Some loopholes to save quantum nonlocality

Luigi Accardi

Centro Vito Volterra, Università degli Studi di Roma "Tor Vergata" – accardi@volterra.uniroma2.it – WEB-page: http://volterra.uniroma2.it

Abstract. The EPR–chameleon experiment has closed a long standing debate between the supporters of quantum nonlocality and the thesis of quantum probability according to which the essence of the quantum pecularity is non Kolmogorovianity rather than non locality.

The theory of adaptive systems (symbolized by the chameleon effect) provides a natural intuition for the emergence of non–Kolmogorovian statistics from classical deterministic dynamical systems. These developments are quickly reviewed and in conclusion some comments are introduced on recent attempts to "reconstruct history" on the lines described by Orwell in "1984"

Keywords: Quantum probability, chameleon effect, adaptive systems, non Kolmogorovian probability
PACS: 01.70

AN HISTORICAL CHALLENGE

In the conclusions of his 1964 paper [11] Bell stated that:

"... In a theory in which parameters are added to quantum mechanics to determine the results of individual measurements, without changing the statistical predictions, there must be a mechanism whereby the setting of one measuring device can influence the reading of another instrument, however remote. Moreover, the signal involved must propagate instantaneously, so that such a theory could not be Lorentz invariant. ... Of course, the situation is different if the quantum mechanical predictions are of limited validity. ..."

To say that the predictions of a theory are "... of limited validity ..." is either a truism (and surely Bell was not a trivial person) or an euphemism to mean that the theory is wrong, and the latter interpretation is the one commonly adopted in the literature.

The theories which allow "... to determine the results of individual measurements ..." are the deterministic theories.

Thus the above statement of Bell is a very strong one and leaves no space for ambiguity: **either quantum mechanics is wrong or any deterministic completion of this theory which leads to the same experimentally measurable predictions must be incompatible with the locality principle.**

A few years later the deterministic theories, compatible with the locality principle, were called "local realistic theories" and Bell's result was synthesized in the statement: **no local realistic theory can reproduce the EPR correlations**.

17 years after, in another widely quoted paper ([10], Sect. 3: "Difficulty with locality"), Bell strengthened his 1964 statement as follows:

"... certainly Einstein could no longer write so easily, speaking of local causality "... I still cannot find any fact anywhere which would make it appear likely that locality will have to be abandoned." ... "

Thus, while in his 1964 paper Bell's emphasis was on the impossibility of a local realistic completion of quantum theory, in his 1981 paper the difficulties with locality are shifted from the completions of quantum mechanics (QM) to QM itself.

This development is not surprising: after all, if one accepts as true the impossibility to reproduce the predictions of QM by means of any local realistic theory, then it is natural to interpret this as an indication that QM should be either "non realistic" or "non local".

On the other hand, the experiments, which in 1981 begun to have a developed history, seemed to support the "realism" assumption (cf. discussion below). In view of this Bell, being convinced that his 1964 argument was correct, had all the rights to conclude that the experimental data, accumulated between 1964 and 1981, justified the above mentioned radicalization of his position on the locality issue. The majority of physicists could not find convincing critiques against these arguments and ended up to accept, although reluctantly, this "quantum non locality": a term that since those years has been recurrent in the literature. For example Stapp [32] says: "... it is widely accepted

today that some sort of nonlocal effect is needed to resolve the problem raised by the works of Einstein, Podolsky and Rosen (EPR) and John Bell. ...” and, in a more specific way, by De Muynk [18]: “ ... It has often been felt that the most surprising feature of Bell experiments is the possibility (in certain sense) of a strict correlation between the measurement results of the two measured observables, without being able to attribute this to a previous preparation of the object (no ‘elements of physical reality’). ...”.

This De Muynk’s statement clearly explains why the experimental results were interpreted as evidence supporting the idea of the existence of ‘elements of physical reality’, thus ruling out the “non reality” assumption and leaving open only one possibility: “non locality”. The argument is already present on the EPR paper and can be summed up as follows.

The strong correlations in the EPR type experiments are incompatible with the idea that the virtual values of the polarization observables, coded into the singlet state, become real by virtue of a random mechanism, triggered locally by the measurement apparatus (orthodox interpretation). Here the term “locally” means that random mechanisms, acting in widely separated regions of space, are independent. This incompatibility is an elementary statistical fact: if I toss two coins independently, I expect that, in many tosses, the fraction of coincidences in the results will be approximately $1/2$. But in the EPR type experiments these fractions can take all values between 0 and 1, thus independence is ruled out. Dependence could rise only either from a “pre–determination”, as argued in the original EPR paper, or by an instantaneous nonlocal “agreement” between the two particles (or the two apparata).

But, if Bell is right, “pre–determination” is ruled out by Bell’s inequality. Thus, if this is the case, we **must**, even reluctantly, accept non locality.

QUANTUM PROBABILITY AND THE NON LOCALITY CHALLENGE

The huge interest in Bell’s inequality was motivated by the widespread belief that this inequality allows to prove experimentally the incompatibility between the fundamental physical theory of our times (QM) and one of the fundamental pillars of physics (the locality principle).

This was indeed an historical challenge and was recognized as such by a huge number of people including some of the best minds in physics, philosophy, mathematics, ...

If an answer has to be given to this challenge, it must stated as clearly and strongly as the above reported statements by Bell. Furthermore this clarity and strength must be the expression, as it was in the Bell case, of a rational combination of theorems and experimental data.

Starting from 1979 the quantum probabilistic approach was developed as an answer to Bell’s challenge. Its main conclusion, after a 25 year confrontation is exactly the opposite of the Bell statement reported at the beginnibg of this section: **there is no principle contradiction between quantum theory, locality and realism**.

The quantum probabilistic point of view is supported by:

(i) new mathematical results (the statistical invariants, the theory of adaptive dynamical systems, ...)

(ii) a new physical intuition: adaptive dynamical systems (symbolized by the chameleon metafora) can produce, by means of local, deterministic interactions, the non classical statistics which is the real core of inequalities such as the two slit one, Bell’s one, etc. ... .

(iii) the construction of a classical, reversible, local deterministic, binary dynamical system which exactly reproduces the EPR correlations. There are no artificial post–selection assumptions in the statement of the theorem which proves that the above mentioned model effectively reproduces the EPR correlations: everything is predetermined (in the chameleon sense) at the source, exactly as required both by EPR and Bell!

(iv) an experiment realizing, with three independent computers, the deterministic model mentioned in (iii). Since the thesis, advocated by Bell and his followers, is that no classical system can reproduce, by means of local independent choices, the EPR correlations, the present experiment, although realized using computers, is not a simulation of an experiment, but a real experiment, because the personal computers used in it are surely macroscopic classical systems.

In addition the protocol of the above experiment faithfully reproduces the protocol actually followed in all EPR type experiments (meaning by this, all experiments in which the separation between the two particles is large enough to guarantee that the independence assumption, crucial in any real EPR type experiment, is satisfied).

The physics community should rejoice because of these results. In fact they lead to a re–evaluation of the orthodox interpretation by making it free of spurious metaphysical statements and by supporting it with new mathematical tools and new physical ideas.

In some sense one might say that, just as von Neumann measurement theory was the first attempt to give a mathematical context to Bohr’s idea on the role of the measurement apparatus in QM, the quantum probabilistic

approach explicitly introduces, in von Neumann measurement theory the appropriate mathematical formulation of the notions of causality, locality and of "observable dependent interactions" (chameleon effect).

DYNAMICAL SYSTEMS: PASSIVE AND ADAPTIVE

Definition 1 *A classical, discrete time, deterministic (passive) dynamical system is a quadruple:*

$$\{\Omega\ ,\ \mathscr{O}\ ,\ P\ ,\ T\} \tag{1}$$

where:
– Ω is the **state space** *(more precisely $\Omega = (\Omega, \mathscr{F})$ is a measurable space)*
– $\mathscr{O}$ is a set of **observables** *(measurable maps from Ω to $\boldsymbol{R}$)*
– P is the preparation of the experiment (a probability measure on Ω - **initial distribution***)*
– $T\ :\ \Omega \to \Omega$ is a discrete time **dynamics** *(measurable map)*
The system is called **reversible** *if T^{-1} exists (and is measurable P–a.e.).*

We do not require that the statistics is invariant under the dynamics (i.e. that $P \circ T^{-1} = P$). Thus, if $\langle\cdot\rangle$ denotes the expectation value at time 1 (if 1 is the instant of measurement, this is precisely what we measure), then for any observable $A \in \mathscr{O}$, one has:

$$\langle A\rangle = \int_S A(Tx)dP(x) = \int_S A(y)dP\circ T^{-1}(y) = \int_S A(Tx)p(x)dx = \int_S A(y)p\left(T^{-1}(y)\right)dy \tag{2}$$

$p(x)$ is the probability density (denoted $\rho(x)$ in Bell's notations)

Definition 2 *A classical deterministic* **adaptive** *dynamical system is a quadruple:*

$$\{\Omega\ ,\ \mathscr{O}\ ,\ \{P_A\}_{A\in\mathscr{O}}\ ,\{T_A\}_{A\in\mathscr{O}}\} \tag{3}$$

where:
– Ω (the state space) and $\mathscr{O}$ (the observables) are as in Definition (1).
– for each $A \in \mathscr{O}$:
(i) P_A is a probability measure (the preparation of an experiment to measure A)
(ii) $T_A\ :\ \Omega \to \Omega$ is an $\mathscr{F}$–measurable map (the adaptive dynamics given that A is measured)
The system is called reversible if T_A^{-1} exists and is measurable for each $A \in \mathscr{O}$.

