No-faster-than-light-signaling implies linear evolutions. A re-derivation

Angelo Bassi[*]

Department of Physics, University of Trieste, Strada Costiera 11, 34151 Trieste, Italy

Istituto Nazionale di Fisica Nucleare, Trieste Section, Via Valerio 2, 34127 Trieste, Italy

Kasra Hejazi[†]

Department of Physics, Sharif University of Technology, P.O. Box 11155-9161, Tehran, Iran

Abstract

There is a growing interest, both from the theoretical as well as experimental side, to test the validity of the quantum superposition principle, and of theories which explicitly violate it by adding nonlinear terms to the Schrödinger equation. We review the original argument elaborated by N. Gisin [1], which shows that the non-superluminal-signaling condition implies that the dynamics of the density matrix must be linear. This places very strong constraints on the permissible modifications of the Schrödinger equation, since they have to give rise, at the statistical level, to a linear evolution for the density matrix. The derivation is done in a heuristic way here and is appropriate for the students familiar with the Textbook Quantum Mechanics and the language of density matrices.

[*]bassi@ts.infn.it
[†]hejazi_kasra@physics.sharif.edu

1

1   Introduction

Quantum Mechanics is a very successful physical theory. It has been revolutionary in many diverse areas, from particle physics to condensed matter and statistical physics, explaining many features of nature in a unified framework.

Quantum Mechanics has introduced many new concepts to physics. Historically it came into play when Max Planck first assumed that the energy of the standing waves in a cavity can only take discrete values. Then he could explain the frequency distribution of the black-body-radiation. As in this example, one of the very first applications of the theory was to make quantities, especially energy, discrete; from which the theory gets its name.

Many other concepts have been introduced by Quantum Theory. As a second example consider the idea of identical particles and indistinguishability; this is crucial in many ways to particle statistics in Statistical Physics. For instance, one cannot track down one electron, among many, and talk about its definite state, an idea which was almost impossible in non-Quantum way of thinking. Just to name some other examples from the large list of new concepts due to Quantum Theory, consider basic and at the same time important notions like spin, tunneling and the peculiar non-locality.

Besides new concepts, Quantum Mechanics has great power in predicting quantitative results; think of the Anomalous Magnetic Moment of the electron. It can be derived using the formalism of the next revolution, Quantum Field Theory, which was invented because of the need for a Relativistic Quantum Theory. The agreement between the theory and experiment is astonishing: more than ten decimal places.

Nevertheless, over the years several people repeatedly suggested to modify the Schrödinger equation, which is the fundamental dynamical law for quantum systems. One can identify three main reasons to do so.

The first and most widespread reason is the so-called measurement process. Quantum mechanics predicts that there are physical states corresponding to the superposition of two other states. Such superpositions are evident experimentally in the microscopic world, but are absent in our everyday experience of the macro-world. One does not see anything in a superposition of eigenstates of position like: $[|I \text{ am in Tehran}\rangle + |I \text{ am in Trieste}\rangle]/\sqrt{2}$. This is especially important when dealing with macroscopic systems interacting with microscopic ones, i.e. in measurements: in such cases the dynamics as given by the Schrödinger equation immediately predicts the formation of macroscopic superpositions corresponding to two or more possible outcomes of the measurement.

To avoid superposition in macroscopic physics, the standard approach is to

assume that a collapse of the state vectors takes place [2–4]. As a result one has a theory which deals with the physical world in two different ways. At the very fundamental microscopic world everything evolves according to the Schrödinger equation, but at the stage of macroscopic interactions with the micro-world, the wave function collapses. Even the border between these two kinds of evolutions is ambiguously defined. Moreover, the interaction between the measurement apparatus and the system can be tracked down to the fundamental interactions between the particles of the apparatus and the system, which are described by the linear Schrödinger equation. But it is not possible to derive the collapse from such a scheme. [5, 6]

There are several ways one can get out of this trouble. One example is *Bohmian Mechanics* [7–14], which essentially does not contain collapse. Despite having a wave function, particles have definite locations in this theory. A measurement is simply an interaction between the measurement apparatus and the system, without any mystery. Another way out, which is a very natural thing to do, is to modify Schrödinger equation in order to have one dynamical evolution both for the micro- and the macro-world. The new equation should, approximately, correspond to the old Schrödinger evolution for microscopic systems, and should yield the collapse of the wave function in the macroscopic domain. Collapse Models are examples of these kinds of modifications [5, 15–20]. The dynamics is modified by adding non-linear and stochastic terms to the usual Schrödinger equation.

A second motivation for modifying the standard quantum dynamics comes from Quantum Field Theory. This is a linear theory (again, the superposition principle) and it is quite natural to think that this is only the first order approximation of a deeper level non-linear theory. The paradigmatic example is Newtonian gravity, which now we understand as the weak field limit of General Relativity. S. Weinberg first suggested such a possibility [21].

As a third motivation, there is the longstanding problem of the unification of quantum and gravitational phenomena. While waiting for a fully consistent and successful unified theory, one can write phenomenological equations. One such equations is the so-called Schrödinger-Newton equation, which is non-linear [22–24]. Different parts of the wave function of a system interact among themselves through the Newtonian potential.

