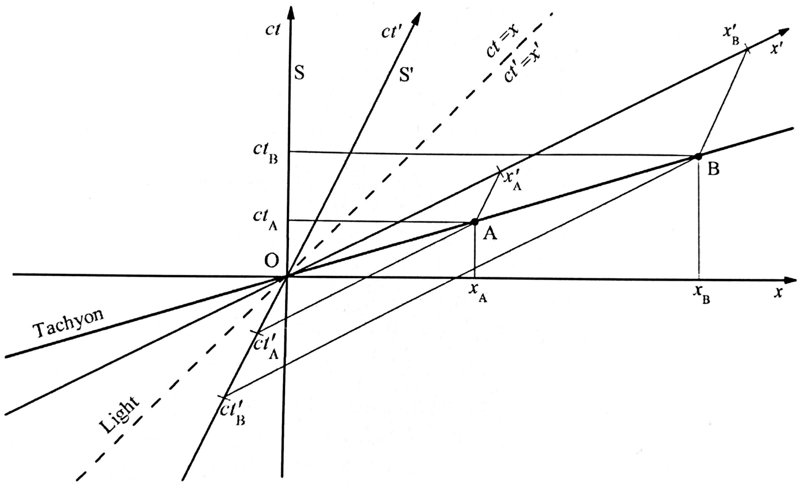

Figure 4.1: *A tachyon worldline.*

occurs. That is, in general, when the world line of the tachyon lies in the region between the light-cone and the $x'y'z'$ hypersurface, the temporal sequence of its space-time points is the same for S and S'; otherwise it is reversed. For example, let's consider Fig.4.2, where we look at the events of emission and absorption of a tachyon in two inertial frames which satisfy the condition of Eq.4.2.

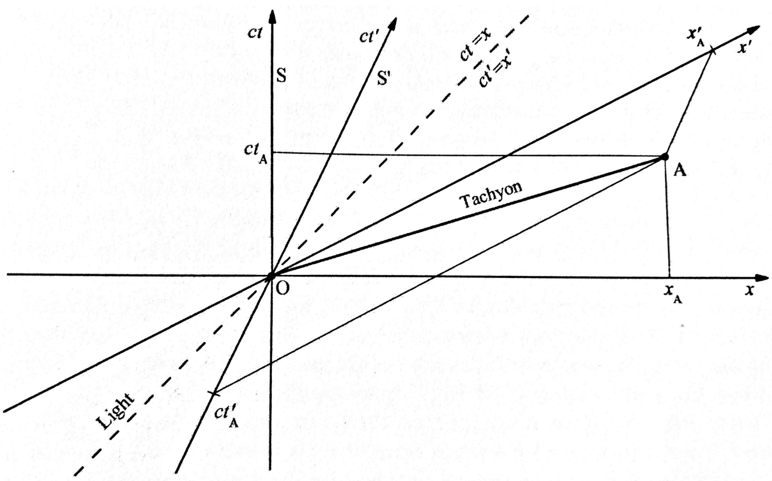

Figure 4.2: *Creation (O) and annihilation (A) of a tachyon as seen in two inertial frames.*

Whereas in S the particle is first emitted (in O) and then absorbed (in A), in S' the absorption precedes the emission! So, in this frame, the tachyon will appear to travel backward in time. This is rather problematic because causality would require that the creation (emission) of a particle occurs before its annihilation (absorption). We will come back to this issue in a later section, after looking how things apparently become bad also in dynamics.

4.2 Dynamics of tachyons

We shall start by recalling the relativistic energy and (magnitude of) momentum for a particle moving at velocity u (in a given Lorentz frame S):

$$E = m\gamma_p c^2 = \frac{mc^2}{\sqrt{1-\frac{u^2}{c^2}}} \quad , \quad p = m\gamma_p u = \frac{mu}{\sqrt{1-\frac{u^2}{c^2}}} \tag{4.3}$$

Since these quantity would become imaginary for a tachyon, i.e. a particle with $u > c$, we assign an imaginary (rest) mass m to it, in order to rescue real dynamical quantities. *Why is this acceptable?* Well, the rest mass is defined as the mass of a particle measured by an observer in the frame in which the particle is at rest; but as we have seen in Sec.4.1, tachyons move superluminally in every inertial frame, so they cannot be brought to rest. Therefore m is not a measurable quantity and thus its being imaginary is not a relevant issue. So, the energy and momentum for a tachyon, if one defines $m = im_i$, $m_i \in \mathbb{R}$, can be written as:

$$E = \frac{m_i c^2}{\sqrt{\frac{u^2}{c^2}-1}} \qquad , \qquad p = \frac{m_i u}{\sqrt{\frac{u^2}{c^2}-1}} \tag{4.4}$$