For adaptive dynamical systems formula (2) becomes

$$\langle A\rangle = \int_S A(T_Ax)dP_A(x) = \int_S A(y)dP_A\circ T_A^{-1}(y) = \int_S A(T_Ax)p_A(x)dx = \int_S A(y)p_A\left(T_A^{-1}(y)\right)dy \tag{4}$$

The following examples illustrate the diffeences between formula (2) [passive] and (4) [adaptive].

Example (Common features) . In a box there are N objects. On each of these objects one can measure two binary $\{\pm1\}$–valued observables A, B, so that the set of observables is

$$\mathscr{O} := \{(A_n), (B_n) : n = 1, 2, \ldots, N\}$$

and one can choose the configuration space to be

$$\Omega := \{\pm1\}^N = \{\text{functions } \omega : \{1,\ldots,N\} \to \{\pm1\}\}$$

Example (Passive) . The objects are macroscopic balls and the observables A, B are:
$A_n :=$ the color of the n–th ball $(n = 1, 2, \ldots, N)$
$B_n :=$ the weight of the n–th ball $(n = 1, 2, \ldots, N)$
It is possible to conceive an apparatus such that a measurement of weight does not alter color and conversely. Therefore in this case it is reasonable to use formula (2).

Example (Active) . The objects are electrons and the observables A, B are:
$A_n :=$ the spin of the n–th electron in direction a $(n = 1, 2, \ldots, N)$

$B_n :=$ the spin of the n–th electron in direction b $(n = 1, 2, \ldots, N)$

where a, b are vectors in 3–dimensional real space.

The spin of an electron in direction x $(x = a, b)$ is defined in terms of the electron's response to a magnetic field in direction x (cf. [FeynLeSa66]).

Thus a question like: *"what is the value, at time t, of the electron spin in direction b if at time t I have measured its spin in direction a?"* has the same meaning as the question *"what is, at time t, the color on the leaf of a chameleon of which we have measured the color on the wood?* (to make a full analogy with electrons, we have to postulate that the box contains an equal number of healthy and of mutant chameleons and that a mutant chameleon becomes brown on the leaf and green on the wood). It is clear that, in this case, one has to use formula (4).

VON NEUMANN'S MEASUREMENT THEORY

Let us rephrase, following [4], von Neumann's measurement theory so to make it neutral with respect to classical or quantum physics.

Consider a **dynamical system** S (say, a particle) with state space S_S and an apparatus A with state space S_M. The state space of the composite system will then be

$$\Omega := S_S \times S_M$$

According to von Neumann's measurement theory a measurement of the system S by means of the apparatus A will be described by a reversible dynamical system

$$T_t : \Omega = S_S \times S_M \to \Omega = S_S \times S_M \tag{5}$$

If time is discretized such a dynamics is described by a single map

$$T : \Omega = S_S \times S_M \to \Omega = S_S \times S_M \tag{6}$$

The preparation of the experiment is described by a probability measure P on $\Omega = S_S \times S_M$

$$P \in \text{Prob}\,(\Omega) = \text{Prob}\,(S_S \times S_M) \tag{7}$$

because a preparation is needed both for system and apparatus.

LOCAL CAUSAL MEASUREMENTS

The two additional requirements of:

(i) locality

(ii) causality

were not discussed by von Neumann. Following [4], let us introduce these notions in von Neumann's measurement scheme. For reasons of space, we only recall the main definitions without explanaining why they effectively are the mathematical expression of locality and causality. For this we refer to [4], [7].

Consider a **classical dynamical system** composed of two particles $(1, 2)$ with state space S_1, S_2 respectively and two apparata A_1, A_2 with state spaces M_1, M_2 respectively.

The state space of the composite system will then be

$$\Omega = S_1 \times S_2 \times M_1 \times M_2 \tag{8}$$

According to von Neumann's measurement theory a measurement of the system$(1, 2)$ by means of the apparatus (A_1, A_2) is described, in discrete time, by a reversible dynamical system

$$T : \Omega \to \Omega \tag{9}$$

and the preparation of the experiment is described by a probability measure P on Ω

$$P \in \text{Prob}\,(\Omega) \tag{10}$$

Definition 3 *A dynamics T on the state space Ω, given by (8), is called local if it has the form $T = T_1 \otimes T_2$ where*

$$T_1 : S_1 \times M_1 \to S_1 \times M_1 \tag{11}$$

$$T_2 : S_2 \times M_2 \to S_2 \times M_2 \tag{12}$$

are dynamics. This means

$$T_1 \otimes T_2(\sigma_1, \lambda_1, \sigma_2, \lambda_2) = T_1(\sigma_1, \lambda_1) T_2(\sigma_2, \lambda_2) \quad , \ \sigma_1 \in S_1, \lambda_1 \in M_1, \sigma_2 \in S_2, \lambda_2 \in M_2 \tag{13}$$

Definition 4 *A probability measure P on the space*

$$\Omega = S_1 \times S_2 \times M_1 \times M_2 \tag{14}$$

is called local and causal if it has the form

$$P(d\sigma_1, d\sigma_2, d\lambda_1, d\lambda_2) = P_S(d\sigma_1, d\sigma_2) P_1(\sigma_1, d\lambda_1) P_2(\sigma_2, d\lambda_2) \tag{15}$$

where
– *$P_S(d\sigma_1, d\sigma_2)$ is a probability measure on $S_1 \times S_2$*
– *$P_1(\sigma_1, d\lambda_1)$ is a positive measure on M_1 for all $\sigma_1 \in S_1$*
– *$P_2(\sigma_2, d\lambda_2)$ is a positive measure on M_2 for all $\sigma_2 \in S_2$*
Such a probability measure is called trivial if one has (P_S–a.e.):

$$\int_{M_1} P_1(\sigma_1, d\lambda_1) = 1 \tag{16}$$

$$\int_{M_2} P_2(\sigma_2, d\lambda_2) = 1 \tag{17}$$

Remark The distinction between trivial and non trivial local, causal probability measures was first pointed out in [6] where it is proved that, using trivial probability measures it is impossible to violate Bell's inequality. All the no–go theorems for local hidden variables are based on the implicit postulate that a local causal probability measure is necessarily trivial. The probability measure used in the EPR–chameleon experiment is local causal but non trivial. In [7] it is described a general method to construct examples of non trivial, local, causal probability measures on the space (8).

VIOLATION OF BI WITH FULL EFFICIENCY, NO REJECTION OF PARTICLES, NO POST-SELECTION

This section is devoted to the discussion of the following question: **"Can we build a natural physical intuition, entirely based on classical physics and in full agreement woth the locality principle, of the EPR correlations?"**

We anticipate the answer to this question, which will emerge from the analysis below. The answer is: **Yes, if we understand the mechanism through which adaptive dynamical systems subject only to local interactions, can produce non classical statistics** (such as the EPR correlations). This is precisely the *chameleon effect*: it gives a simple, intuitive picture of how local interactions may alter global statistics thus showing that the implicit mathematicl assumption in Bell's analysis (a single probability space to describe mutually incompatible experiments) is physically untenable.

In order to be more precise let us first of all state the problem.

We will deal exclusively with clasical systems composed of two particles $(1,2)$ and two apparata (M_1, M_2), which measure binary observables, $S_a^{(1)}, S_b^{(2)}\rangle, \ldots$, labeled by indices $a, b, \ldots$ and called "spin" to emphasize the analogy with the EPR type experiments.

The apparata can make local independent choices among these labels and we use the notation (M_a, M_b) to mean that apparatus M_1 has made the choice a and apparatus M_2 the choice b.

We fix the notations and the assumptions of section (). In these notations a state of the global system is specified by a quadruple $(\sigma_1, \sigma_2, \lambda_1, \lambda_2)$ where (σ_1, σ_2) describe the particle degrees of freedom and (λ_1, λ_2) the apparatus'

ones. These parameters specify tha state of our classical deterministic (reversible) system in the usual sense of phase space, i.e. they are a minimal set of observables whose knowledge at a given time t_0 fully determines the values of all other observables in all times, after or before t_0. These could include observables such as position, momentum, spin, polarization, ..., but their specific nature will not be relevant for the discussion that follows.