All these efforts raise an important question: 'Are there limits in modifying the Schrödinger equation?' The answer is yes. As Gisin convincingly showed [1], *whichever way one modifies the Schrödinger equation for the wave function, the time evolution for the density matrix has to be linear if one wants to keep no-*
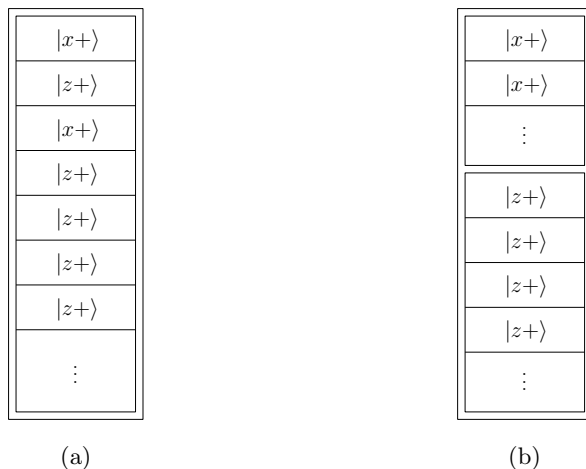
Figure 1: (a) A statistical ensemble, consisting of the states $|z+\rangle$ and $|x+\rangle = \frac{1}{\sqrt{2}}(|z+\rangle + |z-\rangle)$, with probabilities 2/3 and 1/3 respectively. Note that the population of the states in the ensemble is proportional to the probabilities. (b) The same ensemble re-ordered. The ensemble now consists of two pure sub-ensembles. The second sub-ensemble is populated twice as the first sub-ensemble, and so, the probability of the state $|z+\rangle$ is twice as $|x+\rangle$.

faster-than-light-signaling. Here we review this argument with the purpose of giving a pedagogical introduction to the subject.

2   Evolution of the ensembles

We work in a standard quantum framework. The *states* of any given physical system are described by the elements $|\psi\rangle$ of a Hilbert space $\mathscr{H}$. We want also to give a statistical description of the system, for which it is convenient to use an *ensemble (or mixture) of states.* One can identify any such ensemble with a set of state vectors and the corresponding probabilities, i.e. $\{d_i; |\psi_i\rangle\}$. The idea is that the system is in any of such states $|\psi_i\rangle$, but we do not know which one (classical ignorance). We only know the probabilities d_i for the system to be in the states $|\psi_i\rangle$. (figure 1)

The density matrix formalism is the appropriate tool for dealing with statistical mixtures. For a given mixture $\{d_i; |\psi_i\rangle\}$, the *density matrix* is defined as $\rho \equiv \sum_i d_i |\psi_i\rangle\langle\psi_i|$. A density matrix is *pure* if there is only one element in the ensemble (with probability 1), and is *mixed* otherwise.

We name the set of all density matrices as $\mathfrak{B}^+$, and the set of pure ones as $\mathfrak{B}^p$.

Definition. Two mixtures $\{d_i; |\psi_i\rangle\}$ and $\{p_i; |\phi_i\rangle\}$ are equivalent if their cor-

responding density matrices are the same:

$$\sum_i d_i|\psi_i\rangle\langle\psi_i| = \sum_i p_i|\phi_i\rangle\langle\phi_i|. \tag{1}$$

One can easily see that the equivalence of the mixtures is an equivalence relation.

Let us consider a given dynamical evolution for the state vectors of a given system, which in principle has nothing to do with the Schrödinger equation; in particular, it might be non-linear[1]:

$$|\psi_t\rangle = \mathcal{T}_{(t,t_0)}(|\psi_{t_0}\rangle). \tag{2}$$

This dynamical law automatically (and trivially) defines a dynamics also for *pure* density matrices:

$$\rho_{t_0} = |\psi_{t_0}\rangle\langle\psi_{t_0}| \rightarrow \rho_t = |\psi_t\rangle\langle\psi_t|, \tag{3}$$

where $|\psi_t\rangle$ is given by equation (2). So we have the following map defined on the space $\mathfrak{B}^p$ of pure density matrices from time t_0 to time t:

$$\mathcal{E}_{(t,t_0)} : \mathfrak{B}^p \rightarrow \mathfrak{B}^p \tag{4}$$
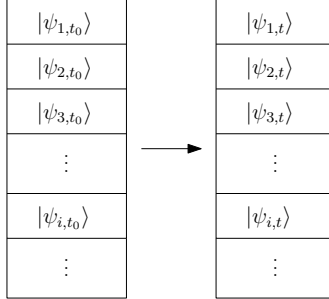
$$\rho_t = \mathcal{E}_{(t,t_0)}(\rho_{t_0}).$$

Figure 2: The evolution of an ensemble; each single state vector evolves independently from the others. Therefore, each pure sub-ensemble evolves according to (4). Note that the statistical weights, being proportional to the number of the copies of each state vector, remain constant.

With the dynamics of pure ensembles known, there remains the question of how *mixed* ensembles evolve. One expects each of the states in the mixed ensemble to evolve according to evolution (2), and independent of the other states.