Further, the 4-velocity of a tachyon $U^\mu = \frac{dx^\mu}{d\tau}$ has again imaginary components because $d\tau$ is imaginary for faster than light particles. But in general the 4-velocity of a particle is not actually a measurable quantity; the important thing is that the dynamical quantity, so the 4-momentum, is indeed real: $P^\mu = mU^\mu$. The energy-momentum relation (Eq.3.11) for a tachyon reads:

$$-m_i^2 c^2 = \frac{E^2}{c^2} - p^2 \tag{4.5}$$

where we see that the 4-momentum is space-like, i.e. it points outside the light-cone. From Eq.4.4, we observe that the E and p have these interesting limits for a tachyon:

$$\begin{aligned} u \to c: &\qquad E \to \infty \qquad p \to \infty \\ u \to \infty: &\qquad E \to 0 \qquad p \to m_i c \end{aligned} \tag{4.6}$$

So, if a tachyon can move at $u = \infty$, its energy goes to zero but it still carries finite momentum.

The graphs in Fig.4.3 show that a tachyon decreases its energy and momentum as it speeds up in a given frame. If $u = \infty$, the worldline of the tachyon lies within a

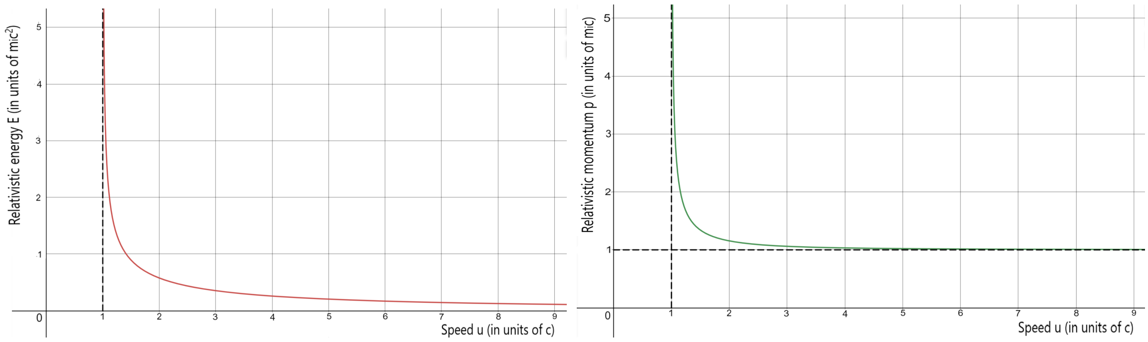

Figure 4.3: *Relativistic energy and momentum of a tachyon.*

hyperplane of simultaneity, i.e. it propagates instantaneously. Note also that Fig.4.3 discloses the nature of the limiting velocity c: we have seen in Sec.3.4 that for a particle which moves (in any frame) subluminally it will take an infinite amount of work to overcome (reach) the light barrier; here, symmetrically, we see that c is also a limit for a superluminal particle, but this time a lower one: *if one wants to slow down a tachyon to speed c, it will be needed infinite energy.*

Now we describe what problem arises in dynamics. From the transformations of energy and momentum of Eq.3.12 (here suppose $p_x = p$), we see that there exist inertial frames where the energy of the tachyon becomes negative. In fact, noting that from Eq.4.4 we have $u = \frac{p}{E}c^2 > c$, then:

$$E' = \gamma(E - vp) \quad \Longrightarrow \quad \frac{E'}{E} = \gamma(1 - \frac{vu}{c^2}) \tag{4.7}$$

so, in a frame which moves with speed v (with respect to S, which measures speed u and $E > 0$ for the tachyon), that satisfies $uv > c^2$, the tachyon has negative energy! This is a big problem: it cannot exist a stable system which emits negative energy particles, since there is no limit on how many such particles it could emit.

Hence, both in kinematics and dynamics very serious issues seem to arise when one tries to study tachyons, namely some Lorentz transformations may cause time reversal of emission-absorption and change in the sign of the energy. However, we will see in the next section that, if we introduce a simple (re-)interpretation of the emission and absorption events, these worries can be neutralized.

4.3 Reinterpretation principle

The key point is to observe that the time reversal and the change of sign of the energy for a tachyon both occur in the same Lorentz frame's transformations: those satisfying the condition of Eq.4.2. In addition, since emission/absorption events are obviously local

(i.e. they are defined at a single point), one can individuate them, *in a given frame*, after observing which one happens first (the emission) and which one later (the absorption). Thus, here is what the reinterpretation principle says: at a (spacetime) point there is no distinction between the emission of a negative-energy tachyon and the absorption of a positive-energy tachyon.
So, back to Fig.4.2, in frame S' there is not a tachyon with negative energy which travels backward in time from absorption in A to emission in O, but instead a tachyon with positive energy, travelling forward in time from (in such a frame) the emission in A to the absorption in O. Actually, we are reasoning in this way: in S' (which satisfies the condition $uv > c^2$), rather than having a 4-momentum $P^{\mu'}$ pointing in the negative time direction (because of $cP^{0'} = E' < 0$), we make use of the freedom to define $d\tau$ as the negative square root of ds^2, in lieu of the positive one. Thanks to this reinterpretation, therefore, the source and the receiver of the tachyon become relative to the frame of reference, while tachyons travel in the forward direction of time with positive energy for all observers. For an example refer to Fig.4.4.