The causality principle requires that (σ_1, σ_2) do not depend on (a, b), because the particle cannot anticipate the choice of the apparata. The locality principle requires that λ_1 (resp. λ_2) does not depend on b (resp. on a) because one apparatus cannot know the choice of the other one. But of course λ_1 may depend on a and λ_2 on b (if I have to measure spin in direction a, I will prepare a magnet in that direction).

The locality principle also requires that the only admissible constraints between the two particles should come from conservation laws (such as the singlet law) and not from direct or mediated interactions and that the interaction between apparatus M_1 (resp. M_2) on particle 2 (resp. 1) are negligible (cf. [4] or [6] for a mathematical formulation).

The "large spacial separation" between the two particles and the two apparata is useful to assure that this condition is verified with very good approximation even in a non relativistic theory and not by chance played a crucial role in both the EPR and the Bell arguments. Any experiment, claiming to be relevant to the EPR problem, should convincingly argue that this condition is verified (the reader familiar with recent experiments made with particles very close to each other, or even constrined by various kinds of traps, will surely understand the motivation of this caveat).

Since in classical dynamical systems any probability measure is a convex combination of δ–functions, what just said for exact preparations also extends to statistical preparations and leads to the local causal structure described by formula (15) (cf. [7] for more details).

For each pair of particles, the experimentalists measure a pair of observables $(S_a^{(1)}, S_b^{(2)})$. **A crucial condition is that each measurement is done on the same pair**, otherwise it makes no sense to speak of conservation laws. In other words, the empirical correlations are **conditioned** on the event that the measurement has been performed on two particles belonging to the same pair.

The two experimentalists agree in advance that each of them will make M measurements. After that, they have to join the collected data so to calculate the empirical correlations

$$\langle S_a^{(1)}(t) S_b^{(2)}(t) \rangle$$

We suppose that the two experimentalists are in widely separated laboratories and that they cannot follow the trajectory of the single particles.

In these conditions, **how can they join the data being sure that two, independently done measurements acted on particles of the same pair?**

The usual answer to this question (and apparently the only viable one, at the moment) is that the two experimentalists sincronize their clocks and decide that two particles belong to the same pair if and only if they are detected simultaneously. In particular the formula, used in all the EPR type experiments including the EPR–chameleon one, for the the empirical correlations is (cf. [8]):

$$\frac{N_{++}(a,b) + N_{--}(a,b) - N_{+-}(a,b) - N_{-+}(a,b)}{N_{++}(a,b) + N_{--}(a,b) + N_{+-}(a,b) + N_{-+}(a,b)} \tag{18}$$

where $N_{++}(a,b)$ is the number of measurements of type (M_a, M_b) which gave the simultaneous result $(+,+)$ (and similarly for the others). Notice that this formula makes sense only for coincidences, so that the the denominator is the total number os simultaneously detected pairs.

Now we will introduce several very strong idealizations and we invite the reader to check that, each time one drops one of them, thus making the model more realistic, the number of time coincidences will decrease, thus making our conclusion stronger.

We suppose that both experimentalists have ideal detectors, with no accidental losses, no spurious counts, no blind times. In other words absolutely all particles of all pairs, and only them, are detected and no errors are done in the measurement of the observables.

We want to prove that, even under these unrealistically strong constraints, the chameleon effect can guarantee a violation of the Bell's inequalities in full respect of locality.

The argument exploits the phase space picture of the dynamics of the measurement process and goes as follows: The state of the global system at time t is

$$\omega(t) := (\sigma_1(t), \sigma_2(t), \lambda_1(t), \lambda_2(t))$$

The detector is made up of two parts, corresponding to the two measurements:

$$D = (D_1, D_2)$$

By the locality principle D_1 only reacts to the varaibles $(\sigma_1(t), \lambda_1(t))$ and D_2 only to $(\sigma_2(t), \lambda_2(t))$ so that:

$$D(\omega(t)) = (D_1(\sigma_1(t), \lambda_1(t)), D_2(\sigma_2(t), \lambda_2(t)))$$

We can also add (and we will assume this in the following) the additional constraint that the two functions are identical (i.e. that the two detectors are perfectly equal):

$$D_1(\sigma, \lambda) = D_2(\sigma, \lambda) = D_0(\sigma, \lambda) \qquad ; \qquad \forall \sigma, \lambda$$

There is a waiting time between the emission of the particle and the triggering of the detector. This means that the detector is activated when the phase space trajectory of the subsystem $(1, M_1)$ (resp. $(2, M_2)$) hits a certain activation region (or "window"), denoted A and depending only on D_0 (so that it is the same for the two detectors). When the phase space trajectory of the subsystem hits the activation region, the particle is absorbed and the detector is instantaneusly (no blind time) ready for a new count. Let us denote

$$t_A(\sigma_1(\,\cdot\,), \lambda_1(\,\cdot\,)) \qquad ; \qquad (\text{resp. } t_A(\sigma_2(\,\cdot\,), \lambda_2(\,\cdot\,)))$$

the first hitting time of the activation region A along the tajectory $(\sigma_1(t), \lambda_1(t))$ (resp. $(\sigma_2(t), \lambda_2(t))$).

We postulate that the probability of a simultaneous arrival, in the activation region, of two or more particles from different pairs either in the same detector or in two different ones, is zero. Moreover we require that the hitting time for each sub–trajectory is finite (i.e. that before or later each particle activates the detector):

$$0 < t_A(\sigma_1(\,\cdot\,), \lambda_1(\,\cdot\,)) < +\infty \qquad ; \qquad (\text{resp. } 0 < t_A(\sigma_2(\,\cdot\,), \lambda_2(\,\cdot\,)) < +\infty)$$

Under these assumptions no particle is lost or artificially rejected.

These conditions are trivially implementable in any discrete time model (in particular in the EPR–chameleon model of [6]) and in the corresponding simulation. By interpolation this implementation can be carried over to continuous time.

In these notations the detector functions can simply be written in the form:

$$D_1(\sigma_1(t), \lambda_1(t)) = \chi_A(\sigma_1(t), \lambda_1(t)) S_a^{(1)}(\sigma_1(t), \lambda_1(t))$$

$$D_2(\sigma_2(t), \lambda_2(t)) = \chi_A(\sigma_2(t), \lambda_2(t)) S_b^{(2)}(\sigma_2(t), \lambda_2(t))$$

where $\chi_A(x)$ is zero if $x \notin A$ and 1 if $x \in A$. This means that the detector does nothing until t becomes equal to the first hitting time of A (along the given sub–trajectory). When this time is reached, it gives the correct value of $S_a^{(1)}$ (resp. $S_b^{(2)}$) at that moment. At this time the measurement process stops and the apparatus becomes ready for the next pair. [There is no problem in introducing a blind time between two consecutive detections. If this is smaller than the emisssion time between two consecutive pairs, then no particle is lost. This more realistic assumption allows to avoid a discontinuity in the detector function.]

At this point the chameleon effect enters crucially into the play: because of it the local dynamics depend on the choice of the measurement. Therefore, corresponding to the choices (M_a, M_b) one will have:

$$(\sigma_1(t), \lambda_1(t)) = T_a^t(\sigma_1(0), \lambda_1(0))$$

$$(\sigma_2(t), \lambda_2(t)) = T_b^t(\sigma_2(0), \lambda_2(0))$$

where T_a^t, T_b^t are the local adaptive dynamics. This means that the two hitting times may be different (cf. the figure below) and, when this happens, the experimentalists will not detect a coincidence.

Conclusion: the early insistencde of quantum probability ([2], [3]) on the fact that every probability is a conditional probability, is just a theoretical formulation of everyday experimental practice and does not require that any particle is lost or artificially rejected.

THE SITUATION WITH QUANTUM NON LOCALITY

Let us summarize the the status of quantum non locality at the light of the four new ideas introduced by quantum the probabilistic approach in the the foundation of QM (some of these critique ideas have now been stably accepted in the foundation debate):

(i) the discovery that the Bell inequality belongs to the wider family of "two-slit type inequalities" and the individuation of the common probabilistic root of all these inequalities [2].

(ii) the notion of "statistical invariant" and the first example of a statistical invariant for a quantum model (the two-slit type inequalities are partial statistical invariants for the Kolmogorov model) [1].

(iii) the information-theoretical axiomatization of probability classical and quantum) and the first (apparently the only one, at the moment) mathematical deduction of the quantum model from physically plausible axioms [Ac82c], [Ac95a].

(iv) the mathematical formulation of the theory of adaptive systems (chameleon effect) [4] and the experimental proof that adaptive systems can violate Bell inequalities by local interactions [6], [7] and without rejection of particles (cf. section () below) .

The common probabilistic root of the "two-slit type inequalities" and of "Bell type inequalities" is the implicit assumption, in both cases, that the statistical data, coming from three mutually incompatible experiments, can be described by a single classical (Kolmogorovian) probabilistic model. The experimental data, in both cases, contradict this assumption.