---

[1]A generic non-linear operator acting on $|\psi\rangle$ will be shown as $O(|\psi\rangle)$, while for a linear operator we will use the notation $O|\psi\rangle$. Note that this will not be done for operators acting on density matrices, these operators in this paper are either linear or just acting on pure states. Moreover, we assume that the evolution of any initial state depends only on the state itself and not on the particular way the state is prepared. [30]
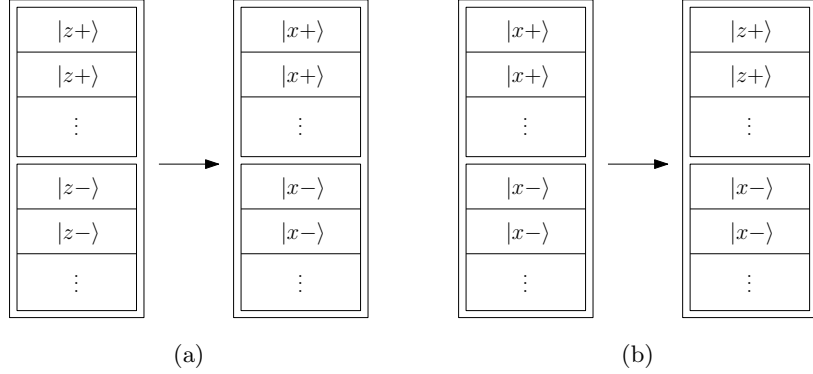
Figure 3: (a) An ensemble of states $|z+\rangle$ and $|z-\rangle$, with 1/2 probabilities. The density matrix for this ensemble is $\rho^a_{t_0} = \frac{1}{2}|z+\rangle\langle z+| + \frac{1}{2}|z-\rangle\langle z-| = \frac{1}{2}I$. Each state evolves according to evolution (2). We assume that the fate of $|z+\rangle$ is $|x+\rangle$, and that the fate of $|z-\rangle$ is $|x-\rangle = \frac{1}{\sqrt{2}}(|z+\rangle + |z-\rangle)$. Hence, the final mixture consists of $|x+\rangle$ and $|x-\rangle$ with probability 1/2 for each. The final density matrix is $\rho^a_t = \frac{1}{2}|x+\rangle\langle x+| + \frac{1}{2}|x-\rangle\langle x-| = \frac{1}{2}I$.
(b) An equivalent mixture, consisting of $|x+\rangle$ and $|x-\rangle$ with equal probabilities (1/2). The initial density matrix is $\rho^b_{t_0} = \frac{1}{2}|x+\rangle\langle x+| + \frac{1}{2}|x-\rangle\langle x-| = \frac{1}{2}I$. Let us assume that the states evolve as: $|x+\rangle \to |z+\rangle$ and $|x-\rangle \to |x-\rangle$. The latter happens due to a non-linearity in the presumed evolution. The final density matrix will be $\rho^b_t = \frac{1}{2}|z+\rangle\langle z+| + \frac{1}{2}|x-\rangle\langle x-| \neq \frac{1}{2}I$, which is clearly different from ρ^a_t. This is an example of two initially equivalent ensembles, becoming inequivalent, due to the non-linear evolution of their state vectors.

(look at figure 2) In other words, different pure sub-ensembles of an ensemble do not interact, i.e. each one evolves independently according to equation (4). Therefore a generic density matrix evolves in the following way:

$$\sum_i d_i |\psi_{i,t_0}\rangle\langle\psi_{i,t_0}| \to \sum_i d_i |\psi_{i,t}\rangle\langle\psi_{i,t}| \tag{5}$$
$$= \sum_i d_i \left[\mathscr{E}_{(t,t_0)}(|\psi_{i,t_0}\rangle\langle\psi_{i,t_0}|) \right],$$

This suggests the possibility of extending the map $\mathscr{E}_{(t,t_0)}$ from the space of pure density matrices to the space of all density matrices:

$$\mathscr{E}_{(t,t_0)} : \mathfrak{B}^+ \to \mathfrak{B}^+ \tag{6}$$
$$\rho_t = \mathscr{E}_{(t,t_0)}(\rho_{t_0}).$$

However, in general this is not possible, because two different decompositions of the same initial density matrix ρ_{t_0} may correspond to different density matrices after evolution. (figure 3) In other words one density matrix can have different destinies according to the ensemble it is representing. This means that the map in (6) is not well defined in general.

However, Gisin's theorem states that if one wants to rule out the possibility for superluminal communication, different mixtures which are equivalent have

to remain equivalent even after the evolution. So that the map $\mathscr{E}_{(t,t_0)}$ in (4) can be really extended to the whole set of density matrices as in (6). The theorem will be proved in the next section.

One can also think of stochastic evolutions of the state vectors, which is the case in Collapse Models, and see what constraints can be put on this type of the evolution. Gisin's theorem immediately generalizes also to such evolutions. (see Appendix A for details.)

3   The theorem

We will be working with finite-dimensional Hilbert spaces. That is reasonable, because in case of infinite-dimensional spaces one can always work with a finite-dimensional subspace. The finite-dimensional Hilbert space describing the state vectors of the system we will be working with, is $\mathscr{H}$.