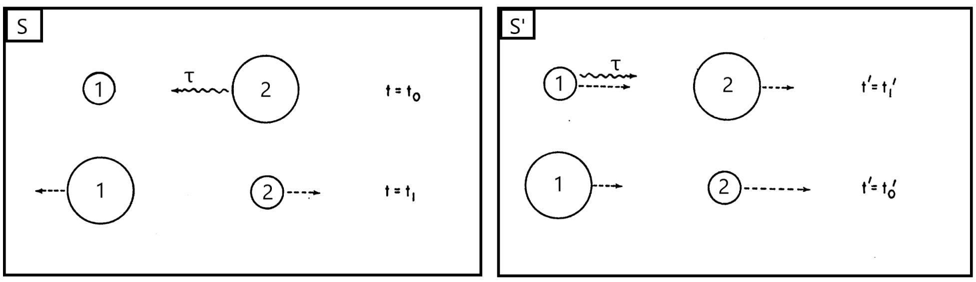

Figure 4.4: *Emission and absorption of a tachyon by two atoms, as seen in two different frames satisfying the condition of Eq.4.2. The dimension of the circles distinguishes the ground state (smaller ones) from the excited state (bigger ones) of the atoms.*

Here we suppose that the sources of tachyons are atoms 1 and 2. In the inertial frame S, both atoms are at rest; initially, 1 is in its ground state, whereas 2 is in an excited state. At $t = t_0$, let atom 2 emit a tachyon τ; this requires some amount of energy and hence 2 will drop to its ground state, kicking back with some speed. When the tachyon is absorbed at $t_1 > t_0$ by atom 1, this latter jumps to an excited state and also recoils in the direction of τ. If S' is an inertial frame such that the condition of Eq.4.2 is satisfied, there $t'_1 < t'_0$ and the processes of emission/absorption are reversed. Again, 1 begins in the ground state, 2 in an excited one, however in this frame they are both moving. Now atom 1 is the one emitting, at t'_1: how can it do that if it is in its ground state? Well, it may convert part of its translational energy into internal one and simultaneously emit the tachyon. This process is known as *tachyonic condensation*. It cannot happen if the

emitted particle has time-like 4-momentum, but it is instead possible for a space-like one, because in this case the emission is seen as absorption in the rest frame of the atom. Finally, 2 will absorb the tachyon at t_0' and it will make a transition to the ground state, by simultaneously increasing its kinetic energy.
So, we have described how the same process occurs for the two observers, according to the reinterpretation principle. Note that the atoms have the same (respective) state in both frames at the beginning and at the end of the events.
We shall now ask if tachyons may indeed be used to send reliable non-local signals. One can easily construct an exemplary situation in which this question appears to have a negative answer.

4.4 Tolman paradox

In this section we will describe a paradox, also known as *tachyonic anti-telephone*, introduced firstly by Einstein himself and Arnold Sommerfeld in 1910, but that owes its name to Richard C. Tolman, who presented it in 1917 [8].
Consider Fig.4.5.

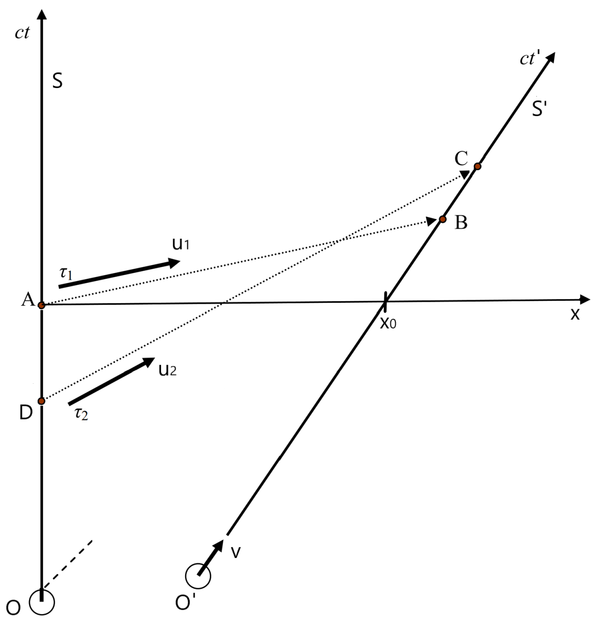

Figure 4.5: *Minkowski diagram showing the Tolman paradox.*

We have two inertial frames, S and S', the latter moving at velocity v along x with respect to the first. Suppose that two observers, O and O' are located in the origins of