Once made explicit this implicit assumption, the problem becomes: is it true that classical physics, starting from the same conditions of quantum physics, i.e. mutually incompatible experimental preparations, cannot reproduce the EPR statistics in full respect of the locality principle?

Stated otherwise: is the emergence of non Kolmogorovian statistics a specific feature of the quantum world or it is a deeper, more general phenomenon, of universal applicability?

The analysis of [4], schematically summarized in section () shows that the classical physics of adaptive systems can produce non Kolmogorovian statistics, not as a consequence of artificial, ad hoc models, but of a general principle which should be included in any serious theory of (classical or quantum) measurement. This principle was baptized "the chameleon effect".

The general analysis of [4] was substantiated in a concrete model in the paper [6] (in [7] some notational misprints in the proof were corrected) and the simulation of this model on independent classical computers provided an experimental proof of the local nature of the model.

The conclusion to be drawn from the above considerations is that Bell's analysis, which is correct when applied to passive systems, is wrong when applied to adaptive systems, where the chameleon effect plays a crucial role. Furthermore, it is wrong for fundamental reasons, not for complicated and artificial mechanisms.

The conclusion of this analysis is that:

the present experimental data cannot distinguish between quantum non locality and an underlying classical, deterministic chameleon effect.

A POSSIBLE LOOPHOLE TO SAVE QUANTUM NON LOCALITY

The analysis of section () is based on the idea that the local interactions with the apparata induce local modifications of the phase space trajectories which alter the global statistics of coincidences and produce non Kolmogorovian correlations.

Now let us imagine (or dream) a situation in which, in the notations of section (), all the phase space trajectories, for all possible choices of the spin directions a, b, hit simultaneously the activation window of the two apparata. In this case the conditioning on coincidences will be trivial because all pairs give a coincidence. Therefore the intuitive picture of section (), explaining the emergence of non Kolmogorovian correlations, would not be applicable.

In fact it is not necessary to have 100% coincidences: a sufficiently high percent is all what is needed to exclude the EPR correlations.

From the point of view of classical physics this is impossible, because different interactions will have generically different hitting times while the above situation would imply that all the hitting times are equal, independently of the interactions, with the exception of a small percent.

This is a very weird behavior, but can one exclude it in principle? Of course not. The only thing that we can say is that at the moment there is neither experimental evidence of such an unplausible behavior nor any theoretical argument

which suggest a plausible scenario for such an occurrence.

But let us even suppose that in the future such a behavior will be effectively detected or such a scenario – proposed. Would this invert the present situation, where quantum nonlocality is totally unnecessary with respect to the classical physics alternative?

Again the answer should be not, because the discussion of section () is only one possible intuitive picture of the emergence of non Kolmogorovian correlations and there is no proof that this interpretation is the only one compatible with the mathematics of adaptive systems.

On the contrary, and this point seems to have been underestimated by some people although it was already emphasized in **[6], the model constructed in that paper contains no assumptions at all on coincidences, conditioning,** ... **.** As the reader can check looking at () this matematical model is just a classical deterministic, reversible, adaptive, local dynamical system reproducing the EPR correlations.

Summing up: the present experimental data are perfectly compatible with the locality principle and the supporters of quantum non locality have yet a long way to go if they want to find arguments which appeal to scientific rationality rather than to mutual consensus.

THE CONSPIRACY OF DETECTORS: WHY A QUARTER OF CENTURY OF SILENCE?

What later was called "the detection loophole" was first considered in an article by Pearle in 1970 [29], possibly anticipated by some comments due to Shimony et al. in 1969 [14]. The main idea of this paper is well expressed by its title (Hidden-variable example based upon data rejection) and by the words of the author:

If *"... the experimenter rejects these data (in the belief that the apparatus is not functioning properly and that, if it had been functioning properly, the data recorded would have been representative of the accepted data) .... "*, then a local hidden-variable model for the EPR correlations can be built.

Pearle's paper is surely mathematically correct, although somewhat involved, however it is a fact that, from 1970 to the early 1990's (i.e. after more than ten years of insistence, from the QP–approach on the role of conditioning), the paper by Pearle, although well known to the experts, is hardly ever quoted.

It is difficult to undertand if the meticulous care with which the experimentalists, after 1970, discuss the efficiency problem and compare the counts with and without polarizers (cf. e.g. [8]) is an indirect trace of the influence of this paper or if comes from other sources.

Certainly the "data rejection" argument would suggest that the denominator in (18) depends on a, b and this dependence was carefully checked in the experiments with the surprising result that no trace of such a dependence was found. This was interpreted as an evidence in favor of the fact that the experimental data were "a fair sample", i.e. in Pearle's words: "representative of the accepted data".

A second objective fact is is the exponential explosion, in the period from 1970 to the early 1990's, of publications on the Bell inequalities: not only from authors supporting Bell's non locality statement, but also from fierce opponents supporting other statements such as the nonvalidity of QM, ...

A third objective fact emerging from the published literature is that neither side considered Pearle's proposal as a serious alternative. If they had, one should expect that the majority of physicists should have concluded against quantum non locality, namely: **... let us wait for more efficient detectors before abandoning such a fundamental notion as locality ...** . After all Pearle even gave an estimate of how efficient the detectors should be to prevent his "data rejection" argument: 86%.

A fourth objective fact, easily documented from the existing literature, is that exactly the opposite happened: a large majority of very good physicists, without taking into serious consideration Pearle's proposal, were convinced that the experimental evidence was in support of Bell's analysis **Why this possibility was discarded ? Why such a silence for so many years?**

I believe that the best answer to this natural question can be found in the final remarks of Pearle's 1970 paper where he explicitly and fairly declares that:

"... it is difficult to take this hypothesis seriously as a physical principle capable of extension to a large group of phenomena ..."

In 1981 Bell made a list of the four alternative possibilities that he conceived as possible outcomes of the challenge mentioned in section () ([10], Sect. 8: "Envoi" [the boldface is not in the original text])

"... **First** ... quantum mechanics must be *wrong* in sufficiently critical situations ...

... **Secondly** ... it is not permissible to regard the experimental settings a and b in the analyzers as independent variables ...

... Apparently separate parts of the world would be deeply and conspiratorially entangled and our apparent free will would be entangled with them. ...

... **Thirdly** ... to admit that causal influences *do* go faster than light.

... **Fourthly** ... there is no reality below some "classical" "macroscopic" level.

It is quite plausible, although difficult to prove or disprove, that the second alternative refers to the "data rejection" hypothesis. Bell's comment on it is a clear indication of his feelings: the idea that two random inefficiencies, in two widely separated and independent apparata, should always (i.e. in thousands different measurements) be so "conspiratorially entangled" to reproduce always the same result (i.e. exactly the EPR correlations), seemed to be so absurd and implausible that for more than 20 years nobody found it fruitful to investigate further this possibility.

An additional reason of the diffidence generated by the "data rejection" idea is that Bell's challenge was a principle and conceptual one, pointing to a principle contradiction between the two basic theories of contemporary physics. Pearle's proposal was not of a principle nature, but contingent to the efficiency level of detectors in a given time.

Moreover this alternative did not open new perspectives with respect to the Bell alternative, discussed in section (). In fact the already quoted statement in [29]: *"... the experimenter rejects these data (in the belief that the apparatus is not functioning properly and that, if it had been functioning properly, the data recorded would have been representative of the accepted data) .... "*, when applied to the EPR correlations, brings beck to the already mentioned Bell alternative, namely:

(i) the experimenter's "belief" is correct. But in this case the rejection procedure is useless because, without rejection, one would find the same statistics which, according to Bell's analysis, if experimentally confirmed, implies non locality.

(ii) the experimenter's "belief" is not correct. But in this case the rejection procedure is a falsification of the real statistics. Consequently the agreement would not be between a local hidden variable theory and QM, but between an arbitrary deformation of a local hidden variable theory and QM. Since, by arbitrarily rejecting empirical data one can cook any type of relative frequencies, this result is not particularly interesting.

EFFICIENCY AND THE CHAMELEON EFFECT: THE REBIRTH OF THE ORTHODOX INTERPRETATION

As shown in the previous section, the "inefficiency loophole" fell into oblivion for decades because it was not able to provide a convincing physical intuition of what was behind the quantum correlations. It was perceived as an ad hoc stratagem rather than (to use again Pearle's words) "... a physical principle capable of extension to a large group of phenomena ...".

It is precisely this natural physical intuition, this general "physical principle capable of extension to a large group of phenomena" that was provided by the idea of adaptive system, symbolized by the chameleon [4].

To illustrate what is new, in the chameleon effect, with respect to the data rejection hypothesis let us begin with a contemporary formulation of this hypothesis given in the paper [23]: *"... The detection loophole is based on the following fact: in real experiments the efficiencies of the detectors are such that the number of detected events is significantly smaller that the set of tested quantum systems.*

One has to assume that the sample over which the statistics is measured is a fair sample. ..."