We will use un-normalized vectors, which will make our work easier and our results simpler to read. So we will change notation. The *square norm of a vector* from now on will show its probability in the ensemble, so that instead of the mixture $\{d_i; |\psi_i\rangle\}$ we will be using $\{|\psi_i\rangle\}$, and:

$$\langle \psi_i | \psi_i \rangle = d_i.$$

We will be dealing with the time evolution of two arbitrary but equivalent mixtures $\{|\psi_i\rangle\}$ and $\{|\phi_j\rangle\}$ throughout, where $i = 1, \ldots, n_\psi$; $j = 1, \ldots, n_\phi$. Because the two ensembles are equivalent we know that one density matrix ρ describes both, i.e. $\rho = \sum_i |\psi_i\rangle\langle\psi_i| = \sum_i |\phi_i\rangle\langle\phi_i|$. Now we have all the ingredients to formulate Gisin's argument.

As in a usual Bell-type experiment, there are Alice and Bob, far apart from each other, who make measurements on entangled particles[2]. We use the Hilbert space $\mathscr{H}$ to describe the states of the particles traveling towards Alice's place and use the Hilbert space $\mathscr{K}$ for the ones traveling towards Bob's region. Suppose $\mathscr{K}$ is large enough for the argument to apply. It can be shown that the following state $|V\rangle$ can be shared between the two parties:

$$|V\rangle = \sum_i |\psi_i\rangle \otimes |\alpha_i\rangle = \sum_i |\phi_i\rangle \otimes |\beta_i\rangle, \tag{7}$$

where $\{|\psi_i\rangle\}$ and $\{|\phi_j\rangle\}$ are the two equivalent mixtures in $\mathscr{H}$ that, we started with, upon which Alice will make measurements, and $\{|\alpha_i\rangle\}$ and $\{|\beta_i\rangle\}$ are two

---

[2]For simplicity we will speak of particles, but the two types of systems on which Alice and Bob make measurements can be generic.

different orthonormal bases of $\mathcal{K}$ (Bob's system). The preparation of such a state is possible for any arbitrary choice of the equivalent mixtures $\{|\psi_i\rangle\}$ and $\{|\phi_j\rangle\}$. (see Appendix B for a proof and details.)
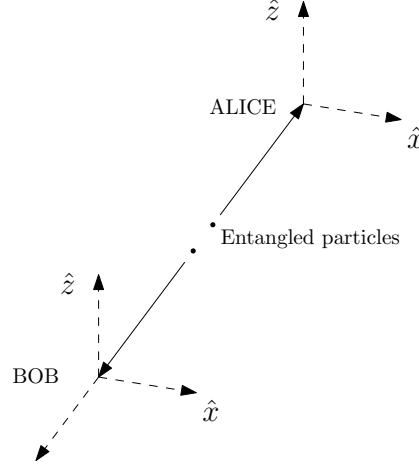

Figure 4: An ensemble of entangled states $|V\rangle$ is shared between Alice and Bob. Take the state $|V\rangle = \frac{1}{\sqrt{2}}(|x+\rangle \otimes |z+\rangle + |x-\rangle \otimes |z-\rangle)$ as an example. Only the spin components of the entangled particles have been shown in the state. Suppose Bob decides to measure S_z on his particles. Measuring the z-component spin of each particle by Bob, makes the state of its entangled particle in Alice's system definite, due to a collapse: either $|x+\rangle$ if Bob gets $+1$ or $|x-\rangle$ if Bob gets -1 in his measurement. So a mixture of states $|x+\rangle$ and $|x-\rangle$ with equal probabilities is prepared for Alice.

What if Bob decides to measure S_x? One can easily see that the state $|V\rangle$ can be represented as $|V\rangle = \frac{1}{\sqrt{2}}(|z+\rangle \otimes |x+\rangle + |z-\rangle \otimes |x-\rangle)$ also. Therefore, this time a mixture of $|z+\rangle$ and $|z-\rangle$ with equal probabilities will be prepared. Note that these two kinds of mixtures are equivalent, i.e. : $\frac{1}{2}\{|x+\rangle\langle x+| + |x-\rangle\langle x-|\} = \frac{1}{2}\{|z+\rangle\langle z+| + |z-\rangle\langle z-|\}$.

Suppose an ensemble of states $|V\rangle$ is shared between Alice and Bob. Bob has two different choices (among the many); either to measure the observable O_α which has the $|\alpha_i\rangle$'s as eigenvectors, or to measure the observable O_β which has the $|\beta_i\rangle$'s as eigenvectors. The first type of measurement will prepare the ensemble of states $\{|\psi_i\rangle\}$ and the second will prepare the ensemble $\{|\phi_i\rangle\}$ for Alice, as a result of the collapse of the state vector. They correspond to different mixtures, but the density matrix and so all of the expectation values Alice can measure are the same. The two ensembles are equivalent. (figure 4) Notice that Bob can prepare any equivalent ensemble for Alice, by measuring a properly chosen observable, without any need to change the previously shared state respectively. (look at Appendix B) Therefore, this first part of the argument shows that *different but equivalent mixtures can be prepared at a distance, by using suitable entangled states and measurements.*