the two frames and have devices capable to emit and absorb tachyons. At a given time observer O emits, in S frame, a tachyon τ_1 with speed u_1 towards the other observer (event A). This tachyon is then absorbed by O' (event B). Now let O' emit, in S' frame, a second tachyon τ_2 moving with velocity u_2' towards O (event C), immediately after the "receipt" of τ_1. τ_2 will be then finally absorbed by O (in S').
We have always specified in which frame we see the emission/absorption because of what we have seen in the previous sections. If the condition of Eq.4.2 does not hold, then the observers agree on the sequence of events, namely A-B-C-D, and no causal paradox arises. But suppose that the condition of Eq.4.2 is met for both u_1 and u_2'. According to the reinterpretation principle, emission and absorption are switched from S to S' and vice-versa. The problem is that the second tachyon, τ_2, "reaches" O (it is spontaneously emitted in S) before the first one, τ_1, is emitted. So, for example, one may imagine that observer O would therefore decide not to emit at all tachyon τ_1. But this was the trigger for the emission of τ_2 by observer O'! It seems we have fallen in a causal anomaly. This is *Tolman paradox*. As we have seen, it arises because O and O' exchange tachyons in such a way that their worldlines form a loop.
Let's compute explicitly; suppose the first emission A occurs at time $t = 0$ in S, when O' is at position x_0 in this frame. Thus, the Lorentz transformation for this problem can be inferred by Eq.3.1 (we set here $c = 1$ for faster calculation):

$$x' = \gamma((x - x_0) - vt) \qquad , \qquad t' = \gamma(t - v(x - x_0)) \tag{4.8}$$

So events A ("emission" of τ_1) and B ("absorption" of τ_1) will occur in the two frames S and S' at times:

$$\begin{aligned} A: \quad & t_A = 0 \qquad\qquad , \qquad t'_A = \gamma v x_0 \\ B: \quad & t_B = \frac{x_0}{u_1 - v} \qquad , \qquad t'_B = \frac{x_0}{\gamma(u_1 - v)} \end{aligned} \tag{4.9}$$

where we have used $x_A = x_0(1 + \frac{v}{u_1 - v}) = x_0 \frac{u_1}{u_1 - v}$. If at t'_B observer O' emits tachyon τ_2 with velocity $-u_2'$, $u_2' > 1$, using the first of Eq.3.8 (here seen as transformation of velocity from S' to S, so we have to substitute $-v$ with v) it will move in S with velocity:

$$u_2 = \frac{-u_2' + v}{1 - u_2' v} \tag{4.10}$$

If the condition $u_2' v > 1$ (Eq.4.2 with $c = 1$) is met, then the events C-D will be reversed in S; thus, the "absorption" of τ_2 will here occur at time:

$$t_D = \frac{x_0}{u_1 - v} - x_0 \frac{u_1(u_2' v - 1)}{(u_1 - v)(u_2' - v)} = \frac{x_0}{(u_1 - v)(u_2' - v)} [u_2'(1 - u_1 v) + u_1 - v] \tag{4.11}$$

where we see that if also $u_1 v > 1$ holds, then $t_D < 0 = t_A$. For example, in the extreme case $u_1 = u_2' = \infty$ (tachyons instantaneously connecting space-like events, with no transport of energy but still having finite momentum, as seen in Sec.4.2) one has:

$$t_D = -v x_0 \tag{4.12}$$

This tachyonic anti-telephone is the major argument against the possible existence of tachyons. One might exclude the violation of causality if tachyons were not exploitable for reliable (superluminal) signaling: that is, their emission/absorption is not completely experimentally manipulable. In fact, each observer O and O' can distinguish a spontaneous emission from the absorption of a negative-energy tachyon sent by the other observer, only by the comparison of their experimental registers, so through classical communication. Moreover, the loop of the paradox suspends tachyon τ_1 in a limbo of existence-non existence. These elements make us reminiscent of quantum mechanics. In the next chapter we will show how this intriguing connection may be achieved.

Part III

Tolman paradox and entanglement

Chapter 5

Paradox or quantum superposition?

> It can be argued that in trying to see behind the formal predictions of quantum theory we are just making trouble for ourselves.
>
> John Stewart Bell

5.1 Entanglement as loop of tachyons

We shall now show how entangled states of particles may be described through time loops of tachyons exchanged between those particles [9]. Let's consider Fig.5.1. We can see the two particles P and Q as two atoms with two possible states, ground state and excited one, in a situation similar to the one we have already encountered in Fig.4.4.