This is a rephrasing of the already twice quoted statement of Pearle and shows how remarkably stable has been the interpretation of this hypothesis in more than 30 years.

Starting from this statement let us show how the notion of **"fair sample"** turns out to be the key one to distinguish the new, quantum probabilistic intuition from the classical (Kolmogorovian) one. This notion has a standard interpretation since centuries: if in a box there are $1 << N$ balls and of these $3/4$ are green and $1/4$ are brown, a **"fair sample"** of this system is a set of $1 << M << N$ balls of which approximately $3/4$ are green and approximately $1/4$ brown. If, instead of a single observable (color) we consider several observables (color, weight, material, ...), labeled with indices $a, b, c, \ldots$ and denoted $S_a, S_b, S_c, \ldots$ respectively, then we can extend the above definition by considering joint properties. The corresponding joint relative frequencies constitute the common preparation of the experiment, denoted C. Since the act of measurement does not change the prepared values (passive systems), denoting $M_a, M_b, M_c, \ldots$ the event that we measure property $a, b, c, \ldots$ respectively, one has:

$$P(S_a = s_a|M_a) = P(S_a = s_a|C)$$

where s_a denotes any value of S_a. This situation of a single common conditioning corresponds to the uniqueness of the probability space.

Suppose now that in the box there are adaptive systems and that only one observable at a time can be measured (incompatibility). The observables are now response observables and they have to be understood in the sense: *if I will interact with the apparatus M_a my response will be s_a*. In this case the preparation of the ensemble will be described by the conditional probabilities

$$P(S_a = s_a|M_a)$$

but now there is no necessity, neither logical nor physical, that these conditional probabilities are compatible with a single conditioning. Yet the experimental measurements can be completely "fair", in the sense that they correctly reproduce the probabilities $P(S_a = s_a|M_a)$ (which are the only observable ones).

Summing up: **in the intuition of classical probability (which is the one implicitly underlying the statements of Bell and followers) "fair sample" is meant as a synonim of "single classical probability space". In the intuition of adaptive systems, a sample can be fair without admitting a single probability space description**. Since the non existence of a single probabiity space is the main point of the violation of the various two–slit type inequalities, we see how the idea of adaptive systems is helpful in building a natural physical intuition of the quantum correlations.

Up to now all what we said is applicable to single as well as to composite systems (e.g. pairs). Now let us come back to the definition of "efficiency" of a detector in the context of pair measurements.

It is common practice, in the calibration of instruments, to separate random from systematic errors. In view of this, in the standard definition of efficiency as

$$\eta := \frac{\text{nr. of coincidences}}{n_1 + n_2}$$

where n_1 (resp. n_2) is the total number of photons revealed by apparatus 1 (resp. 2), it should be made explicit that, for the parameter η to be meaningful, one should filter out from the denominator all spurious counts which have a deterministic origin (systematic errors).

For example, suppose that two basketball players throw simultaneously a ball, with the same velocity, towards two vertical tables at equal distance from each of them. Suppose there is a counter with a clock, connected to each table, and that these counters are used to set up a statistics of coincidences. Suppose that the two players are not equally good and one of them hits the target, say $1/2$ of the times, while the other one always hits the target. Then, if we use the above formula to measure the efficiency of the detector, we conclude that the combined detector has efficiency (less or equal to) $1/2$ even if all the hits have been counted, i.e. even if the real efficiency is 1.

Experimentalists are well aware of this caveat and give it for granted. Some theoreticians are less aware of this fact and this has lead to an amusing confusion between chamaleon effect (which is sytematic and deterministic) and inefficiency of detector (which is what remains when the deterministic effects have been filtered away).

The difference between these two notions is a principle one which can experimentally evidentiated as proved in section (6) of [6].

PHILOSOPHY OF SCIENCE AND HISTORY OF SCIENCE

Philosophy of science is the modern development of the old "theory of knowledge". It deals, among other problems, with the mutual coherence among different levels of description of the natural world (theories). In this sense the problem of the mutual coherence between QM and relativity, from a principle not only technical point of view has been for several years one of the core problems of the philosophy of science.

One of the questions, investigated by history of science, is how different parts of the scientific community react to scientific discoveries and new ideas leading to solutions of old problems.

This is a field where history and sociology of science intersect, sometimes in a really instructive way. In the following we bring some examples.

Euclid invented non Euclidean geometries, ... of course

One of the most interesting aspects of new scientific discoveries is that they allow to look from new perspectives to established results and old ideas.

For example nowadays the date of birth of non Euclidean geometries is generally recognized in Gauss' intuition that the geometry of surfaces is intrinsic, i.e. it can be distinguished from flat geometries by means of measurements not requiring the knowledge of the fact that the surface is or can be embedded into a higher dimensional space.

This intuition was substantiated by Gauss in his "Theorema egregium":

$$\alpha+\beta+\gamma=2\pi+\kappa A$$

where α,β,γ are the inner angles of the triangle, A its area and κ a real constant, now called Gaussian curvature, which was the first example of **geometrical invariant**.

This is not only a generalization of the old famous Euclid's result concerning the sum of the inner angles of a triangle, but also a radical change of perspective in the consideration of this result. After Gauss' theorem, Euclid's relation

$$\alpha+\beta+\gamma=2\pi$$

could no longer be considered a general property of all (physical) triangles but as a **compatibility condition** among the inner angles of a triangle which guarantees the possibility of an "euclidean model" for this triangle.

Now let us imagine a modern geometer who, with the hindsight coming from Gauss' theorem, would argue as follows:

It is obvious that the non euclidean geometries go back to Euclid. In fact he was the first who wrote a compatibility condition which gives the necessary and sufficient condition for three given angles to be the inner angles of an "euclidean triangle".

It is obvious that Gauss did not discover anything new. People like Mercator were using spherical triangles 300 years before him, to design maps of the earth. And what to say of Pascal, who 200 years before Gauss was already working with projective geometry, which is the first historical example of a non trivial manifold ...!

None of the episodes recalled in the above statement is false, but who would agree that the above statement is globally correct?

Any reasonable (and fair) person knows that the (geometric) compatibility conditions, before the "theorema egregium", where considered **inside a single model**. Only after this theorem the conscience emerges that these conditions can be generalized and this generalization used to discriminate between a multiplicity of models.

Similarly, in probability, the compatibility conditions have a long history which begins with the very birth of this discipline, in the seventeenth century, and culminates with the Kolmogorov consistency theorem, of which all these compatibility conditions are corollaries.

The situation with the mathematical formalism of quantum mechanics, before quantum probability, was the probabilistic analogue of the use of spherical trigonometry or projective geometry before the birth of noneuclidean geometry: it was empirically clear that these new models were not fitting with the usual ones, but the precise mathematical formulation of this difference had to wait for centuries in the case of geometry and for decades in the case of probability.

The birth of non Kolmogorovian probability (of which the quantum probabilistic one is a fundamental, but particular case) begins when:

(i) these "compatibility conditions" are no longer considered as results inside a probabilistic model, but as **statistical invariants**, distinguishing the different models. In perfect analogy with what happened in geometry.

(ii) the first example of full **statistical invariant** for a non trivial (i.e. non Kolmogorovian) model is computed [1].

It is wrong, it is impossible, ... I mean ... it is true and I already knew it

When an old standing open scientific problem is solved the first reaction of a fraction of the academic community is to say: *everything is wrong* then, when the theoretical and experimental evidence in favor of the new thesis has become so large to make it impossible to continue to hide it under the carpet, the same people say: **we had always known !**

The history of science is full of examples of the above situation, but it is always a source of fun to discover new ones.

The interested reader, may compare the (present) web page of Reinhard Werner with the following statement, taken from a letter of Werner to Luigi Accardi of 23 Feb 2001, i.e. about one year before the paper [6] which proves the statement below to be flatly wrong.

After this comparison (possibly done before a new feed back change), the eventual interested reader can decide by his/her own if the title of the present subsection is appropriate to the situation or not.

" ... Regarding locality there are two possibilities now, and the choice is not a matter of taste, but is part of the physical model, which has to be judged on physical grounds:
(A) The modification of dynamics [due to the chameleon effect (this clause is mine (LA)] remains local

That is the change of dynamics introduced introduced by Bob's choice of device does not affect Alice's part of the system and vice versa.

Clearly, this will always be only an approximation, but it is often a good one. In fact, it can often be controlled by an estimate showing that the small interactions present will not suffice to explain are highly significant correlation data.

So if the dynamical change remains local, we can summarize the Level 1 description into a specification of the Level 0 observables, and *these will remain local*. Hence the inequalities will hold. ... "

APPENDIX I: THE EPR–CHAMELEON DYNAMICAL SYSTEM

The EPR–chameleon dynamical system is a local, deterministic, reversible, classical dynamical system reproducing the EPR correlations. It was first constructed in [6].