8

Now the crucial point comes. Suppose these two mixtures $\{|\psi_i\rangle\}$ and $\{|\phi_j\rangle\}$, after some time, become inequivalent, because of the dynamics given by Eq. (5). Then Alice can find out the difference between them by making appropriate measurements and computing averages. Therefore she is in the position to understand which observable Bob decided to measure, even if Bob is arbitrarily far apart. There is the possibility for superluminal communication, no matter how long it takes for the two mixtures to become appreciably different from each other. Therefore, if we hold on the assumption that there cannot be faster-than-light signaling, *different but equivalent mixtures have to stay equivalent while time passes.*

The above argument implies that the evolution map $\mathscr{E}_{(t,t_0)}$ defined in (4) can be extended to the whole space of density matrices as in (6), using the rule in (5). *Linearity* of the map is a consequence of using statistical nature of ensembles, i.e. pure sub-ensembles evolve independently. Suppose ρ is a convex sum of ρ_1 and ρ_2:

$$\rho = \lambda\rho_1 + (1-\lambda)\rho_2.$$

If a possible ensemble for ρ_1 is $\{|a_i\rangle\}$ and one for ρ_2 is $\{|b_i\rangle\}$,[3] then one possible ensemble for ρ will be $\{\sqrt{\lambda}|a_i\rangle, \sqrt{1-\lambda}|b_i\rangle\}$. The time evolution for ρ is given by (assuming the extended map (6)):

$$
\begin{aligned}
\mathscr{E}_{(t,t_0)}\rho &= \mathscr{E}_{(t,t_0)}\Big(\sum_i \lambda|a_i\rangle\langle a_i| + \sum_i (1-\lambda)|b_i\rangle\langle b_i|\Big) & (8)\\
&= \sum_i \lambda\mathscr{E}_{(t,t_0)}\Big(|a_i\rangle\langle a_i|\Big) + \sum_i (1-\lambda)\mathscr{E}_{(t,t_0)}\Big(|b_i\rangle\langle b_i|\Big) & (9)\\
&= \lambda\mathscr{E}_{(t,t_0)}\Big(\sum_i |a_i\rangle\langle a_i|\Big) + (1-\lambda)\mathscr{E}_{(t,t_0)}\Big(\sum_i |b_i\rangle\langle b_i|\Big) & (10)\\
&= \lambda\mathscr{E}_{(t,t_0)}\rho_1 + (1-\lambda)\mathscr{E}_{(t,t_0)}\rho_2. & (11)
\end{aligned}
$$

In the fist line, we simply re-wrote ρ in terms of the statistical mixture; in going from the first to the second line, we used (5) applied to ρ; in going from the second to the third line, we used again (5), this time applied to the two mixtures $\sum_i |a_i\rangle\langle a_i|$ and $\sum_i |b_i\rangle\langle b_i|$; in going from the third to the forth line, we use the definition of ρ_1 and ρ_2 in terms of the mixtures defining them.

Now that we know that the evolution of the density matrix is linear, we can make yet further progress by understanding how the state vectors evolve. As is discussed thoroughly in [25], the only deterministic evolutions of the state vectors, which induce linear evolutions at the level of density matrices, are the linear unitary evolutions of the state vectors, i.e. the old Schrödinger evolutions

---

[3]Remember that the norm of the vectors shows their weight in the ensemble.

of the state vectors. As [25] further states, in order to obtain a non-Schrödinger evolution at the level of state vectors, consistent with no-Signaling, one has to let the state vectors evolve stochastically. We will talk more about this in the Conclusion.

Conclusion

As stated in the introduction there are attempts for modifying the time evolution in Quantum Mechanics in order to solve the measurement problem. Gisin's argument restricts the possibilities of modifying the evolution equations up to the border of linearity: *The evolution for the density matrix has to be linear*, no matter how the state vector's evolution is, e.g. deterministic, stochastic, etc.

According to the work of G. Lindblad [26] and of V. Gorini, A. Kossakowsky and E.C.G. Sudarshan [27], a linear evolution for the density matrix can be only of the Lindblad type, if one adds the two further requirements of a quantum-dynamical-semigroup type of equation, and complete positivity:

$$\frac{\mathrm{d}\rho_t}{\mathrm{d}t} = -i[H, \rho_t] + \sum_{k=1}^{n} \left(L_k \rho_t L_k^\dagger - \frac{1}{2} L_k^\dagger L_k \rho_t - \frac{1}{2} \rho_t L_k^\dagger L_k \right). \qquad (12)$$

The first assumption amounts to requiring a Markovian evolution, which is a reasonable starting point for writing down a dynamical equation (all fundamental equations in physics, thus far, are Markovian; non-Markovian processes become relevant at the statistical level, when phenomenological equations are the only reasonable way for analyzing a complex system). The second assumption becomes necessary when entangled systems are taken into account.

Then, according to [28] the only possible way to modify the Schrödinger equation for the wave function, which leads to a Lindblad type of equation for the density matrix, is the way collapse models do it. Therefore, given the above premises, one can conclude that *the one given by collapse models is the only possible way to modify the Schrödinger equation, if one requires no superluminal communication.*

Acknowledgement

A

In this appendix, the possibly stochastic evolutions of the state vectors are discussed.