This figure shows, as we have seen in Sec.4.4, the arising of Tolman paradox when the two tachyons exchanged by P and Q have speeds which satisfy the condition of Eq.4.2. We can consider P, initially in the excited state, emitting τ_1 at t_A. Q, initially in its ground state, absorbs the tachyon, jumps to its excited state and emits (in frame S' where this atom is stationary) the tachyon τ_2, going back to the ground state. The absorption of this latter by P is seen in S as a spontaneous emission of τ_2 (from P). Being $t_D < t_A$, this emission would therefore prevent that of τ_1 at t_A, because atom P would already be in its ground state. So, again, we see the loop situation in which the tachyons τ_1 and τ_2, in a limbo of existence/non-existence, and the atoms P and Q, in a limbo of ground/excited states, appear to be because of superluminal connections. But this "limbo" reminds us quantum superposition. In Sec.1.1, in fact, we have seen that, according to the orthodox interpretation, up to a measurement is performed on the system, its state generally consists in a quantum superposition of eigenstates of the

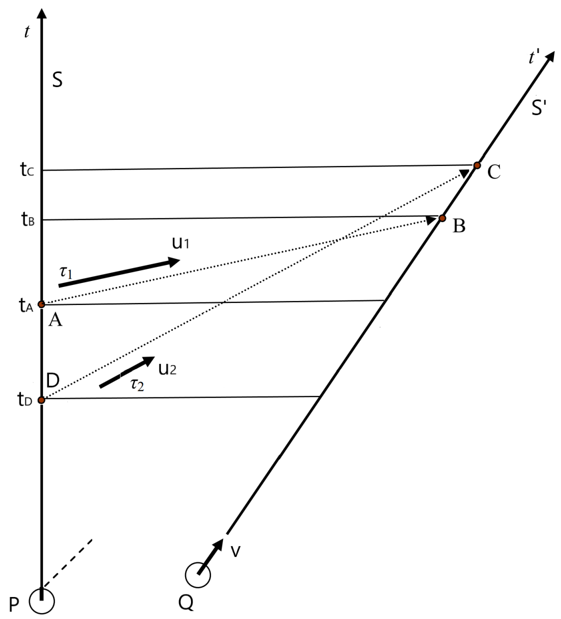

Figure 5.1: *The five temporal regions in Tolman paradox.*

observable to be measured. Moreover, when entangled systems are taken into account, we know from Sec.2.2.3 that only a *superluminal* hidden-variable theory could describe entanglement correlations in agreement with quantum predictions. With Tolman paradox, i.e. time loop of tachyons, it might be possible to describe the superposition of states by means of tachyons exchange, which would establish entanglement between the two parties connected by these faster than light particles.
We can now see how this interpretation may be realized. Refer again to Fig.5.1. We define the following eigenstates:

$$\begin{array}{lll} P: & |P^0\rangle \equiv \text{ground state} & |P^*\rangle \equiv \text{excited state} \\ Q: & |Q^0\rangle \equiv \text{ground state} & |Q^*\rangle \equiv \text{excited state} \\ \tau_j: & |0_j\rangle \equiv \text{no tachyon } j & |\tau_j\rangle \equiv \text{tachyon } j \text{ "emitted"} \\ \multicolumn{3}{l}{|\psi\rangle \equiv \text{state of the overall system}} \end{array} \tag{5.1}$$

where $j \in \{1, 2\}$. The temporal axis can thus be split into five regions, where $|\psi\rangle$ assumes different forms. Let's analyze each one of them.
Region 1:

$$|\psi(t < t_D)\rangle = |P^*\rangle\,|Q^0\rangle\,|0_1\rangle\,|0_2\rangle \tag{5.2}$$

this is simply the initial situation: excited state for atom P, ground state for Q, no

tachyons.
Region 2:

$$|\psi(t_D < t < t_A)\rangle = \frac{1}{\sqrt{2}} |P^*\rangle |Q^0\rangle |0_1\rangle |0_2\rangle \pm \frac{1}{\sqrt{2}} |P^0\rangle |Q^0\rangle |0_1\rangle |\tau_2\rangle = \\ = \frac{1}{\sqrt{2}} (|P^*\rangle |0_2\rangle \pm |P^0\rangle |\tau_2\rangle) |Q^0\rangle |0_1\rangle \tag{5.3}$$

here we have an entangled state between atom P and tachyon τ_2, constructed with the product of, respectively, excited state of the atom with absence of the tachyon and ground state of the atom with presence of the tachyon. Atom Q is in its ground state, τ_1 is absent.
Region 3:

$$|\psi(t_A < t < t_B)\rangle = \frac{1}{\sqrt{2}} |P^*\rangle |Q^0\rangle |\tau_1\rangle |0_2\rangle \pm \frac{1}{\sqrt{2}} |P^0\rangle |Q^0\rangle |0_1\rangle |\tau_2\rangle = \\ = \frac{1}{\sqrt{2}} (|P^*\rangle |\tau_1\rangle |0_2\rangle \pm |P^0\rangle |0_1\rangle |\tau_2\rangle) |Q^0\rangle \tag{5.4}$$

where we have a tripartite entangled state (remember GHZ state of Sec.2.4) between atom P and both tachyons. If P is in its excited state, then τ_1 exists and τ_2 does not; the opposite if P is instead in its ground state. The other atom, Q, is still in the ground state.
Region 4:

$$|\psi(t_B < t < t_C)\rangle = \frac{1}{\sqrt{2}} |P^*\rangle |Q^*\rangle |0_1\rangle |0_2\rangle \pm \frac{1}{\sqrt{2}} |P^0\rangle |Q^0\rangle |0_1\rangle |\tau_2\rangle = \\ = \frac{1}{\sqrt{2}} (|P^*\rangle |Q^*\rangle |0_2\rangle \pm |P^0\rangle |Q^0\rangle |\tau_2\rangle) |0_1\rangle \tag{5.5}$$

here we have another entangled state of three parties, namely atoms P, Q and tachyon τ_2, whereas τ_1 no longer exists. Note that, if one has no detectors for tachyons, this state is just an entangled state between the two atoms, that is, their excited states are always correlated and the same holds for their ground states (i.e. if one atom is measured in a certain state, the other will be find it that very one, once observed).
Region 5:

$$|\psi(t > t_C)\rangle = \frac{1}{\sqrt{2}} |P^*\rangle |Q^*\rangle |0_1\rangle |0_2\rangle \pm \frac{1}{\sqrt{2}} |P^0\rangle |Q^0\rangle |0_1\rangle |0_2\rangle = \\ = \frac{1}{\sqrt{2}} (|P^*\rangle |Q^*\rangle \pm |P^0\rangle |Q^0\rangle) |0_1\rangle |0_2\rangle \tag{5.6}$$

where, finally, we have a bipartite entangled state of P and Q and neither τ_1 nor τ_2 still does exist.

Note that in the displayed example we have assumed equally-weighted superpositions with phase 0 or π, so that the coefficients are just $\pm\frac{1}{\sqrt{2}}$.
Hence we have seen that, at least in the simple situation considered, it is possible to interpret atoms P and Q in Tolman paradox as an entangled Bell pair.

5.2 Final remarks

The shown idea has interesting points, but also several critical ones. Among the former, we can count a renewed interpretation of entanglement: it is established between particles exchanging tachyons in such a way to form a causal loop, exactly the case appearing in Tolman paradox. Moreover, the measurement problem can now be seen in a perhaps more natural way rather than the instantaneous collapse of the wave function: the absorption or deflection of a tachyon by a detector will disrupt the time loop (because the other tachyon does not materialize) and pick the eigenstate of the (entangled) system, in which this latter is found when observed. But, on this point, it will be necessary to understand how tachyons could interact with experimental devices. Further, it is not clear if extensions of the suggested model may account for the entangled correlations of observables with more than two eigenstates. Also, we have seen that one hypothesis is that the two systems (atoms P and Q in the previous example) are in relative motion, with a relative speed v such that Eq.4.2 is satisfied. So, if we suppose that tachyons move with infinite speed, then every value for v will be good. Another question is how a generic entangled state, for example the spin entanglement of the Bohm-EPR pair (Sec.2.1.2), can be described within this theory, i.e. how can the exchanged tachyons have the information about the possible values of spins? Perhaps it might be required the existence of tachyons with non-zero spin. But this would be in conflict with the quantum field theory of tachyons developed by G. Feinberg [7], in which they are spinless particles obeying the Fermi-Dirac statistics. In this mentioned theory, besides, tachyons can be described through waves, and so we may see, for instance, τ_2 as the back-scattering of a primary wave τ_1, when this latter interacts with atom Q.
In conclusion, aware of the multiple issues that the presented interpretation exhibits, one may still be persuaded by the idea that tachyons, not prohibited by Special Relativity, can have a crucial role in that ≪spooky action at a distance≫ that is quantum entanglement.

Appendix A

Measurement problem

We want to go back to the orthodox interpretation and make some thoughts. For example, one question regards the exact status of the wave function (i.e. quantum state) and its instantaneous collapse (i.e. reduction of the state): are these physically real or just purely epistemic? In fact, commonly both readings are used, depending on the physical phenomena to be explained: in the case of interference effects, such as the two-slits experiment, the wave function "goes through" both slits and therefore determines the distribution of electron spots on the shield; in other situations, for instance in Sec.2.1.2 where we have shown how the form of the density matrix of an entangled system is connected on the state of knowledge of *who writes it*, the wave function is just a mathematical device which encapsulates what one knows about the system. This ambiguity is a clear symptom of the nowadays still poor comprehension of what quantum mechanics really means.
Also, we have mentioned in some previous chapters the so called measurement problem; one way it can be stated is the following [15]. These four claims are inconsistent:

1. Wave collapse is merely epistemic.
2. The wave function is complete (i.e. completely specifies the physical state of a system).
3. The wave function always evolves deterministically in accord with the (linear) Schrödinger equation.
4. Measurements always have definite outcomes (i.e. Schrödinger's cat ends up either dead or alive, once observed), picked out from the ensemble of the possible ones.