In this construction one considers 4 classical deterministic local dynamical system

$$(1\ ,\ M_a\ ,\ 2\ ,\ M_b) \tag{19}$$

– 1 and 2 are called particles
– M_a and M_b are called measurement apparata
In the following $1,2$ will be labels for particles and a,b labels for apparata. We suppose that

$$a,b \in [0,2\pi] \tag{20}$$

– the state space of both composite systems $(1,M_a)$ and $(2,M_b)$ is

$$[0,2\pi]\times\mathbf{R} \tag{21}$$

– therefore the **state space** of the whole system $(1,M_a,2,M_b)$ is

$$[0,2\pi]^2\times\mathbf{R}^2 \tag{22}$$

Each of the composite systems $(1,M_a)$ and $(2,M_b)$ has a **local adaptive dynamics**

$$T_{1,a},T_{2,b}\ :\ [0,2\pi]\times\mathbf{R}\to[0,2\pi]\times\mathbf{R} \tag{23}$$

defined by

$$T_{1,a}(\sigma_1,\lambda_1) := \left(\sigma_1, \frac{\sqrt{2\pi}|\cos(\sigma_1-a)|\lambda_1}{4}\right) \in [0,2\pi]\times\mathbf{R} \tag{24}$$

$$T_{2,b}(\sigma_2,\lambda_2) := (\sigma_2,\sqrt{2\pi\lambda_2}) \in [0,2\pi]\times\mathbf{R} \tag{25}$$

Notice that the inverse transformation is:

$$T_{1,a}^{-1}(\sigma_1,\lambda_1) := (\sigma_1,\frac{4\lambda_1}{\sqrt{2\pi}|\cos(\sigma_1-a)|}) \tag{26}$$

$$T_{2,b}^{-1}(\sigma_2,\lambda_2) := (\sigma_2,\frac{\lambda_2}{\sqrt{2\pi}}) \tag{27}$$

For any $x\in[0,2\pi]$, we define the ±1–valued maps (**observables**)

$$S_x^{(1)},S_x^{(2)}\ :\ [0,2\pi]\times\mathbf{R}\to\{\pm1\} \tag{28}$$

so that, $\forall\sigma\in[0,2\pi]$ and $\forall\mu\in\mathbf{R}$

$$S_a^{(1)}(\sigma,\mu)=S_a^{(1)}(\sigma)=\mathrm{sgn}(\cos(\sigma-a)) \tag{29}$$

$$S_b^{(2)}(\sigma,\mu) = S_b^{(2)}(\sigma) = \mathrm{sgn}(\cos(\sigma - b)) = -S_b^{(1)}(\sigma,\mu) \tag{30}$$

$S_x^{(1)}$ is an observable of particle 1; $S_x^{(2)}$ an observable of particle 2.

Finally we have to give the **initial distribution** P of the whole system $(1, M_a, 2, M_b)$.

Since the dynamics $T = T_{a,b}$ is invertible to give P is equivalent to give $P \circ T^{-1}$

Moreover, since the dynamics is local and deterministic, we know that $P \circ T^{-1}$ is local and causal if and only if P is such.

We define $P \circ T^{-1}$ to be the probability measure on $[0, 2\pi]^2 \times \mathbf{R}^2$:

$$p_S(\sigma_1,\sigma_2)p_{1,a}(\sigma_1,\lambda_1)p_{2,b}(\sigma_2,\lambda_2)d\sigma_1\sigma_2 d\lambda_1 d\lambda_2 := \tag{31}$$

$$\frac{1}{2\pi}\delta(\sigma_1 - \sigma_2)d\sigma_1 d\sigma_2 \tag{32}$$

$$\delta((\sigma_2, \frac{\lambda_2}{\sqrt{2\pi}}) - m_a)d\lambda_1 \delta((\sigma_2, \frac{\lambda_2}{\sqrt{2\pi}}) - m_b)d\lambda_2 \tag{33}$$

where m_a, m_b are arbitrary real numbers and

$$(\sigma_1, \sigma_2 \in [0, 2\pi]\ ,\ \lambda_1, \lambda_2 \in \mathbf{R}) \tag{34}$$

Remark (1) Notice the local structure of the initial probability measure.

$p_S(d\sigma_1, d\sigma_2)$ is the initial preparation, i.e. at time $t = 0$, of the composite system $(1, 2)$ and, by the causality principle, it cannot depend on the setting of the apparatus.

In fact at time $t = 0$ the particles cannot know which will be the setting of the apparatus at time $t = 1$ (the first time of interaction with it).

$p_{1,a}(\sigma_1, d\lambda_1)$ and $p_{2,b}(\sigma_2, d\lambda_2)$ are the initial preparations of the local apparata. They are typical "response–type" preparations and must be interpreted in the adaptive sense, i.e.:

if, at time $t = 1$, the particle will arrive to me in the state

$$\sigma_x (= \sigma_1, \sigma_2) \tag{35}$$

then my contribution to the statistics wil be determined by the factor:

$$p_{1,x}(\sigma_x, d\lambda_x) \qquad (x = 1, 2) \tag{36}$$

Theorem

Theorem 1 *The above described dynamical system reproduces the EPR correlations, i.e.*

$$\int S_a^{(1)}\left(T_{1,a}(\sigma_1,\lambda_1)S_b^{(2)}\left(T_{2,b}(\sigma_2,\lambda_2)\right) \cdot p_S(\sigma_1,\sigma_2)d\sigma_1 d\sigma_2 \tag{37}$$

$$p_{1,a}(\sigma_1,\lambda_1)p_{2,b}(\sigma_2,\lambda_2)d\lambda_1 d\lambda_2 = -\cos(a-b) \tag{38}$$

Proof. Under our assumptions the left hand side of (38) (i.e. the correlations) become

$$I = \int\int\int\int S_a^{(1)}(\sigma_1)S_b^{(2)}(\sigma_2)p_s(\sigma_1,\sigma_2)p_{1a}(\sigma_1,\lambda_1)p_{2b}(\sigma_2,\lambda_2)d\sigma_1 d\sigma_2 d\lambda_1 d\lambda_2 = \tag{39}$$

$$= \int\int d\sigma_1 d\sigma_2 S_a^1(\sigma_1)S_b^2(\sigma_2)\frac{1}{2\pi}\,\delta(\sigma_1 - \sigma_2) \tag{40}$$

$$\int d\lambda_1 \delta\left(\frac{4\lambda_1}{\sqrt{2\pi}|\cos(\sigma_1 - a)|} - m_a\right)\int d\lambda_2 \delta\left(\frac{\lambda_2}{\sqrt{2\pi}} - m_b\right) \tag{41}$$

Using the identity

$$\int \delta(a\lambda - m)d\lambda = \frac{1}{a}\int \delta(\lambda - \frac{m}{a})d\lambda \tag{42}$$

one obtains

$$I = \int S_a^1(\sigma_1) S_b^2(\sigma_2) \frac{1}{2\pi} \delta(\sigma_1 - \sigma_2) d\sigma_1 d\sigma_2 \cdot \frac{\sqrt{2\pi}}{4} |\cos(\sigma_1 - a)| \cdot \sqrt{2\pi} = \tag{43}$$

$$= \frac{1}{4} \int S_a^1(\sigma_1) S^2(\sigma_1) |\cos(\sigma_1 - a)| d\sigma_1 = \tag{44}$$

$$= -\frac{1}{4} \int \text{ sgn } \cos(\sigma_1 - u) |\cos(\sigma_1 - a)| \cdot \text{ sgn } \cos(\sigma_1 - b) d\sigma_1 = \tag{45}$$

$$= -\frac{1}{4} \int \cos(\sigma_1 - a) \text{ sgn } \cos(\sigma_1 - b) d\sigma_1 = -\cos(b - a) \tag{46}$$

There is no artificial post–selection in the assumptions of the theorem (no "cospiracy of the detectors"): only the chameleon effect plays a crucial role. An experiment realized with classical macroscopic objects (classical computers) realizes the above mentioned dynamical system.

APPENDIX II: TWO-SLIT TYPE INEQUALITIES

This secion is introduced only because during the conference there was some discussion on the hyerarchy between the Bell inequality and the CHSH one.

The conclusion is that the BI is more general but to be appliczable to to the EPR correlations requires the counterfactual argument. The CHSH is a corollary of Bell's but doen't require it explicitly.

Lemma 1 *For any two numbers $a, c \in [-1, 1]$ the following equivalent inequalities hold:*

$$|a \pm c| \leq 1 \pm ac \tag{47}$$

Moreover in (1) equality holds if and only if either $a = \pm 1$ or $c = \pm 1$.

Proof. The equivalence of the two inequalities (1) follows from the fact that one is obtained from the other by changing the sign of c and c is arbitrary in $[-1, 1]$.