In the case of stochastic evolution of a state vector, initial pure states evolve into mixed mixtures. Think of two identical state vectors present in the ensemble, they will in general evolve into different states. One cannot tell what the future of each state is, but can talk about the statistics of the future of identical copies; something which can be determined based on the type of the stochastic evolution involved. (look at figure 5) One expects a pure density matrix ρ_{t_0} to evolve as:

$$\rho_{t_0} = |\psi_{t_0}\rangle\langle\psi_{t_0}| \to \rho_t = \mathbb{E}\left[|\psi_t\rangle\langle\psi_t|\right],$$

where $\mathbb{E}[\cdot]$ denotes the stochastic average and $|\psi_t\rangle$ has a degree of randomness, despite the fact that $|\psi_{t_0}\rangle$ is known. Accordingly, the map in Eq. (4) changes into:

$$\mathscr{E}_{(t,t_0)} : \mathfrak{B}^p \to \mathfrak{B}^+ \tag{13}$$
$$\rho_t = \mathscr{E}_{(t,t_0)}(\rho_{t_0}).$$

Now each pure sub-ensemble in a mixture evolves as (13). Gisin's argument follows.
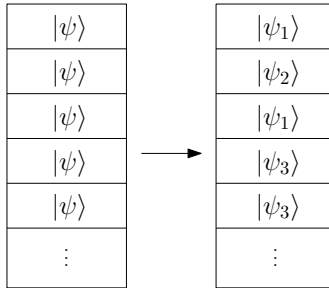

Figure 5: A pure ensemble getting mixed due to the stochastic evolution. The relative population, again shows the statistical weight in the final ensemble.

B

In this appendix it is shown that a state like the one in (7) can be shared between the two parties:

$$|V\rangle = \sum_i |\psi_i\rangle \otimes |\alpha_i\rangle = \sum_i |\phi_i\rangle \otimes |\beta_i\rangle,$$

where $\{|\psi_i\rangle\}$ and $\{|\phi_j\rangle\}$ are two given equivalent mixtures ($i = 1, \ldots, n_\psi$; $j = 1, \ldots, n_\phi$.), and there is no further restriction on them. Let us start at a basic level and prove this via some lemmas. The lemmas are presented in a similar fashion to the results of [29].

We will work in a finite-dimensional space. Any density matrix has a spectral decomposition, i.e. :

$$\rho = \sum_i |\chi_i\rangle\langle\chi_i|, \tag{14}$$

with $|\chi_i\rangle$'s being orthogonal:

$$\langle\chi_i|\chi_j\rangle = \lambda_i\delta_{ij}, \qquad 0 < \lambda_i \leq 1.$$

Lemma 1. *Suppose ρ is the density matrix associated to the ensemble $\{|\psi_i\rangle\}$, then any single vector $|\psi_i\rangle$ can be written as a linear combination of the vectors in the spectral decomposition* (14) *of ρ.*

Proof. Take $\{|\chi_1\rangle, \ldots, |\chi_n\rangle\}$ as the spectral decomposition of ρ. If they are a basis, the theorem is trivial. Suppose they are not. Expand the set of $\{|\chi_1\rangle, \ldots, |\chi_n\rangle\}$ to an orthogonal basis of the whole Hilbert space by adding $\{|\chi_{n+1}\rangle, \ldots, |\chi_{n'}\rangle\}; n < n'$ to it. Note that this can always be done.

Any vector $|\psi_i\rangle$ can be written as a linear combination of these basis-vectors:

$$|\psi_i\rangle = \sum_{j=1}^{n'} m_{ij}|\chi_j\rangle, \tag{15}$$

where m_{ij}'s are complex numbers. Take j such that $n < j \leq n'$. By the orthogonality of $|\chi_i\rangle$'s:

$$\langle\chi_j|\rho|\chi_j\rangle = \sum_{i=1}^{n} \langle\chi_j|\chi_i\rangle\langle\chi_i|\chi_j\rangle = 0, \tag{16}$$

but also:

$$\langle \chi_j | \rho | \chi_j \rangle = 0 = \sum_{i=1}^{n_\psi} \langle \chi_j | \psi_i \rangle \langle \psi_i | \chi_j \rangle \tag{17}$$

$$= \sum_{i=1}^{n_\psi} m_{ij} m_{ij}^*$$

$$= \sum_{i=1}^{n_\psi} |m_{ij}|^2.$$

A sum of non-negative numbers is equal to zero, i.e. they are all zero. So for every i:

$$m_{ij} = 0; \qquad n < j \leq n'. \tag{18}$$

Hence, one can rewrite equation (15) as:

$$|\psi_i\rangle = \sum_{j=1}^{n} m_{ij} |\chi_j\rangle. \tag{19}$$

$\square$

Lemma 2. *Suppose that the density matrix ρ has the spectral decomposition $\{|\chi_1\rangle, \ldots, |\chi_n\rangle\}$. A set of vectors $\{|\psi_i\rangle\}$ has the same density matrix if and only if*

$$|\psi_i\rangle = \sum_{j=1}^{n} m_{ij} |\chi_j\rangle, \tag{20}$$

with m_{ij} satisfying:

$$\sum_{i=1}^{n_\psi} m_{ij} m_{ik}^* = \delta_{jk}. \tag{21}$$

Proof. First assume that density matrix of the mixture is ρ, so:

$$\sum_{i=1}^{n_\psi} |\psi_i\rangle \langle \psi_i| = \sum_{l=1}^{n} |\chi_l\rangle \langle \chi_l|. \tag{22}$$