For instance, if 2,3 hold then there is no way to exclude that a system might be observed in a superposition of states, i.e. the cat does not end up either alive or dead, thus 4 is violated. So, for those interpretations in which the wave function is complete, some sort of wave collapse is needed. In general, any resolution of the measurement problem has to deny at least one of the above assumptions 1,2,3,4. Hence, we can see how the various interpretations of quantum mechanics are constructed. The many-worlds interpretation [16] retains both 2 and 3 and assume that the possible outcomes of a measurement do actually all occur, but in different universes. In the orthodox view, 2 and 4 hold but no description is given about how the wave collapse occurs and why the measurement process breaks the deterministic evolution by Schrödinger equation. A famous proposal for this non-Schrödinger evolution has been put forward by Ghirardi, Rimini and Weber [17]. Another possibility is to retain 3 and 4 but refuse 2: these are the so called no-collapse hidden-variable theories, which *must* be non-local (for Bell Theorem, Sec.2.2), such as Bohm theory (Appendix B).

Appendix B

Bohm - de Broglie theory

We present in this appendix the (non-relativistic) Bohm's interpretation of quantum mechanics, which is the the most famous example of non-local hidden variable theory. As we will see, the variable is not really "hidden", but well specified in the Bohmian framework.
The prototype of this interpretation was Louis de Broglie's "pilot-wave theory" (1927). Bohm developed it starting from 1952 [18]. One key point is the Hamiltonian foundation of wave mechanics, through which Schrödinger derived his famous equation. In Bohm theory, quantum mechanics is not complete; in fact, a complete description of a physical system requires the wave function $\psi(q)$ defined in the configuration space ($\mathbb{R}^{3N}$) of the N-particles system and also the actual configurations $\boldsymbol{Q}_i \in \mathbb{R}^3$, $i \in 1, ..., N$, or positions, of the particles (they are precisely the additional variables, not really hidden!). In Bohmian framework, both the wave function and the particles are real physical entities, the latter always having definite positions at any time, i.e. trajectories. Bohm theory is deterministic, in fact its mathematical heart consists of two equations; Schrödinger equation governs how the wave function changes with time:

$$i\hbar\frac{\partial\psi}{\partial t} = -\sum_{i=1}^{N}\frac{\hbar^2}{2m_i}\nabla_i^2\psi + V\psi \tag{B.1}$$

Here we consider spinless particles. Further, the *guidance equation* specifies the time evolution of the positions of particles, given the wave function:

$$\dot{\boldsymbol{Q}}_i(t) = \frac{\hbar}{m_i}\,\mathrm{Im}(\frac{\nabla_i\psi}{\psi})(Q(t)) = \frac{\nabla_i S(Q(t))}{m_i} \tag{B.2}$$

where $Q(t) \equiv (\boldsymbol{Q}_1(t), ..., \boldsymbol{Q}_N(t)) \in \mathbb{R}^{3N}$ and S is the phase of the wave function: $\psi = |\psi|\exp(\frac{i}{\hbar}S)$. Eq.B.2 can be obtained recalling the link between S, the quantum probability current $j = \frac{\hbar}{m}\,\mathrm{Im}(\psi^*\nabla\psi)$ and density $\rho = \psi^*\psi$ (in the following we consider for simplicity a single particle):

$$\frac{j}{\rho} = \frac{\nabla S}{m} \tag{B.3}$$

Hence, from Eq.B.2, the wave function determines at any time the velocity field, defined in the configuration space of the system. The particles are so "guided by" ψ along continuous trajectories in physical space. Further, the dynamics are local in *configuration space*, because the trajectory, so the velocity at a given point, depends on the value of the wave function around that point, at any given time t. But since each point of configuration space specifies the positions of all the particles of the system, and since the theory is deterministic (the wave function and the configuration of the system fix its time evolution uniquely), then the motion of a particle is linked to the positions of all the others at that time, i.e. an event in one (physical) place can have instantaneous influence on others, space-like separated, points. This is the manifest *non-local* trait of Bohm theory. We have seen, in Sec.2.2.3, that this is unavoidable for hidden-variable theories which provide for a completion of quantum mechanics, in order not to violate Bell's theorem.

How can Bohm theory reproduce the quantum probabilistic predictions if it is deterministic? Well, another hypothesis is necessary, namely the so called *quantum equilibrium*. That is, the initial configuration distribution of the particles in the system is $|\psi(0)|^2$. By the quantum continuity equation:

$$\frac{\partial|\psi|^2}{\partial t} + \nabla \cdot (|\psi|^2 \frac{\nabla S}{m}) = 0 \tag{B.4}$$

one has that, at any time t, the statistical distribution of the particles will satisfy $|\psi(t)|^2$. Therefore, Born rule for position distribution is recovered and, because that of position is one representation of quantum mechanics, all predictions of Bohm theory are then equivalent to standard's ones. So, in Bohmian interpretation we still have probabilistic predictions, but here they arise from *ignorance*, instead of being intrinsic as in orthodox theory (recall Sec.2.2.3).