Since for any $a, c \in [-1, 1]$, $1 \pm ac \geq 0$, (1) is equivalent to

$$|a \pm c|^2 = a^2 + c^2 \pm 2ac \leq (1 \pm ac)^2 = 1 + a^2 c^2 \pm 2ac \tag{48}$$

and this is equivalent to

$$a^2(1 - c^2) + c^2 \leq 1 \tag{49}$$

which is identically satisfied because $1 - c^2 \geq 0$ and therefore

$$a^2(1 - c^2) + c^2 \leq 1 - c^2 + c^2 = 1 \tag{50}$$

Notice that, in (50), equality holds if and only if $a^2 = 1$ i.e. $a = \pm 1$. Since, exchanging a and c in (47) the inequality remains unchanged, the thesis follows.

Theorem 2 *For any 4 numbers $a, b, c, d \in [-1, 1]$ the following inequalities hold:*

$$|ab - cb| \leq 1 - ac \tag{51}$$

$$|ab + cb| \leq 1 + ac \tag{52}$$

$$|ab + cb| + |ad - cd| \leq 2 \tag{53}$$

$$|ab + cb + ad - cd| \leq 2 \tag{54}$$

Proof. For $b \in [-1, 1]$,

$$|ab \pm cb| = |b| \cdot |a \pm c| \leq |a \pm c| \tag{55}$$

so the inequalities (51) and (52) follow from Lemma (1).

Replacing b by d in (52) and adding this to (51) one finds (53). (54) holds because its left hand side is $\leq$ than the left hand side of (53).

Corollary 1 *If $a, b, c, d \in \{-1, 1\}$, then equality holds in all the inequalities (51), (52), (53), (54).*

Proof. The left hand side of (54) is

$$|b(a+c)+d(a-c)| \tag{56}$$

In our assumptions either $(a+c)$ or $(a-c)$ is zero, so (56) is either equal to

$$|b(a+c)| = |a+c| = 2 \tag{57}$$

or to

$$|d(a-c)| = |a-c| = 2 \tag{58}$$

Thus equality holds in (54). Then it must old in (53) because

$$2 = |ab+cb+ad-cd| \leq |ab+cb|+|ad-cd| \leq 2 \tag{59}$$

Therefore it must hold in both (51) and (52) because, if any of these two inequalities is strict, then we obtain the contradiction:

$$2 = |ab-cb|+|ad+cd| < (1-ac)+(1+ac) = 1 \tag{60}$$

Thus the equality sign holds in all the inequalities and this ends the proof.

The Bell inequality

Corollary 2 *(Bell inequality) Let A,B,C,D be random variables defined on the same probability space $(\Omega,\mathscr{F},P)$ and with values in the interval $[-1,1]$. Then the following inequalities hold:*

$$|E(AB-BC)| \leq 1-E(AC) \tag{61}$$

$$|E(AB+BC)| \leq 1+E(AC) \tag{62}$$

and imply

$$|E(AB-BC)|+|E(AD+DC)| \leq 2 \tag{63}$$

$$|E(AB-BC)+E(AD+DC)| \leq 2 \tag{64}$$

where E denotes the expectation value in the probability space of the four variables. Moreover (61) is equivalent to (62) and, if either A or C has values ± 1, then the three inequalities (61), (62), (63) are equivalent.

Proof. Theorem (2) implies the following inequalities (interpreted P–a.e. on Ω):

$$|AB-BC| \leq 1-AC \tag{65}$$

$$|AB+BC| \leq 1+AC \tag{66}$$

$$|AB-BC|+|AD+DC| \leq 2 \tag{67}$$

from which (61), (62), (63), (64) follow by taking expectation and using the fact that $|E(X)| \leq E(|X|)$ and $E(1)=1$. If (63) holds and A has values ± 1 then, choosing $D=A$, (67) becomes

$$|AB-BC| \leq 2-|1+AC| = 2-1-AC = 1-AC \tag{68}$$

from which (61) follows by taking expectations. Finaly (62) follows from (61) by changing C into $-C$.

Theorem 3 *Let $S_a^{(1)},S_c^{(1)},S_b^{(2)},S_d^{(2)}$ be random variables defined on a probability space $(\Omega,\mathscr{F},P)$ and with values in the interval $[-1,+1]$. Then the following inequalities hold:*

$$\left|E(S_a^{(1)}S_b^{(2)})-E(S_c^{(1)}S_b^{(2)})\right| \leq 1-E(S_a^{(1)}S_c^{(1)}) \tag{69}$$

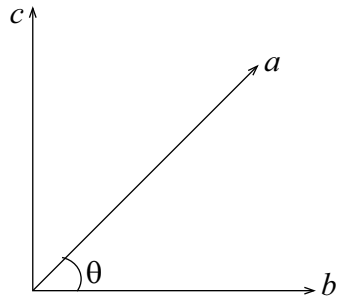
$$\left|E(S_a^{(1)}S_d^{(2)})-E(S_c^{(1)}S_d^{(2)})\right| \leq 1-E(S_a^{(1)}S_c^{(1)}) \tag{70}$$

and imply

$$\left|E(S_a^{(1)}S_b^{(2)})-E(S_c^{(1)}S_b^{(2)})\right|+\left|E(S_a^{(1)}S_d^{(2)})+E(S_c^{(1)}S_d^{(2)})\right| \leq 2 \tag{71}$$

Proof. The thesis is obtained from Corollary (2) by choosing:

$$A = S_a^{(1)} \quad ; \quad B = S_b^{(2)} \quad ; \quad C = S_c^{(1)} \quad ; \quad D = S_d^{(2)} \tag{72}$$



Implications of the Bell's inequalities

Lemma 2 *In the ordinary two–dimensional euclidean plane there exist sets of four unit length vectors a, b, c, d, such that it is not possible to find a probability space* $(\Omega, \mathscr{F}, P)$ *and four random variables*

$$S_a^{(1)} \quad ; \quad S_c^{(1)} \quad ; \quad S_b^{(2)} \quad ; \quad S_d^{(2)} \tag{73}$$

defined on $(\Omega, \mathscr{F}, P)$ *and with values in the interval* $[-1,+1]$*, whose correlations are given by:*

$$E(S_a^{(1)} \cdot S_b^{(2)}) = \varepsilon_{a,b} a \cdot b \tag{74}$$

$$E(S_c^{(1)} \cdot S_b^{(2)}) = \varepsilon_{c,b} c \cdot b \tag{75}$$

$$E(S_a^{(1)} \cdot S_d^{(2)}) = \varepsilon_{a,d} a \cdot d \tag{76}$$

$$E(S_c^{(1)} \cdot S_d^{(2)}) = \varepsilon_{c,d} c \cdot d \tag{77}$$

where

$$\varepsilon_{a,b} \ , \ \varepsilon_{c,b} \ , \ \varepsilon_{a,d} \ , \ \varepsilon_{c,d} \in \{-1,+1\} \tag{78}$$

are arbitrarily chosen and where, if $x = (x_1, x_2)$*,* $y = (y_1, y_2)$ *are two vectors in the plane,* $x \cdot y$ *denotes their euclidean scalar product, i.e. the sum* $x_1 y_1 + x_2 y_2$*.*

Proof. Suppose that, for any choice of the vectors $x = a, b, c, d$ as above, there exist random variables $S_x^{(j)}$ as in the statement of the Lemma.

Then by Theorem (3) they must satisfy the inequality (71) which, in view of (74), (75), (76), (77), becomes

$$|\varepsilon_{a,b} a \cdot b - \varepsilon_{c,b} c \cdot b| + |\varepsilon_{a,d} a \cdot d + \varepsilon_{c,d} c \cdot d| \leq 2 \tag{79}$$

Factorizing $\varepsilon_{a,b}$ in the first term and $\varepsilon_{a,d}$ in the second, and denoting $\varepsilon' := \varepsilon_{a,b}\varepsilon_{c,b}$; $\varepsilon := \varepsilon_{a,d}\varepsilon_{c,d}$ the above inequality becomes equivalent to

$$|a \cdot b - \varepsilon' c \cdot b| + |a \cdot d + \varepsilon c \cdot d| \leq 2 \tag{80}$$

Therefore to prove the statement it will be sufficient to produce four unit vectors $x = a, b, c, d$ whose scalar products violate the inequality (71) for any choice of ε, $\varepsilon' \in \{\pm 1\}$.