Equation (19) takes the form:

$$|\psi_i\rangle = \sum_{j=1}^{n} m_{ij} |\chi_j\rangle,$$

multiplying by $\langle \chi_k|$ we have:

$$\langle \chi_k | \psi_i \rangle = m_{ik} \lambda_k. \tag{23}$$

Then multiplying equation (22) by $\langle\chi_j|$ from left and by $|\chi_k\rangle$ from right we have:

$$\sum_i \langle\chi_j|\psi_i\rangle\langle\psi_i|\chi_k\rangle = \sum_l \lambda_j\delta_{jl}\lambda_k\delta_{kl} \qquad (24)$$
$$= \lambda_k^2\delta_{jk}.$$

And by using (23) one immediately gets:

$$\sum_i m_{ij}m_{ik}^* = \delta_{jk}.$$

Conversely suppose (20) and (21) hold, so we have:

$$\sum_i |\psi_i\rangle\langle\psi_i| = \sum_{i,j,k} m_{ij}m_{ik}^*|\chi_k\rangle\langle\chi_j|$$
$$= \sum_j |\chi_j\rangle\langle\chi_j|$$
$$= \rho.$$

$\square$

Consider the columns of $\boldsymbol{m}$ (the matrix having m_{ij} as components, i showing the row and j the column). They are n vectors in $\mathbb{C}^{n_\psi}$. According to (21) they are orthonormal and so linearly independent. So $n \leq n_\psi$ and also the number of columns of $\boldsymbol{m}$ is less than or equal to the number of its rows.

We will need a square matrix instead of the possibly rectangular one above. For that, one can add $n_\psi - n$ null vectors to the set of $\{|\chi_1\rangle, \ldots, |\chi_n\rangle\}$ to make them $\{|\chi_1\rangle, \ldots, |\chi_n\rangle, |0\rangle, \ldots, |0\rangle\}$ and add columns to $\boldsymbol{m}$ to make them an orthonormal basis of $\mathbb{C}^{n_\psi}$ (the columns were already n orthonormal vectors, look at (21)). So the expanded $\boldsymbol{m}$ will be an $n_\psi \times n_\psi$ square matrix and we will name it $\boldsymbol{M}$. It is easy to see that $\boldsymbol{M}$ is unitary (21). In the matrix notation we will have:

$$\begin{pmatrix} |\psi_1\rangle \\ |\psi_2\rangle \\ \vdots \\ |\psi_{n_\psi}\rangle \end{pmatrix} = \boldsymbol{M} \begin{pmatrix} |\chi_1\rangle \\ \vdots \\ |\chi_n\rangle \\ |0\rangle \\ \vdots \\ |0\rangle \end{pmatrix}. \qquad (25)$$

Note that $\boldsymbol{M}$ acts only on the arrays and is not a Hilbert space operator.

$\boldsymbol{M}$ can be expanded even more. In fact, one can make a square matrix of any size satisfying (21) in this way by adding enough zeros to $|\psi_i\rangle$'s and $|\chi_i\rangle$'s

and adding components to M without changing the old ones, and making sure it remains unitary.

Lemma 3. *Consider two sets of vecotrs $\{|\psi_i\rangle\}$ and $\{|\phi_i\rangle\}$ all belonging to the Hilbert space $\mathscr{H}$ and assume that they correspond to the same density matrix, i.e. $\rho = \sum_i |\psi_i\rangle\langle\psi_i| = \sum_i |\phi_i\rangle\langle\phi_i|$. There will be a Hilbert space $\mathscr{K}$ which has the following property: There is a vector in the tensor product space $\mathscr{H} \otimes \mathscr{K}$ which can be written in this way:*

$$|V\rangle = \sum_i |\psi_i\rangle \otimes |\alpha_i\rangle = \sum_i |\phi_i\rangle \otimes |\beta_i\rangle,$$

where $|\alpha_i\rangle$'s are orthonormal and $|\beta_i\rangle$'s are orthonormal.