The grand merit of Bohm theory is its solution to the measurement problem. A measure is just a physical interaction like many others, there is no need for a measurement postulate (as in orthodox interpretation, Sec.1.1): the outcome will only depend on the state of the system, i.e. the wave function and the configuration $(\psi(t), Q(t))$, at the time the measure is performed, without any wave collapse. Strictly speaking it occurs an "effective" collapse, which means that the parts of the wave function ψ in configuration regions distant to the one our system is in are (almost) irrelevant (recall the locality in configuration space!).

Bohmian mechanics is thus of first-order, in contrast to classical one. In his original work, Bohm introduced the so called "quantum potential", because of the formal analogy of the equation for the phase S:

$$-\frac{\partial S}{\partial t} = \frac{(\nabla S)^2}{2m} + V - \frac{\hbar^2}{2m}\frac{\nabla^2|\psi|}{|\psi|} \tag{B.5}$$

with the classical Hamilton-Jacobi equation ($\hbar \to 0$ and S becomes the action, i.e. Hamilton's principal function). The extra term on the right is the quantum potential; it was supposed to guide the particles along the trajectories, so that to apparently recover a Newtonian-like mechanics. But, while it may be useful to understand how classical mechanics should emerge from Bohmian one, quantum potential formulation is rather misleading. In fact, while in classical (Newtonian/Hamiltonian) mechanics positions *and* momenta are the fundamental variables from which, known the potentials (i.e. forces), one can derive trajectories, instead, in Bohmian quantum theory the fundamental variables are, as we have seen in Eq.B.1 and Eq.B.2, wave function and particles' positions. In conclusion, we remark that in Bohm theory the wave function and the particles are both *real* physical entities. So, for example, in the double-slit experiment, the particle only passes through one slit, whereas the wave function goes through both of them; the two parts of the wave function (one for each slit) overlap in configuration space and influence the trajectory of the particle until the spot on the shield. If many particles are involved, then the interference pattern is consequently explained.

Bibliography

[1] E. Schrödinger. "Die gegenwärtige Situation in der Quantenmechanik". *Naturwissenschaften* **23**, 807–812 (1935).

[2] A. Einstein, B. Podolsky, and N. Rosen. "Can Quantum-Mechanical Description of Physical Reality Be Considered Complete?". *Phys. Rev.* **47**, 777 (1935). `https://cds.cern.ch/record/405662/files/PhysRev.47.777.pdf`

[3] M. Born. "The Born-Einstein Letters". Walker, New York, (ed.) 1971, pp.170-1.

[4] J. S. Bell. "On the Einstein-Podolsky-Rosen Paradox". *Physics* **1** (3), 195-200 (1964). `https://cds.cern.ch/record/111654/files/vol1p195-200_001.pdf`

[5] R. P. Feynman. "Simulating physics with computers". *Int. J. Theor. Phys.* **21**, 467–488 (1982).

[6] O. M. P. Bilaniuk, V. K. Deshpande, E. C. G. Sudarshan. "Meta-Relativity". *Amer. Journ. Phys.*, **30**, 718-723 (1962).

[7] G. Feinberg. "Possibility of Faster-than-Light Particles". *Phys. Rev.* **159**, 1089-1105 (1967).

[8] R. C. Tolman. "The Theory of the Relativity of Motion". University of California Press, Berkeley, 1917, p. 54.

[9] M. Fayngold. "A Possible Resolution of the Tolman Paradox as a Quantum Superposition". arXiv:1104.2531v1 [physics.gen-ph] (2011) `https://arxiv.org/ftp/arxiv/papers/1104/1104.2531.pdf`

[10] P. A. M. Dirac. "The Principles of Quantum Mechanics". Oxford University Press, Oxford, 1958.

[11] D. Bohm and Y. Aharonov. "Discussion of Experimental Proof for the Paradox of Einstein, Rosen, and Podolsky". *Phys. Rev.* **108**, 1070 (1957).

[12] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt. "Proposed Experiment to Test Local Hidden-Variable Theories". *Phys. Rev. Lett.* **23**, 880 (1969).

[13] D. M. Greenberger, M. A. Horne, A. Zeilinger. "Going beyond Bell's Theorem". arXiv:0712.0921 (2007).

[14] A. Einstein. "Zur Elektrodynamik bewegter Körper". *Annalen der Physik* **17**, 891-921 (1905).
`https://www.fourmilab.ch/etexts/einstein/specrel/specrel.pdf`

[15] T. Maudlin. "Three measurement problems". *Topoi* **14**, 7–15 (1995).

[16] H. Everett, J. A. Wheeler, B. S. DeWitt, N. Graham. "The Many-Worlds Interpretation of Quantum Mechanics." Princeton Series in Physics. Princeton University Press, Princeton, 1973.

[17] G. C. Ghirardi, A. Rimini, and T. Weber. "Unified dynamics for microscopic and macroscopic systems". *Phys. Rev. D* **34**, 470 (1986).

[18] D. Bohm. "A Suggested Interpretation of the Quantum Theory in Terms of ≪Hidden≫ Variables". *Phys. Rev.* **85**, 166,180 (1952).