To this goal we choose $a = d$ so that (80) becomes

$$|a \cdot b - \varepsilon' c \cdot b| + |1 + \varepsilon a \cdot c| \leq 2 \tag{81}$$

If the three vectors a, b, c are chosen to be in the same plane and such that b is perpendicular to c and a forms an angle θ with b, as in Figure (1) below,

then the inequality (81) becomes:

$$|a \cdot b| + |1 + \varepsilon a \cdot c| = |\cos\theta| + |1 + \varepsilon \sin\theta| = |\cos\theta| + 1 + \varepsilon \sin\theta \leq 2 \tag{82}$$

According to the sign of $\cos\theta$ this leads to consider the two functions:

$$\cos\theta + 1 + \varepsilon\sin\theta =: f(\theta) \qquad ; \qquad \theta \in [-\frac{\pi}{2}, +\frac{\pi}{2}] \tag{83}$$

$$\cos(\theta) + 1 + \varepsilon\sin(\theta - \pi) = \cos(\theta) + 1 - \varepsilon\sin(\theta) \qquad ; \qquad \theta \in [-\frac{\pi}{2}, +\frac{\pi}{2}] \tag{84}$$

it is sufficient to study (83). The derivative of $f(\theta)$ is zero for $tg(\theta) = \varepsilon \in \{\pm 1\}$ and the maxima correspond to positive $\cos(\theta)$ i.e. $\theta = \varepsilon\frac{\pi}{4}$

Therefore, for θ close to $\pm\pi/4$ or to $\pm\pi/4 - \pi$, according to the sign of ε, the left-hand side of (80) will be close to $1 + \sqrt{2}$ which is strictly larger than 1.

Therefore for such a choice of the unit vectors a, b, c, d, random variables $S_a^{(1)}, S_c^{(1)}, S_b^{(2)}, S_d^{(2)}$ as in the statement of the Lemma cannot exist.

Corollary 3 *In the notations of Lemma (1), there exist sets of four vectors a, b, c, d, such that it is not possible to find a probability space* $(\Omega, \mathscr{F}, P)$ *and four random variables*

$$S_a^{(1)} \quad ; \quad S_c^{(1)} \quad ; \quad S_b^{(2)} \quad ; \quad S_d^{(2)} \tag{85}$$

defined on $(\Omega, \mathscr{F}, P)$ *and with values in the interval* $[-1, +1]$*, whose correlations are either of the form (singlet type)* $E(S_x^{(1)} \cdot S_y^{(2)}) = -x \cdot y \qquad ; \qquad x, y = a, b, c, d$ *or of the form (anti–singlet type)* $E(S_x^{(1)} \cdot S_y^{(2)}) = x \cdot y \qquad ; \qquad x, y = a, b, c, d$

Proof. The singlet type would contradict Lemma (1) with

$$\varepsilon_{a,b} = \varepsilon_{c,b} = \varepsilon_{a,d} = \varepsilon_{c,d} = -1 \tag{86}$$

The anti–singlet type would contradict Lemma (1) with

$$\varepsilon_{a,b} = \varepsilon_{c,b} = \varepsilon_{a,d} = \varepsilon_{c,d} = 1 \tag{87}$$

REFERENCES

1. L. Accardi and A. Fedullo, "On the statistical meaning of complex numbers in quantum theory", *Lettere al Nuovo Cimento*, **34**, 161-172 (1982), University of Salerno preprint May (1981).
2. L. Accardi, "Topics in quantum probability", *Phys. Rep.*, **77** pp. 169–192 (1981).
3. L. Accardi, *The probabilistic roots of the quantum mechanical paradoxes*, In: The wave–particle dualism, S. Diner, G. Lochak, and F. Selleri (eds.) Reidel, Dordrecht (1984) pp. 297–330.
4. L. Accardi, *Einstein-Bohr: one all*, in: The interpretation of quantum theory: where do we stand? Acta Enciclopedica, Istituto dell'Enciclopedia Italiana (1994) pp. 95–115, Volterra preprint No. 174 (1993).
5. L. Accardi and M. Regoli, *Locality and Bell's inequality*, in: QP-PQ XIII, Foundations of Probability and Physics, A. Khrennikov (ed.), World Scientific (2002) 1–28 Preprint Volterra, N. 427 (2000) quant-ph/0007005.
6. L. Accardi, K. Imafuku and M. Regoli, "On the physical meaning of the EPR–chameleon experiment", *Infinite dimensional analysis, quantum probability and related topics*, **5**, N. 1, pp. 1–20 (2002), quant-ph/0112067 Volterra Preprint N. 494 December (2001).
7. L. Accardi, K. Imafuku and M. Regoli, *Adaptive dynamical systems and the EPR–chameleon experiment*, in: Proc. Conf. "Foundations of Probability and Physics-2", Vaxjo, June 2-7 (2002), Vaxjo University Press, Ser. Math. Modelling in Phys., Engin., and Cogn. Sc., vol. 5 (2003) pp. 11–36.
8. A. Aspect, P. Grangier and G. Roger, Phys. Rev. Lett. **49**, 91 (1982).
9. P. Ball, *Exorcising Einstein's spooks*, Nature 29 November (2001).
10. J.S. Bell, "Bertlmann's Socks and the Nature of Reality", *Journal de Physique*, **42**, Colloque suppl. C2, pp. 41–61 (1981).
11. J.S. Bell, On the Einstein Podolsky Rosen Paradox", *Physics*, **1** no. 3, pp. 195–200 (1964).
12. G. Brassard, R. Cleve and A. Tapp, *The cost of exactly simulating quantum entanglement with classical communication*, quant-ph/9901035.
13. T. Brody, (Springer 1994): Compilation Ed. by De La Pena and P. Hudson Real and imaginary non localit in: Brody T., De La Pena *Nuovo Cimento*, **58B**, p. 455 (1979), [citato da A. Rizzi].
14. J.F. Clauser, M.A. Horne, A. Shimony and R.A. Holt, *Phys. Rev. Letters*, **49**, pp. 1804–1806 (1969), in: "J. S. Bell, Speakable and unspeakable in quantum mechanics", Cambridge Univ. Press (1987).
15. M. Czachor, "On classical models of spin", *Foundations of Physics Letters*, vol. 5 n. 3 (1992).

16. M. Czachor, "On some class of random variables leading to violations of the Bell inequality", *Physics Letters A*, vol. 129, n. 5 6 (1988).
17. W. De Baere and W. Struyve, *Subquantum nonreproducibility and the complete local description of physical reality*, Preprint (2001).
18. W.M. de Muynck, *Interpretations of quantum mechanics, and interpretations of violation of Bell's inequality*, Preprint 2001.
19. W.M. de Muynck, W. De Baere and H. Martens, "Interpretations of quantum mechanics, joint measurement of incompatible observables, and counterfactual definiteness", *Foundations of Physics*, **24**, (12) pp. 1589–1663 (1994), Discussed in: [32].
20. R.P. Feynmann, R.P. Leighton and M. Sands, *Lectures on Physics, III*, Addison–Wesley (1966).
21. L.E. Szabo and A. Fine, *A local hidden variable theory for the GHZ experiment*, quant-ph/0007102.
22. R.D. Gill, quant-ph/0301059 *Time, Finite Statistics, and Bell's Fifth Position*, Richard D. Gill in: Proc. of "Foundations of Probability and Physics-2", Ser. Math. Modelling in Phys., Engin., and Cogn. Sc., vol. 5/, Växjö Univ. Press (2003) pp. 179–206.
23. N. Gisin and B. Gisin, "A local hidden variable model of quantum correlation exploiting the detection loophole", *Phys. Lett. A*, **260**, pp. 323–327 (1999), Quant-ph/9905018 6 may 1999, version of April, 22 (2001).
24. K. Hess and W. Philipp, "A possible loophole in the Bell's theorem and the problem of decidability between the views of Einstein and Bohr", *PNAS*, **98**, 25, pp. 14228–14233 (2001).
25. K. Hess and W. Philipp, "A possible loophole in the theorem of Bell", *PNAS*, **98**, 25, pp. 14224–14227 (2001).
26. A.Yu. Khrennikov, *Non-Archimedean analysis: quantum paradoxes, dynamical systems and biological models*, Kluwer Acad.Publ., Dordreht (1997).
27. A. Khrennikov, O.G. Smolyanov and A. Truman, *Kolmogorov probability spaces describing Accardi models of quantum correlations*, to appear in: Open Systems and Information Dynamics (2004).
28. T. Miyadera and M. Ohya, *On Chameleon effect and Bell's inequality*, communication at the: International Conference: Quantum Information 2003, mathematical, physical engineering and industrial aspects (ICQI03IIAS), November, 5-7, (2003), to appear.
29. P. Pearle, "Hidden–variable example based upon data rejection", *Physical Review* D vol. 2 (8) pp. 1418–1425 (1970).
30. S. Notarrigo, "A Newtonian separable model which violates Bell's inequality", *Il Nuovo Cimento*, **83B**, (2) pp. 173–187 (1984).
31. M. Steiner, *Towards quantifying non-local information transfer: finite-bit non-locality*, quant-ph/9902014.
32. H.P. Stapp, "Comments on "Interpretation of quantum mechanics, joint measurement od incompatible observables, and counterfactual definitess", *Foundations of Physics*, **24** (12) (1994).
33. A. Tartaglia, "Is the EPR paradox a paradox?", *Eur. J. Phys.*, **19**, p. 307 (1998).
34. G. Weigh, T. Jennewein, C. Simon, H. Weinfurter and A. Zeillinger, "Violation of Bell's inequalitiy under strict Einstein locality conditions", *Phys. Rev. Lett.*, **81**, pp. 5039–5043 (1998).