Proof. Let $\{|\chi_1\rangle, \ldots, |\chi_n\rangle\}$ be the spectral decomposition of the density matrix ρ. Assume $n_\psi \leq n_\phi$ without loss of generality. Now as stated before, one can add zeros, if necessary, to the sets of vectors to make the number of members of every set n_ϕ. So from (25) we have:

$$|\psi_i\rangle = \sum_{j=1}^{n_\phi} M_{ij}|\chi_j\rangle,$$

M is unitary, so:

$$|\chi_i\rangle = \sum_{j=1}^{n_\phi} M_{ij}^\dagger |\psi_j\rangle. \tag{26}$$

Let $\mathscr{K}$ be $\mathbb{C}^{n_\phi}$, and let $\{|a_i\rangle\}$ be any orthonormal basis for $\mathscr{K}$. Consider the following vector:

$$
\begin{aligned}
|V\rangle &= \sum_i |\chi_i\rangle \otimes |a_i\rangle \tag{27} \\
&= \sum_{i,j} M_{ij}^\dagger |\psi_j\rangle \otimes |a_i\rangle \\
&= \sum_j |\psi_j\rangle \otimes |\alpha_j\rangle,
\end{aligned}
$$

where we defined $|\alpha_j\rangle = \sum_i M_{ij}^\dagger |a_i\rangle$. Let us see if $|\alpha_i\rangle$'s are orthonormal:

$$\langle\alpha_i|\alpha_j\rangle = \sum_k (M_{ik}^\dagger)^* M_{jk}^\dagger = \sum_k M_{ki} M_{kj}^* = \delta_{ij}.$$

We can do the same thing for $|\phi_i\rangle$'s:

$$|V\rangle = \sum_i |\chi_i\rangle \otimes |a_i\rangle = \sum_j |\phi_j\rangle \otimes |\beta_j\rangle, \tag{28}$$

with

$$\langle \beta_i | \beta_j \rangle = \delta_{ij}.$$

So from (27) and (28):

$$|V\rangle = \sum_i |\psi_i\rangle \otimes |\alpha_i\rangle = \sum_i |\phi_i\rangle \otimes |\beta_i\rangle. \qquad (29)$$

$\square$

This concludes the mathematical background necessary to prove Gisin's theorem.

References

[1] N. Gisin, Helv. Phys. Acta **62**, 363 (1989).

[2] S. Weinberg, *Lectures on Quantum Mechanics*, Cambridge University Press (2012).

[3] J. J. Sakurai and Jim J. Napolitano, *Modern Quantum Mechanics*, Addison-Wesley (2010).

[4] R. Shankar, *Principles of Quantum Mechanics*, Plenum Press (1994).

[5] For a detailed discussion of measurement problem and possible ways out look at:
A. Bassi and G. C. Ghirardi, Phys. Rep. **379**, 257 (2003).

[6] A. Bassi and G. C. Ghirardi, Phys. Lett. A **275**, 373 (2000).

[7] D. Bohm, Phys. Rev. **85**, 166 (1952).

[8] D. Bohm, Phys. Rev. **85**, 180 (1952).

[9] D. Bohm, and B. J. Hiley, *The Undivided Universe: An Ontological Interpretation of Quantum Theory*, Routledge, New York (1995).

[10] J. Bub, *Interpreting the quantum world*, Cambridge University Press (1997).

[11] D. Dürr, and S. Goldstein, *Quantum physics without quantum philosophy*, Springer (2012).

[12] D. Dürr, S. Goldstein, and N. Zanghì, Journal of Statistical Physics **67**, 843 (1992).

[13] D. Dürr and S. Teufel, *Bohmian Mechanics*, Springer (2009).

[14] P. R. Holland, *The Quantum Theory of Motion: An Account of the de Broglie-Bohm Causal Interpretation of Quantum Mechanics*, Cambridge University Press (1993).

[15] G.C. Ghirardi, A. Rimini and T. Weber, Phys. Rev. D **34**, 470 (1986).

[16] G.C. Ghirardi, P. Pearle and A. Rimini, Phys. Rev. A **42**, 78 (1990).

[17] L. Diosi, Phys. Rev. A **40**, 1165 (1989).

[18] S.L. Adler and A. Bassi, Science **325**, 275 (2009).

[19] A. Bassi, K. Lochan, S. Satin, T.P. Singh and H. Ulbricht, Rev. Mod. Phys. **85**, 471 (2013).

[20] S. Weinberg, Phys. Rev. A **85**, 062116 (2012). ArXiv:1109.6462v3 (2011).

[21] S. Weinberg, Ann. Phys. **194**, 336 (1989). Phys. Rev. Lett. **62**, 485 (1989).

[22] L. Diosi, Phys. Lett. A **105**, 199 (1984).

[23] R. Penrose, Phil. Trans. Roy. Soc. Lond. **356**, 1927 (1998).

[24] D. Giulini and A. Grossardt, Class. Quant. Grav. **28**, 195026 (2011).

[25] Ghirardi, G.C. and Grassi, R., *Dynamical Reduction Models: some General Remarks*, in *Nuovi Problemi della Logica e della Filosofia della Scienza*, D. Costantini et al. (eds), Editrice Clueb, Bologna. (1991)

[26] G. Lindblad, Commun. Path. Phys. **48**, 119 (1976).

[27] V. Gorini, A. Kossakowski and E.C.G. Sudarshan, J. Math. Phys. **17**, 821 (1976).

[28] A. Bassi, D. Duerr and G. Hinrichs, Phys. Rev. Lett. **111**, 210401 (2013)

[29] For a well-written treatment of equivalent mixtures and applications look at:
L.P. Hughston, R. Jozsa and W. Wootters, Phys. Lett. A **183**, 14 (1993).

[30] Cavalcanti E 2014 private communication